PAGE 1209
Ancient worm tracks, or merely cracks?

1211
What really transpired in Copenhagen

DATA SHARING

DNA Sequencer Protests Being Scooped With His Own Data

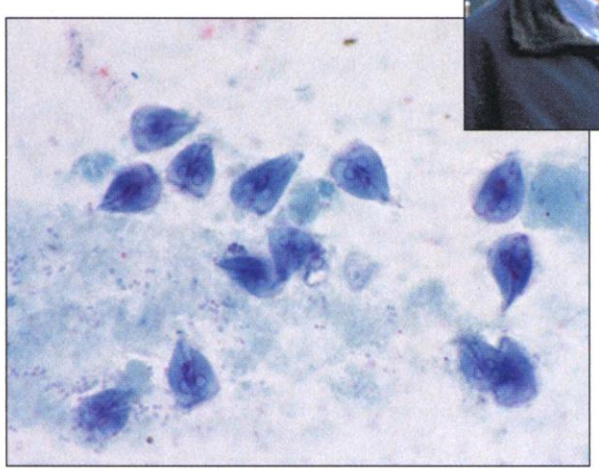
Following what he calls an "egregious" violation of scientific etiquette, a researcher has shut down a public Web site containing his team's raw sequence data for *Giardia lamblia*, a diarrhea-causing protozoan. Mitchell Sogin of the Marine Biological Laboratory (MBL) in Woods Hole, Massachusetts, says he blocked public access 2 weeks ago to this federally supported DNA Web site after discovering that a colleague had published a paper using MBL's sequence information. He intends to restore the site (www.mbl.edu/Giardia) after the rules have been clarified and the data edited.

The dispute is the latest in a string of clashes between those who collect and those who interpret data, and it brings into focus some questions that have been festering in the genome community. Among them: How much control should DNA sequencers wield over the data they gather? And should they be forced to share preliminary results—as many are now required to do—before they publish their own analysis?

In 1997, MBL and a group of collaborators* won a 5-year, $2.6 million award to sequence *Giardia*. The main sponsor, the National Institute of Allergy and Infectious Diseases (NIAID) in Bethesda, Maryland, requires that data be made public as the work progresses. MBL complied by routinely releasing unassembled, raw DNA sequence on its Web site. But MBL intended that its own team should be first to publish an analysis of *Giardia*'s genome. They also asked that others use MBL's preliminary results only to develop reagents or explore mutually agreed-upon projects. These restrictions, Sogin says, "were posted prominently on our Web site."

"I was enraged," Sogin says, when he learned of a paper by Hyman Hartman and Alexei Fedorov in the online early edition of the 5 February *Proceedings of the National Academy of Sciences* (*PNAS*). He claims that the authors used MBL's genome data to publish conclusions about *Giardia*'s evolution. That is precisely what Sogin and his collaborative group had intended to do under their NIAID grant. "This was the intellectual rationale for the *Giardia* grant," Sogin says.

Hartman, an affiliate of the biology department at the Massachusetts Institute of Technology, has been researching the evolution of complex organisms (eukaryotes) since 1984. He denies that he violated MBL's conditions; he was taken completely by surprise by the "huge controversy." He adds that he has known Sogin "for 20 years," although he thinks Sogin "didn't take my work seriously"—at least, not until last month.



Burned. Mitchell Sogin stopped releasing *Giardia lamblia* genome data after another researcher used the information for a paper.

In their *PNAS* paper, Hartman and Fedorov, a Harvard University biologist, argue that eukaryotes inherited some of their most interesting cell structures from a predecessor organism no longer found in nature. They call it a chronocyte and argue that it was distinct from bacteria and archaea. To support this idea, they examine lists of proteins encoded by genomes of yeast, microbes, and other creatures represented in public databases. Among others, they cite Sogin's *Giardia* data set—which, they note in their paper, they "downloaded" with MBL's help in February 2001. Hartman insists that he used the *Giardia* data only as "a filter" to subtract out proteins of primitive organisms. (*Giardia* branched off from other eukaryotes very early.) "I never analyzed [Sogin's] data," he insists.

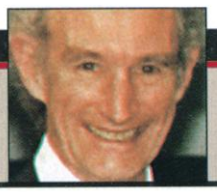
Sogin nonetheless "went ballistic," says Hartman, appealing to the editor of *PNAS*, Nicholas Cozzarelli, for a correction. But Cozzarelli declined to intervene. Sogin "wants to encourage people to use his data but not publish it without his permission," says Cozzarelli, who thinks that Sogin is claiming too much control and that his position is "untenable."

Sogin is not alone in arguing that sequencers should have greater control over the use of unpublished raw data from their own labs. "It's a big concern," says Claire Fraser, president of The Institute for Genomic Research in Rockville, Maryland, which is sequencing many microbial genomes. "We feel that a full download of a database" for publication without "intellectual input" from the DNA sequencers is "not appropriate." Richard Hyman, a researcher at Stanford University's genome technology center, says he has contemplated suing people who use his lab's data in publications without permission. However, he's not convinced that Hartman's paper violated the rules.

The researchers who benefit most from public access to genome data are the computer wizards of bioinformatics. They are sometimes seen as "parasites" because they rely on others for raw material, says Sean Eddy of Washington University in St. Louis, Missouri, a leader in this field. Eddy has championed free access to genome data collections in the past. But he says he has become more aware of the need to protect the publication rights of DNA sequencers. He even suggests that it may be time to "revisit the rules" that demand prompt public release of raw data, laid down during the heyday of the Human Genome Project.

* The *Giardia lamblia* genome project includes teams at MBL; the University of Texas, El Paso; the University of Arizona in Tucson; and the University of Illinois, Urbana-Champaign.

CREDITS: (TOP TO BOTTOM) PAULINE LIM/MBL; M. ABBEY/PHOTO RESEARCHERS



1212
Royal Society faces inquiry

LEAD STORY 1214
Becoming human

Such a reassessment may be about to happen. A National Academy of Sciences panel will soon begin an examination of the "responsibilities of authorship in the biological sciences." Chaired by biologist Thomas Cech, president of the Howard Hughes Medical Institute, the group will hold its first public meeting on 25 February, with a goal of trying to figure out whether there should be "a single set of accepted standards" for data sharing. **–ELIOT MARSHALL**

INFECTIOUS DISEASES

Researchers Crack Malaria Genome

LAS VEGAS, NEVADA—Genome scientists call it the toughest microbe they ever took on. But at a meeting* last Monday, an international group presented the long-awaited first glimpse of the full genome of *Plasmodium falciparum*, the most important malaria parasite. A project to sequence the genome, begun 6 years ago, is now almost finished, and the results could be published this summer.

The announcement marks the beginning of what appears to be shaping up as a milestone year for malaria research. The same groups have now put their teeth into a handful of other *Plasmodium* species, and most could be done before the year is over. To top it off, another international consortium plans to publish the genome sequence of the most important malaria mosquito, *Anopheles gambiae*, also this year. And a new class of inexpensive drugs is showing promise in animal trials (see next story and p. 1311).

Malaria researchers say the flood of new data is a welcome boost for their field. Most of the *Plasmodium* sequence has already been put in public databases and has stirred excitement among scientists, says Dyann Wirth, director of the Harvard Malaria Initiative: "It has really energized the field and brought it together."

The *P. falciparum* genome was cracked in a joint effort by the Sanger Centre in the U.K., The Institute for Genomic Research (TIGR) in Rockville, Maryland, the U.S. Naval Medical Research Center (NMRC) in Silver Spring, Maryland, and Stanford University. When they started in 1996, the teams thought the organism's 30-million-base-pair genome was too hefty to handle for a now-popular technique called whole-genome shotgun sequencing. Instead, they divided up the organism's 14 chromosomes and attacked each of those using the shotgun method.

The project was a "real nightmare," says TIGR's Malcolm Gardner, because adenine and thymine, two of the four building blocks of DNA, together make up 80% of the organism's genome. (The proportion is about 59% in humans.) That made it hard to clone *Plasmodium*'s genetic material in bacteria and to sequence it, says Gardner.

After developing new techniques and software to overcome those problems, the teams now have the sequence almost complete, with more than 10-fold coverage and just over 500 gaps left, most of them in three particularly difficult chromosomes. While some members are trying to close them, others are "working furiously" to annotate the more than 5600 candidate genes found by gene-finding software, Gardner says.

Using the same techniques, the TIGR team has also sequenced the genome of *Plasmodium yoelii,* which causes malaria in rodents, with fivefold coverage; the group is also working on *Plasmodium vivax*, which causes a less severe form of human disease. The Sanger Centre, meanwhile, has taken on several other *Plasmodium* species. Those additional genomes will be particularly useful for researchers using animal models and should help identify what exactly makes *P. falciparum* such a killer to humans.

Already, the raw data posted so far have been "tremendously helpful" for scientists searching for new drugs, says Stewart Shuman of the Sloan-Kettering Institute in New York City, one of many researchers who have dredged the data and come up with new drug targets. One group of German researchers not only found an enzyme that could make a good target, but also discovered that an existing antibiotic could block it, and clinical trials are being planned (*Science*, 3 September 1999, p. 1573).

But new drugs, although desperately needed, won't win the war against malaria, cautions Stephen Hoffman, a veteran malaria researcher at Celera Genomics in Rockville, Maryland, and former head of the Malaria Department at NMRC, who's involved in both the *Plasmodium* and *Anopheles* sequencing projects. Existing drugs can cure malaria, Hoffman points out—but they are too expensive or difficult to get to the people who need them most, and drug resistance is an increasing problem. A vaccine holds the surest hope of preventing the more than 1 million malaria deaths each year, mostly of African children. That's why Hoffman urged his fellow researchers to find such a vaccine, using a concerted, systematic approach—much like the genome project itself.

–MARTIN ENSERINK AND ELIZABETH PENNISI

CREDIT: CHARLES HANLEY/AP

Tragic toll. Malaria kills 1 million children a year, mostly in Africa.

* Second Conference on Microbial Genomes, Las Vegas, Nevada, 10–13 February.

INFECTIOUS DISEASES

Candidate Drug Breaks Down Malaria's Walls

The global picture of malaria is grim: Each year 300 million to 500 million people are infected and more than 1 million die, mostly children under the age of 5. Malaria parasites have become increasingly resistant to the most commonly used and least expensive antimalarial drugs. What's needed, experts say, are new potent and inexpensive drugs, ideally aimed at entirely new therapeutic targets that will make it that much harder for the parasites to acquire multidrug resistance.

On page 1311, a team led by biochemist Henri Vial of the University of Montpellier II and the Centre National de la Recherche Scientifique in France reports on just such a new class of antimalarial drugs, one that has shown remarkable effectiveness in rodents and primates. Unlike most commonly used antimalarials, which target the parasites' hemoglobin metabolism or DNA synthesis, the new compound inhibits the parasites' ability to synthesize protective membranes while sequestered away within red blood cells. The compound "seems to be extremely potent and active against even multidrug-