

Bioinformatics: smartest software is still just a tool

The proliferation of ever-more-sophisticated biological software can lead to misuse.

Sir — A recent Letter to *Nature* by R. D. King and colleagues, describing a ‘robot scientist’ (*Nature* 427, 247–252; 2004), gives us pause for reflection not only on our own humanity, but also on the increasing use of software in biology. A quick scan of publicly available Internet databases provides some idea of the hundreds of software programs, and their manifest applications, available for biology. Unfortunately, this very prevalence of software in biology may be creating a problem.

My own observations of colleagues and students suggest that the difficulties they experience using biological software could lead to misuse. They often only use the ‘default’ parameters, either because they don’t know how to adjust the parameters to refine their analysis, or because they are concerned that such changes may affect their results. This gap in knowledge and skill is understandable, given that it is difficult enough to keep up with advances

in biological science — never mind all the details of computer software. Moreover, there is increasing pressure on the biologist to manage, and analyse, a deluge of biological data.

Unfortunately, it seems that, while relying on software as a biological tool, we are not giving it the same careful, and critical, consideration as other tools we use. In particular, we seem to be selectively ignorant, not only of the need to calibrate, characterize and standardize software, but also of how software works.

We routinely calibrate, characterize and standardize our normal laboratory tools (such as pipettes, and chemical or biological assays) and understand how they work. We do this so that we can be confident in the results of the experiment, and our interpretation of the results.

Why should software be any different? The complexity of the software with its millions of lines of code, the underlying algorithm and the mathematics may be too

daunting for most biologists to fully understand. But we should not be so impressed by the perceived efficacy of software that we are willing to forgo the strict scientific procedures and quality control that we apply to other laboratory tools.

The promise of software in biology is exciting and could lead to greater understanding of biological systems. Informatics is undoubtedly changing how biological science is being done. Consequently, biological software is increasing in complexity and sophistication, which is all the more reason to be stringent in its application and in understanding its limitations. Software is a useful and valuable tool, but it should be treated no differently from any other tool in biological science.

Steven Bottomley

*School of Biomedical Sciences,
Curtin University of Technology, Kent Street,
Bentley, Western Australia 6102, Australia*

Bioinformatics needs a software archive

Sir — In February 2003 the US National Institutes of Health (NIH) released a statement on its data-sharing policy, in which it strengthened its commitment to free exchange of final research data produced by public funds (see *Nature* 421, 877–878; 2003). Little attention has been given, however, to a related question — sharing and protecting the products (software and tools) of bioinformatics research, especially infrastructure generated to support large projects.

As projects and databases evolve, some will inevitably lose funding or be shut down. What happens to the bioinformatics software? In one such case, the Genome Database, an international collaboration set up in support of the Human Genome Project, lost US federal funding in 1998, although it was later reopened with private funding. During the shut-down process, a significant portion of the source code (though not data) was irreparably lost. This database represented an investment of more than US\$50 million by US, European and Asian sources. The loss of the code, a public asset, occurred because there was little supervision during decommissioning.

As a minimum safeguard, we propose the creation of a Bioinformatics Software Archive, in which an archival copy of

bioinformatics software would be maintained in a secure central repository supported by public funding. There are moves in the bioinformatics community to adopt open-source standards for software, which provide a means of online archiving, but these are not suitable for every type of software. As with software released under open-source agreements, a central archive will help stop too many researchers trying to reinvent the wheel (thereby saving research funds).

If funding agencies such as the NIH are serious about data sharing, they should be serious about protecting the software used to produce those data.

A. Jamie Cuticchia, Gregg W. Silk

*Genome Database, RTI International, Research
Triangle Park, North Carolina 27709, USA*

Autistic geniuses? We’re too ready to pathologize

Sir — While I am in sympathy with the thrust of Allan Snyder’s thoughts on ‘autistic genius’ (*Nature* 428, 470–471; 2004), and the book by Michael Fitzgerald, *Autism and Creativity*, which he reviews, I think there is a danger of going overboard on the subject of pathology and creativity. Snyder cites me as saying that the philosopher Ludwig Wittgenstein had autistic traits, but I have never thought or suggested this. Snyder’s misapprehension

arose from a conversation that I had with the autistic scientist Temple Grandin — reported in my book *An Anthropologist on Mars* (Knopf, New York, 1995) — about another researcher’s suggestion that Wittgenstein may have been autistic. Grandin erroneously attributed this opinion to me in her otherwise excellent book, *Thinking in Pictures* (Doubleday, New York, 1995).

I think that pathologizing genius, and pathologizing historical figures, has become an obsession with us; and also that the concepts of ‘autism’ (and ‘Tourette’s syndrome’, etc.) have become so distended that they are vastly overused. It seems to me extremely unlikely, from the evidence we have, that Wittgenstein or Einstein or Newton were significantly autistic.

On the other hand, there is strong historical evidence to suggest that the eighteenth-century chemist Henry Cavendish was autistic (see O. W. Sacks, *Neurology* 57, 1347; 2001). Unlike most other supposed ‘autistic geniuses’, he showed a near-total incomprehension of common human behaviours, social relationships, states of mind, and money, as well as an almost obsessed attention to detail — which led to the great generalizations he was later to erect. But it is important to sift through the evidence with great care before diagnosing or pathologizing.

Oliver Sacks

2 Horatio Street, New York, New York 10014, USA