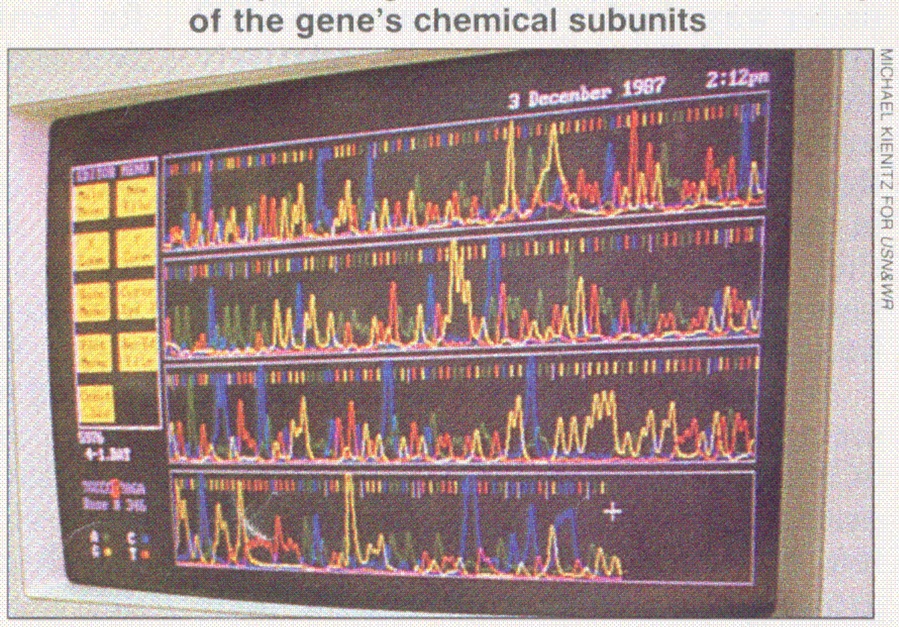
Reading the human blueprint

MEDICINE ■ The goal is to decipher every human gene. But some biologists wonder if it's worth the unprecedented cost

CHRISTOPHER MORROW FOR *USN&WR*

Nobel Prize-winner Walter Gilbert hopes to raise $8 million in private capital to do the job. Below: A partially deciphered gene. Colors show the identity of the gene's chemical subunits



MICHAEL KIENITZ FOR *USN&WR*

■ Physicists have their atom smashers. Astronomers have their telescopes. Now, it's biologists' turn at big science. Only this time the object under scrutiny will not be a distant star or an atom, but ourselves. The ambitious goal is to decipher the 3 billion individual ciphers that together form man's genetic code—what amounts to a complete chemical formula for a human being.

Called the human genome initiative, the scale and scope of the project are unprecedented in biology's history: It will take several decades to complete and could cost between $500 million and $3 billion. Proponents justify the hefty price tag by insisting that the project will guarantee United States leadership in the increasingly competitive pharmaceutical industry. They also point to its huge scientific dividends. By working out the precise functions of genes responsible for genetic diseases such as cystic fibrosis and Huntington's disease, scientists may be able to devise powerful new therapies. Eventually, this expedition into the core of human existence promises answers to some of the most profound questions in modern biology: How does a single fertilized egg grow into an organism as complex as a human baby? What genetic changes turn a healthy cell into a cancerous one? How do genes direct the aging process?

A genetic atlas

The first steps toward unraveling the entire human genetic code already have been taken. That is to determine the location of specific genes on the long strands of DNA that make up the 46 human chromosomes. Each gene, which directs the production of a single protein in the body, is made up of a chain of some 1,500 chemical subunits known as bases. It is the sequence of these bases that specifies the exact chemical structure of the protein. Structural materials that make up blood, muscle and skin, the hormones that course through the blood stream and the enzymes that drive each and every step of metabolism—all these substances are made of protein.

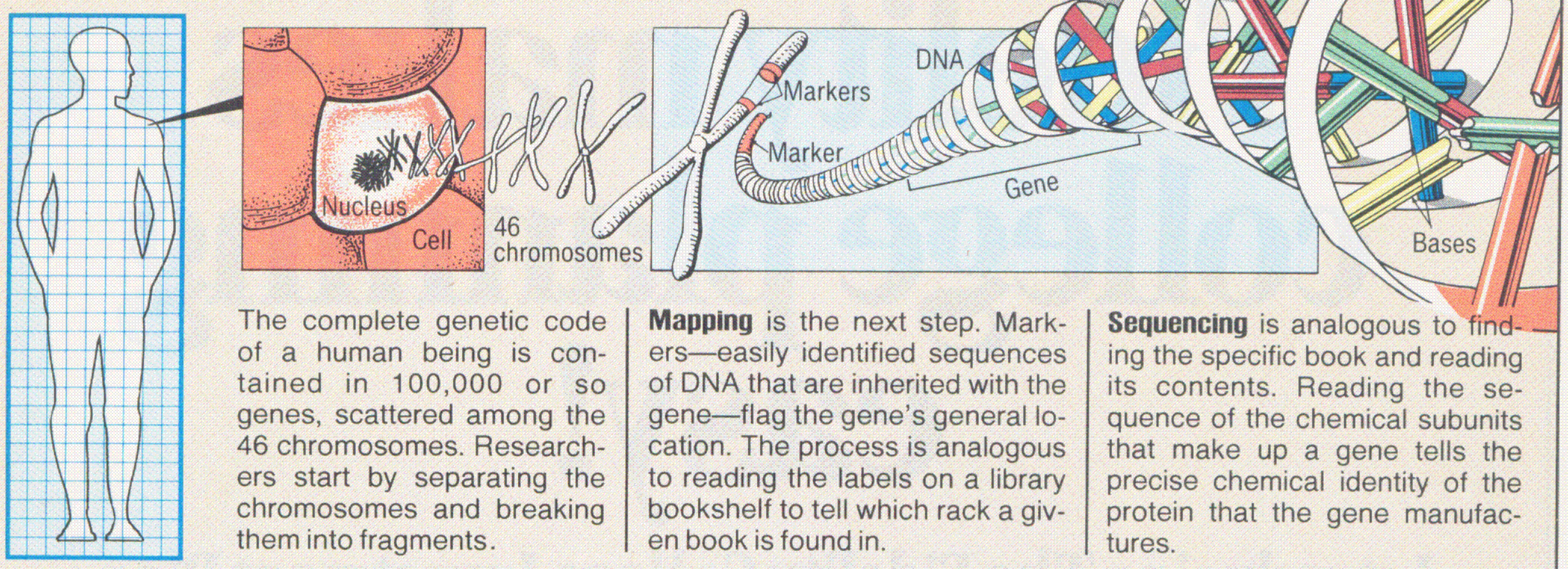
In the past few years, scientists have mapped the approximate locations of at least 400 of the 100,000 or so human genes. The first to be targeted for study are those associated with the 3,000 known hereditary disorders. By reading the sequence of bases in these defective genes, it will be possible to determine exactly why the gene either fails to function or produces a malformed protein. But that's just the beginning, for researchers are now discovering genes that influence the onset of a host of common diseases not usually thought of as hereditary, including heart disease, rheumatoid arthritis and Alzheimer's disease.

Disease fighters

In addition to genes associated with illness, the mapmakers are charting the locations of genes that produce immune proteins such as interferon that fight disease. But the majority of these immune chemicals still remain to be discovered and their genetic blueprints traced to specific chromosome locations. "We've only identified about 1 to 2 percent of all the body's proteins," says Leroy Hood, a molecular biologist at the California Institute of Technology. "Think what powerful [disease-fighting] tools lie ahead when we find the other 98 percent."

Determining the sequence of bases in a gene until recently has been a tedious process, requiring a painstaking chemical analysis to determine the identity of each and every base. Now, an automatic sequencer developed by Hood and colleague Lloyd Smith can "read" an entire gene in a day or so. The process entails tagging fragments of genes with fluorescent dyes, whose colors are then scanned by a laser beam and the information recorded by a computer. New sequencers promise

HOW RESEARCHERS UNRAVEL THE GENETIC CODE



The complete genetic code of a human being is contained in 100,000 or so genes, scattered among the 46 chromosomes. Researchers start by separating the chromosomes and breaking them into fragments.

Mapping is the next step. Markers—easily identified sequences of DNA that are inherited with the gene—flag the gene's general location. The process is analogous to reading the labels on a library bookshelf to tell which rack a given book is found in.

Sequencing is analogous to finding the specific book and reading its contents. Reading the sequence of the chemical subunits that make up a gene tells the precise chemical identity of the protein that the gene manufactures.

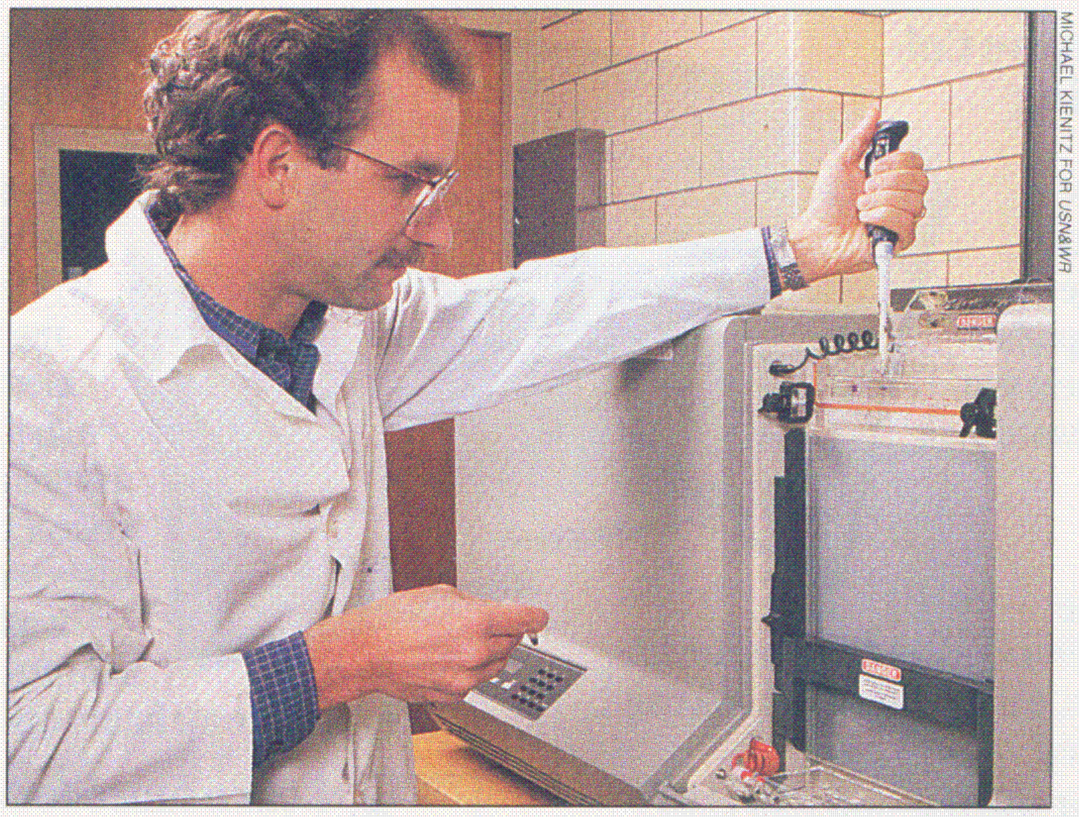
to make the process 10 times faster.

But even assuming such advances, some researchers question whether the brute-force approach of sequencing the entire genome even makes sense. Only about 4 percent of all the DNA found in chromosomes actually contains genes—that is, sequences that direct the manufacture of proteins. The rest seem to be evolutionary driftwood—DNA cast off so long ago in our evolution that the sequences have mutated into total gibberish. Robert Weinberg of the Whitehead Institute at the Massachusetts Institute of Technology is one of several leading geneticists who cannot see the point in "wading through a sea of drivel to emerge dry-shod on a few tiny islands of information."

Another obstacle is a turf battle between the Department of Energy and the National Institutes of Health. DOE is advocating a Manhattan Project approach: A focused program that would develop advanced sequencers and computers in a crash effort to unravel man's entire genetic code. Although DOE would seem an unlikely government body to become involved in such an endeavor, the agency has gained considerable expertise in the field through studies of how genes are damaged by radiation.

Footing the bill

The National Institutes of Health, however, is already spending $300 million a year to study the genes of diverse life forms, including $100 million exclusively on human genes. And NIH prefers to fund smaller groups of researchers working in individual laboratories across the country, rather than managing a single megaproject.



MICHAEL KIENITZ FOR *USN&WR*

Lloyd Smith and his automatic gene sequencer

With government plans still up in the air, at least one group is trying to raise capital to do the job privately. Nobel Prize-winning biochemist Walter Gilbert of Harvard University, a founder of Biogen, one of the pioneer biotech companies, is calling his new venture Genome Corporation. He even has plans to copyright human gene sequences—an unsettling thought to many researchers who would prefer to see the information remain in the public domain. Legal scholars point out, however, that copyright protection is only afforded to authors of original works. Says Susan Rosenfeld, a New York City attorney who specializes in legal issues related to genetics, "About the closest candidate in this case is God."

Gilbert isn't intimidated by the Almighty's competition. If he can't copyright his gene sequences, he will compile his genetic data into a commercial data bank and charge users a fee to gain access to the information through computer-phone links. That won't preclude scientists from gaining the data by other means. But many scientists may prefer the convenience of the data base, just as subscribers to data bases such as Nexus can save a trip to the library by scanning its files for newspaper articles.

As for raising the additional $8 million in venture capital needed to launch his company, perhaps Gilbert and other entrepreneurs could benefit from the advice of biologist David Tepfer of the Institut de la Recherche Agronomique in Versailles, France. In a letter to the British journal *Nature,* Tepfer arrives at a financing scheme after posing the obvious question: *Whose* genome should be sequenced first? After all, each individual's genes are somewhat different. Not to be swayed by nationalistic interests, Tepfer quickly rules out such candidates as Ronald Reagan, Margaret Thatcher and François Mitterrand. "My suggestion," he writes, "is that it go out to tender. Unfortunately, J. P. Getty and H. Hughes are dead, but there must be somebody who can afford to be sequenced." ■

by Kathleen McAuliffe