# Genome Project: An Experiment in Sharing

The Human Genome Project is in many respects a gigantic experiment in data sharing. Around the world, investigators are working on pieces of the same puzzle. And whether the project succeeds will depend in large measure on these investigators making available their data and materials—cell lines, probes, and clones—to their colleagues and competitors.

But will they do it voluntarily, or do they need a nudge—or a kick? While sharing may be the norm in, say, immunology or bacterial genetics, human genetics has always been intensely competitive. So should the National Institutes of Health and the Department of Energy, which both fund the genome project, promulgate rules to govern access to data and sharing of materials? And, whether they are formal or informal, what should the rules be? Some of the issues the genome project raises are brand new. Everyone agrees, for example, that materials should be available at the time of publication, but much of the information generated in the genome project will never be published, at least not in a conventional sense.

A DOE committee has drafted some guidelines, which have yet to be formally endorsed. They stipulate that data and materials must be publicly available 6 months after they are generated or characterized. But at NIH, James Watson, who heads the genome project, is shying away from setting rules. "I hope groups will form their own rules. We are very loathe to impose rules on anyone, unless we are forced to," he says.

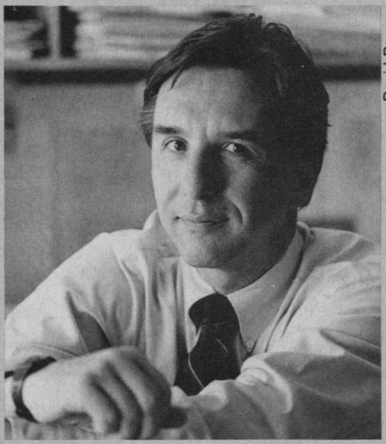
Watson's laissez-faire attitude is not shared by Walter Gilbert of Harvard University. He thinks that rules are clearly in order, and that NIH ought to get on with drafting them. "In my view, the genome center should take a stronger position that materials should be available," he says.

But, like Watson, Maynard Olson of Washington University, a member of the NIH genome advisory committee, is leery of rules. "We are talking about such a vast amount of data and potentially so many biological reagents. If you take a rule-based approach to this venture, then following close on its heels will be a large bureaucracy to enforce it," he says. Instead, "What you would like to do is change the culture."

And that is just what is starting to happen, says Watson, who speaks of a "much greater spirit of cooperation." He points to the new collaborative plans to map chromosome 21 as evidence that the community will develop its own ways of sharing data. Because of its known role in Down syndrome and its suspected role in Alzheimer's disease, chromosome 21 has generated a vast amount of interest. Lots of groups are already hard at work constructing maps of the chromosome—first developing a series of landmarks spaced along the chromosome, and then a collection of ordered DNA fragments. But the maps all these groups generate will be essentially useless unless they pool their data and adopt a common language. For that reason, David Cox, a chromosome 21 mapper at the University of California at San Francisco, called together 30 of the major groups in early April. As Cox hoped, they pooled their data and agreed to a common set of landmarks, or DNA markers, to guide their efforts.

But that was "pretty ho-hum" compared with what happened next, says Cox. To his surprise, these highly competitive groups also agreed to a collaborative venture to obtain a complete set of overlapping fragments or clones, using the very large yeast artificial chromosome (YAC) clones developed by Maynard Olson and his colleagues at Washington University (see article on p. 956). David Patterson of the Eleanor Roosevelt Institute for Cancer Research is getting a complete set of these YAC clones and, provided NIH kicks in some funds, he will run a screening service for anyone who wants to participate. Investigators will send their DNA probes, or a bit of sequence that will identify the probe, to Patterson, who will then search the YAC library for the corresponding clone. In return, the clone "becomes the common property of the entire group," says Cox.

"There are lots of ways the plan could crash and burn," says Cox, who admits that "paranoia is still alive and well." A few people, for example, said they would not send their sequences to Patterson because he might use them for his own ends. "As if he would have time," laughs Cox. Still, Cox is optimistic: "I think it will work. The cooperation is truly remarkable. People aren't fighting to sequence the same small piece of DNA and pull out the same clone."



David Powers

**The 21 club.** *Mapper David Cox helped push plan for chromosome 21.*

If the chromosome 21 plan becomes a model, Watson's hands-off approach to setting rules could be proved right—at least for mapping data. But, as Watson himself concedes, establishing guidelines for sequencing data may be trickier. The issue, at bottom, is how long an investigator gets to ponder the long-sought information in private. "There is no problem about holding onto it for a reasonable time. If you have worked out the sequence, you want the fun, the pleasure, of looking at it," says Watson. But how long is reasonable? Watson has just set up a sequencing subcommittee to advise the genome center on it.

The issue is clearly contentious. "It is unacceptable to get the sequence and then immediately release it," says Craig Venter, an NIH researcher who is one of a handful of biologists embarking on large-scale sequencing—working out 5 million bases or so. It could take several years to sequence that long a stretch, and Venter thinks researchers should have another year to characterize the DNA before making the sequence public. Still others argue that they should be able to hang onto the data until they are done, period.

But some researchers regard even a year as far too long. "My view is that if government pays for it, it can perfectly well demand the sequence be deposited promptly," says Gilbert, who is on the committee Watson has established. He thinks sequence data should be made public within 3 months and suggests that such a requirement be written into grants and contracts.

The debate will likely continue for some time. But in forcing researchers to confront the issues in the first place, the genome project is already beginning to break down some of the intense competitiveness among human geneticists. ■ LESLIE ROBERTS