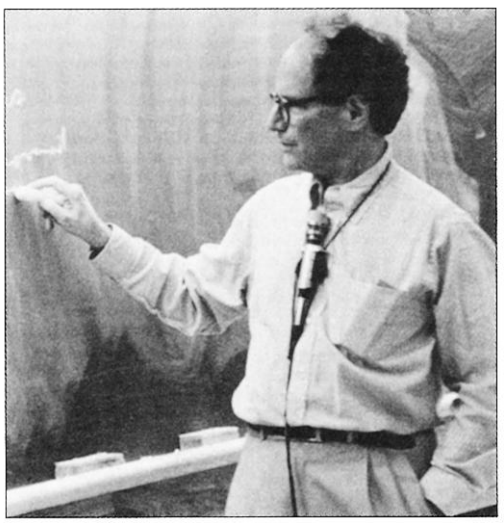
Controversial From The Start

The human genome: the crown jewel of 20th century biology, heralded at the White House, plastered on the covers of countless magazines—and at last spelled out today in intricate detail in both *Science* and *Nature*. Deciphering this string of 3 billion A's, T's, G's, and C's is being hailed as an achievement that will usher in a new era of biology and even alter our understanding of who we are.

That's a far cry from how the idea was greeted when it was first proposed 15 years ago. "Absurd," "dangerous," and "impossible," scoffed numerous critics, who noted that the technology did not exist to sequence a bacterium, much less a human. And even if the project's starry-eyed proponents could by some miracle pull it off, who would want the complete sequence data anyway?

It turns out a lot of people did. This once-ludicrous proposal became one of most hotly contested—and contentious—races in recent scientific history. Although the race has been dominated in the past few years by the acrimonious feud between the public and private teams, tensions go way back. And no wonder, with a prize this great and a project that has transcended and transformed traditional ways of doing biology. "The change is so fundamental, it is hard for even scientists to grasp," notes geneticist Maynard Olson of the University of Washington, Seattle, who ranks decoding the human genome as one of the biggest accomplishments ever in biology.



Walter Gilbert. A crucial early proponent, he later tried to set up a company to produce and sell genome data.

An impossible dream

One of the first to grasp that potential was Robert Sinsheimer, a biologist who was then chancellor of the University of California (UC), Santa Cruz. UC astronomers were already angling to build the world's biggest telescope, and Sinsheimer was looking for a project of similar magnitude in biology. Unraveling the sequence of the human genome might be just the ticket—if he could rally the scientific support and, of course, money. At the time, the largest genome yet sequenced was the minuscule Epstein-Barr virus—and that feat had taken several researchers years to complete. To apply such tools to the human genome, nearly 20,000 times bigger at 3 billion bases, was audacious beyond belief.

In 1985, Sinsheimer assembled some of the best minds in the nascent field of genome analysis to hash over the proposal at his idyllic campus, nestled in the hills above the sleepy beach town of Santa Cruz. John Sulston of Cambridge University and Robert Waterston of Washington University in St. Louis, who were already trying to map the genome of the nematode *Caenorhabditis elegans*, were there, as was Bart Barrell, head of large-scale sequencing at the U.K. Medical Research Council (MRC). So were genetic mappers David Botstein, then at the Massachusetts Institute of Technol-

Objection #1: Big Biology Is Bad Biology

The human genome project was biology's first foray into "big science," and many scientists abhorred the idea at the outset. Researchers feared that a massive sequencing project would siphon precious dollars from investigator-initiated research, destroying the cottage industry culture of biology in the process. And just as bad, the project didn't even amount to hypothesis-driven science at all. Rather, critics charged, it was no more than a big fishing expedition, a mindless factory project that no scientists in their right minds would join. Were they right?

Not exactly, says David Baltimore, president of the California Institute of Technology (Caltech) in Pasadena, who raised some of the early concerns. "One of the things I didn't fully anticipate was the state of progress in automation," he says. In the mid-1980s, gene sequencing was done by hand. Baltimore and others feared that it would take an army of "worker bees" to carry out sequencing on a genomewide scale. But sequencing machines pioneered by Leroy Hood and colleagues at Caltech changed that equation forever. Today, sequencing is nearly completely automated.

The genome project was still a fishing expedition, of course. But the enormous haul of genomic data it netted has changed most minds about such "discovery" research. This once-maligned type of research has enabled teams around the world to explore newfound genes and their links to health and disease. "Discovery science has absolutely revolutionized biology," says Hood, now director of the Institute for Systems Biology in Seattle, Washington. "It's given us new tools for doing hypothesis-driven research," maintains Hood, and these tools help rather than hinder individual investigators.

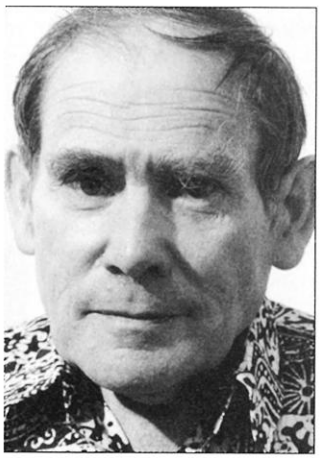
The biggest objection to the audacious proposal was that funding for the genome project would come at the expense of other quality science. "There was a worry that it was a zero-sum game," says Maynard Olson, a genome center leader at the University of Washington, Seattle. "Frankly, it was a gamble that we'd be able to expand the pie [of research dollars]." But the gamble paid off. In a 1998 National Research Council report, a committee led by Bruce Alberts, a former professor at the University of California, San Francisco, recommended that the human genome project be funded separately from traditional science budgets. And Congress happily went along, giving the Department of Energy $10.7 million and the National Institutes of Health $17.2 million for the new project in fiscal year 1988.

By voicing the early concerns, "I think we did what we hoped we would do," says Baltimore. "It helped develop a debate, which set us on a productive course." –ROBERT F. SERVICE

CREDIT: VICTOR A. MCKUSICK

UNSUNG HERO: PHIL GREEN

Phil Green, a mathematician and software designer, wrote the phred and phrap programs at Washington University in St. Louis, Missouri. These became essential tools for evaluating the quality of raw DNA sequence and linking up assemblies. He's now at the University of Washington, Seattle, creating new programs.



Sydney Brenner. Joked that sequencing was so boring it should be done by prisoners.

ogy (MIT), Helen Donis-Keller, then at Collaborative Research Inc., and sequencing aficionados Walter Gilbert and George Church of Harvard University and Leroy Hood of the California Institute of Technology in Pasadena. Their collective conclusion: bold, exciting—but simply not feasible. Sinsheimer's proposal for a genome institute at Santa Cruz died, but not before it had captured Gilbert's imagination.

Gilbert soon became the proposal's biggest champion, and his support meant the idea could no longer be blithely dismissed. A decade earlier, Gilbert and Allan Maxam, also at Harvard University, had invented a brand-new technique that enabled scientists for the first time to determine the genetic sequence of an organism. (Gilbert went on to share the Nobel Prize with Fred Sanger of Cambridge University, who independently invented a similar technique.) And he soon won over another giant of molecular biology: James Watson, who shared a Nobel Prize with Francis Crick and Maurice Wilkins for their 1953 discovery of the double helical structure of DNA.

The ambitious idea had also captivated Charles DeLisi, a cancer biologist who was then head of the Office of Health and Environmental Research at the Department of Energy (DOE). To DeLisi, the genome project was a logical outgrowth of DOE's mandate to study the effects of radiation on human health. Another equally compelling rationale—but one DeLisi did not openly tout—was that a massive new endeavor could provide new focus for DOE's national labs, whose bombmaking skills were in diminishing demand.

At the urging of DeLisi and DOE colleague David Smith, the Los Alamos National Laboratory hosted a workshop in Santa Fe, New Mexico, in March 1986 where the excitement was palpable. The idea quickly gained momentum, dominating discussion at a June meeting at Watson's Cold Spring Harbor Laboratory in New York. By then, biologists were beginning to think the project just might be doable. But whether it was worth doing was another matter (*Science*, 27 June 1986, p. 1598).

To many, like Botstein and Nobel laureate David Baltimore, then at MIT, the project ran counter to the way biology had been conducted for decades. The best work, the mantra went, came from investigator-initiated studies in small labs, not from some massive, goal-driven effort. Moreover, this was technology development, not experimental biology, and it would be mind-numbingly dull. Sydney Brenner of the MRC facetiously suggested that project leaders parcel out the job to prisoners as punishment—the more heinous the crime, the bigger the chromosome they would have to decipher. What was truly horrifying was the price tag, which was quickly estimated at $3 billion—a number that stuck through countless reports ever since. If the National Institutes of Health (NIH) were to foot the bill, the megaproject would rob funds from the rest of biology, the critics asserted. "It endangers all of us, especially the young researchers," warned Botstein.

The scientific value seemed dubious as well. Although many biologists agreed that maps of the chromosomes would be useful for finding genes, what good would come from deciphering every A, T, G, and C, especially since most of them were "junk" that did not code for genes. The sequence might be handy to have, but "was it worth the cost, not in terms of dollars but in terms of its impact on the rest of biological science?"

CREDITS: (TOP TO BOTTOM) HHMI; MRC; ILLUSTRATION: CAMERON SLAYDEN

NEW SCIENCE:

Finding the Talismans That Protect Against Infection

Since 1995, the minigenomes of dozens of pathogenic microbes have been sequenced, including those that cause tuberculosis, cholera, and ulcers. Many others are almost in the bag, including the much larger genome of *Plasmodium*, the malaria parasite. That data flood is helping researchers understand how nefarious microorganisms work—and how they might be stopped.

The giant human genome promises to help solve another poorly understood problem: why some people get sick and die when they encounter a pathogen, whereas others stay healthy as an ox. Such information could eventually help put more people in the latter category.

Researchers have long known that differences in disease susceptibility are partly genetic, the most famous example being the

gene for sickle cell hemoglobin, which offers protection against malaria to those who inherit one copy of it. (Having two copies causes sickle cell anemia.) Several other susceptibility genes have been discovered for various diseases; malaria now tops the list with 14 genes. "We're just beginning to scratch the surface," says Adrian Hill, a geneticist at the University of Oxford in the United Kingdom.

To identify genes that might confer susceptibility or resistance, researchers try to find genetic differences between large groups of patients and healthy controls. Without the complete genome, they could only look for previously discovered genes. Now, they can theoretically take each and every gene into consideration. Eventually, such work will lead to a better understanding of the molecular interaction between a bug and its host. That, in turn, may reveal new drug or vaccine targets.

–MARTIN ENSERINK

Objection #2: Why Sequence The Junk?

Genes and their corresponding proteins get most of the attention, but they make up only a tiny fraction—1.5% or less—of the human genome. The other 98% of DNA sequence that does not code directly for proteins was once dismissed as "junk DNA," and numerous researchers argued that it would be a waste of time and money to include the repetitive, hard-to-sequence regions in the genome project. But scientists have discovered many riches hidden in the junk, and as the project nears completion, several researchers predict that some of the most intriguing discoveries may come from areas once written off as genetic wastelands.

Included among the noncoding DNA, for example, are the crucial promoter sequences, which control when a gene is turned on or off. The repetitive sequences at the ends of chromosomes, called telomeres, prevent the ends of the chromosome from fraying during cell division and help determine a cell's life-span. And several teams have begun to make a strong case that repetitive, noncoding sequences play a crucial role in X inactivation, the process by which one of the two X chromosomes in a female is turned off early in development. Other genes are turning up in areas previously dismissed as barren. Scientists had assumed, for example, that the regions next to telomeres were buffer zones with few important sequences. But in this week's issue of *Nature*, H. C. Reithman of the Wistar Institute in Philadelphia and his colleagues report that these regions contain hundreds of genes. "The term 'junk DNA' is a reflection of our ignorance," says Evan Eichler of Case Western Reserve University in Cleveland.

The human genome has much more noncoding DNA than any other animal sequenced so far. No one yet knows why. At least half of the noncoding DNA seems to be recognizable repeated sequences—perhaps genomic parasites that invaded the genomes of human ancestors. Eichler suspects that such repeats might provide some genomic wiggle room. Long stretches of noncoding DNA provide "a built-in plasticity that may be bad at the individual level, but if an organism is going to evolve, it may be a huge selective advantage," he says.

"There is a rich record of our history" in the repeats, agrees Francis Collins of the National Human Genome Research Institute in Bethesda, Maryland. "It's like looking into our genome and finding a fossil record, seeing what came and went."

–GRETCHEN VOGEL

asked Paul Berg of Stanford University.

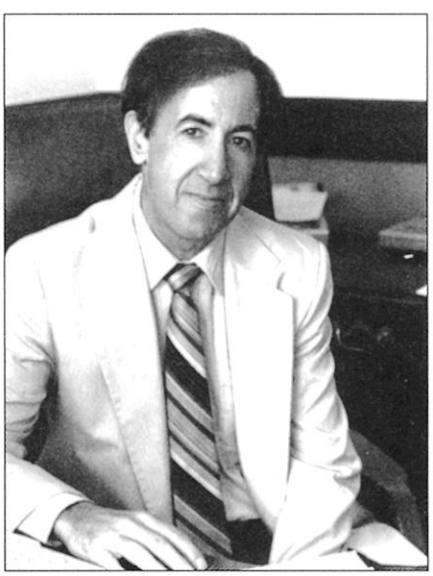
As the biology community wrestled with the merits of the project, NIH staked out a position firmly on the fence. By contrast, DeLisi and Smith were decidedly gung ho. DeLisi aggressively gained support for the project, first from his superiors at DOE and then from Congress, starting a small Human Genome Initiative within DOE in 1986. The following year, a prestigious advisory panel to DOE called for an all-out effort and urged the agency to take the lead. DOE was the logical choice, DeLisi argued, because this was "big science," DOE's stock-in-trade, whereas NIH had never attempted a project of this scope (*Science*, 8 August 1986, p. 620; 31 July 1987, p. 486).

The fact that DOE—not NIH—was lobbying for the project only heightened some biologists' unease, because they put great store in NIH's peer-review system. "The fear is not big science so much as bad science," said Botstein, who in 1986 denounced DOE's proposal as "a scheme for unemployed bombmakers."

Emerging consensus

Political posturing continued until 1988, when a National Research Council (NRC) panel gave the project its official seal of approval (*Science*, 12 February 1988, p. 725). Chaired by Bruce Alberts, then at UC San Francisco, the panel contained some of the project's staunchest advocates, such as Gilbert and Watson, and also some skeptics, including Botstein, mouse geneticist Shirley Tilghman of Princeton University, and yeast expert Olson, then at Washington University in St. Louis. Within a year, the panel endorsed the project unanimously, calling for a rapid scale-up in "new and distinctive" funds to $200 million a year over the next 15 years.

In the process, the panel redefined the project, laying out a phased approach that mollified critics and has guided the initiative ever since. Rather than plunge into sequencing—which no one knew how to do on a massive scale anyway—the project should begin by constructing maps of the human chromosomes. These would greatly speed the search for disease genes, offering immediate medical payoffs. The panel recommended that full-scale sequencing be postponed until new technologies made it faster and cheaper.



Charles DeLisi. An early advocate, he launched the Human Genome Initiative within the Department of Energy in 1986.

But it was the panel's recommendation to analyze the genomes of simple organisms, such as *Escherichia coli*, yeast, and the roundworm *C. elegans*, and eventually the mouse, that proved most persuasive. Tilghman and Botstein, in particular, argued vociferously that biologists had no hope of understanding the human genome if they couldn't compare it to the genomes of experimental organisms. Luckily for biologists, evolution has been remarkably conservative, retaining the same genes over and over again in different organisms, explains Tilghman—and it is far easier to figure out a gene's function by experimenting with it in a fruit fly than in a human. Looking back, Tilghman sees this as one of the panel's smartest decisions: "Model organisms were an extraordinary investment. We learned how to sequence on these simpler organisms. And more important, we got a preview of the human genome by sequencing these organisms."

Gilbert, however, was impatient with the panel's cautious approach and with the interagency dithering. Arguing that the technology was already good enough to sequence the human genome, he left the NRC panel to launch his own company, Genome Corp. His plan, remarkably similar to J. Craig Venter's vision a decade later, was to set up a sequencing factory to churn out the data, which he intended to copyright and sell. "[It will be] available to everyone ... for a price," he explained (*Science*, 24 July 1987, p. 358). The plan infuriated Watson, who rankled at the idea of selling something as

fundamental as data on human DNA. But the debate subsided when Gilbert failed to raise sufficient funds.

NIH makes its move

As the genome project gained congressional funding and scientific respectability, NIH wrested control from DOE. Urged on by a group of advisers who met outside Washington, D.C., in Reston, Virginia, in March 1988, then–NIH director James Wyngaarden announced that NIH would create a special office for genome research (*Science*, 13 May 1988, p. 878). In short order, he nabbed Watson to head it, and with that coup, NIH was firmly ensconced as the lead agency. It has remained so, even as the project gathered international collaborators and Britain's Wellcome Trust took on a prominent role.

Watson proved a shrewd strategist, skilled in the care and feeding of those who controlled congressional purse strings, and a tough taskmaster. "My name was good," he says by way of explanation. Indeed, members of Congress were spellbound when the eccentric Nobel laureate swept in to testify. Watson was eloquent in touting the project's goal: "to find out what being human is." He also had the refreshing quality of saying what he thought, no matter how politically incorrect—an unusual quality in Washington, D.C.

Even as the project began, Watson's advisory panel was still debating the proper balance for the project—how much should be devoted to building tools, like maps and faster sequencing machines, and how much to actually using these tools to find disease genes? (*Science*, 13 January 1989, p. 167) Watson was adamant: Even though disease genes captured the public imagination and kept the dollars flowing, this project was designed to build the equivalent of a particle accelerator: They should not be sidetracked. As Botstein explained at a January 1989 meeting, "We are looking at the production of a set of tools that will enable human geneticists to do what they want. We are the Cray, if you like. We don't write software for your particular applications."

At the same time, Watson relentlessly pushed the first stage of the project and its most tangible goal—building maps of the human chromosomes. Knowing that Congress did not have the patience to wait 15 years for results, Watson staked his reputation on getting the maps done in five. With the maps in hand, genes would fall out in short order, including the putative Alzheimer's gene, which, Watson joked, should be a priority given the age of most members of Congress.

Progress was rapid. By 1990, Sulston and colleagues had nearly completed the physical map of the worm—changing worm biology forever—and Olson and colleagues were proceeding apace on yeast (*Science*, 15 June 1990, p. 1310). Faster and easier ways to clone and map DNA were coming on line, and sequencing trials were beginning. For a short time, the controversy that had dogged the project from the outset seemed to have dissipated.

Venter, round one

That newfound harmony was shattered in June 1991, when Venter, who ran a large sequencing lab at the National Institute for Neurological Disorders and Stroke, went public with an iconoclastic plan: Why not focus on finding the genes—the "real goods" that both scientists and companies were clamoring for—and leave tedious sequencing until later? Venter and colleague Mark Adams had developed a new technique, called expressed sequence tags, that enabled them to find genes at unprecedented speed. Never one of Watson's inner circle, Venter boasted that this new approach "was a bargain in comparison to the genome project" and claimed he could find 80% to 90% of the genes within a few years, for a fraction of the cost (*Science*, 21 June 1991, p. 1618).

Watson dismissed Venter's "cream-skimming approach," but their feud remained subterranean until a few weeks later, when Venter described his work at a

NEW SCIENCE:

Nailing Down Cancer Culprits

A general sending troops out to battle wants as much intelligence about the enemy and its weaknesses as possible. Researchers fighting cancer hope the complete human genome sequence will help provide such information.

The sequence will greatly speed the identification of the genetic underpinnings of cancer. Over the past 15 years or so, researchers have learned that cancers are usually caused by the accumulation of several gene mutations, some of which activate cancer-promoting oncogenes, whereas others inactivate tumor suppressor genes. And though scientists have fingered roughly 100 oncogenes and 30 or so tumor suppressors, that's "only a fraction of the genes that cause cancer," says cancer gene expert Bert Vogelstein of the Johns Hopkins University School of Medicine in Baltimore, Maryland.

In the past, once researchers determined where in the genome a cancer gene resides, they could still spend months, or even years, scouring the region—often a megabase or two long—looking for likely candidate genes to test. Now, Vogelstein says, that can be done "literally with the click of a button. The availability of the sequence enormously simplifies the search for those [missing cancer] genes."

Researchers are also using microarrays and other techniques to measure changes in the expression of thousands of genes at a time—information that provides a very detailed picture of the alterations leading to cancer development and spread. Knowing all the human genes will make this picture more complete. Researchers have already found that tumors that look similar to a pathologist may display different gene expression patterns—and that these differences can reveal potentially lifesaving information about how the cancers will respond to therapy.

–JEAN MARX

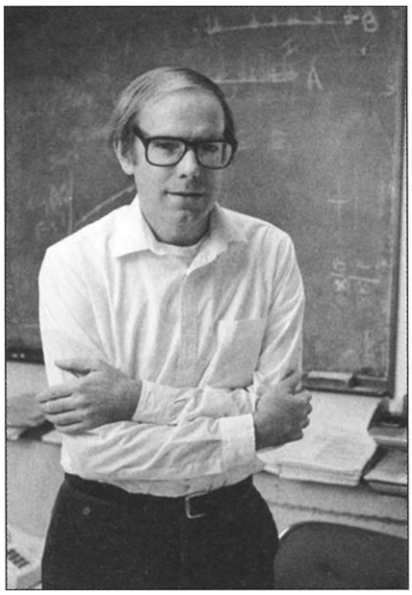
UNSUNG HEROES: MEL SIMON & PIETER DE JONG

Although they were slow to win acceptance, the bacterial artificial chromosomes (BACs) created by geneticist Simon (left) of the California Institute of Technology in Pasadena soon became the "currency of the genome," as he says. These clones' large capacity and stability make them highly efficient. Using BACs, Caltech's de Jong created massive "libraries" of DNA from various human tissues for sequencing.

CREDITS: (LEFT TO RIGHT) ILLUSTRATION: CAMERON SLAYDEN; CALIFORNIA INSTITUTE OF TECHNOLOGY; SUSAN SPANN PHOTOGRAPHY

congressional hearing. NIH was so impressed with his progress, Venter said, that it was filing patent applications on the partial genes he was identifying—at a rate of 1000 a month.

Watson erupted, denouncing the patenting scheme as "sheer lunacy" and noting that "virtually any monkey" could do what Venter's group was doing (*Science*, 11 October 1991, p. 184). What irked him was that Venter and NIH had no clue about the function of the genes from which these fragments came. If the patents held, that meant anybody could lay claim to most of the human genes, undercutting patent protection for biologists who labored long and hard to identify whole genes and figure out what they did. "I am horrified," Watson told Congress.



Maynard Olson. Helped pave the way with work on mapping the yeast genome.

Watson also went to war on this issue with his boss, NIH Director Bernadine Healy. The fight cost him his job. In April 1992 he returned to Cold Spring Harbor Laboratory, muttering that no one could work with that woman (*Science*, 17 April 1992, p. 301).

Venter, too, left NIH in 1991 when he was offered $70 million from a venture capital company to try out his gene identification strategy at a new nonprofit, The Institute for Genomic Research (TIGR).

From tools to medicine

After Watson's sudden departure, NIH picked gene hunter Francis Collins of the University of Michigan, Ann Arbor, to take the helm. Fresh from the heady success of finding several elusive genes—including those involved in cystic fibrosis, neurofibromatosis, and Huntington's disease—Collins was then in a highly competitive race to find the gene involved in a form of inherited breast cancer.

A physician by training, Collins brought a different perspective to the genome project, placing its medical applications front and center. Collins charmed Congress and the media by riding to work on his motorcycle and playing guitar in a pick-up rock band. Whereas Watson and his advisers had spoken of creating a tool, Collins talked about saving children's lives. "The reason the public pays and is excited—well, disease genes are at the top of the list," he explained.

It was a heyday for gene hunters. The early investments in the genome project paid off as increasingly sophisticated maps of the human and mouse genomes were compiled (*Science*, 1 October 1993, p. 20). With these maps in hand, the time it took to track down most disease genes dropped from a decade to perhaps 2 years. Every week, it seemed, another deadly disease gene was discovered. Lost in the hoopla, however, was the fact that finding a gene was a far cry from having a treatment, much less a cure. The consortium was growing as well, fueled by an infusion of funds from the Wellcome Trust, which in 1993 set up a major new sequencing lab, the Sanger Centre near Cambridge, with Sulston as its head.

But sequencing overall was lagging behind. At the existing rate and cost, Collins lamented when he took on the job, there was no chance they could finish the sequencing by 2005. None of the "blue sky" sequencing technologies that had been imagined at the outset materialized, and with U.S. funding tight and much of the money concentrated on mapping, Collins was worried that "we have mortgaged part of our future."

Steady, incremental advances were enabling scientists to spew out longer "sequence reads," and the cost was slowly dropping. Even so, reassembling the DNA fragments in correct order was tricky. To do so, the sequencers looked for similar patterns in the fragments—much like assembling a jigsaw puzzle—but one with lots of missing pieces. Some pieces just wouldn't fit, some "fit" in the wrong place—others "got lost" in the cloning process. Still others refused to be sequenced.

Sequencing clearly needed a shot in the arm—and soon got one, but from an unlikely source. In 1995, Venter surprised the community by announcing that along with Hamilton Smith, then at Johns Hopkins, and TIGR colleagues Rob Fleischmann and Claire Fraser, they had sequenced the first entire genome of a free-living organism, *Haemophilus influenzae*, at 1.8 megabases (*Science*, 28 July 1995, p. 496). What's more, they had done it in just a year using a bold new approach, whole-genome shotgun sequencing, that NIH had insisted wouldn't work and wouldn't fund.

Sequencers in the publicly funded project had adopted a conservative, methodical approach—starting with relatively small chunks of DNA whose positions on the chromosome were known, breaking them into pieces, then randomly selecting and sequencing those pieces and finally reassembling them. Eventually, larger pieces called contigs would be hooked together. By contrast, Venter simply shredded the entire genome into small fragments and used a computer to reassemble the sequenced pieces by looking for overlapping ends.

CREDIT: HERB WEITMAN/WASHINGTON UNIVERSITY

Objection #3: Impossible to Do

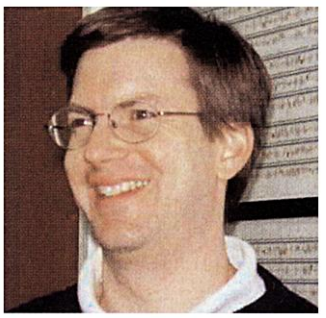
Perhaps the most surprising thing about the human genome project is that it was begun at all. In the mid-1980s, the technology for decoding DNA's sequence of chemical bases was in its relative infancy. State-of-the-art labs could sequence only about 500 bases a day, working day in and day out. And the computer technology that came to play such a vital role in the project wasn't even invented yet. "In retrospect, the optimism that the project could be done on a 15-year timetable was striking," says Maynard Olson, who directs a sequencing center at the University of Washington, Seattle.

Unexpectedly, however, says Stanford University geneticist David Botstein, sequencing technology didn't need a revolution to make the leap in speed. "In the early days, it was believed that a radical new technology would be required" to sequence the full human genome, says Botstein. "But it didn't turn out that way."

Incremental but vital improvements in manipulating DNA and chemical probes enabled researchers to switch from identifying bases with radioactive probes to fluorescent ones. That eased the way for detectors to read and catalog the sequence of bases automatically. That automation was then honed with the advent of high-speed machines that pushed snippets of DNA through dozens of capillaries, reducing the sequencing time and cost of reagents. "It was definitely evolution," says molecular biologist David Baltimore, president of the California Institute of Technology in Pasadena. "But you can go a long way with evolution."

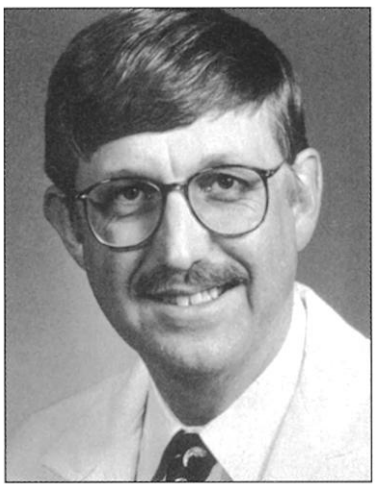
–R.F.S.

UNSUNG HERO: MARK ADAMS



Ever since he teamed up with J. Craig Venter at the National Institutes of Health (NIH) in 1990, Adams has been one of the country's top sequencing gurus. After developing expressed sequence tags with Venter at NIH, Adams followed him to The Institute for Genomic Research (TIGR) in Rockville, Maryland, and then to Celera, also in Rockville, where he is refining methods for whole-genome shotgun sequencing.

NIH's deliberate approach won its spurs a year later, when an international consortium knocked off the yeast genome. Although still tiny, relative to humans, it was a major step up in size and complexity. By April 1996, Waterston and Sulston, who were well into sequencing *C. elegans*, were champing at the bit, urging Collins to let them plunge into all-out sequencing. In the right hands, they argued, the technology was good enough; the only stumbling block was money. "Just do it," Sulston urged at the time. The two also broached the heretical topic of dropping the accuracy goal to speed the process, from 99.99% to 99.9% (*Science*, 12 April 1996, p. 188).



Francis Collins. Favored a deliberate, methodical approach to mapping and sequencing.

But Collins would not be rushed. The goal was to assemble the definitive "book of life," and he insisted it be done to the highest possible quality. He decided to test the water with six pilot projects—a cautious style that earned him praise in some corners and criticism in others. The charge to the labs was to complete a major chunk of sequence while also demonstrating big improvements in cost and speed. After that, he said, the project would home in on its final strategy.

Collins soon abandoned his measured approach—not because of the persuasiveness of Waterston and Sulston's arguments, but because Venter threw down the gauntlet.

Venter redux

Showing a knack for impeccable timing, Venter dropped his bombshell on 9 May 1998, just days before the annual gathering of genome scientists at Cold Spring Harbor Laboratory. Venter announced that he had teamed up with Perkin-Elmer Corp., which was about to unveil an advanced, automated sequencing machine, to create a new company that would single-handedly sequence the entire human genome in just 3 years—and for a mere $300 million (*Science*, 15 May 1998, p. 994). What's more, said Venter, when he was done he would give the data away free to the community by posting it on his company's Web site. The company, soon to be named Celera Genomics and located in Rockville, Maryland, would make money not from the raw data, he explained, but from the analysis it would perform and sell to subscribers. Venter proposed to sequence the genome with the brute-force shotgun technique that had worked so well in *Haemophilus*—but this time, he would be shredding the entire 3-billion-base genome into zillions of fragments.

Leaders of the public project were angry and incredulous. After they had spent years laying the groundwork, could Venter really beat them to the finish and steal the glory? They were also deeply worried that if Congress fell for Venter's bravado, it might pull the plug on the public project. Venter's plan would never work, they countered—the sequence would be riddled with holes and impossible to reassemble.

Yet as they disparaged Venter's claim, they could not dismiss it. Venter had surprised them before. And this time, he had a hefty bankroll and 300 of Perkin-Elmer's sequencing machines, just then rolling off the assembly line at $300,000 a pop. And to reassemble his sequenced fragments, Venter would use one of the world's fastest supercomputers.

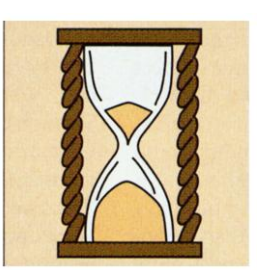
The leaders of the public program wasted no time in increasing the pace and reorienting the game plan in an attempt to beat him to the finish line. Collins announced new goals for the public project in September 1998, just 6 months after Venter's surprise announcement (*Science*, 18 September 1998, p. 1774). First, the consortium would complete the entire genome by 2003—2 years ahead of schedule, but also 2 years behind Venter. And, in a dramatic departure from previous philosophy,

NEW SCIENCE:

A Parakeet Genome Project?

Researcher William Haseltine, head of Human Genome Sciences Inc. in Rockville, Maryland, likes to claim that knowledge from the human genome, combined with a few technology breakthroughs, will someday enable humans to live forever. Most researchers who study aging have more modest expectations—for example, trolling the genome for new insights into genes involved in so-called oxidative damage to cells and genes, which is thought to limit an organism's life-span.

A few in the field have another request: sequence the parakeet. One avian genome, the chicken's, is in progress, but George Martin of the University of Washington, Seattle, and Steven Austad of the University of Idaho says aging research



could gain key insights from comparing the genome of a "real flier" with that of humans. "Good flying birds have remarkably long life spans for their size," he says: Some can live for 20 years or more. At the same time, they use an enormous amount of energy—a process that researchers believe is at the root of oxidative damage. Mice, for example, use much less energy but typically live only 2 years. A parakeet genome project, Martin says, could tell scientists "what the birds are doing that's so great"—and how humans might mimic their secrets.

–GRETCHEN VOGEL

CREDITS: (TOP TO BOTTOM) CELERA; VICTOR A. MCKUSICK; ILLUSTRATION: CAMERON SLAYDEN

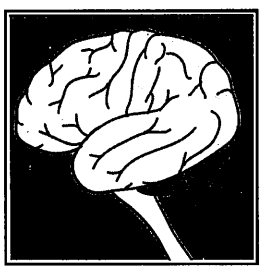
NEW SCIENCE:

Brain Calls Dibs on Many Genes

The human brain is an expensive tool: A huge proportion of human genes are thought to be involved in constructing, wiring up, and maintaining the nervous system. Neuroscientists hope the completed genome will help them to nail down the brain's share. Current estimates range from "a fair chunk" of the genome to "40%" to "most."

No one knows what all these genes do, but placing them on gene chips to see which ones are expressed by developing neurons is like "having a new type of microscope, a new way of looking at cells," says neurobiologist Ben Barres of Stanford University. His team is using such chips, as well as protein analysis, to spot molecular signals passed between neurons and support cells called glia early in development, when neurons start transmitting messages.

The completed genome will also accelerate the search for genes at fault in neurodegenerative diseases. Neurogeneticist



Huda Zoghbi of Baylor College of Medicine in Houston looks for candidate genes in the *Drosophila* genome, then tries to find homologs in the human sequence. Making the jump from fruit fly to human used to take a year of lab time, she says; now she'll be able to search computerized databases to find candidate genes in minutes.

Other neuroscientists hope the genome will help solve otherwise intractable questions about human behavior. For example, psychiatrist Eric Nestler of the University of Texas Southwestern Medical Center in Dallas and computational biologist David Landsman of the National Library of Medicine in Bethesda, Maryland, point out in this week's issue of *Nature* that newly identified genes might help make sense of addiction. Cocaine acts on certain dopamine transporters, which differ between people; correlating people's transporter subtypes with their propensity for cocaine addiction might reveal why some people are more vulnerable to the drug than others, they suggest.

–LAURA HELMUTH

the project would produce a "rough draft," covering 90% of the genome, by the spring of 2001. Scientists were clamoring for the data even in rough form, Collins said by way of explanation. Yet he also admitted that producing a rough draft and making it public was a strategic move to undercut any patent position Celera or other businesses might claim.

In a crucial test of the shotgun strategy, Celera first tackled the 180-megabase genome of the fruit fly *Drosophila melanogaster*. Venter teamed up with a publicly funded team headed by Gerald Rubin of UC Berkeley, and by March 2000, they had pulled it off. This proved that the shotgun methods could work on a big, complex genome, said Venter (*Science*, 25 February 2000, p. 1374).

The race was on, punctuated by dueling press releases. First Venter announced in October 1999 that his crew had sequenced 1 billion bases of the human genome—a feat pooh-poohed by NIH, which noted that Celera hadn't released the data for other researchers to check. Then NIH jumped into the game, announcing in November that it had completed 1 billion bases, holding a "birthday" party at the National Academy of Sciences, complete with balloons and T-shirts emblazoned with the double helix. Venter countered in January 2000 that his crew had compiled DNA sequence covering 90% of the human genome, the public consortium asserted in March that it had completed 2 billion bases, and so on. Issues of data access heated up too, with the public consortium denouncing Venter for his plan to release his data on the Celera Web site rather than in GenBank, the public database. The feud became increasingly ugly, with each side disparaging the other's work and credibility in the press. Leaders in the scientific community urged them to stop squabbling and work together.

The two had, in fact, begun talking about a possible collaboration in December 1999. Eric Lander, who runs the Whitehead/MIT Genome Center, was the main go-between. The two approaches are complementary, he said, and collaborating would speed the process.

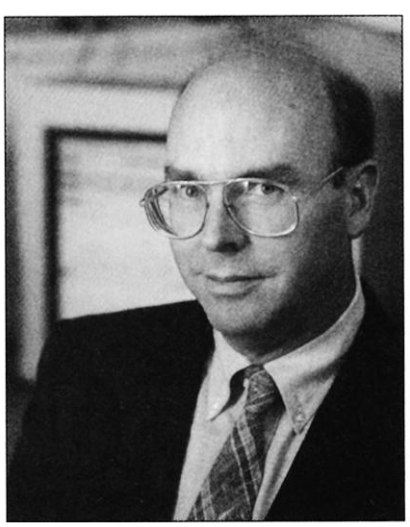
But in March, the discussions foundered amid considerable acrimony when the Wellcome Trust leaked to the press a letter from Collins to Venter, citing irreconcilable differences (*Science*, 10 March 2000, p. 1723). The sniping, seemingly at its peak, escalated further, until many considered it an embarrassment. "If they were my children, I would give them both a time out," said one leading scientist at the time.

Behind the scenes, Ari Patrinos of DOE played intermediary, finally brokering a truce under which both groups would announce their drafts at the same time, thereby sharing the glory. Venter still would not deposit his data in GenBank, as the consortium wanted, but he did concede that the public data has been useful in his own work. Defusing the issue of priority and credit, the two agreed to publish simultaneously, perhaps even in the same journal. Collins and Venter granted an exclusive interview to *Time*, which heralded, "The race is over," and pictured the beaming duo side by side in their lab coats. They were all smiles, too, at a White House ceremony in June where President Clinton lauded both scientists for their phenomenal achievement, and Collins and Venter lavished praise on one another (*Science*, 30 June 2000, p. 2294).

The façade held for 5 months—longer than many would have predicted—before all hell broke loose over plans to publish their papers (see p. 1189). At issue, again, was Venter's refusal to deposit his data in GenBank and the terms he might impose on commercial or academic users (*Science*, 15 December 2000, p. 2042). The two did manage to achieve simultaneous publications—but in separate journals.

In their magnanimous moments, both concede that their race has speeded the project, to everyone's benefit. "Ten, 15 years from now, nobody is going to care about all this fuss and bother," says Collins. "They're going to care that we got the fly sequence done, and shortly after that we got the human sequence done, and shortly after that we got the mouse sequence done. And all this back and forthing over who did what and what strategy was used and which money was public and which was private is probably going to sink below the radar screen. And hallelujah."

–LESLIE ROBERTS



J. Craig Venter. Threw down the gauntlet with his commercial plan to shotgun sequence the human genome.

CREDITS: (TOP TO BOTTOM) ILLUSTRATION: CAMERON SLAYDEN; NINDS