

The DNA data libraries

SIR—A leading article in your journal recently commented on the sequencing of the human genome (*Nature* 331, 465; 1988) after a report from the US National Research Council on this topic. First, *Nature* recommends that this project should be an informal part of the scientific endeavour without any coordination and in the next column the author paradoxically suggests that an internationally coordinated mega-databank should be introduced. More careful research would have revealed that the National Institutes of Health (NIH), the European Molecular Biology Laboratory (EMBL) and the Japanese funding agencies have taken several steps towards this goal. They have also recognized the inevitable obsolescence of the current databases and, with this in mind, began research into novel approaches with an international workshop as long ago as February 1987.

A unified nucleic acid sequence database may be available within 2–3 years, based mainly on data transmitted directly by researchers and integrated with other data collections such as protein data, physical and genetic gene maps and protein 3-D structures. This collaboration involves Europe, the United States and Japan and 50 per cent of the submissions already come from the originators of the sequence information, that is, the scientists themselves. Three days after your leading article was published, the international advisory board of the databanks, comprising three experts from both the United States and Europe and two from Japan, met in Bethesda, Maryland, to discuss priorities for future tasks and to review collaborative projects between the databanks. The recent collaboration in a number of areas, including the adoption of a common data entry form and the successful launch of a direct submission system, was recognized and praised. The advisory panel finally strongly recommended that the outstanding problem in unifying some of the more crucial aspects of annotation should be solved in the next 6–9 months. Integration with other kinds of databases should also be pursued as rapidly as possible.

The procedure introduced on 1 January 1988 in a collaboration between the journal *Nucleic Acids Research* and the EMBL data library requests authors to submit sequences to the EMBL data library in computer-readable form, before journal submission and review. This will not only speed up the entry of sequences into the banks but will also provide an entirely new type of publication. It is desirable that this approach be accepted by all journals in the field. The advisory board therefore expressed a strong recommendation that the scientists should be

primarily responsible for the data presented, with any information added by the databank or other experts being clearly separated from the original submission.

Additional resources have recently been made available to the data libraries. Genbank has received a \$17 million contract from NIH for the next 5 years and EMBL has, through the establishment of a science-led biocomputing programme, provided the hardware and software base. EMBL has also requested funds from the European Community and the European pharmaceutical and chemical industry for establishing a European network with regional nodes to coordinate the anticipated traffic of nucleic acid sequences.

It suffices to conclude that the anomaly with the databanks is on its way to being rectified and that we have already come a long way towards taking care of the data from a human genome project if launched. We are close to a joint international data source in biology. International coordination in science is more dependent on hard undisturbed work than on bombastic and uninformed statements. *Nature's* journalists would profit from doing their homework before formulating their regretfully influential and sweeping statements.

LENNART PHILIPSON

*European Molecular Biology Laboratory,
Postfach 10.2209,
Meyerhofstrasse 1,
6900 Heidelberg, FRG*

Phd by thesis

SIR—Beverly Halstead's swingeing attack (*Nature* 331, 497; 1988) on established practice may focus attention on an important issue, but he goes too far in equating the doctoral degree with publications. Is he seeking to fit untidy researchers into a neat package of academic measurability?

A firm distinction must be made between (1) aptitude for research (the PhD degree) and (2) an end-product of research (publications). There are numerous reasons why the results of much painstaking research fail to appear in print. Negative results are notoriously difficult to publish and poor use of English may be a bar to publication while casting no light on research aptitude. (And publishing houses are more amenable to economic and political pressures than are degree-awarding bodies.)

But approaches of novelty are often suppressed purely for their originality. Although most students may fit into Halstead's box, what of those individuals whose radical thinking becomes an inspiration to us all? It can be quite acceptable to award a PhD degree even if one suspects that the conclusion of the disserta-

tion is wrong: scholarly disagreement alone does not constitute grounds for refusal. The opposite is true for publication.

In fairness to Halstead's argument, if good publications are indeed an adequate expression of research aptitude, a dissertation might well consist of a group of publications bound together with a general introduction and conclusion (raising, of course, the question of what precisely constitutes a publication). Unfortunately, it is often all too easy for the mediocre student in a dynamic research group to accumulate a series of publications.

A publication record can be neither a necessary nor a sufficient condition for the award of the PhD degree. May it not be that the tried and tested method of thesis submission and review, in good faith, by a panel of established researchers is so far the only adequate procedure for assessing aptitude for research?

RICHARD LATHE

*CNRS-LGME & U184-INSERM,
11 Rue Humann,
67085 Strasbourg Cedex,
France*

SIR—The PhD degree has three main variables, the student, the supervisor and the project. It is the supervisor who can influence the current trend in "ever-expanding monsters". The writing of a PhD thesis should be an enjoyable experience if the student receives guidance on its proper function and preparation.

T.K. BRADSHAW

*Sittingbourne Research Centre,
Sittingbourne, Kent ME9 8AG, UK*

US budget deficit

SIR—Your advice ("Transatlantic tax dilemmas", *Nature* 331, 643; 1988) to increase US income tax implicitly assumes that the US Congress will use the expected revenue increase to reduce the deficit. Nothing could be further from the truth. There are numerous proposals now before that august body to appropriate all kinds of goodies for special interest groups, ranging from free day-care centres down to funding of airports for billionaire playboys. Any additional tax receipts will go toward those expenditures, not to reducing the deficit. A year or so ago, a study was conducted to determine the experience from past 'solutions' of the proposed type. It was found that the deficits did not go down; they went up. The truth is that the current US Congress is a deeply flawed institution run by professional politicians interested in only one thing: reelection. In the words of the immortal Vince Lombardi, "Winning isn't everything; it is the only thing!"

ROBERT C. WHITTEN

*1117 Yorkshire Drive,
Cupertino, California 95014, USA*