One by one, new colors appear on the computer screen—red, blue, green, yellow—ticking off the previously unknown code of human DNA. Each color is one chemical base—one "letter" in the 3 billion-plus letter script of the human genetic code. Soon the screen is awash in colors, a pointillist painting that holds the secrets of human heredity.

Color by color, base by base, that scene is being played out in a handful of laboratories in the United States and England as scientists race to finish the Human Genome Project—the most audacious endeavor undertaken in biology. When they are done, perhaps as early as the end of 2000, they will have accomplished what many said was impossible just 10 years ago: uncovering the complete genetic instructions of a human being.

This revolution will bring changes no less sweeping than those wrought by the microchip in the 1990s. As with information technology, imaginations and business energies will be fixed on biotech breakthroughs that promise to unfold in the 21st century.

The consumer will keenly feel the effects: simple blood tests, for instance, that reveal one's risk of developing cancer or Alzheimer's, or custom-made drugs that work without side effects. Genetic knowledge may be harnessed to cure diseases or even slow the aging process.

Exploring the unknown. Scientists got a peek of what's to come in December, when the first human chromosome was decoded. Francis Collins, who directs the project from the National Institutes of Health, confessed that he got chills when he glimpsed the "landscape of a complete human chromosome," likening it to seeing the surface of another planet for the first time. Equally thrilling, said Collins, were the implications: that "this most ambitious of human endeavors, a journey into the human genetic instruction book," will succeed.

This vast project is designed to unlock the secrets in a genome—an organism's complete genetic code, which determines whether a fertilized egg develops into a mouse, a fly, a human, or some other living species (gatefold, Page 39). A simple misspelling in a gene can wreak havoc—

CRAIG VENTER

turning a normal cell cancerous, for instance. To some as yet unknown degree, genes also influence our personalities and behavior. Having the full human instruction manual in hand promises profound changes in science, medicine—and even in our understanding of ourselves.

The instruction manual will give scientists glimpses into the molecular changes that underlie a host of diseases, from cancer to diabetes to mental disorders. But such benefits do not come without a cost. Will insurers and employers have access to an individual's genetic profile? Will the medical benefits go to the rich and not to the poor? Will genes be manipulated to design smarter kids? “The consequences are so far-reaching as to touch every aspect of society,” says Eric Lander, a mathematician turned geneticist who runs one of the biggest sequencing labs, at the Massachusetts Institute of Technology.

Biology's moonshot. Aside from the electronic whirrings from row upon row of boxy PE 3700 sequencing machines, Lander’s lab is eerily quiet. Hour after hour, the machines toil, deciphering DNA at a rate of millions of bases a day. If the genome project is biology's moonshot, then the MIT megalab and four others like it are the Cape Canaverals, albeit ones where there is little work for humans, save for tending the machines. Like the race to the moon, the scramble to decipher the genome underscores man's drive to explore the unknown. With the stakes so high, the race has also seen its share of acrimony and egos as outsized as the endeavor.

One of the first mavericks was Walter Gilbert, a brilliant and irreverent biologist at Harvard who championed the idea—the "Holy Grail" as he called it—15 years ago. He was soon joined by another giant in the field: James Watson, codiscoverer of the structure of DNA. The two touted the idea, but most biologists were not impressed. At the time, the biggest organism that had been deciphered was a minuscule virus, just 240,000 bases long—and that took several researchers years to complete. At 3 billion bases, the human genome would consume thousands of scientists for countless years. And with its $3 billion price tag, it would divert funds from other biological studies. "It endangers us all," warned MIT geneticist David Botstein in 1986.

Even NIH, the nation's major funder of biomedical research, wanted no part of it, leaving the U.S. Department of Energy to lobby Congress for the project. But the audacious idea prevailed. The turning point was a 1988 National Research Council (NRC) report, which unanimously endorsed the project—a surprise, as the NRC panel contained tough critics, including Botstein and mouse geneticist Shirley Tilghman of Princeton University, as well as Gilbert and Watson. They compromised on a phased approach that would begin not with sequencing but by constructing "maps" of the genome. Not only were maps relatively simple to construct, but they would greatly speed the search for disease genes. As for sequencing, the panel recommended that it be postponed until new technologies made it faster and cheaper. It was the panel’s recommendation that the project also map and sequence the genomes of simple organisms, such as yeast, a bacterium, a roundworm, and a fruit fly, that carried the day. Most geneticists had cut their teeth working on these creatures, which share many of their genes with humans. Biologists had no chance of understanding the human genome without insights from these organisms, Tilghman argued.

As the genome project gained congressional funding, NIH wrested control of the project away from DOE. NIH set up a special genome office and, in a coup that brought instant credibility, nabbed Watson to head it. "My name was good," recalls Watson. He set a deliberate plan: The
goal was to build a set of power tools—the map and the sequence—akin to the particle-smashing accelerators that had transformed physics. And he relentlessly pushed the project's first tangible goal: the maps that would speed the search for genetic defects that cause devastating hereditary diseases. Knowing full well that Congress did not have the patience to wait 15 years for results, Watson staked his reputation on getting the maps done in five. Once the maps were complete, genes would fall out in short order—including the putative Alzheimer's gene, which Watson joked should be a high priority, given the age of most congressmen.

Map building entails looking for distinctive pieces of DNA and using them as landmarks along the chromosomes. The more landmarks, the better the map. But first scientists had to find those landmarks, a task that entailed scouring through base after base of DNA. Slowly, researchers found landmarks and placed them on the map. Even before they were done, the landmarks proved invaluable, enabling investigators to find the cystic fibrosis gene in 1989. Genes for other diseases, including neurofibromatosis and Huntington's, quickly followed, years earlier than would have been possible without the maps.

At the same time, biologists and engineers were trying to speed the sequencing process. Until the mid-1980s, the task consisted of breaking DNA into fragments, tagging them with radioactivity, and then placing them in a slab of clear semi-solid gel, sandwiched between two pieces of glass. The gel was zapped with an electric current, which pulled the fragments along at different speeds, depending upon their size. Finally, technicians photographed the gel and "read" the sequence—which looked like a series of gray blobs. With these techniques, a skilled worker could sequence a mere 15,000 bases a year, but by 1986, Caltech biologist Leroy Hood and colleagues had found a way to up the speed to 15,000 bases a day—thanks to the first automated sequencing machine.

Enter Venter. Meanwhile, other scientists were working out the kinks in the new machines by testing them on the genomes of yeast and worms. Although the hurdles remained high, the project was on course—until it was challenged by a relatively unknown NIH biologist, J. Craig Venter.

In 1991, Venter came forward with an iconoclastic idea: Rather than painstakingly decipher each and every base of DNA, why not concentrate first on sequencing the genes? What's more, Venter boasted that he had a clever way to find the genes, which he touted as a shortcut to the genome project. Venter simply looked in DNA for the telltale signs of an active gene and then sequenced a small bit of it. The partial sequence served as a "tag" to identify the gene. Thumbing his nose at the genome establishment, Venter asserted that he could find most of the 100,000 human genes within a few years.

Watson dismissed Venter's cream-skimming plan, but Venter pursued it anyway, and NIH began filing patent applications on Venter's partial genes at a rate of 1,000 each month. When Watson found out, he erupted, denouncing the patenting scheme as "sheer lunacy." He went to war with both Venter and then NIH Director Bernadine Healy, who sided with Venter. At issue was whether scientists could stake a claim on a gene, or a partial gene, without even knowing what it did. If they could, that claim could discourage other researchers from laboring long and hard to understand that gene and develop it into a product, say, a diagnostic test. The fight cost Watson his job.

After Watson's abrupt departure, NIH picked Francis Collins, a gene hunter extraordinaire at the University of Michigan, to take the helm. Dedicated and ambitious, Collins was fresh off the heady successes
of finding several disease genes. While Watson had talked about building tools, Collins, a physician by training, talked about saving lives.

Soon the early investments began to pay off as maps were completed and genes were discovered at breakneck speed. But the project's ultimate goal, sequencing the genome, was lagging behind. Collins, like Watson before him, had hoped for radical new technologies that would make sequencing a breeze. None had materialized.

Investigators continued to fine-tune the machines so they could handle longer stretches of DNA. But finishing the job—piecing the fragments together like a giant jigsaw puzzle—was proving tough. Some pieces of DNA simply refused to be sequenced, leaving troubling gaps.

Shot in the arm. Genome sequencing sorely needed a shot in the arm. It came from an unlikely quarter. In 1995, Venter, then head of the Institute for Genomic Research in Rockville, Md., published the entire DNA sequence of the microbe Haemophilus influenzae—the first sequence of a free-living organism. The microbe was small—just 1.8 million bases—but the feat was stunning because Venter's crew blasted through it in just one year. More startling still, Venter did it using a new approach, whole genome shotgun sequencing, that NIH researchers had insisted wouldn't work. Venter's technique involved chopping up the entire genome into small pieces, sequencing each piece, and then reassembling them in correct order. NIH-funded scientists, on the other hand, were starting with one relatively small chunk of the genome and then breaking it into pieces—an approach that left them far fewer pieces to reassemble.

NIH's approach was soon vindicated when a Stanford University team polished off the genome of baker's yeast, a milestone not only because of its size—15 million bases—but because of its complexity: Its cell resembles those of humans. It provided stunning insights, as Tilghman predicted. Buoyed by these successes, researchers were increasingly eager to plunge into the human genome. In April 1996, Collins gave six labs the charge to sequence a long stretch of human DNA while simultaneously improving the technology. If all went well, the six groups would complete 3 to 5 percent of the human genome by 1999. Then the project would scale up, with completion in 2005, a daunting task since that entails sequencing the genome 10 times over to ensure accuracy.

That cautious game plan was turned on its head when Venter dropped a bombshell. He announced in May 1998 that he had teamed up with Perkin-Elmer Corp. of Norwalk, Conn., to create a new company, Celera Genomics, that would single-handedly sequence the human genome in just three years, and for a mere $300 million. The claim was not far-fetched: Venter had several innovations to rely on and a hefty bankroll to fuel them. He had 300 of the world's fastest sequencing machines, the PE 3700, each capable of accurately processing 400,000 bases a day, and one of the fastest supercomputers in existence. And he had the whole-genome shotgun approach that had worked so well on Haemophilus. But this time, Venter was going to shred all 3 billion bases of the human genome.

Although Collins maintained a polite front, he and other scientists in the public consortium felt angry and threatened. If Congress believed Venter's bravado, would it pull the plug on the federal effort? In public, and even more in private, scientists denounced Venter's headline-grabbing style and insisted the shotgun technique could not work on the human genome. They accused Venter of conspiring to lock up the genome in patents and deny the public the fruits of the genome project.

But all the while, the publicly funded re-
searchers were placing orders for the PE 3700, and—more important—Collins was revamping the program to try to beat Venter. The public consortium decided to concentrate money in the five fastest sequencing labs—at MIT, Washington University in Missouri, Baylor College of Medicine in Texas, DOE's genome institute in California, and the Sanger Centre in England. Collins vowed to finish the sequence by 2003, two years ahead of schedule. And, in a major shift, he promised to produce a "rough draft" of the genome by spring 2000, a year ahead of Venter—a strategy designed to undercut any attempt by Celera to patent most of the genes.

Feats with DNA. The race was on. First Venter said he had sequenced the fruit fly with Gerry Rubin at the University of California–Berkeley. Next, Venter boasted that 1 billion bases of human DNA were done, a feat dismissed by NIH because he did not release the data. Soon NIH proclaimed that it, too, had sequenced 1 billion bases and put the data on the Internet. Venter countered at 3 billion.

Just last month, the rivals reached a détente and began talking of collaboration. By then it was clear that Venter was unstoppable, the two approaches were complementary, and Congress was not going to pull the plug on the public effort.

And surely there was enough in the scientific future to cover both entities in glory: Already, genetic knowledge is illuminating the processes at work in many diseases. By using new "DNA chips," dime-size wafers containing samples of DNA, scientists can see which genes are switched on in a cancer cell—which suggests myriad new ways of attacking the disease. Lander, for one, is convinced that cancer will eventually be cured, though not in the next 10 or 20 years. "The reason I do this work is so my kids' kids will never die of cancer. And that is a pretty darn good goal."

Well before then, doctors should be able to use genetic profiles to predict how individuals will respond to drugs. Further out is the prospect of "personalized medicine," in which drugs are tailor-made to a patient. Some of the most profound changes will emerge when scientists uncover the genes that increase the risk of common diseases—not just cancer, but diabetes, heart disease, Alzheimer's, and mental disorders. Even before new drugs are ready, tests will reveal an individual's risk and show ways, such as dietary change, to prevent it.

Stunning as the achievement is, scientists no longer call the sequence the "Holy Grail." "It's a parts list," Lander says. "If I gave you the parts list for the Boeing 777 and it had 100,000 parts, I don't think you could screw it together and you certainly wouldn't understand why it flew." But with the parts in hand, scientists can begin to figure out what each does and how they build a human—a task that should keep biologists busy through the new century. "This is absolutely the beginning," says Venter, "which is why we are in such a hurry to get there."

In a less tangible but equally profound way, the genome project may change how people view themselves. The sequence will provide tools to probe which personality traits are hardwired in the genes, which are more amenable to environmental influences. The danger is that people will begin to see themselves—and others—as no more than their genes. But as Collins notes, "free will will not go out of style" once the project is done.

**GENETIC TESTING**

**Many benefits, many perils**

Scientists mapping human DNA are rushing us headlong into a genome century, but the creaky American health care system is unprepared. A little genetic knowledge can be a dangerous thing. Already, there are 44 million uninsured, and the ill and disabled often find it hard to get coverage. With genetic testing certain to become more common, people with a genetic predisposition to diseases such as breast cancer or Alzheimer's could be vulnerable to losing health or life insurance. That is particularly troubling because the new tests will predict only the likelihood of inheriting a disease. Finding cures will lag behind. That's why fewer than 15 percent of those at risk for deadly Huntington's disease have sought the testing already available. And if a prenatal test shows risk, will insurers pressure a woman to abort a fetus? Can they deny coverage for the child if she doesn't? Says Boston University ethicist George Annas: "The right not to know is going to become as important as the right to know."

Also worrisome: In a nation of health care have-haves and have-nots, gene-based treatments may become an option only for the wealthy.

---

**ERIC LANDER**


---

**Joseph P. Shapiro**
Decoding the Future

The goal of the Human Genome Project is to decipher the genetic code—the instruction book for building a human being. That means working out the exact order of the roughly 3 billion chemical "letters" that make up the entire complement of DNA, or genome, including the 100,000 or so genes. Understanding this human instruction manual promises to transform biology and medicine in the 21st century.

GENETICS 101
The genetic code is written in DNA contained within each of our cells. In 1953 Francis Crick and James Watson cracked the code by revealing the "double helix" structure of the DNA molecule and how it directs the manufacture of proteins, which in turn orchestrate the body's physical processes.

THE MECHANICS OF SEQUENCING
To decipher the human genetic code, researchers have to determine the exact order of the 3 billion chemical letters that make up our DNA. Here is how they are doing it:

1 Whose DNA?
Researchers for the Human Genome Project collected semen or blood samples from six to 10 anonymous volunteers and extracted the DNA.

2 The book of life
The genome, the complete set of genetic instructions, is often called the "book of life." This book is arranged into 23 distinct volumes, or chromosomes.
1 Inside the nucleus
Within the cell nucleus, the DNA is organized into 23 pairs of chromosomes. Along the chromosomes, arrayed like beads on a necklace, are the genes—segments of DNA that contain instructions for making proteins.

2 Twisted pairs
The DNA molecule consists of two intertwined strands that look like a twisted ladder—the double helix. Each strand is composed of four chemical units: the nucleotide bases A, T, C, and G. The strands are held together by bonds between the bases: A always pairs with T, and C with G.

3 A recipe in code
The exact order of these letters in a gene determines the structure and function of the resulting protein—whether it is an oxygen-carrying hemoglobin, a germ-zapping antibody, or one of approximately 50,000 other “workhorses” of the cell.

3 Shredding the text
The goal is to figure out the exact order of every single letter in all of the volumes. Researchers are using two approaches. In one, they start with a volume, or chromosome, and then select one page from it. They shred that page into smaller pieces, making lots of DNA fragments. In the other approach, researchers shred all 23 volumes at once, making many more fragments. The sequencing process is the same either way.

4 Copying
Researchers need billions of copies of each DNA fragment for efficient analysis. To make them, they insert the human DNA fragments into bacteria, which act like copying machines. When fed the right nutrients, the bacteria make billions of copies of themselves—and the human DNA fragments—every day.
**Protein synthesis**
The DNA strand temporarily unwinds, and the instructions for making a protein are copied onto RNA, a DNA-like strand that serves as a template for building proteins.

**Tiny factories**
When the RNA is complete, it leaves the cell nucleus and enters the cytoplasm, where it attaches to a ribosome, a little protein factory. The ribosome reads the instructions on the RNA and connects amino acids together to build a protein. Proteins are involved in everything from metabolizing food to fighting disease.

**Genes**
- Genes are the basic units of heredity. A gene is a segment of DNA that contains the recipe for making a protein.

**Protein power**
The 20 different amino acids are the building blocks of proteins, which provide the structural components of cells and tissues.

**Tagging the DNA**
The human DNA is then extracted and treated with special fluorescent dyes. These dyes make each DNA base glow a different color under ultraviolet light.

**The sequencing machine**
Fast-moving robotic arms then drop the DNA fragments into 96 thin tubes, or capillaries, inside a sequencing machine. When an electric current is run through the DNA, the negatively charged DNA fragment is pulled toward a positive charge at the end of each tube. As DNA fragments emerge at the other end, a laser beam excites the fluorescent dyes, and a camera records the colors. A portion of the sequence has been captured.

**Putting it back together**
Powerful computers read the DNA fragments and look for overlapping patterns in the sequence. These patterns are used to assemble the fragments in the correct order, just like piecing together a torn page out of a book. When all the pieces are reassembled, scientists will have deciphered, for the first time, the complete genetic instructions for a human being—an achievement with profound implications for biology and medicine.
USING THE GENE SEQUENCE

Even before the Human Genome Project is complete, the information buried in the genes is being used to trace human origins and to uncover clues to human diseases. Many more applications are in the wings.

Human origins

By examining tiny changes in the DNA of modern-day populations, molecular anthropologists are tracking where humans came from, the routes they traveled, and how they are related. Clues hidden in DNA reveal a family tree in which all 6 billion people now living trace their ancestry back to a group of 50,000 humans who lived in Africa 150,000 years ago. DNA evidence is also illuminating more-recent history, such as the peopling of Europe.

DNA fingerprinting

On Aug. 8, 1986, in Leicestershire, England, Richard Buckland was arrested for the rape and murder of a 15-year-old girl. About three months later he was released—the first accused murderer to be freed on the basis of a DNA test. Since that time DNA tests have been used in a growing number of criminal investigations worldwide. By comparing the DNA left at a crime scene—in blood, semen, hair, or flecks of skin—with that of a suspect, scientists can determine with amazing precision whether the two match.

Fitting the drug to the patient

How a person responds to drugs is determined, in part, by tiny variations in the DNA. By analyzing a patient’s genetic profile, doctors may be able to determine in advance whether a drug will work, or whether it will cause dangerous side effects. This approach is already being applied to breast cancer. These techniques could usher in an era of “personalized” medicine, with drugs tailor-made for an individual’s genetic profile.

Preventive medicine

Understanding our genes and the role they play in common diseases should make possible a new type of preventive medicine. Most diseases, from cancer and heart disease to Alzheimer’s, have a genetic component, with several genes interacting to increase a person’s risk. As these genes are identified, it should be possible to design tests to determine a person’s chance of developing, say, diabetes or colon cancer. Knowing that risk could be a powerful impetus for taking steps to avoid the disease, perhaps by changing diet or getting regular screening tests.

A genome comparison

The amount of genetic material in living creatures varies widely. In this chart, each volume represents 50 million base pairs.

Rational drug design

Most pharmaceutical drugs are blunt weapons. Many were identified by trial and error, and scientists often know little about how they work. Cancer therapy, for instance, involves bombarding the body with toxic chemicals that damage normal cells along with the cancer cells, and cause dangerous side effects. Increased understanding of the genetic changes at work in cancer cells may make it possible to design “smart bomb” drugs that attack only cancerous cells.

Wings or arms?

Biologists are excited about their ability to compare the human DNA sequence with that of other organisms, from lowly fruit flies to worms or to mice. Key genes, such as those that regulate cell division, appear remarkably similar from species to species. By manipulating the genes in simpler creatures, scientists gain insights into fundamental biological mechanisms. And by comparing gene sequences among species, they can probe the mysteries of embryonic development.

Beyond the genes

The Human Genome Project is deciphering a “consensus” genome, composed of the DNA of a few individuals. But in reality there is no generic human genome. Each person is unique, with millions of distinguishing genetic changes that influence how we look, what diseases we get, and how we behave. As scientists delve into these variations, they will gain insights into what makes each person unique at the genetic level. But the mass of genetic data cannot explain each person’s personality or potential—those are factors far bigger than the genes.