Letters

Seismic Monitoring in the Soviet Union

I am compelled to clarify issues regarding the U.S. Geological Survey (USGS) raised in R. Jeffrey Smith's article "Soviets agree to broad seismic test" (News & Comment, 1 Aug., p. 511). The impression given in the article, that Jack Evernden recently spoke as a USGS official on behalf of "a USGS proposal" to install seismographic equipment in the Soviet Union, is erroneous and misleading. Evernden, a USGS employee, during a recent private trip to the Soviet Union and at his own initiative, discussed his desire to see seismographic equipment installed in the Soviet Union to pursue his own research. Evernden apparently anticipated support from other agencies for this proposal and its ultimate execution under a scientific exchange agreement on earthquake prediction between the United States and the Soviet Union. It is unfortunate that Evernden represented himself as a USGS official to Soviet officials while on a private visit.

DALLAS L. PECK
Office of the Director,
U.S. Geological Survey,
Department of the Interior,
Reston, VA 22092

It is necessary that your readers be made aware of an erroneous impression given in the recent article "Soviets agree to broad seismic test." This article states that the Defense Advanced Research Projects Agency (DARPA) expressed a willingness to support a proposal to establish a seismic monitoring network inside the Soviet Union which was discussed in Moscow with the Soviet Academy of Sciences in May 1986. We had no knowledge of plans for these unofficial discussions, nor were any prior indications or commitments made to fund the establishment of a monitoring network in the U.S.S.R.

In February 1986, Jack Evernden, a USGS employee, submitted a proposal directly to DARPA for basic research on "High frequency noise measurements and Q determinations in the U.S.S.R. and U.S.A." This proposal called for the temporary deployment of several seismic instruments at a number of sites in the Soviet Union to collect the data necessary for the research. It is DARPA policy that support for projects with foreign countries be only on an approved government-to-government basis. This would be especially true of a project involving the Soviet Union. It was determined that this proposal was inappropriate since the established agreement between the USGS and the Soviet Union Institute of Physics of the Earth involves cooperation in earthquake prediction research, not test ban monitoring research. We therefore could give this effort no further consideration. It was totally inappropriate for Evernden to convey a willingness of our support in his subsequent private meetings with officials of the Soviet Academy of Sciences.

ROBERT C. DUNCAN
Defense Advanced Research Projects Agency,
1400 Wilson Boulevard,
Arlington, VA 22209-2308

R. Jeffrey Smith incorrectly suggests that the Department of Energy supports the Evernden proposal and that the Department has "expressed a willingness to support the proposal with appropriate funds and equipment." While individuals associated with the Department or with the National Laboratories may very well have discussed the Evernden proposal with some of its sponsors, they were not expressing the Department's view.

The Department of Energy certainly supports obtaining more seismic data about the Soviet Union; but the fact remains that the Evernden proposal is being promoted not so much for its scientific merits as for its role as a potential step toward a comprehensive nuclear test ban. The Administration has stated that a test ban or moratorium is not now in the national security interests of this country, and the Department of Energy would not support or fund any effort that is contrary to Administration policy.

ANSON FRANKLIN
Office of Communications,
Department of Energy,
Washington, DC 20585

Response: My remembrance of events is, of course, quite different from the history sketched in these three letters, but I believe there to be no purpose in using the pages of *Science* for what might well degenerate into endless quibbling about the details of historical fact and fancy. I will only say that I feel Smith's article to be an accurate report and that every act of mine, whether or not within the context implied by the three letters, had only one purpose, that purpose being to serve in the best way I know how the U.S. government and the American people.

JACK EVERNDEN
Post Office Box 174,
Davenport, CA 95017

Response: Before the publication of my article, I spoke with a DARPA expert who assured me of the agency's enthusiasm for the Evernden proposal, and of its previous, verbal commitment to contribute funds if the Soviets accepted it. Similar expressions of enthusiasm were made by persons employed at or affiliated with DOE.

I also deliberately raised the issue of USGS involvement with one of Evernden's superiors at the agency's headquarters in Reston, Virginia. The official specifically said that it would be correct to describe the Evernden plan as an "informal USGS proposal," and this is exactly how it was described in my article.—R. JEFFREY SMITH

Human Genome Sequencing

Roger Lewin (Research News, 8 Aug., p. 620) summarized a recent meeting organized by the Howard Hughes Medical Institute at which the proposal to "sequence the human genome" was discussed. During this meeting sentiment seemed to shift away from "sequencing" toward "mapping" the genome for a variety of reasons, some political, some technical. Sequencing was seen as too expensive ($3-billion estimate), likely to divert funds from other worthy projects, likely to give the Department of Energy too much control, and better delayed a few years until it could be done more efficiently. These are important considerations, but they assume that sequencing the entire human genome is a worthwhile project. There is a fundamental reason for doubting this assumption: most of the DNA in the human genome does not code for proteins and may have no sequence-dependent function at all, or at least none that will be revealed by random sequencing. The evidence comes from a variety of experimental and theoretical considerations.

The human genome contains 3.5 picograms of DNA or about 3×10^9 base pairs. Nearly every fragment that has been sequenced contains some noncoding regions; the question is simply how much of the total is noncoding? It is instructive to begin with a theoretical calculation of how much DNA is needed to code for a reasonable number of proteins. Let us assume that there are 20,000 to 30,000 different proteins in the human body with an average molecular weight of 70,000 daltons. These require only 5×10^7 bases out of the total of 3×10^9 or about 2%. Even using a rather unlikely assumption of 100,000 different proteins, one comes up with a figure of no more than 10%.

Because something is theoretically possible doesn't mean it is true or even likely. However, we know that some complex organisms get by with very small amounts of

DNA. The metazoan with the smallest known genome is the worm *Caenorhabditis*, which has 8×10^7 base pairs; another small genome is that of the fly *Drosophila*, with 1.7×10^8 base pairs. These examples show that a complex organism *can* be constructed with only 3 to 6% of the DNA found in humans. Interestingly, even in these organisms a considerable fraction of the DNA does not code for proteins.

At this point one might be tempted to say flies and worms are one thing, but obviously humans have more DNA because of their greater complexity. In fact, among eukaryotic organisms genomic DNA content has virtually nothing to do with complexity. Here we must deal briefly with what is known in the chromosome field as the C-value paradox (C-value is another name for genome size). The paradox has two parts (*1*). The first is that organisms of similar morphological complexity for evolutionary relatedness often have vastly different C-values; the second is that most eukaryotes have much higher C-values than can be accounted for by protein-coding needs, humans being merely one of many examples. Among vertebrates the highest C-values belong to some salamanders, which have about 30 times as much DNA as humans; surely salamanders are not 30 times more complex than humans! On the other hand, some fish manage with a genome less than a third that of humans. Among invertebrates one finds the same puzzling variation. As already noted, *Drosophila* has a small genome, but this is not because it is an insect; some grasshoppers have two to three times as much DNA as humans. Plants have a similar range of values, again not related to evolutionary or morphological criteria (lettuce has much less DNA than humans, but corn and lilies much more). In summary, therefore, the human genome is large relative to its protein-coding needs, but as genomes go, it is neither very small nor very large.

Over the past 20 years the question of genome organization, including the C-value paradox, has attracted enormous attention, both experimentally and theoretically. Perhaps the most important generalization is that variations in genome size are *not* due to variations in the reiteration frequency of protein-coding genes. Thus, the idea that organisms with high C-values have many copies of each gene, whereas those with low C-values have only one or a few, is certainly false. There *are* a great many reiterated sequences in organisms with high genome contents, but few of these code for protein.

Where is the noncoding DNA? Most of it is in "spacer" regions between genes, although a minor and variable amount is within genes as introns. Neither the spacers nor introns (with a few exceptions) code for proteins, and there is no evidence that their specific sequences are important, as opposed to their length, position, secondary structure, or some other feature. If one wants to argue that we should sequence 90 to 98% of the human genome in hopes of discovering some new sequence-dependent function of introns and spacers, the answer is simply that that is bad science. There are already plenty of such sequences stored in computers, and if one wanted another million or so bases for analysis, they could be had cheaply without sequencing the whole genome. Until the spacer and intron DNA's are shown to have some sequence-dependent role, there is no intellectual justification for sequencing them at random.

Although I strongly oppose the sequencing project in its simplistic version, I do believe that knowledge about the human genome is intrinsically interesting and certain to be of medical value; furthermore, we have the techniques and an adequate theoretical framework to justify greater effort in this area. I believe we should proceed simultaneously along two lines. First, mapping studies could begin, using as a guide what Alan Coulson and John Sulston have already accomplished with the worm *Caenorhabditis* (*2*). Even this task will be heroic, since the human genome is 40 times larger than the worm's (and mapping requires all of the genome, coding and noncoding). Preliminary chromosome sorting would reduce the problem a great deal. Second, individual investigators should continue to sequence whatever genes appear to be of greatest interest. If a larger scale project is undertaken, then it should begin with complementary DNA (cDNA) clones. In these clones most of the DNA codes for protein and therefore is currently interpretable. Furthermore, the cDNA clones could be matched to their appropriate places on the physical map by nucleic acid hybridization. If it seemed valuable one could then sequence the genomic regions corresponding to the cDNA's. An enormous advantage of this approach is that one would already know the limits of each gene as well as the correct reading frame, information that is difficult to extract from raw and inevitably inaccurate sequence data in an uncharted region of the genome.

The mapping and cDNA sequencing would be expensive. After the initial strategy was worked out in detail, the intellectual challenges might not seem so alluring. Thus the work might well require some kind of contractual or programmatic aspect outside the usual investigator-initiated grant system. However it may be organized, my plea is simply that we think about this project in light of what we already know about eukaryotic genomes and not set in motion a scientifically ill-advised Juggernaut.

JOSEPH G. GALL
Department of Embryology, Carnegie Institution of Washington, Baltimore, MD 21210

REFERENCES

1. J. G. Gall, *J. Cell Biol.* **91**, 3s (1981); T. Cavalier-Smith, Ed., *The Evolution of Genome Size* (Wiley, New York, 1985).
2. A. Coulson, J. Sulston, S. Brenner, J. Karn, *Proc. Natl. Acad. Sci. U.S.A.*, in press.

Underground Storage Tanks

One critical point about the recent briefing by Marjorie Sun "EPA grapples with regulating underground storage tanks" (News & Comment, 1 Aug., p. 518) should be clarified.

The Environmental Protection Agency (EPA) national survey on underground petroleum storage tanks, released on 24 June, makes no estimate and draws no conclusion about the amount of the nation's ground water, including drinking water, that may be at risk from tank leaks.

The EPA survey specifically emphasized that while it found 35% of the tanks tested failed a tank tightness test, this does *not* indicate those tanks are leaking under normal operating conditions.

A tightness test is a screening mechanism. It must be followed with corroborating testing procedures to avoid confusing leaks with other factors that could cause a test failure—often loose fittings or worn gaskets at or above the top of an underground tank.

At least three major oil companies that have used the same type of underground tank tightness test as the EPA found, when they completed follow-up testing procedures, that actual leak rates were vastly lower than test failure rates. Leak rates for these three companies ranged from 0.97 to 2.6%, whereas original tightness test failures ranged from 10 to 19%. The EPA, unfortunately, did not undertake any follow-up procedures.

Even when leaks occur, they typically are detected and corrected before ground water is affected and usually are confined to the property of the tank system owner. The data from the member companies of the American Petroleum Institute make it clear that the vast majority of their tanks—95% or more—are not leaking and do not represent a major threat to drinking water.

WILLIAM F. O'KEEFE
American Petroleum Institute, 1220 L Street, NW, Washington, DC 20005