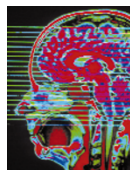




Hard cheese
Final sequencing of the mouse genome 'could be delayed'
p602



Getting ahead
Gateway to brain data promised for the Internet
p603



Hard choice
Spain faces dilemma over whether to join French synchrotron
p604



Ruled out
Tough new regulations for xenotransplants
p606

Biologists challenge sequencers on parasite genome publication

Colin Macilwain, Washington

Rival approaches to the way in which genome sequences are published are creating growing tension between scientists at the sequencing centres and those who want to use the sequencing data to further their study of the organisms involved.

International consortia that are sequencing *Plasmodium falciparum*, a parasite that causes malaria, and *Trypanosoma brucei* (see right), which causes sleeping sickness, are each currently engaged in heated arguments over the wisdom of publishing preliminary, annotated sequences in advance of the completion of full sequences.

In each case, prominent biologists who specialize in the organism are pushing for early publication. But genome sequencing centres say the publication of preliminary data could deprive their teams of the proper credit for the full, annotated sequence when it is completed.

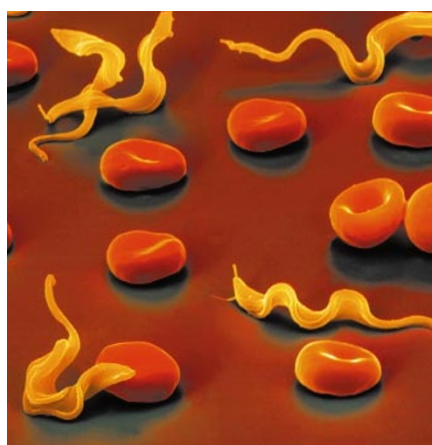
Claire Fraser, president of The Institute for Genomic Research (TIGR) at Rockville, Maryland, which is part of both consortia, believes that if outside scientists publish preliminary annotations (proposed function of a stretch of DNA) based on raw sequencing data made available voluntarily by the sequencing centres, the final complete sequence may never be published.

"It would be unlikely that the sequencing groups would come up with enough data, over and above those already published, to warrant publication," she says. "So it becomes not just an issue of credit for publication, but an issue of data piracy."

The problem Fraser describes is not confined to teams working on parasites involved in tropical diseases. Dozens of projects are under way to sequence the full genomes of interesting organisms, and most expect to take between three and five years to sequence, annotate and publish the complete genome.

The issue is common to most genome projects, says Michael Gottlieb, programme officer for parasitology at the National Institute of Allergies and Infectious Diseases, which is supporting work in both consortia.

"Our concern is to achieve an appropriate



Deadly sequence: *Trypanosoma brucei*, the protozoan responsible for sleeping sickness.

balance between the community's interest in getting the information out as soon as possible, and the sequencer's interest in getting the first opportunity to publish the annotated sequence," he says.

The small communities of researchers specializing in these organisms want information about the organisms' genes as quickly as possible. And although the sequencing centres usually release the raw sequence data

on the Internet as they obtain them, these data are of limited use to microbiologists who lack the bioinformatics capacity needed to interpret them. Some are teaming up with bioinformaticists to partially annotate the public data themselves. The trouble starts when they try to disseminate these partial annotations to their colleagues.

The issue came up last November at the regular meeting of the *T. brucei* consortium, which rejected a request from George Cross of Rockefeller University in New York to publish an annotation of part of the genome. "We're not allowed to make the knowledge we have developed available to anyone else in the field," complains Cross.

"There is a tremendous and increasing tension" between the sequencing centres and the microbiologists, he claims. "The key issue is that the sequencing centres want the sequence to be 99.9 per cent complete, but most biologists can get tremendous input from interim data."

The same issue was set to arise in Britain this week at a meeting of the malaria genome consortium at the Sanger Centre in Cambridge. Lou Miller, a microbiologist at the National Institutes of Health in Bethesda, Maryland, was expected to ask to publish a

New Russian science head named

Carl Levitin, Moscow

Ilya Klebanov, a deputy prime minister with close links to Russia's military industries and its arms-exporting companies, has been appointed the country's new head of industrial and scientific policy. As such, Klebanov will control the new Ministry for Industry, Science and Technologies, which 'swallowed' the former Ministry of Science and Technologies.

Former science minister Mikhail Kirpichnikov says the decision to found the super-ministry (see *Nature* 405, 384; 2000) was taken "virtually at the moment of signing the presidential decree on the new

cabinet", and neither he, Yuri Osipov — the president of the Russian Academy of Sciences — nor Klebanov were consulted.

Kirpichnikov has expressed his concern to prime minister Mikhail Kasyanov about whether the new cabinet structure is the best way of managing Russian science. He says he was told this was a chance to unite scientific research 'scattered' over many programmes.

Osipov says Aleksandr Dondukov, the new minister of industry, science and technology, has promised to consult the academy before taking decisions on basic science, and hopes to build a powerful 'scientific unit' within the ministry.

preliminary annotation of the *P. falciparum* sequence he has prepared with Eugene Koonin of the National Centre for Biotechnology Information.

Miller's proposal was met with indignation at the sequencing centres when he approached them in March. "What he's done has clearly gone against the interests of the funding agencies and sequence centres," says Fraser. "The sequencing groups felt they were being held hostage," she says, because Miller was offering to publish in collaboration with them, "but saying he would do it anyway".

Fraser says Miller backed off once the sequencing groups' hostility became apparent. Miller declines to speak on the record about his plans. He says public discussion is unlikely to improve relations between the parties involved. However, it is understood that he has no plans to publish his annotation without the consent of the consortium.

"He's not doing it to get the credit," says Malcolm Gardner, who heads the *Plasmodium* sequencing effort at TIGR. "He simply believes that it is in the best interests of the community that the information gets out there." However, Gardner adds, "people who have invested four years of work in this should have the privilege of publishing it."



Malaria: lethal in children.

David Roos, a microbiologist at the University of Pennsylvania, Philadelphia, has obtained funds from the Burroughs Wellcome Fund to present a preliminary annotation of the malaria parasite sequence on the Internet. But he says he encountered some hostility when he first became involved with the consortium two years ago.

"I ran into anxiety, sometimes bordering on paranoia, at the sequencing centres," he recalls. Roos published some important genes from the parasite in the *Proceedings of the National Academy of Sciences*. "At the time it caused a tremendous uproar. There was anxiety that we'd skimmed the cream from the project — but in fact what we did added value to what the sequencing centres do."

Roos thinks Miller's work should be included in the Internet portal his team is creating to make the malaria parasite sequence accessible to microbiologists. A first version of the portal will go live this week, at [http://](http://e2kroos.cis.upenn.edu/PlasmoDB.html)

e2kroos.cis.upenn.edu/PlasmoDB.html.

Some researchers, however, will continue to press for publication of preliminary annotations of organism genomes. Most — but perhaps not all — will adhere to the data-release policies, posted beside the preliminary sequence data, which tell researchers to obtain permission before publishing results based on the data.

Separate genome projects will consult leading scientific journals about whether the release of preliminary annotation on the web will prejudice subsequent publication of complete, annotated chromosomes.

But ultimately these tensions could change the way biology is published. Roos points out that high-energy physicists, who publish in teams of hundreds, are named on papers in alphabetical order, with no lead author. "My prediction is that the same will happen in genomics," he says.

Roos sits on a couple of departmental appointment boards at his university, and jokes that they'll have to start assessing candidates by their true contribution, rather than inferring it from the order in which their name has appeared on papers. "We'd need to think about what they've done," he says. "And that wouldn't be such a bad thing." ■

Drive for more genomes threatens mouse sequence

Alison Abbott

Mouse geneticists are expressing concern that completion of the mouse genome sequence could be delayed because of increasing pressure to sequence the genomes of other animals. The doubts were voiced at a recent workshop organized by the National Institutes of Health (NIH) to set priorities for resources for mouse genetics and genomics.

The mouse was selected two years ago as the second mammalian species, after humans, to have its genome sequenced both by a public consortium led by the NIH and by the private company Celera. This is because the mouse is widely considered to be the best-understood species genetically.

The number of mutant mouse strains is growing rapidly. These are being generated using technologies that 'knock out' targeted genes and by random chemical-mutagenesis screens, where mice are exposed to a mutagen and those displaying interesting new traits are selected for breeding. Many of these mutant strains are potentially important models for human disease, and accurate information about their gene sequences will help to unravel the molecular bases of these diseases.

Within the past year, the NIH has set up 10 centres to sequence the mouse genome, and plans to divert capacity in human genome sequencing centres over to the mouse as the human genome nears

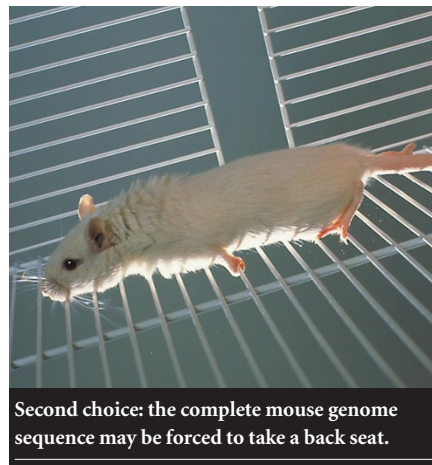
completion. But many other species, from the zebrafish to the rat, are also vying for the attention of sequencers.

"We always understood that the mouse genome would be finished, but learnt only at the workshop that a firm decision had been taken only for a working draft," says Rudi Balling, director of the Institute for Mammalian Genetics at the National Research Centre for Environment and Health in Munich.

So far, the mouse sequence is more advanced in the private than in the public domain. Last week Celera announced that after only two months it has sequenced more than a billion base pairs of the 129/SvJ mouse strain, the strain most often used to generate knockouts. This is estimated to be around one-third of the full genome.

Celera's president, Craig Venter, says the company will have the finished, assembled sequence by next summer. Correct assembly of the vast number of base pairs will be relatively easy, he says, as it can be directed by the blueprint of the human genome. This 'humanized mouse genome', as Venter calls it, will help to pinpoint genes on the human genome and help to identify gene-regulatory areas, he says.

But Celera's mouse genome sequence will be restricted to the company's subscribers. The public consortium, with its slower but



Second choice: the complete mouse genome sequence may be forced to take a back seat.

more exact approach, is aiming for a 'working draft' of the genome of the C57BL/6 mouse strain — widely used in genetic and immunological studies — well before its original target date of 2003, says Elke Jordan, deputy director of the NIH's National Human Genome Research Institute.

She declines to put a date on finishing, saying that in principle the NIH's commitment to finish "remains unchanged: it's what everybody wants". But she says there are "too many uncertainties about funding" for a date to be fixed.

This view is echoed by Eric Lander, director of the Whitehead Institute at the