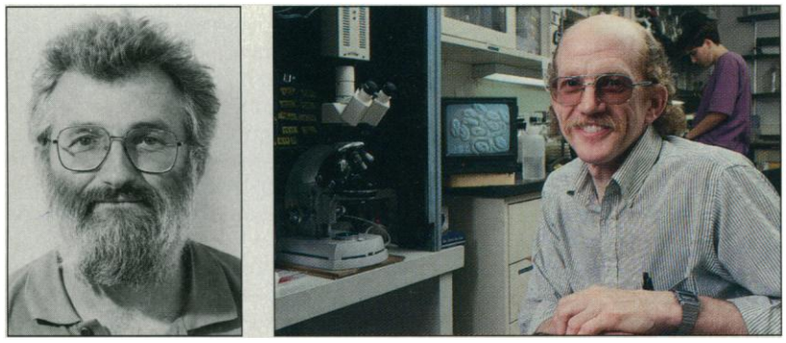
HUMAN GENOME PROJECT

A Strategy for Sequencing the Genome 5 Years Early

In meetings over the past 6 weeks, two respected gene sequencers have been delivering a startling message: The chief goal of the Human Genome Project—obtaining a complete sequence of the 3 billion bases in human DNA—can be achieved as early as 2001, 5 years ahead of schedule. What is more, they say, it can be done without any fancy new technology. The two optimists—John Sulston, director of the Sanger Center in Cambridge, United Kingdom, and Robert Waterston, director of the Genome Sequencing Center at Washington University in St. Louis—have sketched a plan that they think could deliver the Holy Grail of genomics for $300 million to $400 million over 5 years. The U.S. share might be $40 million to $54 million a year, or just over a third of the present budget of the National Institutes of Health's National Center for Human Genome Research (NCHGR).



TOM HEINE/WASHINGTON UNIV. SCHOOL OF MED.

Full speed. John Sulston *(left)* and Robert Waterston have been floating a proposal to shift into high gear on sequencing, using current technology.

Researchers who have studied the plan believe it is feasible. But it is generating controversy. To accommodate this speedup, federal program directors might have to revamp their budgets and reorder priorities. Unless they get a huge funding increase from Congress—which seems unlikely—they would probably have to de-emphasize other plans for technology development and genetic map-building to promote DNA sequencing. At a meeting last week, some researchers worried that such a shift would benefit big labs while leaving many of the smaller ones out in the cold. Yet the idea is being widely debated because everyone agrees that the main objective of the genome program, after all, is to sequence the genome. And, as Sulston asks: If it can be done now, "why fiddle around"?

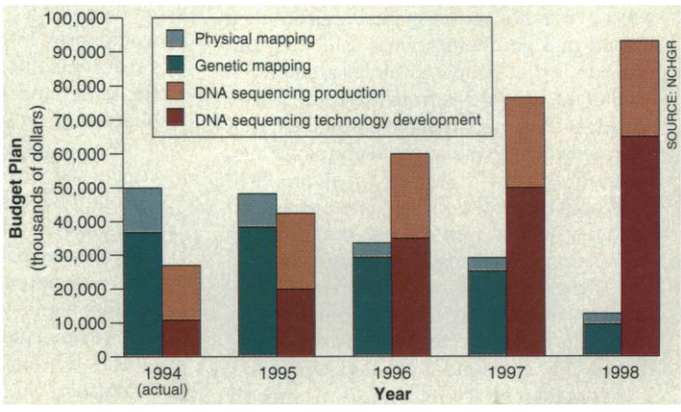
Until now, the government's strategy has been to spread its money around and let a thousand flowers bloom. The two federal backers of the Genome Project—NCHGR and the Department of Energy—have been investing in scores of small projects. Some have focused on finding unique signposts and mapping discrete pieces of the human genome; others on making laboratory robots that might take over some of the tedious labor in sequencing; and a few have begun to process short stretches of human DNA. The project's leaders hoped that all this experimentation would lead to a breakthrough that would sharply cut the cost of obtaining sequence data. But that hasn't happened.

Now, Waterston and Sulston are arguing that instead of waiting for the ideal technology, it is time to move pragmatically into the final phase of the program—sequencing the genome—using tools already at hand. Maynard Olson, the University of Washington, Seattle, geneticist and DNA sequencing expert, agrees, noting that even if a technological breakthrough appeared today, it might take longer to polish it for mass production than it would take to sequence the genome with existing technology. Mark Guyer, NCHGR's assistant director, says that he has sensed "a lot of excitement" as well as apprehension about this suggestion. Certainly, he said, "it is the first time that anybody has come up with a scheme even on paper that looks like a plausible way to approach the 15-year goal of the genome project."

The idea of going into high gear on sequencing was officially broached at a meeting organized by NCHGR on 16 December in Reston, Virginia. Guyer had summoned all the genome barons—the directors of a score of Genome Science and Technology Centers, or GESTECs—to a strategy session. In his letter of invitation, Guyer wrote that it was time to "address the next phase" of the program. He warned that NCHGR is being asked to take on more and more tasks, but that "there is likely to be a shortfall" between goals and funds. "It has become clear," he wrote, "that the calculus of what NCHGR should do over the next few years to meet its scientific goals is highly dependent upon what can be expected with regard to DNA sequencing." He hoped the Reston meeting would take a critical look at the state of the art in sequencing and help the agency evaluate the credibility of plans for surging ahead.

Waterston was invited, and he laid out the Sulston-Waterston plan for the first time. Encouraged by NCHGR, he told the other directors that he and Sulston—partners in an effort to sequence the genome of the worm *Caenorhabditis elegans*—had grown more and more confident over the past year about their ability to process large quantities of DNA sequence. So confident, he said, that they are now ready to move from worm to man. He said that his own lab is now able to process 15,000 "reads" (samples of 400 to 500 DNA bases) per week. He thought his lab could manage a five-fold scale-up to 84,000 reads per week. If three labs running in tandem could achieve this rate, Waterston projected it would take 5 years to cover 99% of the entire human genome, with an overall accuracy of 99.9%. Areas of the genome that seemed biologically important would be sequenced to 99.99% accuracy. Furthermore, he said, his experience with the worm suggests that the sequencing itself could be done for 10 cents per base, not counting the cost of developing raw materials for sequencing, which might cost an additional 2 cents per base. He guessed it might cost $20 million to $25 million per lab, per year.

Waterston and Sulston say that they hit on two critical insights in the past year that



Change of emphasis. A draft budget plan for NIH's genome project shows a shift from mapping to sequencing.

led them to make this proposal. First was the fact that no quantum leap in DNA technology was on the horizon. And second, Sulston says, was the realization that the entire human genome need not be deciphered to the 99.99% accuracy he and Waterston are achieving on the worm's DNA. A small relaxation in standards to 99.9% accuracy would permit a tremendous saving in time and money. Moreover, says Sulston: "The entire bill for biomedical research would be lower in 10 years' time if we start the sequence now than if we delay. If we don't start now, there will be innumerable other gene hunts and sequencing projects going forward, taking collectively an enormous amount of money. ... It would be far better if that money could be spent on biology" based on sequence data "which could be collected efficiently by the genome project."

Since December, the Waterston-Sulston team has presented its scenario twice more in public—at a meeting of the international Human Genome Organization in London on 24 to 25 January, and at the meeting of the NCHGR advisory council in Washington, D.C., on 30 January. At the council meeting, NCHGR also unveiled a draft 5-year budget plan that seemed to dovetail with Waterston's talk. It showed that the NCHGR staff would like to shift sharply away from "mapping" genes (locating them on chromosomes) and spend more on sequencing the genome (see chart on p. 783).

The scientists who went to these meetings didn't formally review the proposals for human DNA sequencing, but they have given the idea a broad vote of confidence—although with some reservations. In a report to the NCHGR advisory council, Lloyd Smith, a developer of sequencing technology at the University of Wisconsin, Madison, said he found it remarkable that the budget plan with its "complete reorientation" of the genome project had been accepted at the Reston meeting without demur. Olson made a related point in a report he prepared on the Reston meeting at NCHGR's request. "The basic feasibility of the Waterston scenario," he wrote, "was not seriously challenged." And Guyer, in a phone interview, said, "My impression is that people think it's credible."

Part of the "news from Reston," Guyer added, is that "people have confidence that even current sequencing technology can take us a long way" toward the goals proposed by Waterston and Sulston. The biggest uncertainty, in fact, is in the step that precedes sequencing. Before researchers can sequence DNA, they must clone it into a well-organized set of bacterial vectors (cosmids) suitable for use in a production lab, something that hasn't been done as yet. As Olson wrote in his notes for NCHGR, there is "no broadly usable source of mapped cosmid clones ... with which to support even the early years of such an initiative. Early implementation of the Waterston scenario would require an immediate, relatively massive commitment to produce such a cosmid resource."

The problem arises because most vectors used to clone human DNA are too unstable, too sparsely developed, or too little tested to be fully reliable. Yeast artificial chromosomes belong in the first category, while cosmids and new cloning systems such as bacterial artificial chromosomes belong in the second and third, according to Smith.

Waterston agrees that "someone's going to have to create a resource for sequencing; no one's done it for human DNA." However, he thinks there is enough raw material already in government-funded labs, covering half-a-dozen chromosomes, to sustain a pilot project. In fact, Sulston reports that his lab is now sequencing a 2.2-million-base stretch of human DNA that includes the Huntington's disease region on chromosome 4—the longest contiguous piece of human DNA ever sequenced—and the preliminary results look good. Next, Sulston and Waterston hope to join with others to attack chromosome 22.

Smith and other members of the NCHGR council also raised some questions at last week's meeting that went beyond the technical issues of cosmid acquisition. Smith, for example, was concerned that the genome program clearly define standards for accuracy and completeness it expects to achieve before taking a leap into large-scale DNA sequencing. He worried that as pressure to cut costs increases, quality might be sacrificed. And many researchers—both at the council meeting and at Reston—were concerned that the NCHGR might be narrowing its technological options too early. In his report to the agency, Olson expressed misgivings about the "innovation-suppressing effects" of a heavy investment in one particular system—in this case, traditional gel-based sequencing technology—before the alternatives have been fully explored.

One of the thorniest issues, however, has more to do with politics than technology. It is the possibility, mentioned on 30 January by council member Daniel Camerini-Otero, that it "will be hard on the community" if the decision to go into large-scale sequencing creates "just a few large centers." At present there are 20 GESTECs. Camerini-Otero, chief of the genetics branch at the National Institute of Diabetes and Digestive and Kidney Diseases, was concerned that many would not be able to survive the transition from mapping to sequencing. David Botstein of Stanford University responded that the program should stick with "our policy of the past"—to follow the best science and let the stragglers fend for themselves. Botstein added that his own "optimistic view" was that most of the centers will be able stay in the game.

As for NCHGR, Guyer says the agency will soon make available a new "pot of money" for sequencing. The agency will revise and reissue a $3 million request for applications (RFA) aimed at developing better gel-based sequencing equipment. This will be combined with a new RFA along the lines of a draft plan that suggested spending $10 million to $15 million a year on pilot projects in human DNA sequencing. The aim is to get researchers like Waterston and Sulston to refine their ideas and submit them in scholarly form for close scrutiny. Guyer hopes these new ideas can be peer reviewed late this year and—if all goes well—funded in early 1996. If so, the finishing line for the marathon Human Genome Project might be a lot closer than it seemed a year ago.

–Eliot Marshall

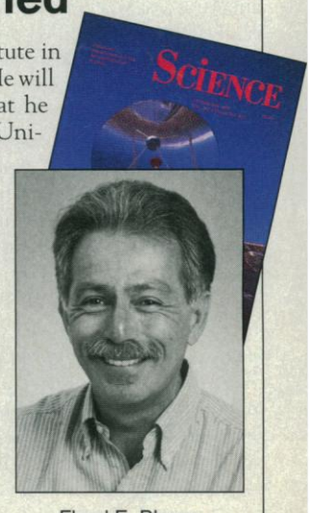
PUBLISHING

Science Editor-in-Chief Named

Floyd E. Bloom, a neuroscientist at the Scripps Research Institute in La Jolla, California, has been named editor-in-chief of *Science*. He will replace Daniel E. Koshland Jr., who announced last year that he would retire from *Science* to return to full-time research at the University of California, Berkeley. Bloom, age 58, will continue to run a lab and serve as chair of the Department of Neuropharmacology at Scripps.

Bloom has spent 33 years in basic research, focusing on the chemical control of neuronal activity. He previously worked at the National Institute of Mental Health and the Salk Institute for Biological Studies. He was elected to the National Academy of Sciences in 1977, is a past president of the Society for Neuroscience, and has served on the council of the Institute of Medicine and on the board of the American Association for the Advancement of Science (AAAS).

Bloom's appointment was announced this week by the AAAS, *Science*'s publisher. He will officially assume his new role in the spring of 1995.



Floyd E. Bloom