HUMAN GENOME PROGRAM

# NIH Launches the Final Push To Sequence the Genome

It would be an exaggeration to say that the U.S. Human Genome Project has adopted Nike's slogan: "Just Do It." But the message being sent to the U.S. genome community this week comes close. As *Science* went to press, Francis Collins—director of the National Center for Human Genome Research (NCHGR) at the National Institutes of Health (NIH)—was preparing to announce six research awards that will launch the final phase of the effort to decipher the human genetic code.

The new grants—costing more than $20 million this year and $60 million over the next 3 years—are intended to speed up the pace of automated DNA sequencing, the brute-force technique that will be used to determine the exact order of the 3 billion nucleotide base pairs in the human genome. And they come with a requirement that the sequence data they generate be made rapidly available to the research community.

RICK KOZAK

**Encouraging openness.** Francis Collins wants data published rapidly.

The announcement reflects a growing sense that it's time to go beyond "mapping"—preparing indexes that locate genes scattered along the vast length of the chromosomes—to the final task of determining the actual DNA sequence of the chromosomes. At a meeting in Bermuda in February, the academic genome community endorsed a decision to move forward tentatively with pilot projects that will begin this arduous final task. At the same time, Collins and others sought to win agreement at the meeting on a few general principles: Those who win the new grants will freely share all their sequence data with the community, they will refrain from seeking early patents on whatever they produce, and the quality of data will be kept very high.

By 1999, if all goes well, the six pilot projects funded this week (see table) should yield precise data (99.99% accurate) for stretches of DNA on several chromosomes, covering about 3% of the total human genome. This work will complement another major sequencing effort already launched in Europe: Last November, the Wellcome Trust—a private British philanthropy—announced that it will put about $75 million over 7 years into a concentrated DNA sequencing project run by John Sulston at the Sanger Center in Cambridge.

The U.S. program will have several objectives, reflecting NCHGR's desire to hedge its bets and circulate its limited funds to a variety of labs. According to NCHGR, the new program will test the claim made by Sulston and Robert Waterston of Washington University in St. Louis that today's DNA sequencing technology is already good enough to do the job cheaply and accurately (*Science*, 10 November 1995, p. 903) But the money will also be spent on improving the state of the art in sequencing; most of the grants planned this week include support for developing new technology, such as improving the quality of fluorescent dyes used in chemical reactions, automating the processing of gels containing DNA, and creating more sophisticated software to interpret data from gel-reading robots.

In addition to testing the capacity of labs to step up the output of sequence data, NCHGR is also testing its own powers to persuade investigators to share information. New grantees will be required to release all their sequence data on the World Wide Web "within a few days or weeks" of discovery, and NCHGR is encouraging them not to take out patents on raw sequence data. Also, by late May, the Human Genome Organization plans to set up an international Web page on which each sequencing team will describe its research territory and maintain an electronic link to its data.

Collins says the principles of quick release and no early patents were endorsed "unanimously" by 30 to 40 leading scientists in the field—including the winners of these new NCHGR grants and the Sanger Center—at a closed meeting sponsored by the Wellcome Trust in Bermuda on 25 to 28 February. One NCHGR grantee who attended that meeting, Richard Gibbs of the Baylor College of Medicine in Houston, says that Wellcome Trust staffers are now circulating a manifesto, signed by participants in the meeting, that asks other researchers to agree to release sequence data as soon as feasible and to refrain from patenting "raw" DNA sequences.

Collins recognizes that some institutions may not be ready to embrace the "Bermuda principles," but he says investigators who receive NCHGR funds should expect to comply. And he notes that NCHGR will monitor grantees carefully to see that they make raw data available quickly. "That's going to be a major item" to be examined when a grant comes up for review in a couple of years, Collins says. Anyone who hoards or patents data will risk losing future federal support.

NCHGR's new sequencing program also reflects a consensus reached at the Bermuda meeting on another important issue: quality control. In the past, Sulston and Waterston had argued that it would be sufficient to aim for moderately precise data—which they defined as 10 errors or gaps in every 10,000 bases sequenced (99.9% accuracy). They said it would be worth accepting some degree of imprecision in exchange for speed and lower cost. Many of their colleagues were skeptical, however, and some openly questioned their cost estimate of about 15 cents per base sequenced.

When the sequencers debated these issues at the Bermuda meeting, Collins says, the group decided that contrary to the original Sulston-Waterston proposal, the "general sense" was that "we should start out trying to be really accurate." All six recipients of the NCHGR grants will therefore be asked to shoot for only one error per 10,000 bases sequenced (99.99% accuracy).

Collins and Gibbs credit Phil Green of the University of Washington, Seattle, for developing a computer program that should help make this accuracy possible without substantially increasing the cost. Green's program objectively determines the statistical confidence value of ambiguous base pairs as it

**GENOME SEQUENCING AWARDS**
($ millions)

| Investigator | Institution | Funding | Research |
|---|---|---|---|
| Mark Adams | The Institute for Genomic Research | 3.2 | Chromosome 16 sequencing, software development |
| Richard Gibbs | Baylor College of Medicine | 1.3 | Chromosome X sequencing dyes, DNA purification methods |
| Eric Lander | Whitehead Institute–MIT Genome Center | 4.1 | Chromosomes 9 and 17, robotics, maps for sequencing |
| Richard Myers | Stanford Univ. | 2.5 | Chromosomes 4 and 21, directed sequencing, DNA chips |
| Maynard Olson | University of Washington, Seattle | 1 | Streamlined high-accuracy sequencing, chromosome 7 |
| Robert Waterston | Washington Univ. | 6.7 | Large-scale sequencing on chromosomes 22 and X |

tries to assemble a finished sequence from overlapping stretches of raw DNA data. Then it includes in the final sequence only data that meet a minimum standard.

Waterston, acknowledging that "there have been some changes in what the community would like," says he will now aim for the goal of 99.99% accuracy. He says his team at Washington University, which got the largest sequencing grant ($6.7 million), calculates that this high-precision goal "is going to raise costs and slow work down a little." But he adds: "We still think we can be pretty cost-effective." He estimates that it will cost around 30 cents per base for the first year's operations. Collins is looking for considerable improvements over that figure in subsequent years: "If we can't get the price down to 20 cents a base, it's going to be hard to get this project done," he says.

Some researchers, Gibbs among them, worry that with those kinds of pressures, some sequencing teams will cut corners to meet production goals as the deadline for reviewing these grants draws near in 1998. Gibbs is concerned that quality of data will suffer as sequencers get trapped on this "slippery slope," allowing more and more ambiguity to creep into their data in order to keep output high. To offset that possibility, Collins says NCHGR may ask an independent team to sequence random lengths of DNA already generated by the grantees, to check for gaps and errors. He admits, however, that this kind of monitoring would be expensive.

The 1998 reviews will mark another critical turning point in the program: Some—but perhaps not all—of these labs will make it into the third year of funding. After that, NCHGR may pick one or more of the most successful strategies they have developed to receive support for an even more ambitious ramp-up of genome sequencing. Collins promises, however, that the competition will not be restricted to the six labs that won the grants announced this week: If other sequencing efforts improve upon what is being done, they could eventually win support.

Indeed, NCHGR has already reached beyond the established sequencing community for this first round of grants, funding as sequencers several groups that had focused primarily on gene mapping. These include Eric Lander's mapping team at the Whitehead Institute–Massachusetts Institute of Technology Center for Genome Research in Cambridge, Massachusetts, Maynard Olson at the University of Washington, Seattle, and Richard Myers at Stanford University.

Gibbs, who says he looks forward to the competition, observes: "It's going to be a very interesting 2 years."

**–Eliot Marshall and Elizabeth Pennisi**

SCIENCE AND THE LAW

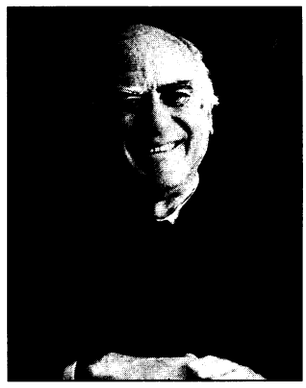
# New York Courts Seek 'Neutral' Experts

Facing a flood of claims that turn on disputed scientific evidence, New York courts have taken an unusual step: They are bringing in their own experts. Out of the hundreds of thousands of women seeking redress for health problems they believe may be linked to silicone breast implants, thousands have filed claims in New York; one plaintiff's attorney, Perry Weitz of Weitz & Luxenberg, himself represents about 6000 such women. To manage this impending barrage of litigation and referee the combat among hired experts, three judges, led by Jack Weinstein, chief judge for the federal court for eastern New York, decided to create an independent panel of scientific experts to give them neutral advice.

The panel's first task—according to an order signed on 3 April by Weinstein, federal Judge Harold Baer Jr., and state Supreme Court Judge Joan Lobis—will be to identify and recruit other nonpartisan experts willing to help the court. Then, once this larger advisory body is in place, it will consider general principles for establishing cause and effect in these cases, giving "particular attention ... to claims respecting immune system dysfunction and connective tissue and rheumatic disease" as described in a summary of complaints prepared for the court by Weitz's law firm.

The scientific battles are likely to be intense. Three major epidemiological studies have found little or no evidence that implants have caused serious immune system or other diseases. But Weitz and his experts charge that the studies were manipulated to find no health effects. Weitz also claims that the studies used erroneous and outdated definitions that caused them to understate the real incidence of disease.

According to Margaret Berger, a professor at Brooklyn Law School whom Weinstein has consulted on this and other technically complex cases, the goal will be to establish a "dialogue among the experts" that will clear away confusing side issues before substantive disputes are presented to the judge and jury. Berger is one of the three outsiders who have been asked to help out. The others are Joel Cohen, a population expert at Rockefeller University, and Fred Alan Wolf, a physicist and lawyer at the Cooper Union for the Advancement of Science and the Arts in New York. Berger and Cohen have some idea of what they are getting into: Both provided advice to Weinstein on the Johns Manville bankruptcy case after the company was hit with thousands of lawsuits alleging health problems linked to asbestos.



**Seeking scientific help.** Judge Jack Weinstein.

Judges have been empowered to use such panels since 1975, after Congress adopted Federal Rule of Civil Procedure 706. It says that any federal court "may appoint expert witnesses of its own selection," to be compensated in civil cases by the parties to the suit. But courts have rarely taken advantage of the rule, says Joe Cecil, a researcher at the Federal Judicial Center in Washington, D.C. The reason, says Paul Carrington, director of the Center for Private Adjudication associated with Duke University, is that "the judges don't like it, and the reason the judges don't like is that the lawyers don't like it."

Lawyers are against the rule because it forces them to pay for expert testimony they can't control and which may even go against them. The other problem, Carrington says, is that judges don't know how to find the relevant experts or which ones to trust. Two groups are now trying to develop information centers that would help the courts find reliable experts: Carrington's center and a group coordinated by Deborah Runkle at the American Association for the Advancement of Science (*Science*'s publisher).

As might be expected, plaintiff's attorney Weitz argued strenuously against the creation of an independent panel in the breast implant cases. Weitz had objected that judges who use Rule 706 are "usurping the adversarial process" and denying juries an opportunity to weigh all the evidence. But now that Weinstein has ruled against Weitz's objections, Weitz says he's focusing on the positive aspects of the decision. Perhaps, he says, this panel "could be very useful" if it allows his side to expose the alleged manipulation of data in the epidemiological studies.

As for the court's new experts, they're keeping mum. All that Wolf would say about the litigation, for example, is that it may give him a rare opportunity to make use of his own field of expertise—chaos theory.

**–Eliot Marshall**