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:// 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

Massachusetts Institute of Technology. "There is tremendous enthusiasm from all quarters to finish the mouse genome, and scientists won't have to wait very long," he says.

But he admits that "the importance of a lot of different projects will have to be juggled, and it'll all get done in an order that is maximally useful to scientists as a whole". Like the NIH, Germany and France are waiting to assess the scientific promise of different species before committing their national sequencing efforts to finishing the mouse.

Jordan points out that the public consortium, concurrent with providing a working draft, will finish selected areas of the mouse genome judged to be biologically important. Although primarily intended to support priority areas of mouse genetics, this "will also give us a better idea of how much better it really [would be] to have the finished sequence", she says.

Mouse geneticists need no such convincing. "A draft sequence will be a poor tool for mouse geneticists," says Phil Avner, head of the mouse molecular genetics unit at the Pasteur Institute in Paris. It will allow no more than 80 per cent of genes to be identified, and will not provide key information about areas of the genome that regulate the genes, he says.

Even some zebrafish geneticists agree that finishing the mouse genome is a high priority because of the uniquely advanced level of genetic tools available. "But zebrafish biologists will also need a working draft within the next year or two," says Wolfgang Driever, a zebrafish geneticist at the University of Freiburg. "So ultimately it will be a question of timing and funding." ■

Internet gateway planned for neuroinformatics data

Paul Smaglik, Washington

Data on the human brain will soon be available over the Internet via an electronic gateway. Currently being planned by an international consortium, the portal will give researchers access to data at various levels of detail and sophistication.

Backed in principle by the Organisation for Economic Cooperation and Development (OECD), the consortium aims to devise common standards and procedures for the various neuroinformatics databases and software programs scattered around the world.

An OECD working group set up to develop the gateway met for the first time in Genoa last month, and announced that it plans to make decisions on a launch date at its next meeting in Washington in September.

"We'd like to create a portal that would provide access to all the resources that are out there," says Stephen Koslow, director of the Office on Neuroinformatics with the US National Institute of Mental Health, and chair of the OECD working group.

Koslow points out that many individual programs already exist, each working at their own levels of analysis. Finding them and linking them to one site will be the first step, he explains. The next step will be to make them work with each other.

The project resembles earlier efforts to built computational tools for studying genomes. Many groups independently creat-
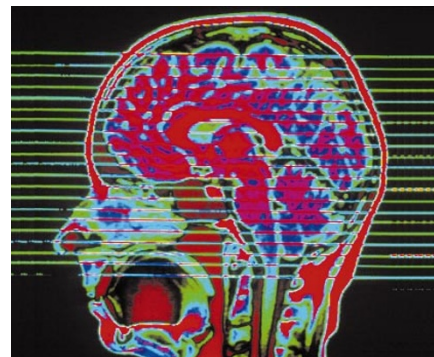
Head first: brain data, such as this MRI scan, will soon be brought together on the Internet.

ed computer programs that eventually became consolidated into software suites.

But neuroinformatics covers a broader spectrum of tools than computational genomics, ranging from brain imaging to the study of genes and proteins. And tying together data dealing with different levels of analysis is a huge challenge. Data can be in many forms, including brain scans showing the development of Alzheimer's disease and genetic databases detailing susceptibility to the disease.

Scientists involved in the Human Brain Project — a US government-backed effort to develop a variety of neuroinformatics tools — are already working to make the various pieces of software more accessible.

Jonathan Cohen, professor of psychology at Princeton University, said last week that although independence in the early days of neuroinformatics software development had resulted in many useful programs, it had also created an electronic 'Tower of Babel'.

"Since a lot of people build these things on their own, they are not always in a format that everyone can use," Cohen told a conference on the Human Brain Project held last week in Bethesda, Maryland.

Cohen's group is providing 'wrappers' to make different programs look the same on a computer screen. So far, this 'FisWidgets' project has provided interfaces for 43 public-domain neuroimaging programs.

Another problem is that data can be in different forms, or images in different resolutions. This is especially daunting when developing multiple-scale models of the nervous system, Nigel Goddard, a bioinformatician at the University of Edinburgh, told the meeting. Such modelling provides a broad map of a system, along with the ability to zoom in on selected parts for greater detail, and retrieve data associated with that image. "We need to have some set of standards so we can integrate the efforts," said Goddard. ■

Software spend boosts Israeli R&D

Haim Watzman, Jerusalem

A recalculation of Israel's spending on research and development (R&D) shows that it spends 3.5 per cent of its gross domestic product (GDP) on civilian research — more than any other OECD country. The figures are the first in Israel to include accurate data on spending by small software companies in Israel.

The new picture comes from a survey carried out by Israel's Central Bureau of Statistics. Previous assessments had placed this figure at just under 3 per cent of GDP (military R&D figures remain classified). Such estimates for high-tech R&D spending were based primarily on information from large companies, explains Simcha Bar-Eliezer, the statistician responsible for the survey.

Shlomo Herskovic, the director of planning and information for the Planning and Budgeting Committee of the Council

for Higher Education, which oversees and allocates the country's higher-education budget, says the new figures change the overall picture of Israel's spending on research.

The survey shows that 21 per cent of the country's civilian research spending is at the universities. Herskovic says the data could be used to argue that the universities should get more research funds. Given the extra spending in software companies, the government is now responsible for only 8 per cent of civilian research expenditure.

The high level of research spending in the software sector could also have implications for computer science programmes at universities. Even current plans to increase the number of computer science graduates by a factor of five over the next three years may be inadequate to supply software firms with the professionals they need, says Herskovic. ■