Withdrawal/Redaction Sheet
Clinton Library

DOCUMENT NO. AND TYPE	SUBJECT/TITLE	DATE	RESTRICTION
~~001a. note~~	~~re: Genome (1 page)~~	~~04/11/2000~~	~~P5~~ Duus 4/6/11 S
001b. letter	Thomas Schneider to President Clinton (5 pages)	03/28/2000	P5, P6/b(6)
002. letter	Thomas Schneider to President Clinton (3 pages)	03/13/2000	P5, P6/b(6)

COLLECTION:
Clinton Presidential Records
Domestic Policy Council
Devorah Adler
OA/Box Number: 20464

FOLDER TITLE:
Genome

2012-0463-S
rc758

RESTRICTION CODES

Presidential Records Act - [44 U.S.C. 2204(a)]

P1 National Security Classified Information [(a)(1) of the PRA]
P2 Relating to the appointment to Federal office [(a)(2) of the PRA]
P3 Release would violate a Federal statute [(a)(3) of the PRA]
P4 Release would disclose trade secrets or confidential commercial or financial information [(a)(4) of the PRA]
P5 Release would disclose confidential advice between the President and his advisors, or between such advisors [a)(5) of the PRA]
P6 Release would constitute a clearly unwarranted invasion of personal privacy [(a)(6) of the PRA]

C. Closed in accordance with restrictions contained in donor's deed of gift.
PRM. Personal record misfile defined in accordance with 44 U.S.C. 2201(3).
RR. Document will be reviewed upon request.

Freedom of Information Act - [5 U.S.C. 552(b)]

b(1) National security classified information [(b)(1) of the FOIA]
b(2) Release would disclose internal personnel rules and practices of an agency [(b)(2) of the FOIA]
b(3) Release would violate a Federal statute [(b)(3) of the FOIA]
b(4) Release would disclose trade secrets or confidential or financial information [(b)(4) of the FOIA]
b(6) Release would constitute a clearly unwarranted invasion of personal privacy [(b)(6) of the FOIA]
b(7) Release would disclose information compiled for law enforcement purposes [(b)(7) of the FOIA]
b(8) Release would disclose information concerning the regulation of financial institutions [(b)(8) of the FOIA]
b(9) Release would disclose geological or geophysical information concerning wells [(b)(9) of the FOIA]

THE PRESIDENT HAS SEEN
9-7-99

cc: CHRIS, ERIC, SARAH + file

The New Y ay, August 29, 1999

Money&I

Section 3

Chris –
Is there anything we should do on this? It dwarfs stem cells as a potential issue. 2000 SOTU? BR

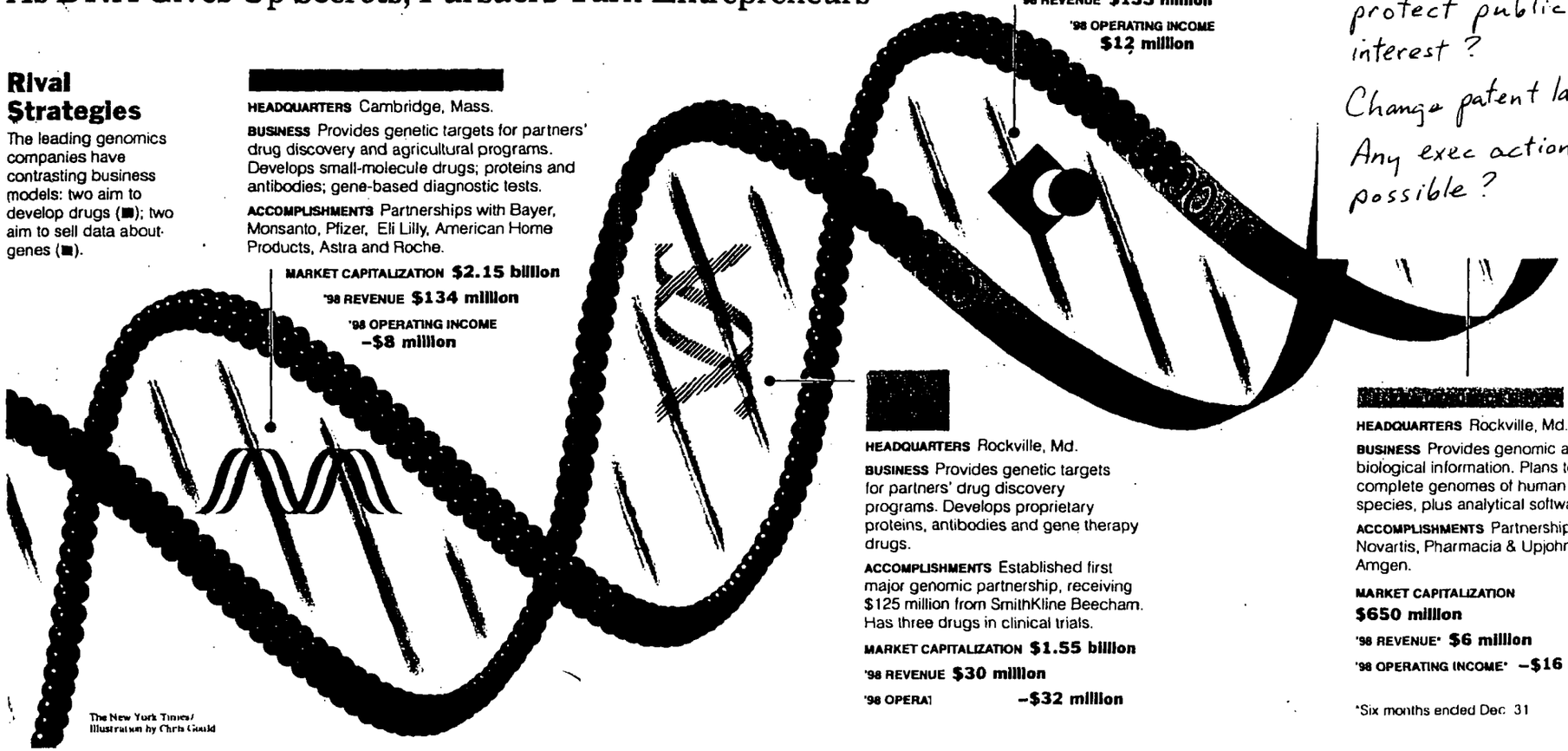
Shouldn't we do something to protect public interest? Change patent laws? Any exec action possible?

The Race to Cash In On the Genetic Code

As DNA Gives Up Secrets, Pursuers Turn Entrepreneurs

Rival Strategies

The leading genomics companies have contrasting business models: two aim to develop drugs (■); two aim to sell data about genes (■).



HEADQUARTERS Cambridge, Mass.

BUSINESS Provides genetic targets for partners' drug discovery and agricultural programs. Develops small-molecule drugs; proteins and antibodies; gene-based diagnostic tests.

ACCOMPLISHMENTS Partnerships with Bayer, Monsanto, Pfizer, Eli Lilly, American Home Products, Astra and Roche.

MARKET CAPITALIZATION **$2.15 billion**

'98 REVENUE **$134 million**

'98 OPERATING INCOME **–$8 million**

HEADQUARTERS Palo Alto, Calif.

BUSINESS Provides data bases of gene sequences and related analytical software by subscription. Manufactures and sells drug discovery tools.

ACCOMPLISHMENTS More than 25 major pharmaceutical and biotechnology subscribers. First genome company to report earnings from operations.

MARKET CAPITALIZATION **$713 million**

'98 REVENUE **$135 million**

'98 OPERATING INCOME **$12 million**

HEADQUARTERS Rockville, Md.

BUSINESS Provides genetic targets for partners' drug discovery programs. Develops proprietary proteins, antibodies and gene therapy drugs.

ACCOMPLISHMENTS Established first major genomic partnership, receiving $125 million from SmithKline Beecham. Has three drugs in clinical trials.

MARKET CAPITALIZATION **$1.55 billion**

'98 REVENUE **$30 million**

'98 OPERAT **–$32 million**

HEADQUARTERS Rockville, Md.

BUSINESS Provides genomic and biological information. Plans to offer complete genomes of human and other species, plus analytical software.

ACCOMPLISHMENTS Partnerships with Novartis, Pharmacia & Upjohn and Amgen.

MARKET CAPITALIZATION **$650 million**

'98 REVENUE* **$6 million**

'98 OPERATING INCOME* **–$16 million**

*Six months ended Dec. 31

The New York Times/
Illustration by Chris Gould

By LAWRENCE M. FISHER

ROCKVILLE, Md.

BEFORE there was an Internet bubble, before biotechnology left a bad taste in money managers' mouths, capital was ventured and shares were sold in a group of companies that aimed to leapfrog the Government's multibillion-dollar effort to map the human genome. By sequencing and patenting genes ahead of the publicly financed laboratories, the investors thought, these companies would control the next era of health care — and make their founders and shareholders rich.

But now, as the international race to spell out the precise sequence of the three billion letters in the human genetic code sprints to the finish line, the markets are demanding that the gene merchants deliver the goods. Investors are starting to demand not just science, but a viable business model, too.

To the victors will belong an asset of almost unfathomable value: the source code of life itself. By sequencing the DNA in all 100,000 genes in the human body — that is, by divining the order of the four chemicals that make up the genes — scientists hope to treat diseases in much the same way that software engineers fix faulty computer programs: by isolating flaws in the code.

If the effort is successful, health care will shift from a paradigm of detect and treat, typically with toxic drugs that sometimes do no more than mask symptoms, to predict and prevent, with therapies of exquisite specificity aimed at the causes of disease. By identifying the genetic roots of illnesses like cancer and heart disease, some experts say, the science of the genome, or genomics, may make it possible for a child born today to live to 150 — or, some say, much longer.

"Death is a series of preventable diseases," said Dr. William A. Haseltine, chairman and chief executive of **Human Genome Sciences** here, declaring the philosophy behind the industry he is helping to invent.

Other genomics leaders have less trouble with that assertion than with Dr. Haseltine's contention that his company already has sequences for 95 percent of the human genes and is amassing a patent portfolio that will make it the gatekeeper for the entire biopharmaceuticals industry.

PLUNGING INTO THE DOUBLE HELIX

For investors, buying shares in genomics companies still requires a leap of faith. But the stocks could become great long-term holdings. Page 6.

Assertions like that have always irritated Dr. Francis S. Collins, the director of the Human Genome Project at the National Institutes of Health in Bethesda, Md. For some of the companies, he said, "financial viability appears to be tied to claims of exclusivity," while the Government's goal is to sequence all the human genes and place them in the public domain for the common good.

Indeed, the whole concept of patenting the bits of biochemical information known as genes offends many from religious leaders to scientists, including James Watson, a co-discoverer of DNA, who resigned as the original head of the Human Genome Project over this issue. While the Patent Office has established rules meant to clarify what kind of genetic discoveries can be patented, the issue is by no means resolved.

The genomics companies have filed patents by the thousands, but only a few hundred have been published, fewer still issued and none of those yet defended in court. So any claims of strong patent positions remain unproven.

But genomics is all about big claims. Last year, Dr. J. Craig Venter, a former N.I.H. researcher, shook up the field when he started a new company, now called

Continued on Page 12

A Race of Entrepreneurs to Cash In on the Genet

Continued From Page 1

Celera Genomics Group, a separately traded subsidiary of the **PE Corporation**, which he said would complete the sequencing of the human genome in 2001, about two years ahead of the public labs.

Dr. Venter's announcement was greeted with considerable skepticism, but it prompted Dr. Collins to move up his goal for completion, too; he now expects a rough draft of the human genome by next spring and a complete genome a year later.

The acceleration of the public and private efforts has already led to a rough winnowing of the genomics field, with some companies disappearing in mergers and others trading at such low share prices that their ability to raise capital and remain independent is doubtful.

A couple of broad strategies have emerged, meanwhile, among the best-capital companies.

Human Genome Sciences and **Millennium Pharmaceuticals** Inc., of Cambridge, Mass., intend to sell drugs, developed with capital raised in a few high-value deals with pharmaceutical companies. Two others, **Incyte Pharmaceuticals** Inc., of Palo Alto, Calif., and Celera, in Rockville, sell only data that others use to identify potential drug targets.

Shares in Human Genome and Millennium have soared this year, while Incyte, once the market leader, has fallen, primarily because of the perceived threat from Celera, which has deep pockets and a similar business model.

"The human genome is going to be sequenced, with absolute certainty, within the next two years," said Dr. Eric Lander, director of the federally financed Whitehead/M.I.T. Center for Genomic Research in Cambridge and a founding scientist at Millennium. "The game has moved on. Now it's: How do you add value to the genome?"

The Code as a Cure

Perhaps the most obvious way to add value to the genome is to turn genetic discoveries into drugs, but so far Human Genome Sciences is the only company that has done so. The stainless-steel fermentation vats at its 80,000-square-foot manufacturing plant here in Rockville would be a common site at a traditional biotechnology company, like **Amgen** or **Genentech**, but they are unique among the genomics players.

The vats produce the proteins that are Human Genome's first two drugs. One, intended to protect blood-forming cells from the toxic effects of chemotherapy, is being tested in women with breast and ovarian cancer. The other is being tested as a wound-healing agent. Both are in the second of three phases of trials typically required of new drugs by the Food and Drug Administration; that means that they are at least two or three years away from the market.

> Prospects for patent battles, as well as miracle cures.

But it is Human Genome's third drug candidate that provides a glimpse of the real promise of genomics. Using a gene injected directly into a diseased area, it prompts the body to grow new blood vessels. The drug is being tested as an alternative to bypass surgery for treating blocked arteries in the heart and limbs.

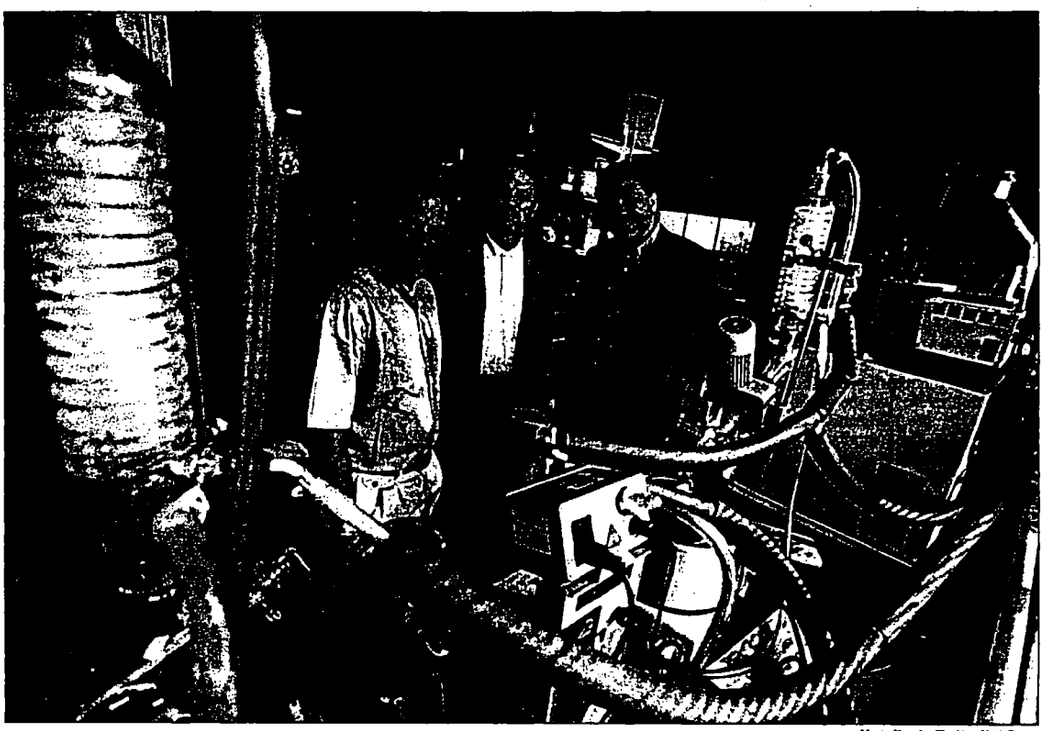
"At this point, we are the only company to create genomics assets and to use those assets to move drugs into the clinic," said Dr. Haseltine, who did pioneering work in sequencing the genome of HIV, the virus that causes AIDS, as a scientist at Harvard's Dana-Farber Cancer Institute.

But these three drugs, and three others that Human Genome hopes to begin testing next year, are only the tip of a very large iceberg. Dr. Haseltine said the company has discovered — and, more importantly, filed patents on — a vast data base of genes.

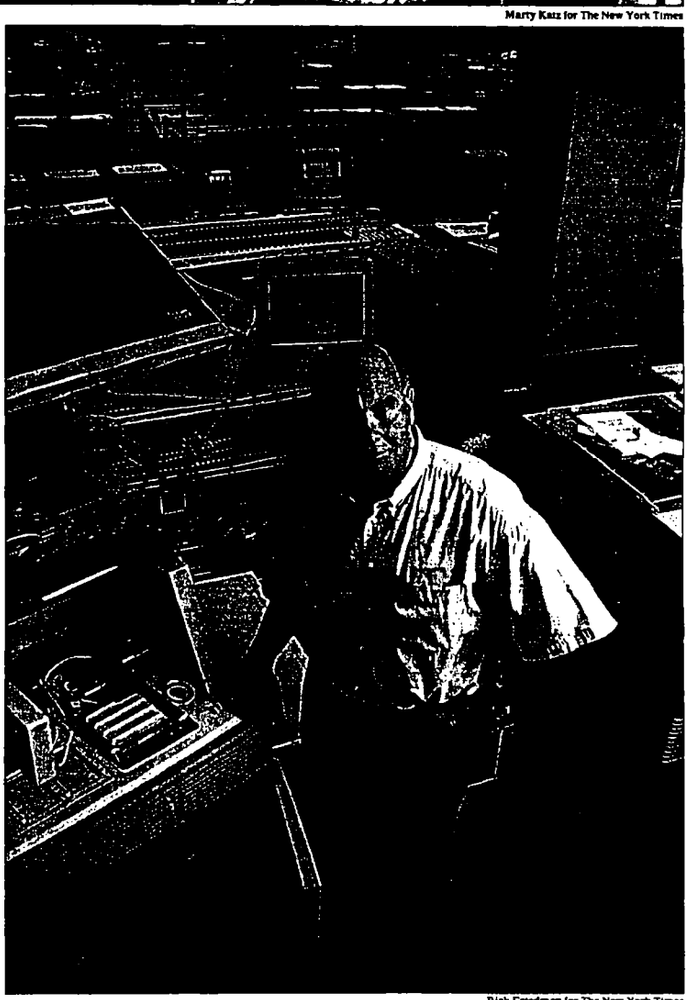
"Any company that wants to be in the business of using genes, proteins or antibodies as drugs has a very high probability of running afoul of our patents," Dr. Haseltine said. "From a commercial point of view, they are severely constrained — and far more than they realize."

Some drug industry executives dismiss this is as self-promotion, though as Human Genome's patents begin to be published, it is clear that the company is assembling an arsenal of intellectual property.

Still, Human Genome is itself constrained by the terms of a deal it struck in 1993 with **SmithKline Beecham** P.L.C., the British pharmaceuticals giant. For $125 million, then the largest biotechnology deal ever, SmithKline obtained exclusive rights to use Human Genome's gene discoveries to produce so-called small-molecule drugs, the kind that can be taken in pill form. The agreement expires in June 2001.



Marty Katz for The New York Times



Rick Friedman for The New York Times

Two companies working to make profitable use of the decoding of human DNA are Millennium Pharmaceuticals, led by, from left above, Steven H. Holtzman, Mark Levin and Dr. Eric Lander, and Celera Genomics Group, whose founder is Dr. J. Craig Venter, at right.

The deal left Human Genome free to develop drugs based on proteins — the easiest kind to take quickly to the clinic. But such drugs are commercially problematic, because they must be injected, and few patients want to take a shot if a pill is available. The company can also develop gene therapies, like its heart drug. But gene therapy has yet to produce an approved drug for anyone.

Nonetheless, Dr. Haseltine said the SmithKline deal was worthwhile, because it gave Human Genome the resources to grab a big lead in genomics. And when it expires, he will be free to strike new partnerships based on the same data base of genes. "Two years from now, we have a big opportunity to remonetize our asset and still have plenty of intellectual property to ourselves," he said.

Partners in the Marketplace

For now, however, the undisputed champion of deal making is Millennium, which has raked in over $1 billion in financing from partners like **Roche Holding** A.G., **Eli Lilly & Company**, **Monsanto** and **Bayer** A.G.

The arrangements are exclusive only for specific diseases (or, in Monsanto's case, for use in agricultural products), so Millennium remains free to pursue its own small-molecule drugs. It has also created subsidiaries for biotechnology drugs and for predictive medicine, like tests to identify genetic conditions showing a predisposition to diseases or adverse drug reactions.

Millennium's executives don't claim a lock on the intellectual property of the human genome, but their deal-making shows the value that drug makers see in the company's work.

A year ago, for example, Bayer agreed to pay Millennium up to $460 million over five years for 225 of what the industry calls validated genetic targets — that is, genes plus the accompanying biological data to show their relevance in areas like cardiovascular disease, cancer, osteoporosis, pain, liver fibrosis and viral infections. Targets that Bayer doesn't choose for drug development revert to Millennium.

"We did a huge search before we went with Millennium," said Wolfgang Hartwig, Bayer's executive vice president of pharmaceutical research. He said Bayer was attracted to Millennium's continual refreshing of its technology and to its freedom from the encumbrances of exclusive deals.

Millennium sprawls around Cambridge in leased space, most of it near the Massachusetts Institute of Technology. In its early days, the company made much of its connection to Dr. Lander and M.I.T.'s genome center. But today, Dr. Lander's role is more clearly advisory. The personalities who define the company are those of its chief executive, Mark Levin, a soft-spoken former venture capitalist, and his head deal maker and chief business officer, Steven H. Holtzman, a Rhodes Scholar with a degree in philosophy.

Mr. Levin tends not to make big claims, but he is building a big company, drawing on experiences early in his career at Genentech, the biotechnology pioneer, where he worked in process engineering in the mid-1980's.

"Genentech put in place the largest molecular biology lab in the world at the time," he said. "We've gone out and put together the largest critical mass in biotech today."

Later this year, Roche is expected to get permission to begin trials of an anti-obesity drug developed from a genetic target supplied by Millennium. **Becton Dickinson** & Company in Franklin Lakes, N.J., will soon offer a melanoma test using technology from Millennium's predictive medicine unit.

But for all the depth and breadth of its science and technology, Millennium is conspicuously lacking a drug of its own in clinical trials.

"We're developing genes and targets and leads," Mr. Levin said. "By the end of 1999, going into 2000, you will see us make additional acquisitions, and these acquisitions will take us into the clinic."

Analysts note that while Millennium has the cash and high-value shares to acquire a drug — and many struggling biotech companies have drugs for sale — the best drug prospects attract multiple bids, and there are many far larger companies out shopping.

tic Code

The Knowledge Merchants

If Millennium and Human Genome aim to be the next Genentech and Amgen, Incyte and Celera have a different role model in mind: **Bloomberg** L.P. Just as Bloomberg sells information to the securities industry but isn't in the brokerage business, these companies sell data bases of genetic information to drug companies but never intend to produce drugs.

Incyte is based in Palo Alto, Calif., deep in Silicon Valley, and it is no coincidence that the heart of its headquarters is a vast, glass-enclosed room full of powerful computers. "At the end of the day, it's the information that matters," said Randy Scott, the president and chief scientific officer. "We are all about the application of Moore's Law to biology," he said — a reference to the observation that computer processing power doubles every 18 months. Applying that exponential growth to genomics should produce similar gains for drug discovery, Dr. Scott said.

Until recently, Incyte had something unique among genomics companies: operating income from sales of its products. Instead of signing huge deals to bankroll its own drug programs, Incyte sold subscriptions to Lifeseq, its genomic data base. Some 25 companies subscribed, including most of the leading pharmaceutical and biotechnology concerns, and Incyte showed steady earnings growth without big outlays on drug development.

But the creation last year of Celera — cash-rich and in no hurry to show a profit — put Incyte on the spot. While Incyte's data base offered gene fragments that point the way to full-length genes, Celera said it would provide whole genomes at a competitive price.

At the same time, the Government's Human Genome Project, now on an accelerated timetable, was placing thousands of genes in the public domain. And companies like **Pangea Systems** of Oakland, Calif., were supplying software tools to analyze this data, at a fraction of the cost of a Lifeseq subscription.

Facing a classic squeeze, Incyte management moved rapidly to improve its product, adding data from the Government's work, cutting the price and introducing new data bases. But the moves cost money, and as the company began reporting losses, investors fled, cutting the share price in half.

"I've been hearing this story — that Lifeseq's dead, it's all in the public domain — and maybe we even started to believe it ourselves," said Roy Whitfield, Incyte's chairman and chief executive. "But it's just not true." No customer has failed to renew its subscription, he said, and he predicted that profits — and investors — would return in the second half of next year.

Celera is the spoiler in genomics, or intends to be. The company plans to sequence the entire human genome and offer the data to subscribers. Other genomes of interest to scientists — like that of drosophila, a fruit fly used in many experiments — will be offered in a similar fashion. Dr. Venter's critics say Celera's hurry-up human genome will be full of gaps and errors, but the company has attracted three blue-chip subscribers: **Novartis** A.G., Amgen and **Pharmacia & Upjohn**.

As a sibling of **PE Biosystems Group**, also part of the PE Corporation, Celera received first access to PE Biosystems' latest automated gene-sequencing machines. (The two companies were created this spring in a reorganization of the old Perkin-Elmer Corporation.) Celera now has 300 of these $300,000 devices, some arrayed in neat rows, some lining the hallways still in their wooden shipping crates, some piled in elevators. The largest public genome center, Dr. Lander's lab at M.I.T. has 120 such machines, and most genomics companies have just a handful.

Combining the powerful equipment with an army of robots and vast amounts of computing power makes it practical and cost-effective to sequence whole genomes, Dr. Venter said.

The idea is that subscribers will pay Celera for partial genetic information as it becomes available, just as they have paid Incyte — but with the kicker that the company will ultimately supply complete genomes for many organisms. Companies can use this information royalty-free, except where it involves a gene sequence patented by Celera. Dr. Venter said he ultimately expects a mixture of revenues — from subscriptions, royalties on drug sales and even individual researchers paying to use the Celera Web site.

"They say I'm going to be the Bill Gates of biology. I'm not sure that's meant to be flattering," he said. Still, "the upside to this business is very, very large, and different from what anybody else is doing."

Many Eager Competitors

Wall Street loves to anoint winners, and given the impact of capital flow, these prophecies can become self-fulfilling. Despite talk that Celera or Human Genome may achieve a Microsoft-like dominance of biological information, many smaller companies are also pursuing opportunities.

The first real product of genomics was a diagnostic kit developed by **Myriad Genetics** Inc. of Salt Lake City, drawing on its discovery of two genes responsible for hereditary breast cancer. **Genome Therapeutics** Inc. of Waltham, Mass., has built a subscription business based on bacterial genomes. **Genset** S.A. of Paris has signed up two partners for what is called pharmacogenomics — the enterprise of applying genetic data to existing drugs to screen out people more likely to have side effects.

There are many more, and their continued proliferation suggests that genomics remains a fertile field.

"The current effort is rather like building the periodic table," said Dr. Collins, head of the Government's project, comparing the work to the fundamental plow work of chemistry.

"The real challenge of genomics — the real fun, the real payoff — is being able to take the periodic table and figure out how it works," he said. "All of that will be poorly understood and will be grist for research for decades to come."

Donna Bise for The New York Times

The Rev. Charles Wilson received a genetic treatment in a clinical trial.

A Vitalizing Gene, Straight to the Heart

Born with a genetic defect that caused his body to overproduce cholesterol, the Rev. Charles Wilson started having heart problems when he was 27. Over the next 30 years, he suffered a heart attack and had surgery many times. But by last summer, the blood supply to his heart was so reduced by atherosclerosis that he could barely get out of bed without severe chest pains.

"My doctor said to me, 'Charles, there's nothing else we can do; you've just got to sit still until technology catches up with you,' " said Mr. Wilson, 57, a Presbyterian minister who lives in Charlotte, N.C.

One day, though, he read a newspaper article about research on coronary gene therapy.

"I went looking on the Internet to find the actual report," Mr. Wilson said. "From that, I found the name of the doctor and the hospital."

He got in touch with Dr. Jeffrey M. Isner, chief of cardiovascular research at St. Elizabeth's Medical Center of Boston, and ultimately entered one of the first clinical trials of a drug developed through genomics.

Physicians in Boston made an incision in Mr. Wilson's chest and injected — directly into his heart — a gene that prompts the creation of new blood vessels. (Subsequently, the gene has been administered through a catheter.)

Although his recovery was slow at first, Mr. Wilson gradually gained strength.

"On the 42d day, as I was walking through the mall with my wife, when we had walked about a quarter-mile, I said, 'Honey, I haven't had any discomfort today,' " he said. "From that point on, I've been able to walk increasingly longer distances, faster and without pain." He now walks a mile and a half a day.

The gene he receives, VEGF-2, is being tested as a drug by **Human Genome Sciences** Inc. and is a product of **Vascular Genetics** Inc., which is a joint venture of Human Genome, Dr. Isner, St. Elizabeth's and Cato Holdings, a contract research organization. The drug is at least three to five years away from the market.

Although Mr. Wilson still has scar tissue and will have to take a variety of medications for the rest of his life, electronic scans of his heart show that it has grown new blood vessels in about 60 percent of the area that was previously blocked.

LAWRENCE M. FISHER

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

National Institutes of Health
National Human Genome
Research Institute
31 Center Drive MSC 2152
Building 31, Room 4B09
Bethesda, MD 20892-2152
(301) 496-0844
FAX (301) 402-0837

David, Sarah
FYI. I think we
should seriously consider
this.

October 25, 1999

TO: The Honorable Christopher C. Jennings, Deputy Assistant to the President For Health Policy

FROM: Francis S. Collins, M.D., Ph.D.
Director, NHGRI, NIH

SUBJECT: Event to Celebrate Major Human Genome Project Milestone

PROGRESS OF THE HUMAN GENOME PROJECT

The Human Genome Project is arming scientists with powerful new information and techniques to unravel the mysteries of disease. This knowledge will dramatically accelerate new strategies for diagnosis, prevention, and treatment of disease, not just for single gene disorders, but for the host of more common complex diseases (e.g. diabetes, heart disease, schizophrenia, and cancer) where genetic differences may contribute to disease risk and response to particular therapies.

The Human Genome Project (HGP) was initiated in 1990 and, from its beginning, has enjoyed significant success. Several of the project's initial goals have been achieved, including building maps to localize and order the position of genes in both the human and mouse genomes, and sequencing the genomes of the bacterium *E. coli*, bakers yeast, and the roundworm *C. elegans*. In addition, sequencing the genome of the fruit fly (*Drosophila melanogaster*) is nearly complete. In 1996, in preparation for tackling the much larger and more complex human genome, pilot projects began to test new technologies and strategies for human DNA sequencing. In March 1999, an international consortium supported by NIH, the Department of Energy (DOE) and international collaborators launched the full-scale effort to sequence the estimated 3 billion basepairs that make up the human genetic instruction book. The consortium, with the U.S. taking the lead, and with important participation by the U.K., France, Germany, Japan and China, expects to produce at least 90 percent of the human genome sequence in a "working draft" form by the spring of 2000, years earlier than expected. The final, high quality genome sequence will be completed by 2003 or earlier. Relevant to the purpose for this Memorandum, the deposition of the first billion base pairs of the human genome sequence into public databases is imminent.

THE IMPORTANCE OF PUBLIC ACCESS TO THE DATA

All sequence data produced by the international consortium is deposited in public databases for free access within 24 hours. In fact, this practice was adopted in 1996, and is now a condition of continued funding for recipients of grants from NIH, DOE, and the U.K. Wellcome Trust. This policy is based on the assumption that broader, faster access to data will stimulate a larger number of products in a shorter timeframe and better serves the public interest. The rapid free availability of the sequence is invaluable to academic scientists studying the molecular basis of human health and disease, as well as corporate researchers engaged in drug development.

The vast majority of scientists in both the public and private sectors have applauded this openness. A recent dramatic example of the importance of such free accessibility to information about the human genome is the founding of "The SNPs Consortium" by ten pharmaceutical companies and the Wellcome Trust. This Consortium is collaborating with NIH in developing a catalog of information about human genome variation, and putting all of the data into the public domain.

Human Genome Project scientists are committed to this policy to ensure findings are used in the public interest. This stance contrasts markedly with the approach taken by a few private sector organizations that aim to keep genomic sequence information secret or constrained by patents and licenses. One such company announced just this week that they have filed 6,500 provisional patent applications on genetic sequences.

A POSSIBLE U.S. - U.K. STATEMENT ON THE DESIRABILITY OF PUTTING GENOMIC SEQUENCE IN THE PUBLIC DOMAIN

For almost a year, Neal Lane and Sir Robert May have been discussing the possibility of jointly issuing a Proclamation endorsing the policy of the Human Genome Project to make DNA sequence information widely and freely available. A copy of the draft statement follows. Joint issuance of this Proclamation by President Clinton and Prime Minister Tony Blair would be timely, and would also be a gratifying example of international harmony on a matter of great importance to the future of science. Note that the draft Proclamation is carefully worded to endorse free access to "fundamental data on the human genome, including the human DNA sequence and its variations"; it does not criticize the use of intellectual property claims to protect discoveries about the USE of such sequences.

PROTECTING THE PUBLIC FROM GENETIC DISCRIMINATION

The rapid acceleration of the timetable for the Human Genome Project also makes it imperative that the necessary protections are put in place to prevent misuses of this information. The public, while supportive of the Human Genome Project and optimistic about the medical benefits of this research, is concerned that new genetic technologies and information may be used for harm, rather than for benefit. This concern is already affecting the choices individuals make about their own health care and their decisions whether to participate in research. The Clinton Administration has endorsed two important policy positions to prevent the misuse of genetic information. In June 1997 President Clinton called on Congress to enact legislation to prohibit the use of genetic information

in health insurance. Subsequently, on January 20, 1998, Vice President Gore announced the Administration's support for federal legislation to bar employers from discriminating against workers in hiring or promotion because of their genetic makeup, and released a Department of Labor report summarizing the legislative recommendations.

The Department of Health and Human Services has worked with the Department of Labor, the EEOC, and the White House to develop an Executive Order to apply the recommended protections to the federal workforce. The Order would prohibit federal agencies from collecting, using, or disclosing the protected genetic information of their employees, except under limited circumstances and with the employee's informed consent. Signing the executive order would provide significant protections to federal workers and provide an opportunity to renew active deliberations of workplace discrimination legislation in Congress.

ISSUE

The International Human Genome Sequencing Consortium is approaching an important milestone: completion of the first billion basepairs of the estimated 3 billion basepairs in the human genome. The Human Genome Project is planning an event to celebrate this important achievement and to recognize the important contributions of the hundreds of individual scientists whose hard work, skill, and commitment have brought us this far, and whose continued dedication will be essential to completing this awesome project.

The event will be held at the National Academy of Sciences (NAS) soon after completion and deposition into the public data base of the billionth basepair of human DNA sequence. This is estimated to be completed in late November or early December. To facilitate the involvement of scientists across the US and around the world, each major genome center is planning a local celebration that will be connected via satellite or internet to the NAS event. Members of voluntary health organizations, genetic professional societies, genetic support groups, biotechnology and pharmaceutical industry representatives, interested congressional members and staff, and the media will be invited to attend.

This event provides an opportunity for the President and high level Administration officials to be involved in a significant accomplishment in biomedical research. It also provide an fitting occasion to make a statement, ideally in a joint proclamation with Prime Minister Blair, about the value of public access to genome sequence information. Furthermore, based on the President's strong support for the Human Genome Project and his administration's commitment to resolving many of the ethical, legal and social policy issues presented by this wealth of new genetic information, this occasion may be a fitting opportunity for him to announce an Executive Order prohibiting genetic discrimination in the workplace.

RECOMMENDATION

I suggest that your office together with the OSTP recommend the President's participation in the event, and his announcement at that event of his strong support for the Human Genome Project's data access policy. In addition, DPC may wish to evaluate the feasibility of issuing an Executive Order prohibiting genetic discrimination in the federal workplace.

cc: Dr. Neil Lane
Dr. Rachel Levinson
Dr. Harold Varmus
Dr. Kathy Hudson

DRAFT

PROCLAMATION TO ENSURE THAT DISCOVERIES FROM THE HUMAN GENOME ARE USED TO ADVANCE HUMAN HEALTH

In the last decade of the twentieth century, scientists from around the world initiated one of the most significant scientific projects of all time: to determine the DNA sequence of the entire human genome, the human genetic blueprint. Progressing ahead of schedule, human genome research is rapidly advancing our understanding of the causes of human disease and will serve as the foundation for development of a new generation of effective treatments, preventions, and cures. To realize this promise, fundamental data on the human genome, including the human DNA sequence and its variations, must be made freely and broadly accessible to scientists everywhere. The human genome is fundamental shared heritage of all humankind. Unencumbered access to this information is essential to enable discoveries that will reduce the burden of disease and promote health around the world. We applaud the decision by scientists working on the Human Genome Project to release information about the human DNA sequence and its variants rapidly into the public domain, and we call on all scientists worldwide to adopt this policy.

THE WHITE HOUSE

WASHINGTON

December 13, 1999

MEMORANDUM FOR THE PRESIDENT

FROM: NEAL LANE Neal

SUBJECT: State of the Union Theme/FY01 Budget Ideas

Science and technology, particularly research and education, should figure prominently in your State of the Union Address to underscore your vision for the 21st century. The emphasis should be on two overarching needs: 1) restoring balance to our science and technology research portfolio by focusing on university-based research in the physical, social, and life sciences and in engineering; and 2) providing funding for the breakthrough research that will serve America's goals for the new millennium: investments that secure the continued prosperity of our economy, improvements in health care and in standards of living, and better education and training for America's students and workers.

Increasing Balanced Investments in Long-Term Research and Development

You have been consistent and eloquent about your commitment to investing in America's future even in times of constrained resources. In every year of your Administration, you have increased the S&T budget. The 21st Century Research Fund, an unprecedented statement of the importance of broad support of S&T, has fared extremely well in Congress – showing an 8 percent increase in FY 2000.

When you make remarks on the eight years of record peacetime expansion in the U.S. economy, you should – as Alan Greenspan has – credit science and technology for much of the outstanding performance. Innovations in science and technology have caused a chain reaction in the economy that:

- has resulted in improved productivity, which yields increases in wages or increases in benefits without letting loose higher inflation;
- has led businesses to continue to invest in high-tech capital equipment, making our capital stock more productive and profitable; and
- has provided U.S. businesses with payoffs from technology's efficiencies and from greater education in school and on the job that U.S. workers have acquired to keep pace with the new technology.

I also hope you will remark on the progress toward a goal you and the Vice President established in 1994: total national (public and private) R&D investment of 3 percent of GDP. By current projections, total annual R&D expenditures in the United States will be $247 billion in 1999 – an 8.8 percent increase over the $227 billion estimated for 1998, and the third year in which R&D spending outpaced the growth in the economy as a whole. R&D as a share of GDP will reach 2.79 percent in 1999. This 1999 forecast of R&D as a share of GDP is the highest since 1967's 2.80 percent, and in 2000 the U.S. R&D/GDP ratio could exceed the all-time high ratio of 2.87 percent in 1964.

As you noted in your economic speech of December 3, however, we need to make greater progress next year in funding a balanced research portfolio. NIH and biomedical research received over two-thirds of the FY 2000 increase in the 21st Century Research Fund. Non-defense research outside of NIH increased only 2.4 percent in FY 2000. Since FY 1994, non-defense R&D outside of NIH has not kept pace with inflation, and is down 5 percent in real terms. The physical and social sciences, engineering, and non-biomedical life sciences have barely kept pace with inflation over the last several years. Their failure to thrive poses a risk to our leadership in such critical areas as information technology (where we have made a good start with IT^2), nanotechnology, and energy. It even undermines biotechnology and medical research, which depend on advances in chemistry, biology, physics, mathematics, and several fields of engineering.

Theme: Innovation for Continued Prosperity

I propose a $1.5 billion initiative (above passback) in Science and Technology for the 21st Century for the FY 2001 budget. Such an increase would allow your Administration to achieve four important objectives: 1) increase university-based research (outside of NIH, government support for university-based research is only $8-9 billion); 2) correct the growing disparity between the scientific disciplines and restore the all-important balance between biomedical research and the physical, social, and non-biomedical life sciences and engineering; 3) sustain the U.S. position as the world leader in all fields of science and technology; and 4)increase support for research and development that support your priorities, including:

- Information technology – Further spur the transformation of the way we obtain services ranging from entertainment and commerce to education and health care; quicken the pace of discoveries across the scientific disciplines.
- Nanotechnology – Find ways to manipulate matter at the atomic and molecular level to make novel and useful devices and materials. Nanotechnology could have a future impact on our economy and our society that is as significant as the development of the transistor and the Internet.
- Education research – Find out what *really* works in teaching math and science to kids and apply that knowledge to the task of making sure all Americans are given a chance to participate in science. As the PCAST noted, "less than 0.1 percent of our nation's expenditures for elementary and secondary education . . . were invested to determine which educational techniques actually work, and to find ways to improve them."
- Biomedical research -- Win the war on cancer and declare war on asthma, mental illness, and heart disease; provide vaccines for the debilitating diseases of the developing world; produce an AIDS vaccine.
- Critical Infrastructure Protection and Weapons of Mass Destruction – Protect Americans from cyberthreats and from chemical and biological warfare; lower the anxiety level.
- Aviation safety – Increased avionics research can halt the trend toward weekly fatal airline accidents and reduce the aviation fatal accident rate by a factor of five within 10 years
- Climate and Energy – Create a clean energy future to reduce climate change and the health impacts of pollution, while growing the economy through energy technology exports; produce high-value chemicals, fuel, and power from new cash crops to boost farm income and add good jobs to rural economies.
- Environmental Research – Better understand the biology and ecology of life on earth and learn how to protect it while also meeting the needs of the American and world's people.

By addressing significant gaps in our knowledge base, these initiatives will support greater equity and prosperity in our nation. I have attached agency-by-agency estimates of possible allocations of the FY 2001 funds for your consideration (Tab A). The outyears should reflect sustained annual growth of at least 5 percent.

Doubling Option: Millennium Fund for University Research

If we seek an enduring legacy of your commitment to investing in science and technology, I believe that the Administration could gain bipartisan support for supplementing and sustaining the FY 2001 funding with a proposal to double university-based research in five years (beginning in FY 2002). By directing that these funds be used to strengthen the integration of research and education, the Administration can underscore the core principles of the Government-University Research Partnership. This Millennium Fund for University Research would require an additional investment of approximately $28 billion in the outyears (on top of the 5% increase recommended above). Such a fund would provide a practical alternative to Congressional proposals that authorize doubling all civilian research and development.

$1.5 Billion: Investment in 21st Century Initiative

21st Century Research Fund Accounts	FY1999	FY 2000 Final*	Change from FY 99 Amount	Change from FY 99 Percent	Passback	Proposed Addback		Initiatives
Total	36,932	39,915	2,983	8.1%	40,478	1,500		
					11.5%	5.1%		
NIH	15,612	17,865	2,253	14.4%	17,933	421	**	Biomedical Research/Vaccines
NSF	3,672	3,897	225	6.1%	3,912	420		IT; Nanotech; Envir Research
DOE	3,547	3,673	126	3.6%	4,050	100		IT; Nanotech; Climate/energy
NASA	4,996	4,885	-111	-2.2%	5,001	100		IT; Nanotech;
DOD	4,259	4,580	321	7.5%	4,072	300	***	IT; Nanotech; CIP/WMD
USDA	1,536	1,575	39	2.5%	1,581	75		Climate/energy; food safety
DOC	827	831	4	0.5%	801	50		IT; Nanotech
DOI (USGS)	798	821	23	2.9%	829	33		Climate/energy
EPA	671	663	-8	-1.2%	759			Climate/energy
VA	316	320	4	1.3%	334	2		Biomedical research
Ed	210	235	25	11.9%	427			Education research
DOT	488	570	82	16.8%	779			Aviation Safety

numbers are in millions of dollars

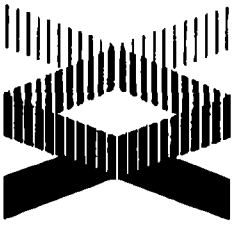
*numbers reflect .38 percent cut

** NIH numbers represent an increase equal to the CPI (2.4%) plus $60 M for Vaccine Initiative.
If NIH were to increase to the BRDPI (3.6%) an additional $154 M to the proposed addback would be needed .

*** DOD numbers assume passback at FY 00 requested levels.

PASSBACK = $563 M above FY 00 Enacted Levels
ADDBACK = $1,500 M above FY 00 Enacted Levels

NATIONAL CENTER FOR HUMAN GENOME RESEARCH, NIH



FAX TRANSMITTAL SHEET

TO: Devorah Adler

FAX #: (202) 456-5557

FROM: Kathy Hudson

DATE: November 23, 1999

# of pages including cover sheet: 16

If there are problems, call (301) 402-2205

Comments:

Things we are on the cusp of doing:
1) decoding genome sequence
2) cataloguing genetic variation
3) understanding gene function
4) individual preventative medicine (see attached article)

Please call if you have questions!

Kathy

Kathy Hudson, Ph.D.
Assistant Director
National human Genome Resaerch Institute

The New England Journal of Medicine

Established in 1812 as THE NEW ENGLAND JOURNAL OF MEDICINE AND SURGERY

VOLUME 341 JULY 1, 1999 NUMBER 1

ORIGINAL ARTICLES

IMAGES IN CLINICAL MEDICINE

SPECIAL ARTICLE

REVIEW ARTICLE

EDITORIALS

CORRESPONDENCE

Reprinted from
THE NEW ENGLAND JOURNAL OF MEDICINE (ISSN 0028-4793).
Published weekly from Editorial Offices at
10 Shattuck Street, Boston, Massachusetts 02115-6094 USA – Fax: (617) 734-4457

Business, Subscription Offices
860 Winter Street, Waltham, Massachusetts 02451-1412 USA – Fax: (781) 893-8103

The New England Journal of Medicine

Special Article

SHATTUCK LECTURE — MEDICAL AND SOCIETAL CONSEQUENCES OF THE HUMAN GENOME PROJECT

FRANCIS S. COLLINS, M.D., PH.D.

THE history of biology was forever altered a decade ago by the bold decision to launch a research program that would characterize in ultimate detail the complete set of genetic instructions of the human being. The idea captured the public imagination, perhaps less in the manner of America's wars on cancer and the acquired immunodeficiency syndrome than in the manner of the great expeditions — those of Lewis and Clark, Sir Edmund Hillary, and even Neil Armstrong. Scientists wanted to map the human genetic terrain, knowing it would lead them to previously unimaginable insights, and from there to the common good. That good would include a new understanding of genetic contributions to human disease and the development of rational strategies for minimizing or preventing disease phenotypes altogether.

The endeavor was both awesome and chancy. The instruction book — the human genome — was vastly larger than any genetic endowment tackled so far, and in 1990, the tools were not yet powerful enough to perform the task. Critical social questions, such as whether the new technologies for reading our genetic constitution would challenge our identities, our fundamental right to privacy, or our freedom from discrimination, loomed without answers.

Yet, a public science initiative focused so sharply on the molecular essence of humankind was too intriguing and too promising to forgo. Since the 1970s, nearly all avenues of biomedical research have led to the gene, for genes contain the basic information about how a human body carries out its duties from conception until death. In between, of course, our bodies struggle to survive in a challenging environment. Largely, but not entirely, at the behest of our genes, we fare better or worse.

Not surprisingly, disease researchers wanted to bag their leading gene suspects as soon as possible and at the least expense. This was no small order — the 80,000 or so human genes are scattered throughout the genome like stars in the galaxy, with genomic light-years of noncoding DNA in between. The billions and billions of uncharted DNA units (approximately 3 billion base pairs in humans) frustrated searches regularly in the late 1980s, often at great health and financial cost. If gene hunters were to mine miracles from the human genome, they needed more powerful tools and more ambitious strategies.

ACCELERATED GOALS FOR SEQUENCING THE HUMAN GENOME

In 1988, Congress appropriated funds to the Department of Energy (DOE) and the National Institutes of Health (NIH) to begin planning the Human Genome Project. Planners set a 15-year time frame, estimated that the price tag would be $3 billion, and laid out formal goals to get the job done.[1] On October 1, 1990, the Human Genome Project officially began.[2] According to early plans, the human race would witness its own blueprint in fine detail in the year 2005.

In the fall of 1998, however, improvements in technology, success in achieving early mapping goals (see below), emerging research opportunities, and a growing demand for the human DNA sequence prompted project leaders in the United States and abroad to promise the blueprint — the complete DNA sequence of the human genome — two years ahead of schedule, in 2003.[3] The technology to accomplish such a task is at hand. Indeed, only six months later, in March 1999, 15 percent of the sequence was in a finished or nearly finished state. The largest centers participating in the Human Genome Project received new grants to begin full-scale sequencing of the human genome, and the timetable was moved up yet again. Pilot sequencing projects had been so successful that the planners of the Human Genome Project now felt confident that at least 90 percent of the human sequence could be completed in "working draft" form by the spring of 2000, considerably earlier than expected.

The NIH and DOE genome programs expect to contribute 60 to 70 percent of the sequence. Scientists funded by the Wellcome Trust at the Sanger Centre in Cambridge, England, along with other international partners, will produce the remainder.

Presented as the 109th Shattuck Lecture to the Annual Meeting of the Massachusetts Medical Society, Boston, May 8, 1999.

From the National Human Genome Research Institute, National Institutes of Health, 31 Center Dr., MSC 2152, Bldg. 31, Rm. 4B09, Bethesda, MD 20892-2152, where reprint requests should be addressed to Dr. Collins.

Reprinted from THE NEW ENGLAND JOURNAL OF MEDICINE (ISSN 0028-4793) Vol. 341: 28-37 (July 1, 1999).

Until the entire human genome sequence has been completed, in 2003 or perhaps earlier, the working-draft sequence will be very useful, especially for finding genes, exons, and other genomic features. But because the working draft will contain gaps and will not be entirely accurate, it will not be as useful as the finished sequence for studying DNA features that span large regions or require a high degree of accuracy over long stretches of the sequence.

The final product must have four characteristics — the four A's of the Human Genome Project. First, the sequence must be *accurate* — that is, the DNA spellings must have an accuracy of 99.99 percent or better. If this is the Book of Life, we should not settle for a rough draft over the long term but should remain committed to producing a final, highly accurate version. Second, large-scale sequencing requires that the shorter lengths of sequenced DNA be accurately *assembled* into longer, genomic-scale pieces that reflect the original genomic DNA. Third, the human DNA sequence must also be *affordable*, and technology development will aim to reduce the cost as much as possible. Finally, the high-quality, finished human DNA sequence should be *accessible* within 24 hours through public data bases.

With regard to this last point, it is imperative that the genome sequence, the fundamental building block of biomedical research, be made rapidly available at no cost to scientists around the world. The human DNA sequence arms scientists seeking to understand disease with new information and techniques to unravel the mysteries of human biology. This knowledge will dramatically accelerate the development of new strategies for the diagnosis, prevention, and treatment of disease, not just for single-gene disorders but for the host of more common complex diseases (e.g., diabetes, heart disease, schizophrenia, and cancer) for which genetic differences may contribute to the risk of contracting the disease and the response to particular therapies.

In 1996, at the Second International Strategy Meeting on Human Genome Sequencing, a group of scientists involved in large-scale genomic DNA sequencing passed a unanimous resolution that all sequence data produced by the Human Genome Project should be freely available and in the public domain, in order to encourage research and development and to maximize its benefit to society. These scientists believed that public availability of these data would encourage coordination and collaboration among scientists while preventing duplication of efforts.

As a result, planners of the Human Genome Project adopted a policy of releasing data every 24 hours to a free, publicly accessible data base. In the United States, GenBank (accessible at http://www.ncbi.nlm.nih.gov), run by the National Center for Biotechnology Information, serves as the public repository of sequence information. The value of this data base as a resource for medical scientists around the world is incalculable. GenBank receives over 200,000 queries a day for information on gene sequences. It took 16 years to get the first billion bases into the data base, but it took only 15 months to add the second billion bases. Over 39,000 species are represented, and over 60,000 sequence-comparison searches are conducted each day.

Because researchers around the world participating in the public sequencing effort have agreed to submit their data to free, public data bases within one day after verification, any scientist, whether based at a university, a corporation, or a government laboratory, can log on by computer and have full and rapid access to the sequence data. The sooner the publicly funded effort to sequence the human genome is complete, with the data deposited in a free and publicly accessible data base, the sooner it will benefit all scientists working to understand and treat disease at a molecular level. In addition, scientists understand human genes by comparing human DNA with DNA from mice, flies, roundworms, and yeast. These comparisons make the human sequence much easier to understand.

As Human Genome Project plans for the period between 1998 and 2003 were being developed, two corporate ventures announced initiatives to sequence a major fraction of the human genome, using strategies that differ fundamentally from the publicly funded approach. The stated intention of one of these ventures to release its data publicly creates the possibility of synergy with the federal effort. If the corporate data and the public data can be merged, the depth of sequence coverage will be even greater. And because the public sequence data will be obtained from mapped DNA, it will provide critically needed anchoring to sequence data generated by private-sector efforts.

MORE THAN JUST THE SEQUENCE

Many people assume that since 1990 most of the work of the Human Genome Project has been devoted to large-scale sequencing. But this activity is actually a relatively recent arrival on the scene, since many other kinds of genome information were needed before full-scale sequencing could be undertaken. Before the 1998–2003 plan, the Human Genome Project's scientific goals mostly addressed genome mapping, technology development, and work to characterize the genomes of certain laboratory organisms. One type of map, known as a genetic map, is particularly helpful for following the inheritance of a disorder through several generations of a family. Genetic maps consist of a series of sequence-based markers that can be used to pinpoint the likely neighborhood of an altered gene responsible for a disease phenotype or other trait. Such maps have been invaluable in identifying and isolating highly penetrant

gene mutations with mendelian inheritance patterns. The goal was to establish markers close enough to give a gene hunter a high likelihood of placing a gene in a reasonably searchable interval. A year ahead of schedule, an international consortium of leading researchers in France and the United States published a genetic map containing almost 6000 markers, spaced less than 1 million bases apart.[4] Such detail was four to six times greater than the 1990 goals called for.

A second type of map, known as a physical map, provides cloned and ordered sets of contiguous DNA that represent regions of a chromosome, or even a whole chromosome. Once genetic markers define the region containing the sought-after gene, cloned pieces from the physical map provide a resource from which investigators can then isolate the gene. A copied replica of 98 percent of the human genome, consisting of thousands of linked pieces of DNA, is complete and meets the genome project's goal for physical mapping.[5] This map contains over 41,000 DNA markers, known as "sequence-tagged sites," or STSs, that properly align the pieces. With this density of markers, most genes in the human genome should lie within fewer than 100,000 bases of an STS.

The Human Genome Project also recognized from its inception its responsibility not only to develop gene-finding and analysis technology, but also to address the broader societal implications of these newfound abilities to decipher genetic information. So the project commits 5 percent of its annual research budget to a program that addresses the ethical, legal, and social implications of genome research (the ELSI program). This program has focused on four high-priority areas: the use and interpretation of genetic information, clinical integration of genetic technology, issues surrounding genetics research, and public and professional education about these issues.

The plan covering the years 1993 through 1998 expanded the mapping goals and explicitly included new objectives for identifying and mapping not just systems of anonymous markers but the genes themselves.[6] Scientists at the National Library of Medicine, at genome research centers, and in private industry later began a program to position expressed DNA from gene regions, or expressed sequence tags (ESTs), on the physical map. After two editions, this gene map represents the most extensive effort so far to locate and identify the 80,000 genes in the human genome. Over 38,000 gene tags have been placed on the map, giving disease-gene hunters a ready list of candidate genes residing in the chromosomal neighborhood they know is involved in a disease.[5]

Before 1996, the goals of DNA sequencing were largely targeted toward genomes of model organisms. From those efforts, scientists have sequenced the genomes of *Saccharomyces cerevisiae*, a species of yeast valuable to biologists and commonly used by bakers and brewers,[7] and *Escherichia coli*, a mainstay of basic biology and the biotechnology industry.[8] The first complete genome sequence of a multicellular organism, the roundworm *Caenorhabditis elegans*, was completed in late 1998.

The yeast genome, which was the first genome of a eukaryote to be completely deciphered, contains 12,057,500 base pairs in its nuclear DNA. Containing some 6000 genes arranged on 16 chromosomes, yeast has already provided biologists with a valuable resource for determining the function of individual human genes involved in medical disorders such as cancer. Now, the complete sourcebook for that organism will allow scientists, for the first time, to piece together information that will provide a comprehensive look at how all the genes in a eukaryotic cell function as an integrated system.

The 100-million-base-pair *C. elegans* genome is distributed among six chromosomes and contains over 19,000 genes. Although barely visible with the naked eye, the tiny roundworm has become an invaluable tool for studying biologic processes such as development, neurobiology, and aging. The worm project began in 1990, as a collaboration between researchers at Washington University, in St. Louis, and at the Sanger Centre, in Cambridge, England. Computer comparisons between the worm sequence and the sequences of other organisms have shown that about 74 percent of named human genes have a definite homologue in *C. elegans*.

The 1998–2003 plan can be thought of as the development of a new and more diverse set of power tools, all of which are to be given away. These tools include the development of the catalogue of variations in the human DNA sequence; new forms of technology and new strategies for studying gene function on a whole-genome scale; and new areas of research in the ELSI program (the "safety goggles"), such as identifying and addressing issues that link genetics to personal identity and racial or ethnic background and examining the implications of these links for philosophical and religious traditions.

IMPLICATIONS FOR UNDERSTANDING GENETIC ILLNESS

Maps and other forms of genome technology provide the tools for a gene-isolation technique known as positional cloning.[9] This technique allows a researcher to confirm the genetic basis of a disease and identify the responsible gene, even when little is known about the gene's function. So far, over 100 disease-linked genes have been isolated with the use of the positional-cloning technique. Whereas gene discovery by this route once took years to decades, an investigator using these powerful tools can now sometimes map and isolate a gene in a matter of weeks. Increasingly, gene hunters are combining positional-cloning techniques with information in EST data bases to narrow their gene searches to rational

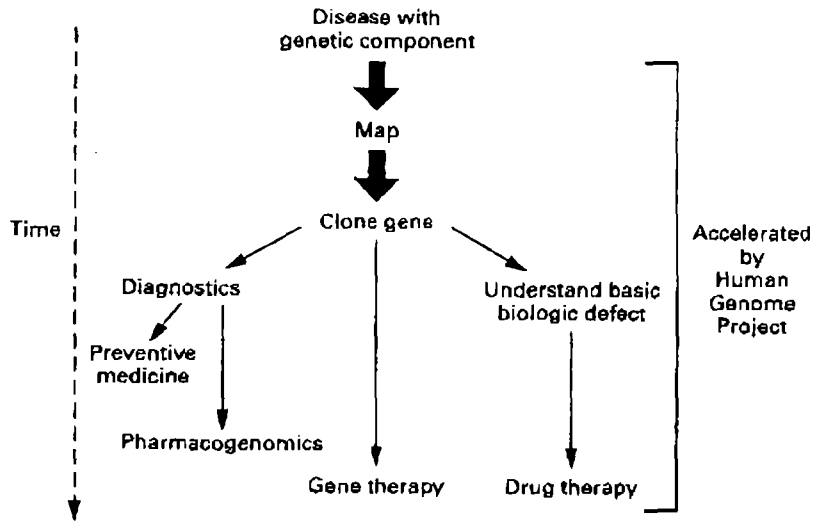


Figure 1. Steps Involved in the Genetic Revolution in Medicine.

Uncovering the genetic contributions to an illness is accomplished by cloning the gene for the disease, with the use of the tools of the Human Genome Project. Once the contributing genes and their disease-predisposing variants have been identified, diagnostic tests can be developed to predict future risk — but these tests are most effective when a preventive strategy is available to reduce the risk in persons found to be predisposed to a particular disease. Another rapidly developing application of diagnostics is pharmacogenomics, the prediction of responsiveness to drugs. Ultimately, the real payoff of genetic research will be the development of new gene therapies and drug therapies, but they will generally require many more years of intensive research.

candidates. This method, called positional candidate cloning,[10] has been used to isolate many altered genes associated with human disease.

Gene isolation provides the best hope for understanding human disease at its most fundamental level (Fig. 1). Knowledge about genetic control of cellular functions will underpin future strategies to prevent or treat disease phenotypes. The recent isolation of genes for Parkinson's disease,[11-14] for example, has greatly advanced molecular research on this baffling disease. In one study of families with early-onset Parkinson's disease, gene hunters mapped a suspect gene to a region of chromosome 4.[11] The region contained approximately 100 genes, among which was 1 known to encode the protein α-synuclein. Earlier research had shown that α-synuclein accumulates in brain cells of people with Alzheimer's disease,[15] and people with Parkinson's disease have similar deposits in the substantia nigra. In just a few months, the researchers showed conclusively that a missense mutation in the α-synuclein gene caused Parkinson's disease in the study families. Further research has shown that a mutation in a gene encoding a protein critical to the breakdown of α-synuclein and other proteins also results in the Parkinson's disease phenotype in a different family.[14] Although most cases of Parkinson's disease appear to have limited heritability, studying rare families of this sort provides crucial clues to the pathways involved — α-synuclein is found in the Lewy bodies in virtually all cases of Parkinson's disease. An understanding of the genetic control of the proteolytic processes of brain proteins may provide new targets for interventions in a number of related neurodegenerative disorders characterized by the accumulation of protein deposits, including Alzheimer's disease, Huntington's disease, and spinocerebellar ataxia.

Even before a gene's role in disease is fully understood, diagnostic applications can be useful in preventing or minimizing the development of health consequences (Fig. 1). DNA tests that look for the presence of disease-linked mutations, for example, are proving to be the most immediate commercial application of gene discovery and the one now used most frequently by clinicians. These tests may help establish the diagnosis of a genetic disease, foreshadow the development of disease later in life, or identify healthy heterozygote carriers of recessive diseases. Genetic tests can be performed at any stage of the human life cycle, and the sampling procedures are becoming less invasive. Whereas genetic testing was once sought almost exclusively by couples with a family history of early-onset disease, for the purpose of family planning, information about genetic status is increasingly sought by persons who wish to learn about their own predisposition to adult-onset illness.

In a growing number of instances, strategies can be implemented to reduce or prevent illness when a genetic cause or predisposition is known. Successes in reducing disease through treatment have been achieved for the hereditary disorders hemochromatosis, phenylketonuria, and familial hypercholesterolemia, among others. Risk reduction through early detection and lifestyle changes may be possible in the case of disorders associated with predisposing mutations, such as some cancers. As therapies build on knowledge gained about the molecular basis of disease, increasing numbers of illnesses that are now refractory to treatment may yield to molecular medicine in the future.

The recent discovery of an altered gene (*HFE*) that leads to hereditary hemochromatosis,[16] a common disorder of iron metabolism, provides an interesting example of the potential for using information about mutations to prevent an adult-onset disease phenotype. A recessive condition, hereditary hemochromatosis affects about 1 in 300 persons of northern European descent and is easily treatable if diagnosed early. Its major symptoms — liver cirrhosis, heart failure, diabetes, arthritis, and other organ damage — do not occur until midlife and are easily misdiagnosed. Untreated, the disease causes early death, but treatment by phlebotomy to remove excess iron allows people with hereditary hemochromatosis to live a normal life span. A single substitution of the amino acid tyrosine for cysteine at codon 282 accounts for the majority of cases.[17]

At first glance, hereditary hemochromatosis seems to be an ideal target for public health approaches to the prevention of hereditary disease: the disorder is common, the number of disease-linked mutations in the gene are few, and effective treatment can minimize or eliminate the effects of the disease. But closer examination reveals a number of complexities that have so far militated against the rapid introduction of this genetic test as a tool for disease prevention.[16] Because the penetrance of the altered *HFE* gene is reduced, especially in females, clinical signs can range from none that are detectable to severe organ damage from iron overload. At the moment, simple detection of the mutation does not predict the most likely clinical course. Before population testing for *HFE* mutations is considered, further research is needed to explain the variations in phenotype among mutation carriers and to correlate the genotype more closely with health outcomes.

IMPLICATIONS FOR THE STUDY OF COMMON DISORDERS

The rather straightforward mendelian rules that govern the inheritance of disease traits have been worked out for many rare disorders that result from highly penetrant changes in a single gene. But teasing out the genetic components of the so-called complex disorders — diabetes, heart disease, most common cancers, autoimmune disorders, and psychiatric disorders — that result from the interplay of environment, lifestyle, and the small effects of many genes remains a formidable task. Most of the successful efforts to identify genes associated with common diseases have focused on highly heritable subgroups, including the *BRCA1*[18] and *BRCA2*[19] genes in breast cancer, the gene for hepatocyte nuclear factor 4α (*HNF-4α*) in maturity-onset diabetes of the young (MODY) type 1,[20] the gene for glucokinase (*GCK*) in MODY type 2,[21] the gene for hepatocyte nuclear factor 1α (*HNF-1α*) in MODY type 3,[22] the gene for human Mut S homologue 2 (*hMSH2*)[23,24] and the gene for human Mut L homologue 1 (*hMLH1*)[25] in hereditary nonpolyposis colon cancer, and the gene for α-synuclein in Parkinson's disease.[11] Linkage analysis and positional-cloning techniques are well suited to discovering genes with such strong influences. But these strategies are not as easily applied to the multiple, low-penetrance variants, which in the aggregate account for a larger percentage of illnesses. Identification of weakly penetrant alleles that contribute to common disorders requires new and more powerful approaches.

To assist in these efforts, the Human Genome Project is initiating new studies of genetic variation in the human population to provide a dense map of common DNA variants. DNA sequence variations include insertions and deletions of nucleotides, differences in the copy number of repeated sequences, and single-nucleotide polymorphisms, or SNPs (pronounced "snips"), which occur most frequently throughout the human genome. About 1 in every 300 to 500 bases in human DNA may be a SNP.

SNPs can be used as markers in whole-genome linkage analysis of families with affected members, as well as in association studies of individuals in a population. Association studies may directly test a variant with potential functional importance or may take advantage of the phenomenon of linkage disequilibrium — in which a marker and a gene are inherited together — to map gene variants associated with disease. Because the human species consists of relatively few generations, recombination events have not disrupted linkage disequilibrium over distances of 3000 to 100,000 bases in most populations. Consequently, association studies view large human populations as evolutionary families and do not rely on studies of nuclear families for gene mapping.[26,27]

Some SNPs may contribute directly to a trait or disease phenotype by altering function. Though most SNPs are located outside protein-coding sequences, those within coding sequences, called cSNPs, are of particular interest because they are more likely to affect gene function. A large, well-characterized collection of SNPs will become increasingly important for the discovery of DNA sequence variations that affect biologic function. Work is already under way

with NIH support to develop a catalogue of 60,000 or more SNPs. A recently formed pharmaceutical consortium will support the production of 300,000 more, with the work being done at the publicly funded genome centers and all the data deposited in the public domain. This is a wonderful example of a public–private partnership to develop a powerful set of research tools that all can use.

NEW FORMS OF TECHNOLOGY FOR GENETIC ANALYSIS AND RISK ASSESSMENT

The transition from genetics to genomics marks the evolution from an understanding of single genes and their individual functions to an understanding of the actions of multiple genes and their control of biologic systems. Whereas the tools of the Human Genome Project initially advanced research on single genes, they are now forming the basis for genomic-scale analysis of the human organism.

The so-called DNA chip currently provides one promising approach to genome-scale studies of genetic variation,[28] detection of heterogeneous gene mutations,[29] and gene expression.[30] The result of an adaptation of dot blot hybridization techniques, DNA chips, also called microarrays, generally consist of a thin slice of glass or silicon about the size of a postage stamp on which threads of synthetic nucleic acids are arrayed. Sample probes are added to the chip, and matches are read by an electronic scanner. As with semiconductors, the capacity of DNA chips has doubled about every two years, so chips that held a few hundred arrays not so long ago now hold hundreds of thousands.

Microarray technology has been applied to the detection of DNA variations as well as expression of messenger RNA in individual cells and tissues. Microarrays are used clinically to detect human immunodeficiency virus sequence variations, p53 gene mutations in breast tissue, and expression of cytochrome P450 genes. In the laboratory, microarray technology has also been applied to genomic comparisons across species,[31] genetic recombination,[32] and large-scale analysis of gene copy number and expression, as well as protein expression, in cancerous tissues.[33]

Use of microarrays and other new technologies to detect DNA variations holds promise, along with family histories and data from large population studies, for establishing a person's risk of contracting common, adult-onset disorders. A base-line genome scan could provide helpful information about a person's risk profile and point to the prevention strategies — if available — that should be used.

GENETIC KNOWLEDGE AND INDIVIDUALIZED MEDICINE

Identifying human genetic variations will eventually allow clinicians to subclassify diseases and adapt therapies to the individual patient. There may be large differences in the effectiveness of medicines from one person to the next. Toxic reactions can also occur and in many instances are likely to be a consequence of genetically encoded host factors. That basic observation has spawned the burgeoning new field of pharmacogenomics, which attempts to use information about genetic variation to predict responses to drug therapies.

For example, researchers discovered that patients with Alzheimer's who have the $\epsilon 4$ subtype of the gene for apolipoprotein E (*APOE* $\epsilon 4$), which affects cholinergic function in the brain, are less likely to benefit from the cholinomimetic drug tacrine than are patients without this subtype.[34] This finding will help in the analysis of data from clinical trials of Alzheimer's therapies and will promote the development of new therapies specifically designed for *APOE* $\epsilon 4$ carriers.

In another example, cholesteryl ester transfer protein (CETP) plays an important part in the metabolism of high-density lipoprotein, a lipoprotein associated with lowered susceptibility to atherosclerosis. A certain genetic variant of the *CETP* gene is correlated with higher plasma CETP levels and lower levels of plasma high-density lipoprotein. One study showed that in men who carried this genetic variant, treatment with pravastatin slowed the progression of coronary atherosclerosis.[35] This finding may allow physicians to predict which patients with coronary artery disease will benefit from treatment with pravastatin.

In a third example, the formation of venous blood clots in the brain and legs is a rare but serious side effect of birth-control pills. One study has shown a dramatically increased risk of cerebral-vein thrombosis among women taking oral contraceptives who also carry the blood-clotting variant factor V Leiden.[36] The risk of other venous thrombotic events is also increased in this group. Foreknowledge of the presence of this variant and consideration of alternative forms of birth control might be useful in minimizing the risk of thrombosis in these women.

Not only will genetic tests predict responsiveness to drugs on the market today, but also genetic approaches to disease prevention and treatment will include an expanding array of gene products for use in developing tomorrow's drug therapies. Since the Food and Drug Administration's approval of recombinant human insulin in 1982, over 50 additional gene-based drugs have become available for clinical use. These include drugs for the treatment of cancer, heart attack, stroke, and diabetes, as well as many vaccines.[37]

Not all therapeutic advances for gene discovery will be genes or gene products. In other instances, molecular insights into a disorder, derived from gene discovery, will suggest a new treatment. Sodium phen-

ylbutyrate, for example, which is approved for the regulation of blood ammonia levels, is being tested in clinical trials for the treatment of cystic fibrosis.[38] The main clinical phenotype in people with cystic fibrosis results from a mutation in the gene for cystic fibrosis transmembrane conductance regulator (CFTR) protein. The mutation prevents normal amounts of CFTR protein from crossing the cell membrane, diminishing the ability of chloride and water to enter and exit the cell. Sodium phenylbutyrate apparently stimulates expression of CFTR protein, allowing more of it to reach the correct location.

ETHICAL, LEGAL, AND SOCIAL IMPLICATIONS

One of the most active areas of the ELSI program has been policy development related to the privacy and fair use of genetic information, particularly in health insurance, employment, and medical research. Debates in this area focus largely on the potential of genetic information to predict an increased likelihood of the eventual development of a disease phenotype in a currently healthy person.

Although many states have attempted to address "genetic discrimination" in health insurance and employment, federal legislation would provide the most comprehensive protection. Concern about the confidentiality of genetic information may make people reluctant to volunteer for studies involving disease-linked gene mutations or genetic therapy, for fear that the results could result in the loss of a job or the loss of insurance coverage.

Largely on the basis of recommendations formulated in workshops held by the Human Genome Project and the National Action Plan on Breast Cancer,[39,40] the Clinton administration endorsed the need for congressional action to protect against genetic discrimination in health insurance and employment. In 1996, Congress enacted the Health Insurance Portability and Accountability Act, which represented a large step toward protecting access to health insurance in the group-insurance market but left several serious gaps in the individual-insurance market that must still be closed.

In the area of workplace discrimination, the Equal Employment Opportunity Commission has interpreted the Americans with Disabilities Act as covering on-the-job discrimination based on "genetic information relating to illness, disease or other disorders."[41] But no claims of genetic discrimination have been brought to the commission, and the guidance has yet to be tested in court, so the degree of protection actually provided by the act remains uncertain.

In the area of privacy, as part of the partnership between the National Action Plan on Breast Cancer and the Human Genome Project, medical researchers, policy makers, and representatives of law, government, the insurance industry, and public health have recently assessed current policies and practices designed to protect confidentiality in genetics research and have identified areas where new or modified policies or practices might enhance the protection of privacy and promote the conduct of research. The group is developing a set of principles for researchers, research institutions, state and federal agencies, and policy makers to consider in formulating measures to protect the privacy of research information.

Other important steps have been taken to ensure the responsible integration of genetic tests into clinical practice. For the most part, genetic testing in the United States has developed successfully, providing options for avoiding, preventing, and treating inherited disorders. But the rapid pace of test development combined with the rush to market new products may create an environment in which the tests are made available before they have been adequately validated. On the recommendation of the Task Force on Genetic Testing,[42] assembled by the Human Genome Project's NIH–DOE ELSI Working Group, the Secretary of the Department of Health and Human Services has established an advisory panel to ensure the safe introduction of genetic tests into clinical practice.

Completion of the first human-genome sequence and the expansion of human genetic research to include studies of genetic variation among subpopulations have raised new questions about ethical, legal, and social issues. The 1998–2003 plan includes an examination of these issues as well as the integration of genetic technology and information into health care and public health activities; the use of knowledge about genomics and gene–environment interactions in nonclinical settings; examination of a variety of philosophical, theological, and ethical perspectives on new genetic knowledge; and consideration of the ways in which racial, ethnic, and socioeconomic factors affect the use, understanding, and interpretation of genetic information, the use of genetic services, and the development of policy.[3]

A HYPOTHETICAL CASE IN 2010

General visions of gene-based medicine in the future are useful, but many health care providers are probably still puzzled by how it will affect the daily practice of medicine in a primary care setting. A hypothetical clinical encounter in 2010 is described here.

John, a 23-year-old college graduate, is referred to his physician because a serum cholesterol level of 255 mg per deciliter was detected in the course of a medical examination required for employment. He is in good health but has smoked one pack of cigarettes per day for six years. Aided by an interactive computer program that takes John's family history, his physician notes that there is a strong paternal history of myocardial infarction and that John's father died at the age of 48 years.

To obtain more precise information about his risks of contracting coronary artery disease and other illnesses in the future, John agrees to consider a battery of genetic tests that are available in 2010. After working through an interactive computer program that explains the benefits and risks of such tests, John agrees (and signs informed consent) to undergo 15 genetic tests that provide risk information for illnesses for which preventive strategies are available. He decides against an additional 10 tests involving disorders for which no clinically validated preventive interventions are yet available.

A cheek-swab DNA specimen is sent off for testing, and the results are returned in one week (Table 1). John's subsequent counseling session with the physician and a genetic nurse specialist focuses on the conditions for which his risk differs substantially (by a factor of more than two) from that of the general population. Like most patients, John is interested in both his relative risk and his absolute risk.

John is pleased to learn that genetic testing does not always give bad news — his risks of contracting prostate cancer and Alzheimer's disease are reduced, because he carries low-risk variants of the several genes known in 2010 to contribute to these illnesses. But John is sobered by the evidence of his increased risks of contracting coronary artery disease, colon cancer, and lung cancer. Confronted with the reality of his own genetic data, he arrives at that crucial "teachable moment" when a lifelong change in health-related behavior, focused on reducing specific risks, is possible. And there is much to offer. By 2010, the field of pharmacogenomics has blossomed, and a prophylactic drug regimen based on the knowledge of John's personal genetic data can be precisely prescribed to reduce his cholesterol level and the risk of coronary artery disease to normal levels. His risk of colon cancer can be addressed by beginning a program of annual colonoscopy at the age of 45, which in his situation is a very cost-effective way to avoid colon cancer. His substantial risk of contracting lung cancer provides the key motivation for him to join a support group of persons at genetically high risk for serious complications of smoking, and he successfully kicks the habit.

This vision of genetically based, individualized preventive medicine is exciting, and it could make a profound contribution to human health. For this vision to be realized, however, protections against the misuse of genetic information will need to be firmly in place. And another critical challenge must be met: physicians, nurses, and other health care providers will need to become familiar with the emerging field of genetic medicine. The need for medical genetic specialists who can sort out the most complex cases will be considerable, but there will not be enough of them to go around, and genetic medicine will be practiced for the most part by primary care providers. Numerous surveys have shown that we are not currently prepared for this — most practitioners in primary care have not had a single hour of instruction in genetics as part of their formal training.[43,44]

TABLE 1. RESULTS OF GENETIC TESTING IN A HYPOTHETICAL PATIENT IN 2010.

CONDITION	GENES INVOLVED*	RELATIVE RISK	LIFETIME RISK (%)
Reduced risk			
Prostate cancer	*HPC1, HPC2, HPC3*	0.4	7
Alzheimer's disease	*APOE, FAD3, XAD*	0.3	10
Elevated risk			
Coronary artery disease	*APOB, CETP*	2.5	70
Colon cancer	*FCC4, APC*	4	23
Lung cancer	*NAT2*	6	40

**HPC1, HPC2,* and *HPC3* are the three genes for hereditary prostate cancer; *APOE* is the gene for apolipoprotein E; *FAD3* and *XAD* are hypothetical genes for familial Alzheimer's dementia; *APOB* is the gene for apolipoprotein B; *CETP* is the gene for cholesteryl ester transfer protein, *FCC4* is the hypothetical gene for familial colon cancer; *APC* is the gene for adenomatous polyposis coli; and *NAT2* is the gene for *N*-acetyltransferase 2.

To meet this urgent need for education, the National Human Genome Research Institute, the American Medical Association, and the American Nurses Association have recently banded together to form the National Coalition for Health Professional Education in Genetics (NCHPEG). NCHPEG (accessible at http://www.nchpeg.org) is a national effort to promote professional education and access to information about advances in human genetics. An interdisciplinary group, NCHPEG comprises the leaders of approximately 100 diverse health professional organizations, consumer and voluntary groups, government agencies, private industry, managed-care organizations, and professional genetics societies. By facilitating frequent and open communication, NCHPEG seeks to capitalize on the collective expertise and experience of its members and to reduce duplication of effort.

CONCLUSIONS

Since the beginning of the Human Genome Project, the increasing detail and quality of genome maps have reduced the time it takes to find a gene from years to months to weeks. We have learned the location of approximately half the 80,000 or so genes packaged on human chromosomes. Genome researchers are at the forefront of cross-disciplinary technological development for the identification and analysis of not just single genes but also whole genomes.

Most recently, we have begun to get the first glimpses of the human genome at its most detailed level: about 15 percent of the 3 billion bits of DNA code that spell out instructions for every function a human body carries out is now available in public

data bases. In just 12 months, a highly useful working draft of 90 percent of the genome will be available, and by 2003, the full DNA sequence of the human will give us unprecedented opportunities to observe and understand the literal Book of Life.

As hoped, genome maps, sequence data, and analytic tools are providing a robust technological infrastructure for biomedical research well into the next century. The Human Genome Project's commitment to make these powerful tools freely available means that any scientist with access to the Internet can already apply genomic approaches to individual research problems. High-technology research tools developed by the Human Genome Project will offer unprecedented opportunities to study human biology and disease in entirely new ways.

As genome technology moves from the laboratory to the health care setting, new methods will make it possible to read the instructions contained in an individual person's DNA. Such knowledge may foretell future disease and alert patients and their health care providers to undertake better preventive strategies. In the wrong hands, however, that same information could be used to discriminate against or stigmatize a person. In response to this concern, the Human Genome Project has catalyzed the development of policy options for lawmakers to consider in their efforts to prohibit genetic discrimination and to protect the privacy of genetic information. The stage is set to solve these vexing problems with effective federal legislation, but this window of opportunity will not stay open indefinitely.

Writing 97 years ago, Sir William Osler described the goals of medicine this way: "To wrest from nature the secrets which have perplexed philosophers in all ages, to track to their sources the causes of disease, to correlate the vast stores of knowledge, that they may be quickly available for the prevention and cure of disease — these are our ambitions."[45] The Human Genome Project, with its audacious goal of providing the tools to uncover the hereditary factors in virtually every disease, has become a major modern component of Osler's vision. The genetic revolution in medicine is under way.

I am indebted to Leslie Fink and Dr. Margaret Bouvier for their assistance in the preparation of the manuscript.

REFERENCES

1. National Research Council Committee on Mapping and Sequencing the Human Genome. Mapping and sequencing the human genome. Washington, D.C.: National Academy Press, 1988.
2. Department of Health and Human Services, Department of Energy. Understanding our genetic inheritance: the U.S. Human Genome Project: the first five years: FY 1991-1995. Washington, D.C.: Government Printing Office, 1990. (NIH publication no. 90-1590.)
3. Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, Walters L. New goals for the U.S. Human Genome Project: 1998-2003. Science 1998;282:682-9.
4. Murray JC, Buetow KH, Weber JL, et al. A comprehensive human linkage map with centimorgan density. Science 1994;265:2049-54.
5. Deloukas P, Schuler GD, Gyapay G, et al. A physical map of 30,000 human genes. Science 1998;282:744-6.
6. Collins FS, Galas D. A new five-year plan for the U.S. Human Genome Project. Science 1993;262:43-6.
7. The yeast genome directory. Nature 1997;387:Suppl:5.
8. Blattner FR, Plunkett G III, Bloch CA, et al. The complete genome sequence of Escherichia coli K-12. Science 1997;277:1453-74.
9. Collins FS. Positional cloning: let's not call it reverse anymore. Nat Genet 1992;1:3-6.
10. *Idem.* Positional cloning moves from perditional to traditional. Nat Genet 1995;9:347-50. [Erratum, Nat Genet 1995;11:104.]
11. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. Science 1997;276:2045-7.
12. Gasser T, Muller-Myhsok B, Wszolek ZK, et al. A susceptibility locus for Parkinson's disease maps to chromosome 2p13. Nat Genet 1998;18:262-5.
13. Jones AC, Yamamura Y, Almasy L, et al. Autosomal recessive juvenile parkinsonism maps to 6q25.2-q27 in four ethnic groups: detailed genetic mapping of the linked region. Am J Hum Genet 1998;63:80-7.
14. Leroy E, Boyer R, Auburger G, et al. The ubiquitin pathway in Parkinson's disease. Nature 1998;395:451-2.
15. Ueda K, Fukushima H, Masliah E, et al. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. Proc Natl Acad Sci U S A 1993;90:11282-6.
16. Burke W, Thomson E, Khoury MJ, et al. Hereditary hemochromatosis: gene discovery and its implications for population-based screening. JAMA 1998;280:172-8.
17. Feder JN, Gnirke A, Thomas W, et al. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. Nat Genet 1996;13:399-408.
18. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 1994;226:66-71.
19. Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene BRCA2. Nature 1995;378:789-92. [Erratum, Nature 1996;379:749.]
20. Yamagata K, Furuta H, Oda O, et al. Mutations in the hepatocyte nuclear factor-4α gene in maturity-onset diabetes of the young (MODY 1). Nature 1996;384:458-60.
21. Froguel P, Zouali H, Vionnet N, et al. Familial hyperglycemia due to mutations in glucokinase: definition of a subtype of diabetes mellitus. N Engl J Med 1993;328:697-702.
22. Yamagata K, Oda N, Kaisaki PJ, et al. Mutations in the hepatocyte nuclear factor-1α gene in maturity-onset diabetes of the young (MODY3). Nature 1996;384:455-8.
23. Fishel R, Lescoe MK, Rao MR, et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell 1993;75:1027-38. [Erratum, Cell 1994;77:167.]
24. Leach FS, Nicolaides NC, Papadopoulos N, et al. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell 1993;75:1215-25.
25. Papadopoulos N, Nicolaides NC, Wei YF, et al. Mutation of a mutL homolog in hereditary colon cancer. Science 1994;263:1625-9.
26. Schafer AJ, Hawkins JR. DNA variation and the future of human genetics. Nat Biotechnol 1998;16:33-9.
27. Collins FS, Guyer MS, Charkravarti A. Variations on a theme: cataloging human DNA sequence variation. Science 1997;278:1580-1.
28. Wang DG, Fan J-B, Siao C-J, et al. Large-scale identification, mapping and genotyping of single-nucleotide polymorphisms in the human genome. Science 1998;280:1077-82.
29. Hacia JG, Brody LC, Chee MS, Fodor SP, Collins FS. Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-colour fluorescence analysis. Nat Genet 1996;14:441-7.
30. DeRisi J, Penland L, Brown PO, et al. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat Genet 1996;14:457-60.
31. Hacia JG, Makalowski W, Edgemon K, et al. Evolutionary sequence comparisons using high-density oligonucleotide arrays. Nat Genet 1998;18:155-8.
32. Winzeler EA, Richards DR, Conway AR, et al. Direct allelic variation scanning of the yeast genome. Science 1998;281:1194-7.
33. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844-7.
34. Poirier J, Delisle M-C, Quirion R, et al. Apolipoprotein E4 allele as a predictor of cholinergic deficits and treatment outcome in Alzheimer disease. Proc Natl Acad Sci U S A 1995;92:12260-4.

35. Kuivenhoven JA, Jukema JW, Zwinderman AH, et al. The role of a common variant of the cholesteryl ester transfer protein gene in the progression of coronary atherosclerosis. N Engl J Med 1998;338:86-93.
36. Martinelli I, Sacchi E, Landi G, Taioli E, Duca F, Mannucci PM. High risk of cerebral-vein thrombosis in carriers of a prothrombin-gene mutation and in users of oral contraceptives. N Engl J Med 1998;338:1793-7.
37. New medicines in development: biotechnology 1998. Washington, D.C.: Pharmaceutical Research and Manufacturers of America, 1998:35.
38. Rubenstein RC, Zeitlin PL. A pilot clinical trial of oral sodium 4-phenylbutyrate (Buphenyl) in deltaF508-homozygous cystic fibrosis patients: partial restoration of nasal epithelial CFTR function. Am J Respir Crit Care Med 1998;157:484-90.
39. Hudson KL, Rothenberg KH, Andrews LB, Kahn MJ, Collins FS. Genetic discrimination and health insurance: an urgent need for reform. Science 1995;270:391-3.
40. Rothenberg K, Fuller B, Rothstein M, et al. Genetic information and the workplace: legislative approaches and policy challenges. Science 1997; 275:1755-7.
41. Compliance manual. Section 902. Washington, D.C.: Equal Employment Opportunity Commission, March 15, 1995.
42. Holtzman NA, Watson MS, eds. Promoting safe and effective genetic testing in the United States: final report of the Task Force on Genetic Testing. Baltimore: Johns Hopkins University Press, 1998.
43. Scanlon C, Fibison W. Managing genetic information: implications for nursing practice. Washington, D.C.: American Nurses Publishing, 1995:1-50
44. National Center for Genome Resources. National survey of public and stakeholders' attitudes and awareness of genetic issues. Washington, D.C.: Schulman, Ronca, and Bucuvalas, 1996.
45. Osler W. Chauvinism in medicine. Montreal Med J 1902;31:684-99.

Contact: Mary Prescott 1-312-397-6604
Jeanne Corrigan 1-312-397-6610

THE SNP CONSORTIUM MAKES FIRST SET OF GENE MARKERS PUBLICLY AVAILABLE

- More than 2,000 Novel Single Nucleotide Polymorphisms (SNPs) Now Accessible to Medical Researchers Worldwide -

- SNP Map To Aid Discovery and Development of Safer and More Effective Therapeutics -

Chicago (November 23) – The SNP Consortium Ltd., a collaborative effort to create a database of genetic markers called single nucleotide polymorphisms (SNPs), today released into the public domain approximately 2300 newly identified and characterized SNPs. These data thus become available for the free and unrestricted use of biomedical researchers worldwide.

The Consortium's SNP data set can be accessed via the Internet, both through The SNP Consortium's website (http://snp.cshl.org/), and "dbSNP" (http://www.ncbi.nlm.nih.gov/), a public database maintained by the National Institutes of Health's National Center for Biotechnology Information.

"Members of the SNP Consortium believe that a high-quality SNP map will prove to be an essential tool for understanding the genetic basis of disease, and as such, should not be subject to intellectual property restrictions," says Arthur Holden, chairman and chief executive officer of The SNP Consortium. "Today's data release is the first of many we plan over the next two years as we work to construct a SNP map that is extensive, reliable, widely accepted, and freely available to medical researchers in industry, academic, government and independent laboratories. We are very excited about this contribution and its impact on speeding up effective pharmacogenomic research within the life sciences sector."

- more –

The SNP Consortium 2-2-2

Single nucleotide polymorphisms – or SNPs (pronounced "snips") – are minute genetic variations that occur throughout human DNA. Their value as genetic markers resides in their simplicity, frequency, and relatively even distribution throughout the genome (the full set of genetic instructions, encoded in long strands of DNA). Of the roughly 3 billion nucleotide pairs (i.e., the "letters") that make up the genetic code, it is estimated that a SNP – a single letter variation in the code from one person to the next – occurs every thousand or so nucleotide pairs.

A high-density map of these genetic signposts will allow researchers to navigate the genome more quickly and efficiently, in search of genes associated with disease. By comparing SNP patterns in various patient and control populations, medical researchers hope to identify genetic differences that predispose some but not others to diseases, and that underlie variability in individual responses to treatment. In turn, this knowledge is expected to enhance understanding of disease processes and facilitate discovery, development and delivery of safer and more effective treatments. Ultimately, SNP technology may lead to the development of "personalized medicine," in which disease prevention strategies and treatments are more closely tailored to an individual's genetic profile.

The SNP Consortium intends over two years to identify 300,000 and map at least 150,000 SNPs, evenly distributed throughout the genome. To date, the Consortium has identified over 15,000 candidate SNPs. The data set released today comprises 2279 SNPs that have been mapped to the locations where they occur on the chromosomes. Protocols for quality assurance indicate a greater than 95 percent confidence level in the validity of the data (that is, that the genetic code indeed varies from one person to another by a single letter at the site of each SNP). Subsequent releases of data will occur as additional SNPs are mapped and their authenticity verified.

As it is constructed, the SNP map will be made publicly available to medical researchers worldwide at the same time it becomes available to members of the Consortium, according to Lincoln Stein, MD, PhD, of the bioinformatics program at Cold Spring Harbor Laboratory, the group responsible for managing the Consortium's SNP database.

The SNP Consortium 3-3-3

The SNP Consortium – an innovative collaboration among industry, the world's largest medical research charity, and several leading academic institutions – is organized as a non-profit entity whose goal is to create and make publicly available a high-quality SNP map of the human genome. The Consortium's members include the Wellcome Trust, Motorola BioChip Systems, and ten pharmaceutical companies: AstraZeneca PLC, Bayer AG, Bristol-Myers Squibb Company, F. Hoffmann-La Roche, Glaxo Wellcome PLC, Hoechst Marion Roussel AG, Novartis, Pfizer Inc, Searle, and SmithKline Beecham PLC.

Academic centers including the Whitehead Institute for Biomedical Research, Washington University School of Medicine in St. Louis, the Wellcome Trust's Sanger Centre, and the Stanford Human Genome Center are involved in SNP identification and mapping. The bioinformatics program at Cold Spring Harbor Laboratory organizes, analyzes, and manages the resulting SNP database, and oversees dissemination of the information contained in the database.

# # #

American College of Medical Genetics

Position Statement on Gene Patents and Accessibility of Gene Testing

The fruits of the human genome project are rapidly redefining the medical community's views of patient care and moving us from a focus on the treatment of disease to a broader perspective in which prevention and diagnosis play equally critical roles. The American College of Medical Genetics (ACMG) believes that gene testing – which is vital to all three approaches -- must remain widely accessible and affordable, and that the development and improvement of safe and effective genetic tests should not be hindered. The decision of the Patent and Trademark Office (PTO) to permit the patenting of naturally occurring genes and disease-causing mutations has produced numerous difficulties. While the ACMG disagrees with the PTO over this fundamental issue, we have further concerns over current patterns of enforcement of patents on genes that are important in the diagnosis, management and risk assessment of human disease.

This statement is directed at current practices related to such enforcement. Enforcement has been effected in one or more of these ways: monopolistic licensing that limits a given genetic test to a single laboratory, royalty-based licensing agreements with exorbitant up-front fees and per-test fees, and licensing agreements that seek proportions of reimbursement from testing services. These limit the accessibility of competitively priced genetic testing services and hinder test-specific development of national programs for quality assurance. They also limit the number of knowledgeable individuals who can assist physicians, laboratory geneticists and counselors in the diagnosis, management and care of at-risk patients.

Further, restricting the availability of gene testing has long-term implications beyond patient care. It affects the training of the next generation of medical and laboratory geneticists, physicians, and scientists in the area enveloped by the patent or license. It also retards the usually very rapid improvement of a test that occurs through the addition of new mutations or the use of new techniques by numerous laboratories that have accumulated samples from affected individuals over many years.

Therefore, it is the ACMG's position that:

- Genes and their mutations are naturally occurring substances that should not be patented.
- Patents on genes with clinical implications must be very broadly licensed.
- Licensing agreements should not limit access through excessive royalties and other unreasonable terms.

August 2, 1999

Cc Eric, Chris, + return

-BR

THE WHITE HOUSE

WASHINGTON

October 1, 1999

MEMORANDUM FOR THE PRESIDENT

FROM: NEAL LANE Neal

SUBJECT: GENOME PATENTING

This memorandum responds to your query regarding the merits of genome patenting. Genome patenting is a complex issue that has been debated since large-scale DNA sequencing became feasible in the late 1980s.

What *is* patentable?

Patents have been granted to applicants who can show that they have obtained a genomic DNA sequence, can describe the protein product encoded by that gene, and can demonstrate the function of the protein in a living organism. Substantial uncertainty exists regarding the patentability of partial or full gene sequences when the applicant does not disclose information regarding the function or utility of the gene and its encoded protein product. The U.S. Patent and Trademark Office (PTO) has granted at least one patent on a partial gene sequence but is rethinking its position. PTO convened a public hearing on Nov. 5, 1998 to obtain public comments on draft guidelines describing the necessary elements for obtaining a patent on genome sequences. I understand PTO plans to issue new draft guidance by November 1.

In 1991, NIH filed patent applications claiming over 6,000 partial gene sequences obtained in the laboratory of Dr. J. Craig Venter but has never appealed PTO's denial of these patents. Under Harold Varmus, NIH adopted a policy promoting open access to research tools, including genomic sequence data.

PTO and NIH have been at loggerheads over the criteria necessary to obtain patents on genomic DNA sequence data, although they agree more generally on the value of patents on the results of later-stage genome research in stimulating investment in health care product development.

What *should* be patented?

NIH discourages grantees from seeking patents on so-called raw human genomic DNA sequence data. In 1997, NIH, DOE and the U.K.'s Wellcome Trust, partners in the international Human Genome Project, adopted a policy (attached) requiring grantees funded specifically to generate large-scale human genomic DNA sequence to deposit such data in a public database within 24 hours of its generation, thus rendering it unpatentable. A September 20th *London Guardian* article inaccurately portrayed this as a new policy designed to preclude all gene patents. (Press guidance on this article is attached.)

The rationale for freely disseminating raw human genomic DNA sequence data is that such broad public access will *promote innovation* and encourage its use in the development of downstream products. Moreover, the non-profit supporters of genomic research believe patents on early-stage genome data are a *disincentive to innovation*. That is, drug companies and other innovators may be reluctant to invest in product development on the suspicion that, having done so, they will learn belatedly that a patent has been granted on a gene sequence critical to their product. This will necessitate negotiation of a license for marketing to proceed. This is not an unreasonable fear, as some companies describe such patenting as their principal goal, and it may be years before the existence of such patents becomes known. While most agree on the value of patents for genes about which there is known function or clear utility (for example, diagnostic genetic tests), exclusive licenses on these patents are controversial due to their potential for restricting the availability of new health care advances.

Not surprisingly, industry views on patenting the early results of genome sequencing are mixed, depending on the company's business strategy. A few companies hope to be able to obtain patents on large numbers of raw gene sequences. One significant occurrence is the April 15th announcement of a consortium formed among 10 of the world's foremost pharmaceutical companies to create a public database of the genetic markers that distinguish one person from another. These markers will be valuable for identifying specific genes involved in common and rare diseases, developing new diagnostic tests, and creating new "personalized" medicines, based on an understanding of minute genetic variations. Members of this consortium believe that it is preferable to have these indicators of human genetic diversity in the public domain, as opposed to dealing with the licensing issue at a later date. They clearly do not feel that making this information public hampers their efforts to develop downstream products based on genetic information or threatens their future intellectual property rights.

The Bayh-Dole Act and other technology transfer statutes give Federal grantees and contractors the right to apply for patents on the results of government-funded research. The NIH-DOE policy requiring rapid access to raw human genomic DNA sequence data precludes the patenting of such information, but once an investigator has determined the biological significance of such information, obtaining a patent is straightforward. It seems a reasonable course of action to continue the policy of making raw human genomic DNA sequence widely available to the largest possible number of inventors, thus stimulating further innovation and product development.

I believe PTO's guidelines will reflect a bright line approach allowing patents on gene sequences only when the inventor can demonstrate how such sequences might be used. I believe this approach is the right one for advancing science and for ensuring rapid development of products based on that science. The current climate of uncertainty regarding the boundaries of patentability is a disincentive to genomic product development, and issuing the guidelines will help resolve this problem.

Attachments

cc: John Podesta
Bruce Reed
David Beier

The National Human Genome Research Institute

NHGRI Policy Regarding Intellectual Property of Human Genomic Sequence

POLICY ON AVAILABILITY AND PATENTING OF HUMAN GENOMIC DNA SEQUENCE PRODUCED BY NHGRI PILOT PROJECTS (FUNDED UNDER RFA HG-95-005)

April 9, 1996

This document describes the policy of the National Center for Human Genome Research with respect to availability and patenting of human genomic DNA sequence produced under grants funded as a result of RFA HG-95-005. In conformity with the existing spirit and philosophy of the Human Genome Project and in response to the recommendations of advisors and the expressed wishes of the community, NHGRI seeks to make DNA sequence information available as rapidly and freely as possible.

Background

The Human Genome Project (HGP) is an international research effort, begun in 1990, which has the scientific goals of generating maps of the human genome[1] and producing the complete sequence of the human DNA by the year 2005. The project was undertaken in the U.S. following the advice of several scientific committees that emphasized its importance in creating a resource that "will facilitate research in biochemistry, physiology and medicine"[2], "have a major impact on health care and disease prevention"[2] and provide ""enormous scientific and technological advances..., having both basic and commercial applications" [3]. At NIH, the National Human Genome Research Institute (NHGRI) was founded to implement the HGP.

The HGP has progressed rapidly, even beyond optimistic expectations. The initial mapping goals are nearly completed and recent improvements in DNA sequencing technology and capacity have led many scientists, including NHGRI advisors, to conclude that complete sequencing of human genomic DNA should begin. Early in 1995, NHGRI issued RFA HG-95-005 to solicit grant applications for pilot projects to test strategies that can potentially scale up to sequence the human genome. The applications received in response to the RFA were peer reviewed in the fall of 1995 and approved by the National Advisory Council for Human Genome Research in January 1996. A set of grants will be funded by April 1996.

At the inception of the HGP, the planners emphasized that, in order to reap the maximum benefit from the HGP, human DNA sequence should be freely available in the public domain. The NIH Ad Hoc Program Advisory Committee on Complex Genomes stated that "Distribution of and free access to the databases (containing the sequence data) must be fully encouraged. Thus, the data must be in the public domain, and the redistribution of the data should remain free of royalties."[3] Similarly, the National Research Council stated: "...access to all sequences and material generated by these publicly funded projects should and even must be made freely available..."[2]. Most recently, an international group of scientists, from both the public and private sectors, who are already involved in genomic DNA sequencing, passed a unanimous resolution that "all human genomic DNA sequence information, generated by centers funded for large-scale human sequencing, should be freely available and in the public domain in order to

encourage research and development and to maximize its benefit to society"[4].

There are very strong scientific arguments that human genomic DNA sequence should be freely available and in the public domain:

o The human genomic DNA sequence is unique. Although there are many other types of information that contribute to the understanding of human biology, e.g., DNA sequence of model organism genomes, in the end, the only source of definitive information about the human is the human sequence.

o The human genomic DNA sequence is a vast resource. It contains a very large number of genes and an enormous amount of additional biological information. It is anticipated that the sequence resource will be the basis for many useful inventions and patentable products. It will take many researchers years to find and characterize all of the genes and other functional elements within the sequence and to use that information to develop products and other approaches that will improve the health of the American people.

o The human genome is a bounded resource. Once the genome has been sequenced, few or no opportunities will exist for discovery of new information that will not make reference to, or be dependent on, that first sequence. Thus, it is important to ensure maximum access of a large number of parties to the initial genomic DNA sequence as it is generated, to provide a broad opportunity for development of new products.

Policy

It is therefore NHGRI s intent that human genomic DNA sequence data, generated by the projects funded under RFA HG-95-005, should be released as rapidly as possible and placed in the public domain where it will be freely available. In order to implement this policy, NHGRI will require that grantees under RFA HG-95-005 adopt a policy of rapid release of data to public databases. This policy will be made a condition of the award.

In NHGRI s opinion, raw human genomic DNA sequence, in the absence of additional demonstrated biological information, lacks demonstrated specific utility and therefore is an inappropriate material for patent filing. NIH is concerned that patent applications on large blocks of primary human genomic DNA sequence could have a chilling effect on the development of future inventions of useful products. Companies are not likely to pursue projects where they believe it is unlikely that effective patent protection will be available. Patents on large blocks of primary sequence will make it difficult to protect the fruit of subsequent inventions resulting from real creative effort. However, according to the Bayh-Dole Act, the grantees have the right to elect to retain title to subject inventions and are free to choose to apply for patents should additional biological experiments reveal convincing evidence for utility. The grantees are reminded that the grantee institution is required to disclose each subject invention to the Federal Agency providing research funds within two months after the inventor discloses it in writing to grantee institution personnel responsible for patent matters. NHGRI will monitor grantee activity in this area to learn whether or not attempts are being made to patent large blocks of primary human genomic DNA sequence.

During this pilot period, NHGRI will be soliciting opinions and collecting evidence from the broad scientific and commercial sectors to allow an evaluation of whether the approach described above is sufficient to ensure that sequence generated by these grants is maximally useful to the research and commercial sectors. If not, NIH will consider a determination of exceptional circumstance to restrict or eliminate the right of parties, under future grants, to elect to retain title.

The National Human Genome Research Institute

NHGRI Policy on Release of Human Genomic Sequence Data

March 7, 1997

At the Second International Strategy Meeting on Human Genome Sequencing (Bermuda, 1997), attendees affirmed the principle that was set out at the First (1996) International Strategy meeting, that primary genomic sequence should be rapidly released. Specifically, the report of the first meeting stated that "sequence assemblies should be released as soon as possible; in some centres, assemblies of greater than 1 kb would be released automatically on a daily basis." The discussions at the 1997 meeting confirmed NHGRI's conclusions that it is extremely important for its large-scale sequencing program to be functioning in a manner consistent with this principle, that such rapid release is technically feasible, and that such unfinished DNA sequence data have already been found to be useful by the larger scientific community. NHGRI has determined, therefore, that its grantees engaged in large-scale genomic DNA sequencing should now be automatically releasing sequence assemblies of 2 kb or larger within 24 hours of their generation. (the trigger for data release is 2 kb, instead of 1 kb, in order to ensure that the released sequence be comprised of at least two sequence reads. Investigators who wish to release smaller assemblies may do so.) Any laboratory funded by NHGRI for large-scale human genomic sequencing must develop and submit to NHGRI a plan to implement such a data release program, which must be implemented within one month of its being approved by NHGRI. No non-competing or competing renewal will be funded until an acceptable plan has been approved. Mandatory data release as described above will be made a condition of the award for any grant funded by NHGRI for large-scale human sequencing.

Genomic & Genetic Data | Human Genome Project
Grant Information | About NHGRI
Policy & Public Affairs | Intramural Research
Offsite Resources | Search

webmaster@nhgri.nih.gov

Genomic and Genetic Data | Grant Information | Policy and Public Affairs |
Offsite Resources | The Human Genome Project | About NHGRI | Intramural Research | Keyword Search

webmaster@nhgri.nih.gov

Human Gene Patents
September 20, 1999

Context: *Today's* London Guardian *ran a story stating that Prime Minister Tony Blair and President Clinton are negotiating a deal to obtain intellectual property rights on the human genome, thereby precluding private sector ventures in this area. U.S. and U.K. government officials have been discussing the possibility of issuing a statement endorsing data access principles already adopted by an international consortium of human genome researchers, but the article's references to government efforts to block private sector patents are inaccurate. The President and Prime Minister have not been involved in these discussions.*

Background: The international Human Genome Project is funded primarily by the National Institutes of Health and Department of Energy in the U.S., and the Wellcome Trust, a non-profit philanthropy in the U.K. Researchers funded by these organizations developed the so-called "Bermuda Principles" under which human genome sequence data is deposited in a publicly-accessible database within 24 hours of being generated. For the past year not-for-profit partners have been discussing the possibility of a memorandum of understanding with Celera Genomics, a U.S. company beginning human genome sequencing. However, data access has remained a sticking point in development of the M.O.U. The public Human Genome Project does not maintain intellectual property rights to the data.

Is the President discussing a deal with the U.K. Prime Minister to stop entrepreneurs from profiting from gene patents?

A. No. The President has not been engaged in these discussions. Government officials in both countries are in general agreement that DNA sequenced data developed under the Human Genome Project ought to be made freely available so as to promote progress in biomedical research and the rapid application of such information to the development of treatment, prevention and cure of human disease. We believe that a joint statement in support of this data access policy might be valuable.

Q. Who will own the data from the Human Genome Project?

A. All data are to be deposited daily in publicly accessible databases to ensure that scientists around the world have unfettered access, free of charge, to the wealth of information and can use it in their research to understand human disease, develop new diagnostics and therapeutic interventions, and improve the health of our citizens.

Because data are deposited daily in publicly accessible sites, no patent protection on raw fundamental genome information arising from the Human Genome Project will be sought by participants.

Gwin Smith

Bill Clinton

Neil Lane can you reply? John

Genome patenting

Bruce

Do we have a position on Genome patenting. There have been several articles on whether it's fair to Medical Research. Is concerned the Human Big Issue —

B

9-23-99

copied Podesta Reed

President of the United States

News Home Page
Photo Galleries
Politics
Nation
World
Metro
Business/Tech
Sports
Style
Travel
Health
Opinion
Weather
Weekly Sections
News Digest
Classifieds
Print Edition
Archives
News Index
Help

Partner Sites:
Newsweek.com
BRITANNICA.COM

Nation Toolbox
On the Web
Census Information
Federal crime data
Economy by region
Stateline.org

Scientists Decode Layout Of Human Chromosome

By Rick Weiss
Washington Post Staff Writer
Thursday, December 2, 1999; Page A01

Scientists from three countries announced yesterday that they had jointly determined the piece-by-piece order of virtually all of the 34 million chemical "letters" that spell out the genetic code of an entire human chromosome.

The achievement provides the first complete molecular script for a single human chromosome, 23 pairs of which carry the estimated 80,000 genes that provide the instructions for constructing a human body.

Researchers around the world praised the advance as a technical tour de force, a triumph of international scientific collaboration, and evidence that the publicly funded Human Genome Project is on track to sequencing by 2002 all of the 3.2 billion units of genetic code that together describe how to make a person.

Scientists hope that by noting "spelling errors" in the chromosomes of people with various diseases, they will be able to understand the molecular underpinnings of those ailments and develop new therapies. But first they must determine the normal sequences for each chromosome, as they now have for chromosome 22, the second smallest of the human chromosomes.

"For the hundreds of genes that are on this chromosome, we now have their anatomy," said Francis Collins, chief of the National Human Genome Research Institute, which funded most of the U.S. contribution to the project. "And for all the medical conditions [that are related to genes on that chromosome], we'll be that much further along understanding them."

About three dozen diseases are known to be caused by faulty genes on chromosome 22. Most are rare, such as DiGeorge syndrome, which involves abnormal development of the immune system, and there is indirect evidence that a gene on chromosome 22 may contribute to the risk of schizophrenia. But with scientists expecting to find a total of 700 or 800 genes embedded in the chromosome's code, scores of other diseases may eventually be shown to have their roots there.

"I can tell you it is no fun chasing down a disease gene when you're in territory no one has ever been in before," Collins said at a Washington news conference yesterday. For scientists hunting for important genes on chromosome 22, he said, having the complete sequence in hand is "a dream come true."

dream come true."

See and hear it all firsthand

Chromosomal sequence analysis also may provide insights into human evolution, because large chunks of molecular text in human chromosomes are very similar to certain DNA chunks in animal chromosomes. A comparison of those molecular scripts could reveal how and when various organisms branched off from one another and diverged over millions of years.

"It's like seeing the surface or the landscape of a new planet for the first time," said Mark Patterson, an editor at Nature, the scientific journal that is publishing the chromosome report today. "It's allowing us to say something about its geography, and it's also allowing us to say something about its history, . . . how the chromosome evolved."

A chromosome, like a skein of yarn, is a single long strand of DNA folded and refolded into a plump, cinch-waisted bundle inside a cell. The DNA strand is made of four chemical subunits--the four "letters" of the genetic alphabet--strung together by the millions. Anywhere from 1,000 to a half-million of those letters typically make up a gene, and the order of the letters within the gene determines what the gene does in the body.

The exact meaning of chromosome 22's molecular sequence remains mostly a mystery. In some ways, it's like a long run-on sentence in a foreign language for which only a modest number of words have been translated into English. Scientists understand what the translated words mean, and they now know the remaining letters, but most of it remains untranslated and therefore the meaning remains unknown. Moreover, only about 3 percent of the script is believed to encode working genes, with most of the rest of the sequence being "junk" DNA, the function of which remains unknown.

But with chromosome 22's complete script in hand--and with similar texts already being translated in other organisms, such as the mouse and the fruit fly--scientists said it won't be long before they become fluent in the language of human genes and begin to make profound discoveries about embryo development, genetic diseases and aging.

The scientific report lists more than 200 scientists as co-authors, an astonishing number that Nature editors said was probably a record. Four institutions cooperated in the venture: the Sanger Center in Cambridge, England; Keio University School of Medicine in Tokyo; the University of Oklahoma in Norman; and Washington University in St. Louis.

Technically, the sequence is not 100 percent complete. Small gaps, constituting a total of about 3 percent of the chromosome's "letters," have proven impossible to decode, in some cases because they contain confusing strings of repeating motifs. And one specialized end of the chromosome, the DNA of which helps cells make another kind of genetic material called RNA, is not included in the new analysis.

But according to a 1996 international convention, a chromosome that is at least 95 percent sequenced and whose remaining gaps are small and well delineated can be deemed "complete." Various analyses showed that the completed sequence contains less than one error for every 50,000 letters.

Researchers on the chromosome 22 project emphasized yesterday that the results of their work were made freely available to other scientists on the World Wide Web every day as they accumulated. And in stark contrast to some privately financed efforts--most notably, that of controversial geneticist and entrepreneur J. Craig Venter of Celera Genomics Corp. in Rockville--they said no one on the team had filed for patents on the genes they had found.

The Genetic Alphabet

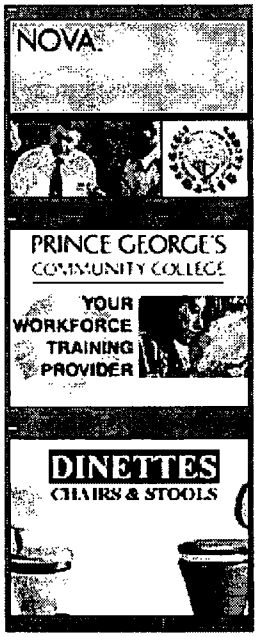
Every cell of every living organism contains genetic material made up of the compounds adenine, thymine, cytosine and guanine, represented as A, T, C and G. The precise order of these compounds regulates every aspect of an organism's life, from conception to death.

E-Mail This Article

From The Washington Post

- **Scientists Decode Layout Of Human Chromosome**
- **Loopholes Lift Candidates' Spending Caps**
- **BEYOND 2000**
- **CORRECTIONS**

Back to the top



Genome Project Reaches Milestone

1 Billion Units of Information Decoded

By Justin Gillis
Washington Post Staff Writer
Wednesday, November 24, 1999; Page E01

Comparing the effort to unravel all human genes to the push a generation ago that put men on the moon, government leaders yesterday celebrated a milestone in the project: the decoding of 1 billion units of genetic information, roughly a third of the total.

When the project is done, no later than 2003, "we will better understand the miracle of how a single cell develops into a unique human being," Donna Shalala, the secretary of health and human services, said in a speech praising the thousands of researchers around the world who are part of the effort. "We will better understand the history of life on earth. And we will better understand what binds us all together as human beings."

Managers of the Human Genome Project, as the international collaboration is known, called a quick celebration at the National Academy of Sciences yesterday to mark a milestone that was actually reached on Nov. 17. On that day, the billionth unit, or base pair, of human genetic information was deposited in GenBank, the giant Internet database that holds the results of gene research.

Because the entire human genetic code is estimated to contain about 3.2 billion base pairs, researchers believe they're about a third of the way toward finishing a draft gene map. Francis Collins, director of the National Human Genome Research Institute, said the draft should be completed by next spring, with two to three years of polishing to follow.

Yesterday's celebration emphasized the strong international flavor of the work, which is being done by 16 laboratories in the United States, Britain, France, Germany, Japan and China. The effort, expected to cost at least $3 billion by the time it is complete, is being supported in this country by the Department of Health and Human Services and the Department of Energy.

Researchers joked yesterday about the difficulty of explaining their work. In one sense, they are simply reading the order of letters in a highly repetitive template that regulates human biology. There are only four letters: A, T, C and G, representing the chemicals adenine, thymine, cytosine and guanine. They appear in monotonous strings that look like GAATTCCATA CATTTCTTCT, and go on and on for billions of letters.

But the order of those letters is the key to many of the great mysteries

of life. Every human cell has a copy of the master template, and the particulars of genetic sequence determine how that cell behaves. The Human Genome Project promises to shed light on such puzzles as how people respond to infectious disease, why some people get cancer and others don't, how arteries in the heart clog up, how babies develop and what can go wrong in the process, and why people grow old and die.

The Human Genome Project was on a leisurely course for completion in 2005 when its leaders were shocked last year by a series of developments. Several private firms, benefiting from technology developed through years of public investment in the project, announced large, rapid gene-sequencing ventures. Most notably, Celera Genomics Corp. of Rockville said it would unravel the entire human genome faster than the Human Genome Project.

The competition sparked a notable acceleration in the public venture. The goal of a rough draft is new, and the final deadline was moved up two years. Collins said yesterday that most of the billion base pairs now in GenBank were unraveled in the past seven months, indicating that the project is progressing rapidly.

Several institutions were praised yesterday for making particularly large contributions to the project. They included gene-sequencing units at the Baylor College of Medicine in Houston, at Washington University in St. Louis, and at the Massachusetts Institute of Technology in Cambridge, Mass.; a Department of Energy unit called the Joint Genome Institute in Walnut Creek, Calif.; and the Sanger Centre, a huge sequencing center in Cambridge, England.

THE WHITE HOUSE

WASHINGTON

Date 4/6/00

To: PODESTA, LANE

From: The Staff Secretary /SEAN

ANOTHER LETTER FROM SCHNEIDER – WE PLAN TO HAVE A REPLY FOR THE PRESIDENT READY TODAY THAT WILL NOTE OUR EFFORTS TO CLARIFY THE MEANING OF THE US-UK STATEMENT AND OUR POSITION ON PATENTS (I.E. WHAT POTUS SAID YESTERDAY) FYI.

Who else should be quoted on this issue – Francis to lower level – their impact has more than we understand

needs to defer

of course we'll defend

Kathy Hudson forced him to cancel public engagements.

RESTRUCTURING ASSOCIATES INC.

Suite 350
1050 Seventeenth Street, N.W.
Washington, D.C. 20036

(202) 775-8213
FAX: (202) 223-0346

EMAIL: raicentral@restructassoc.com

DATE: March 28, 2000 Time: 6:30pm

TO: President Bill Clinton

FAX NO: 202-456-2883

SUBJECT:

FROM: Thomas J. Schneider

TOTAL NO. OF PAGES (INCLUDING COVER SHEET): 5

MESSAGE:

PLEASE ADVISE BY FAX OR TELEPHONE
IF THIS FAX IS ILLEGIBLE OR INCOMPLETE

Withdrawal/Redaction Marker
Clinton Library

DOCUMENT NO. AND TYPE	SUBJECT/TITLE	DATE	RESTRICTION
001b. letter	Thomas Schneider to President Clinton (5 pages)	03/28/2000	P5, P6/b(6)

COLLECTION:
Clinton Presidential Records
Domestic Policy Council
Devorah Adler
OA/Box Number: 20464

FOLDER TITLE:
Genome

2012-0463-S
rc758

RESTRICTION CODES

Presidential Records Act - [44 U.S.C. 2204(a)]

P1 National Security Classified Information [(a)(1) of the PRA]
P2 Relating to the appointment to Federal office [(a)(2) of the PRA]
P3 Release would violate a Federal statute [(a)(3) of the PRA]
P4 Release would disclose trade secrets or confidential commercial or financial information [(a)(4) of the PRA]
P5 Release would disclose confidential advice between the President and his advisors, or between such advisors [a)(5) of the PRA]
P6 Release would constitute a clearly unwarranted invasion of personal privacy [(a)(6) of the PRA]

C. Closed in accordance with restrictions contained in donor's deed of gift.
PRM. Personal record misfile defined in accordance with 44 U.S.C. 2201(3).
RR. Document will be reviewed upon request.

Freedom of Information Act - [5 U.S.C. 552(b)]

b(1) National security classified information [(b)(1) of the FOIA]
b(2) Release would disclose internal personnel rules and practices of an agency [(b)(2) of the FOIA]
b(3) Release would violate a Federal statute [(b)(3) of the FOIA]
b(4) Release would disclose trade secrets or confidential or financial information [(b)(4) of the FOIA]
b(6) Release would constitute a clearly unwarranted invasion of personal privacy [(b)(6) of the FOIA]
b(7) Release would disclose information compiled for law enforcement purposes [(b)(7) of the FOIA]
b(8) Release would disclose information concerning the regulation of financial institutions [(b)(8) of the FOIA]
b(9) Release would disclose geological or geophysical information concerning wells [(b)(9) of the FOIA]

THE [redacted] OUS[redacted]
W N

Date 3/14/00

To: NEAL LANE

From: The Staff Secretary

NEAL - NANCY THINKS POTUS MAY WANT TO SEE BEFORE TODAYS ANNOUNCEMENT. GIVEN THE ISSUES WITH VENTNER/CELERA, THOUGHT YOU MIGHT WANT TO HANDLE IN THE BRIEFING.

SEAN

CC: PODESTA

RESTRUCTURING ASSOCIATES INC.

Suite 350
1050 Seventeenth Street, N.W.
Washington, D.C. 20036

(202) 775-8213
FAX: (202) 223-0346

EMAIL: raicentral@restructassoc.com

DATE: March 13, 2000 Time: 6:30pm

TO: President Bill Clinton

FAX NO: 202-456-6703

SUBJECT:

FROM: Thomas J. Schneider

TOTAL NO. OF PAGES (INCLUDING COVER SHEET): 13

MESSAGE:

PLEASE ADVISE BY FAX OR TELEPHONE IF THIS FAX IS ILLEGIBLE OR INCOMPLETE

Withdrawal/Redaction Marker
Clinton Library

DOCUMENT NO. AND TYPE	SUBJECT/TITLE	DATE	RESTRICTION
002. letter	Thomas Schneider to President Clinton (3 pages)	03/13/2000	P5, P6/b(6)

COLLECTION:
Clinton Presidential Records
Domestic Policy Council
Devorah Adler
OA/Box Number: 20464

FOLDER TITLE:
Genome

2012-0463-S
rc758

RESTRICTION CODES

Presidential Records Act - [44 U.S.C. 2204(a)]

P1 National Security Classified Information [(a)(1) of the PRA]
P2 Relating to the appointment to Federal office [(a)(2) of the PRA]
P3 Release would violate a Federal statute [(a)(3) of the PRA]
P4 Release would disclose trade secrets or confidential commercial or financial information [(a)(4) of the PRA]
P5 Release would disclose confidential advice between the President and his advisors, or between such advisors [a)(5) of the PRA]
P6 Release would constitute a clearly unwarranted invasion of personal privacy [(a)(6) of the PRA]

C. Closed in accordance with restrictions contained in donor's deed of gift.
PRM. Personal record misfile defined in accordance with 44 U.S.C. 2201(3).
RR. Document will be reviewed upon request.

Freedom of Information Act - [5 U.S.C. 552(b)]

b(1) National security classified information [(b)(1) of the FOIA]
b(2) Release would disclose internal personnel rules and practices of an agency [(b)(2) of the FOIA]
b(3) Release would violate a Federal statute [(b)(3) of the FOIA]
b(4) Release would disclose trade secrets or confidential or financial information [(b)(4) of the FOIA]
b(6) Release would constitute a clearly unwarranted invasion of personal privacy [(b)(6) of the FOIA]
b(7) Release would disclose information compiled for law enforcement purposes [(b)(7) of the FOIA]
b(8) Release would disclose information concerning the regulation of financial institutions [(b)(8) of the FOIA]
b(9) Release would disclose geological or geophysical information concerning wells [(b)(9) of the FOIA]

TON

isel to BI files

l issue a final report
ite House's handling

October, will submit
: oversees his office,
the investigations of

insel, Bob Ray, can
estigation has been
ay's spokeswoman,

Ray: Took over
from Kat Starr.

's role in the firings
loyees in 1993. The
non-political careers.
they were cleared,
pay a settlement"
d that led to the first
investigate the Clin-

element still under
tried to intimidate
Willey into recant-
d her in 1993.
he public has a right
the completion" of
in a public way dur-
—Judy Keen

e House said it will

Feds may have tried to bend law for gene map

Health agency denies any 'conspiracy' to recruit biotech company in race to finish project

By Tim Friend
USA TODAY

Federal officials and a British charity secretly attempted to enlist biotechnology firm Incyte Pharmaceuticals in a race to complete the Human Genome Project and win a bitter two-year battle against another company, Celera Genomics of Rockville, Md.

Critics of the plan said that the National Institutes of Health (NIH), was attempting to bypass laws that require federal agencies to place contracts with private companies up for public bidding.

The $3 billion, international Human Genome Project was begun 10 years ago to create the first genetic blueprint of a human being. The NIH funds it with the Wellcome Trust of Cambridge, England. The project has been losing ground against Celera since May 1998, when Celera executives announced intentions to conduct their own private genome project. A possible collaboration between NIH-Wellcome Trust and Celera was scuttled last week when they couldn't agree on terms.

At stake in the contentious battle is scientific glory and potential control over access to information about virtually all human diseases and possible cures. Scientists predict that the first group to complete the human genetic blueprint will win a Nobel Prize and a pedestal in the halls of history.

1999 photo by M. Dan Goldberg, USA TODAY

Competition: J. Craig Venter is president of Celera Genomics. The company is embroiled in a bitter race with the publicly funded Human Genome Project.

NIH's Francis Collins, director of the National Human Genome Research Institute, denied any "conspiracy" to enlist Incyte of Palo Alto, Calif., in a competition against Celera or to circumvent any laws.

Experts say that enlisting Incyte in the genome war would have given the Human Genome Project access to additional gene sequencing and genome assembly technologies similar to those that Celera uses.

NIH officials asked a pharmaceutical company consortium three weeks ago to collaborate on the Incyte strategy. The consortium's leaders said it was inappropriate for the consortium to side with the government against a private company and refused initially to cooperate.

NIH's strategy called for Incyte to assemble a large volume of genetic data for the Human Genome Project over the next three months. Because NIH could not hire Incyte itself without competitive bidding, Collins and Michael Morgan of Wellcome Trust apparently sought out the SNP Consortium to hire Incyte for them, says Alan Roses of Glaxo Wellcome, a member of the consortium, which includes 10 leading drug companies.

Roses says a byproduct of the deal would be the generation of a "windfall" of single nucleotide polymorphisms (SNPs, pronounced snips), which the consortium is spending millions of dollars to mine from the Human Genome Project's database. SNPs can lead drug companies to genes that cause disease or influence how people respond to drugs. The deal would double the SNPs available to the consortium and at the same time boost the capacity of the Human Genome Project to assemble its data by about 20%.

"Somebody at the National Institutes of Health thought up the idea," Roses says.

"I said I would not agree to go along with any proposal that picked out Incyte over others," Roses said in an interview, "because it potentially puts the SNP Consortium in the position of taking sides on this race to get the genome done."

Collins first declined to discuss the existence of the plan, then downplayed the role Incyte or any other company might play in accelerating progress of the Human Genome Project. He said the primary reason for approaching the SNP Consortium was to help it gather data beneficial to drug discovery.

Incyte CEO Randy Scott confirmed NIH's plan to enlist his company in the race.

"We had entered into a dialogue with them (NIH and the Wellcome Trust)," Scott says. "We've indicated over the last year that if there was any way we could help accelerate completion of the Human Genome Project we would."

After some debate, Collins and the SNP Consortium have agreed to put the contract out for bid and then work together, Scott says. Incyte has made a bid and would do the job at a loss. If awarded the contract, he says, Incyte would probably pay "a few million dollars" of its own money to complete the work, which involves gene sequencing and assembly of the data into a useable form.

Arthur Holden, chief executive officer of the SNP Consortium, says requests for bids were sent to several companies last week and that a company, possibly Incyte, will be selected this week.

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

National Institutes of Health
National Human Genome
Research Institute
31 Center Drive MSC 2152
Building 31, Room 4B09
Bethesda, MD 20892-2152
(301) 496-0844
FAX (301) 402-0837

February 28, 2000

~~CONFIDENTIAL~~

J. Craig Venter, Ph.D.
Mr. Tony White
Arnold Levine, Ph.D.
Mr. Paul Gilman
Celera Genomics Corporation
45 West Gude Drive
Rockville, Maryland 20850

Dear Dr. Venter, Mr. White, Dr. Levine and Mr. Gilman:

We appreciated the chance to meet with you on December 29, 1999, to discuss possible grounds for a collaboration between the public Human Genome Project (HGP) and Celera Genomics, on deriving the sequence of the human genome. Prior to the December 29 meeting, there had been numerous discussions about a possible model for collaboration, most notably those among Eric Lander, Craig Venter, Harold Varmus, and Arnie Levine. The five largest genome centers (Baylor, the JGI, Sanger, Washington University, and the Whitehead Institute, referred to as the G5) had been fully briefed on those discussions, and had asked the four of us to represent the public HGP in any negotiations. Other members of the international sequencing consortium were also briefed in general on these developments. We had prepared a proposed statement of "Shared Principles" (enclosed) which we thought accurately reflected the conclusions of the earlier conversations with Celera. From the previous contacts, we did not expect these to be particularly controversial, and were thus disappointed to learn the day before the meeting that some of those principles might not be shared by Celera. Nonetheless, we attempted to negotiate on December 29 in good faith.

A number of important points of agreement between ourselves and Celera were reached on December 29. These included the conclusion that an active collaboration to merge the two data sets could produce a substantially better product, since it would allow bilateral sharing of electrophoretic traces and a joint effort to resolve discrepancies. At a technical level, this kind of data sharing did not seem to present a substantial hurdle in the implementation of a joint effort.

However, several fundamental differences emerged. These are outlined below:

1) Since its inception, the public HGP has been committed to producing human genomic sequence for all to use without any barriers or restrictions, in the expectation that this policy would maximize the benefit to humankind. That policy was explicitly codified in 1996. If a collaboration were initiated, however, Celera indicated they would expect exclusive commercial rights of distribution of the merged product, most likely even beyond the current projected date of completion of the

CONFIDENTIAL

human genome sequence by the public consortium in 2003 (2005 was mentioned). While we understand Celera's business imperatives, the extent of this claim seemed inappropriate to us. The product of a sequence data merge between the HGP and Celera in the summer of 2000 would reflect the collection of shotgun data at roughly 10x redundancy, and would not be of equivalent quality to the finished product expected from the HGP in 2003. Many sequence gaps and regions of ambiguity would be present in the proposed Celera/HGP merged product. By current plans, if no collaboration were initiated, the public HGP expects to have completed shotgun coverage on available genomic clones by no later than December 2000. This would argue that the period of exclusive rights to commercial distribution of the merged data set by Celera should be limited to no more than 6 - 12 months after the initiation of the collaboration.

2) In view of the public HGP's commitment to finishing the human sequence, it would be expected that ongoing efforts would be needed after the merge to close gaps and clear up ambiguities. The design of such finishing strategies would optimally make full use of the merged data set. On December 29, however, Celera argued that the output of any further improvements in the sequence that resulted from the use of the merged data set would also fall under their exclusive rights for commercial distribution. Such a restriction would put the public HGP in an ongoing position of producing data that was less than fully accessible, which would be inconsistent with the internationally agreed principle of open access. To avoid this, the public HGP might actually be forced to continue to collect additional independent and redundant shotgun data in parallel with the collaborative effort, although this position was apparently also unacceptable to Celera. Given Celera's stated plans to make their own version of the data available to the academic community, even in the absence of a joint effort, it is not clear why or how Celera would expect to prevent the public HGP from utilizing the complete merged data set in directing finishing efforts.

3) The public HGP was also concerned by Celera's stated desire to have their proposed commercial rights over the joint product extend beyond databases to experimental applications, such as the construction of genome chips, large primer sets, or applications to proteomics and analysis of regulatory sequences. We were not previously aware of the intention for those commercial rights to extend beyond databases. While establishing a monopoly on commercial uses of the human genome sequence may be in Celera's business interests, it is not in the best interests of science or the general public.

4) While both parties agreed that wide distribution via both a DVD and the internet was desirable to facilitate access to the sequence, there appeared to be substantial differences in how this might be implemented. The public HGP feels that the sequence, along with jointly derived annotation, must be available from one or more noncommercial sites, whereas Celera wishes any internet access to be exclusively through a Celera portal.

5) We were concerned about Celera's possible intention to publish the results of their merge of the public data and their own data, in the event that no agreement is reached. Although the public effort has explicitly made preliminary sequence available for all to use for both academic and commercial research purposes, the public consortium reasonably reserves the right to be the first to publish its own data in a peer-reviewed journal. Publication of other groups' primary data without consent is considered to be a breach of scientific ethics. In addition, it is counter to accepted scientific practice for authors to present results without having examined the primary data.

CONFIDENTIAL

These are substantive issues. We left the December 29 meeting grateful that the exchange had been so open and wide-ranging, but uncertain whether Celera's stated positions were hardened or still open to negotiation. At the conclusion of the meeting, all parties agreed that a follow-up meeting would probably be needed.

One of us (FSC) attempted to reach Dr. Venter to set up a follow-up meeting, but was directed instead to Mr. White. In a phone call on January 22, Mr. White reiterated the points outlined above, and indicated that a follow-up meeting would only make sense if the public consortium agreed to grant Celera commercial protection for the sequence in their database well beyond the time it would have taken the public HGP to generate an equivalent product. Mr. White stated that he expected commercial protection for at least 3 to 4 years, and preferably longer. He also held fast to the demand that finishing efforts by the public consortium, carried out after the merge, would need to be under the same restrictions. When it was pointed out that this was particularly puzzling, since if there were no collaboration the HGP could presumably use Celera's proposed DVD to aid finishing and then place those additional new reads in public databases, Mr. White indicated that they were rethinking their DVD plans, and that the sequence data might only be available on their web site, perhaps with some restriction on this sort of finishing activity by others.

When asked about the extent to which Celera would see their restrictions applying to uses of the genome sequence other than commercial databases, Mr. White reiterated that Celera would expect chip companies to pay a license for this use. As to other genome-wide uses of the human sequence, he said that Celera has not yet defined which types of experiments would require a license, but left the door open to that being broader than just chips.

When asked whether there was some flexibility in the positions outlined by Celera on Dec. 29, Mr. White did not encourage that view.

Despite this discouraging interaction, one of us (FSC) has been attempting for more than two weeks to reach Dr. Venter to see whether there are any remaining grounds for pursuing the collaborative model. After several unanswered phone calls and e-mails outlining the reason for the attempted contact, Dr. Venter's assistant indicated that he is too busy with other matters right now to discuss this until some time after March 6.

A clear conclusion seems to arise from Celera's actions on and subsequent to December 29: that there is no real interest on the part of Celera in continuing to pursue this particular collaborative model. This is disappointing, since much work has gone into exploring such an option, and we had been hopeful in light of the success of the collaborative effort on the Drosophila sequence that something similar could be worked out for the human project.

Continued ambiguity is not conducive to early completion of the human genome sequence. Therefore, unless Celera indicates in the next week (by March 6, 2000) that the conclusions of this letter are substantively incorrect, we will conclude that the initial proposal whereby the data from the public HGP and Celera are collaboratively merged is no longer workable.

~~CONFIDENTIAL~~

We stand ready and willing to engage in further conversations at any time.

With best regards,

Sincerely yours,

Francis S. Collins, M.D., Ph.D.

Harold Varmus, M.D.

Martin Bobrow, CBE DSC FRCP

Robert H. Waterston, M.D., Ph.D.

CONFIDENTIAL

December 28, 1999

SHARED PRINCIPLES

There seems to be general agreement that:

o The public HGP and Celera Genomics each have the capacity to generate substantial coverage of the sequence of the human genome.

o Humankind will be better served if we can find a viable way to join forces to produce a better product in a more timely fashion.

o The methods being used (clone-based and whole-genome shotgun) are, in fact, complementary. They provide much opportunity for cross-checking.

o A collaboration would offer the opportunity for joint optimization of experimental strategy and analytical methods.

o The current antagonism and excessive competition should be replaced with a more collaborative spirit.

o The public HGP is committed to the complete sequence of the human genome being freely available in the public domain.

o The public HGP understands that Celera Genomics will be making its consensus sequence of the human genome broadly available.

o Celera Genomics is able to see assembled data generated by the public labs and is free to use it for its proprietary databases, but a scientific publication that combines substantial data from both sources should, according to accepted scientific practice, be a joint publication involving authors from both groups.

The outline of a collaborative effort might be as follows:

1. Joint Publication.

The two sides would agree to joint publication of a collaborative paper reporting joint analysis of both the public and Celera data, co-authored by appropriate scientists from both sides. An additional paper reporting analysis of the public data from the public HGP labs might also be simultaneously submitted.

The collaborative paper would necessarily involve that co-authors from the public side have sufficient access to the Celera data and any joint analysis during the preparation of a

~~CONFIDENTIAL~~

paper to be able to comfortably sign their names to a paper.

There is some concern about the appearance of public HGP scientists having access to Celera's data before it is made public. The best solution might be to limit the period of access, perhaps to 90 days before submission of a paper.

The most realistic timeframe would be to aim for the paper(s) to appear near the end of 2000 (perhaps in Science), which would require submission in the late summer or fall and would involve collaboration beginning in the late spring or summer.

It is important that the public consortium's participation in this collaboration does not contribute directly to development of intellectual property for Celera, whether in terms of sequence or SNPs.

2. Data Release.

The public HGP will continue to release their data immediately.

Celera and other companies would continue to be free to use these data to create and sell databases.

Upon publication of the papers, the human sequence produced from the analysis of the public and Celera data will also be released in accessible "media". The ideal medium (because of its universal accessibility) would be an internet-accessible database. Access solely through Celera's database has substantial disadvantages from the public consortium's perspective.

3. Public Relations

Consideration will be given to preparing a brief statement indicating that the possibility of collaboration is being explored. During the negotiations no information will be provided to the press by either party.

If agreement is reached, the two sides will jointly prepare a document and press releases describing the rationale for and nature of the collaboration. Both sides will make all efforts to support this agreement. This includes honest and straightforward responses to scientific inquiries, but avoiding disparaging the other party, sowing discord or undermining the collaborative spirit. It will require a generosity of spirit in acknowledging the legitimate and important role of both parties.

Celera Genomics Corp.
45 West Gude Drive
Rockville, MD 20850
Tel: 240 453-3000
Fax: 240 453-4214
www.celera.com

March 7, 2000

Francis Collins, M.D., Ph.D.
Martin Bobrow
Robert H. Waterston, M.D., Ph.D.
Harold Varmus, M.D.
National Institutes of Health
National Human Genome Research Institute
31 Center Drive MSC 2152
Building 31, Room 4B09
Bethesda, MD 20892-2152

CELERA

Dear Francis, Robert, Martin, and Harold:

I am responding to your letter of February 28, 2000. A preliminary question involves your current motivations. While I would like to believe that the purpose for your letter involves furthering good faith efforts toward collaboration, various aspects of the letter indicate the opposite. It presents a one-week ultimatum, reflecting an apparent disregard for the fact that, as my office informed you, I have been entirely unavailable travelling during the past few weeks (the exception being the one day I was in Washington during that period when you were unable to meet with me). While I do not understand the need for your deadline, Celera has no interest in obstructing or delaying the public effort.

I am also concerned by the implications of the release of your letter to the press prior to my return and response. This obviously conflicts with what I thought we both agreed would be important to reaching agreement in this complex area -- avoiding public posturing and contention. While Celera has no issue with publicly standing behind its position, I trust you anticipated the consequences of releasing to the press your selective perspectives on our discussions. If you disfavor collaboration, a simple "no thank you" will suffice.

Speaking for Celera, we continue to be interested in pursuing good faith discussions toward collaboration. Assuming that there is a reciprocal interest in good faith discussions and given the recent misinformation, let me reiterate Celera's consistent position.

As documented in the June 5, 1998 issue of *Science*, Celera's goal is to both discover and broadly disseminate the human genome sequence. Given the costs and value of this effort, we attempted to clarify a dissemination model that is fair and reasonable. Particularly for pure research applications, we foresee information being released at little or no cost to the end user. For those looking to use our information to their financial benefit, we are unapologetic in seeking a reasonable return for our efforts. At the same time, and as distinguished from the business models of many of our competitors, we anticipate broadly disseminating information without the inherent deterrent of requiring database users to pay onerous royalties on the discoveries they make with our data (often referred to as "reach-through" royalties). We have already entered into several third-

party agreements that bind us to this. We will continue to react unfavorably to claims that Celera's intends to withhold information and delay progress, particularly when our fundamental mission is to accelerate the dissemination of information.

As you know, our meeting with you on December 29, 1999 grew out of a series of discussions with Eric Lander of the Whitehead Institute. These discussions produced some reasonable goals and the potential basis of a sound agreement. Unfortunately, Dr. Lander was not a member of your negotiating team and the scope of the December 29 meeting changed dramatically. Much of the discussion was based on hypothetical scenarios that had little to do with the earlier discussions. As a result, a number of points in your letter dramatically misstated Celera's position. For example, your discussion of our needs for intellectual property protection is quite distorted. We stated a need for data protection in response to your assertion that your researchers should have total access to all of Celera's data at their laboratories, including our electrophoretic files. These data are of significant commercial importance and we naturally responded that we would need sufficient protection for them. You have now asserted that Celera's intellectual property protections made in response to your requirement are inhibitions to the collaboration.

Our collaboration with Dr. Gerald Rubin of Berkeley on the *Drosophila* genome has been a model that has been remarkably productive. The terms of that agreement were reviewed and approved by Dr. Collins, Dr. Varmus, and representatives of the Wellcome Trust. Despite your recent actions we remain hopeful that collaboration on similar terms can be achieved for the human genome.

Let me restate Celera's position on the terms of collaboration:

- As with the *Drosophila* genome, Celera will assemble the human genome with its own data. Public data will be used for comparative purposes. As demonstrated with *Drosophila*, excellent assembly was achieved without the public rough draft data. However, Celera is prepared to assemble a consensus human genome using its data and the data produced by the public human genome project. As with the *Drosophila* genome, scientists from the public sequencing labs would be welcome to come to Celera and verify these data and to analyze the resulting genome.

- As with the *Drosophila* genome, joint publications would be produced describing the finished human genome.

- The finished genome consensus sequence would be available to all researchers. We believe a web-based version at Celera's site is an appropriate balance between an open distribution and a protection of our commercial interests. We have also discussed a DVD version and we are open to that option.

- Researchers would be free to use the published data in their research at no cost.

- The only restriction that Celera has ever requested is that other database providers would be prohibited from providing or selling Celera's data as their own. We

certainly agree that if the public project wanted to publish a version of the genome that did not depend in any way on Celera data we would have no objection to your distribution of those data.

While it is not our preference to negotiate this collaboration through the media, we feel compelled to release this letter to the media in order to correct the misconceptions created by your released letter.

We believe that speed is essential for completing the genome. Its completion is the beginning of a new era in medicine that many lives depend on. The research that follows is far more important than squabbles over credit for its completion. Our consistent goal has been to pursue alternatives that will benefit science and the public while at the same time fulfilling our obligations to our shareholders and fairly rewarding them for their investment.

Sincerely,

J. Craig Venter, Ph.D.
President

cc: E. Lander
M. Hunkapillar
A. Levine
T. White

DRAFT

PROCLAMATION TO ENSURE THAT DISCOVERIES FROM THE HUMAN GENOME ARE USED TO ADVANCE HUMAN HEALTH

In the last decade of the twentieth century, scientists from around the world initiated one of the most significant scientific projects of all time: to determine the DNA sequence of the entire human genome, the human genetic blueprint. Progressing ahead of schedule, human genome research is rapidly advancing our understanding of the causes of human disease and will serve as the foundation for development of a new generation of effective treatments, preventions, and cures. To realize this promise, fundamental data on the human genome, including the human DNA sequence and its variations, must be made freely and broadly accessible to scientists everywhere. The human genome is the fundamental shared heritage of all humankind. Unencumbered access to this information is essential to enable discoveries that will reduce the burden of disease and promote health around the world. We applaud the decision by scientists working on the Human Genome Project to release information about the human DNA sequence and its variants rapidly into the public domain, and we call on all scientists worldwide to adopt this policy.

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

National Institutes of Health
National Human Genome
Research Institute
31 Center Drive MSC 2152
Building 31, Room 4B09
Bethesda, MD 20892-2152
(301) 496-0844
FAX (301) 402-0837

February 28, 2000

~~CONFIDENTIAL~~

J. Craig Venter, Ph.D.
Mr. Tony White
Arnold Levine, Ph.D.
Mr. Paul Gilman
Celera Genomics Corporation
45 West Gude Drive
Rockville, Maryland 20850

Dear Dr. Venter, Mr. White, Dr. Levine and Mr. Gilman:

We appreciated the chance to meet with you on December 29, 1999, to discuss possible grounds for a collaboration between the public Human Genome Project (HGP) and Celera Genomics, on deriving the sequence of the human genome. Prior to the December 29 meeting, there had been numerous discussions about a possible model for collaboration, most notably those among Eric Lander, Craig Venter, Harold Varmus, and Arnie Levine. The five largest genome centers (Baylor, the JGI, Sanger, Washington University, and the Whitehead Institute, referred to as the G5) had been fully briefed on those discussions, and had asked the four of us to represent the public HGP in any negotiations. Other members of the international sequencing consortium were also briefed in general on these developments. We had prepared a proposed statement of "Shared Principles" (enclosed) which we thought accurately reflected the conclusions of the earlier conversations with Celera. From the previous contacts, we did not expect these to be particularly controversial, and were thus disappointed to learn the day before the meeting that some of those principles might not be shared by Celera. Nonetheless, we attempted to negotiate on December 29 in good faith.

A number of important points of agreement between ourselves and Celera were reached on December 29. These included the conclusion that an active collaboration to merge the two data sets could produce a substantially better product, since it would allow bilateral sharing of electrophoretic traces and a joint effort to resolve discrepancies. At a technical level, this kind of data sharing did not seem to present a substantial hurdle in the implementation of a joint effort.

However, several fundamental differences emerged. These are outlined below:

1) Since its inception, the public HGP has been committed to producing human genomic sequence for all to use without any barriers or restrictions, in the expectation that this policy would maximize the benefit to humankind. That policy was explicitly codified in 1996. If a collaboration were initiated, however, Celera indicated they would expect exclusive commercial rights of distribution of the merged product, most likely even beyond the current projected date of completion of the

~~CONFIDENTIAL~~

human genome sequence by the public consortium in 2003 (2005 was mentioned). While we understand Celera's business imperatives, the extent of this claim seemed inappropriate to us. The product of a sequence data merge between the HGP and Celera in the summer of 2000 would reflect the collection of shotgun data at roughly 10x redundancy, and would not be of equivalent quality to the finished product expected from the HGP in 2003. Many sequence gaps and regions of ambiguity would be present in the proposed Celera/HGP merged product. By current plans, if no collaboration were initiated, the public HGP expects to have completed shotgun coverage on available genomic clones by no later than December 2000. This would argue that the period of exclusive rights to commercial distribution of the merged data set by Celera should be limited to no more than 6 - 12 months after the initiation of the collaboration.

2) In view of the public HGP's commitment to finishing the human sequence, it would be expected that ongoing efforts would be needed after the merge to close gaps and clear up ambiguities. The design of such finishing strategies would optimally make full use of the merged data set. On December 29, however, Celera argued that the output of any further improvements in the sequence that resulted from the use of the merged data set would also fall under their exclusive rights for commercial distribution. Such a restriction would put the public HGP in an ongoing position of producing data that was less than fully accessible, which would be inconsistent with the internationally agreed principle of open access. To avoid this, the public HGP might actually be forced to continue to collect additional independent and redundant shotgun data in parallel with the collaborative effort, although this position was apparently also unacceptable to Celera. Given Celera's stated plans to make their own version of the data available to the academic community, even in the absence of a joint effort, it is not clear why or how Celera would expect to prevent the public HGP from utilizing the complete merged data set in directing finishing efforts.

3) The public HGP was also concerned by Celera's stated desire to have their proposed commercial rights over the joint product extend beyond databases to experimental applications, such as the construction of genome chips, large primer sets, or applications to proteomics and analysis of regulatory sequences. We were not previously aware of the intention for those commercial rights to extend beyond databases. While establishing a monopoly on commercial uses of the human genome sequence may be in Celera's business interests, it is not in the best interests of science or the general public.

4) While both parties agreed that wide distribution via both a DVD and the internet was desirable to facilitate access to the sequence, there appeared to be substantial differences in how this might be implemented. The public HGP feels that the sequence, along with jointly derived annotation, must be available from one or more noncommercial sites, whereas Celera wishes any internet access to be exclusively through a Celera portal.

5) We were concerned about Celera's possible intention to publish the results of their merge of the public data and their own data, in the event that no agreement is reached. Although the public effort has explicitly made preliminary sequence available for all to use for both academic and commercial research purposes, the public consortium reasonably reserves the right to be the first to publish its own data in a peer-reviewed journal. Publication of other groups' primary data without consent is considered to be a breach of scientific ethics. In addition, it is counter to accepted scientific practice for authors to present results without having examined the primary data.

~~CONFIDENTIAL~~

These are substantive issues. We left the December 29 meeting grateful that the exchange had been so open and wide-ranging, but uncertain whether Celera's stated positions were hardened or still open to negotiation. At the conclusion of the meeting, all parties agreed that a follow-up meeting would probably be needed.

One of us (FSC) attempted to reach Dr. Venter to set up a follow-up meeting, but was directed instead to Mr. White. In a phone call on January 22, Mr. White reiterated the points outlined above, and indicated that a follow-up meeting would only make sense if the public consortium agreed to grant Celera commercial protection for the sequence in their database well beyond the time it would have taken the public HGP to generate an equivalent product. Mr. White stated that he expected commercial protection for at least 3 to 4 years, and preferably longer. He also held fast to the demand that finishing efforts by the public consortium, carried out after the merge, would need to be under the same restrictions. When it was pointed out that this was particularly puzzling, since if there were no collaboration the HGP could presumably use Celera's proposed DVD to aid finishing and then place those additional new reads in public databases, Mr. White indicated that they were rethinking their DVD plans, and that the sequence data might only be available on their web site, perhaps with some restriction on this sort of finishing activity by others.

When asked about the extent to which Celera would see their restrictions applying to uses of the genome sequence other than commercial databases, Mr. White reiterated that Celera would expect chip companies to pay a license for this use. As to other genome-wide uses of the human sequence, he said that Celera has not yet defined which types of experiments would require a license, but left the door open to that being broader than just chips.

When asked whether there was some flexibility in the positions outlined by Celera on Dec. 29, Mr. White did not encourage that view.

Despite this discouraging interaction, one of us (FSC) has been attempting for more than two weeks to reach Dr. Venter to see whether there are any remaining grounds for pursuing the collaborative model. After several unanswered phone calls and e-mails outlining the reason for the attempted contact, Dr. Venter's assistant indicated that he is too busy with other matters right now to discuss this until some time after March 6.

A clear conclusion seems to arise from Celera's actions on and subsequent to December 29: that there is no real interest on the part of Celera in continuing to pursue this particular collaborative model. This is disappointing, since much work has gone into exploring such an option, and we had been hopeful in light of the success of the collaborative effort on the Drosophila sequence that something similar could be worked out for the human project.

Continued ambiguity is not conducive to early completion of the human genome sequence. Therefore, unless Celera indicates in the next week (by March 6, 2000) that the conclusions of this letter are substantively incorrect, we will conclude that the initial proposal whereby the data from the public HGP and Celera are collaboratively merged is no longer workable.

~~CONFIDENTIAL~~

We stand ready and willing to engage in further conversations at any time.

With best regards,

Sincerely yours,

Francis S. Collins, M.D., Ph.D.

Harold Varmus, M.D.

Martin Bobrow, CBE DSC FRCP

Robert H. Waterston, M.D., Ph.D.

December 28, 1999

SHARED PRINCIPLES

There seems to be general agreement that:

o The public HGP and Celera Genomics each have the capacity to generate substantial coverage of the sequence of the human genome.

o Humankind will be better served if we can find a viable way to join forces to produce a better product in a more timely fashion.

o The methods being used (clone-based and whole-genome shotgun) are, in fact, complementary. They provide much opportunity for cross-checking.

o A collaboration would offer the opportunity for joint optimization of experimental strategy and analytical methods.

o The current antagonism and excessive competition should be replaced with a more collaborative spirit.

o The public HGP is committed to the complete sequence of the human genome being freely available in the public domain.

o The public HGP understands that Celera Genomics will be making its consensus sequence of the human genome broadly available.

o Celera Genomics is able to see assembled data generated by the public labs and is free to use it for its proprietary databases, but a scientific publication that combines substantial data from both sources should, according to accepted scientific practice, be a joint publication involving authors from both groups.

The outline of a collaborative effort might be as follows:

1. Joint Publication.

The two sides would agree to joint publication of a collaborative paper reporting joint analysis of both the public and Celera data, co-authored by appropriate scientists from both sides. An additional paper reporting analysis of the public data from the public HGP labs might also be simultaneously submitted.

The collaborative paper would necessarily involve that co-authors from the public side have sufficient access to the Celera data and any joint analysis during the preparation of a

CONFIDENTIAL

paper to be able to comfortably sign their names to a paper.

There is some concern about the appearance of public HGP scientists having access to Celera's data before it is made public. The best solution might be to limit the period of access, perhaps to 90 days before submission of a paper.

The most realistic timeframe would be to aim for the paper(s) to appear near the end of 2000 (perhaps in Science), which would require submission in the late summer or fall and would involve collaboration beginning in the late spring or summer.

It is important that the public consortium's participation in this collaboration does not contribute directly to development of intellectual property for Celera, whether in terms of sequence or SNPs.

2. Data Release.

The public HGP will continue to release their data immediately.

Celera and other companies would continue to be free to use these data to create and sell databases.

Upon publication of the papers, the human sequence produced from the analysis of the public and Celera data will also be released in accessible "media". The ideal medium (because of its universal accessibility) would be an internet-accessible database. Access solely through Celera's database has substantial disadvantages from the public consortium's perspective.

3. Public Relations

Consideration will be given to preparing a brief statement indicating that the possibility of collaboration is being explored. During the negotiations no information will be provided to the press by either party.

If agreement is reached, the two sides will jointly prepare a document and press releases describing the rationale for and nature of the collaboration. Both sides will make all efforts to support this agreement. This includes honest and straightforward responses to scientific inquiries, but avoiding disparaging the other party, sowing discord or undermining the collaborative spirit. It will require a generosity of spirit in acknowledging the legitimate and important role of both parties.

fits patent proxy

Swept away in a gold rush to cash in on our genes

By Ellen Goodman, Globe Staff, Globe Columnist, 3/5/2000

I have to tip my cap to Donna MacLean. The British poet and waitress walked into a patent office and filed an application for a very unusual invention: "Myself."

"It has taken 30 years of hard labor for me to discover and invent myself," she wrote, "and now I wish to protect my invention from unauthorized exploitation, genetic or otherwise."

MacLean was motivated by a sense of humor and a sense of outrage. It seems she'd been reading about companies trying to patent genes and decided she'd better hang on to her own private property.

Hers is more of a statement than a solution; it's better performance art than patent law. But she's onto something.

We are close to deciphering all the human genes. Scientists expect to complete sequencing the 3 billion units that make up our DNA by the end of the year.

Most of these 3 billion bits are, I hate to say, pretty worthless, no matter how attractively they may be arranged in Ms. MacLean. We share 30 percent of our genes with lettuce. But some genes may hold secrets that cure cancer, Alzheimer's, and other scourges of our genetic inheritance. So the scientists working on the Human Genome Project are putting sequences into the public venue - the Internet - for all researchers as fast as they can.

Meanwhile, a few companies like Celera, run by the flamboyant J. Craig Venter, are racing to stake a private claim on whole stretches of DNA. The president of Incyte has bragged that whoever wins the patenting race will become the "eBay for genes." They can offer them for commercial use to the highest bidder.

The gold rush for patents on every gene, whether it has a known use or not, has gotten the scientific community and at least one poet-waitress

COLUMNS
Ellen Goodman
Derrick Z. Jackson
Jeff Jacoby
Robert Jordan
Martin F. Nolan
David Nyhan
Thomas Oliphant
Jack Thomas
Joan Vennochi

YOUR VIEWS
Write a letter to the editor

Sections
PAGE ONE
NATION | WORLD
METRO | REGION
BUSINESS
SPORTS
LIVING | ARTS
EDITORIALS | OP-ED

Weekly
Health | Science (Tues.)
Food (Wed.)
Calendar (Thu.)
At Home (Thu.)
Picture This (Fri.)

Sunday
Automotive
Focus
Learning
Magazine
New England
Real Estate
Travel
City Weekly
South Weekly
West Weekly
North Weekly
NorthWest Weekly
NH Weekly

Features
Archives
Book Reviews

Columns
Comics
Crossword
Horoscopes
Death Notices
Lottery
Movie Reviews
Music Reviews
Obituaries
Special Reports
Today's stories A-Z
TV & Radio
Weather

Classifieds
Autos
Classifieds
Help Wanted
Real Estate


Merlin-Net
GET THE PICTURE
Buy a Globe photo

Help
E-mail addresses
Send us feedback

Alternative views
Low-graphics version
Acrobat version *(.pdf)*

Search the Globe:
◉ Today
○ Yesterday
SEARCH

Search the Web
Using AltaVista:
SEARCH

worried.

It's not just the philosophical questions - Who owns my body? Who owns the entire genetic repository of the human species? - that spook people. (Let's remember that eBay was just asked to auction off a human soul.) There is also a question about the public good. What's the best, fairest, fastest way to go from mapping the human genome to curing human disease?

There is something to be said for patenting. In theory, it gives private companies the incentives to get the medicine to the marketplace. Scientists may be satisfied with Nobel Prizes, but stockholders want returns.

On the other hand, patenting can inhibit progress by making it prohibitively expensive. The marketplace does a dicey job of regulating medicine. Just compare the falling cost of a Palm Pilot and the rising cost of medicine.

When can patenting help and when can it hurt? At the moment, Ellen Clayton, director of Vanderbilt's Genetics and Health Policy Center, expresses a view shared by many that "patents are being awarded for too little intellectual work."

Consider the current feud over rights to a gene that may be important in AIDS therapy. A group of AIDS researchers discovered a promising use for CCR5. But a biotech company that did little more than list the letters and number on the code already had a patent on its commercial development. It was, rued researcher Robert Gallo, as if the biotech company said, "I found a fungus, therefore I should get credit for penicillin."

The questions being raised now are: Should the people - or even the computer - doing the relatively routine work of identifying a gene get to be the landlord charging whatever the market will bear? Or should that property go to the person who figures out what to do with it?

Francis Collins, head of the Human Genome Project, uses a tollbooth analogy. "Medical research is a road to discovery," he says. You start out on a bumpy path with a lot of ruts, wrong turns, and dead ends. You get along far enough until you see a clear way to turn knowledge into a product. "At that point it's OK to have a toll. You need to do the upkeep on the road. The patent maintains the road."

But if the tolls start piling up too early, too often, and too expensively, nobody will take that road. "The tendency in biotechnology has been to put the tollbooths earlier and earlier," says Collins. Business is hogging the freeway. Patent law and the Patent Office were set up long before anyone heard of DNA. Now the whole blueprint is in play.

In France you can't patent genes at all. In Britain, our gal Donna is trying to get a poetic grip on "myself." Before someone corners the American market on chromosomes, we'd better protect public access to a nice, clean, affordable set of genes.

Ellen Goodman is a Globe columnist.

This story ran on page C7 of the Boston Globe on 3/5/2000.

[Send this story to a friend | Easy-print version | Add to Daily User]

THE WHITE HOUSE

WASHINGTON

February 11, 2000

MEMORANDUM FOR THE PRESIDENT

FROM: NEAL LANE Neal

SUBJECT: GENE PATENTS

L.A. Times article

Reaction to today's L.A. Times article has been extremely positive from the scientific community. We had very strong supportive statements from NIH and AAAS. Industry support depends on the business strategy of the company. Some companies whose financial plan is based on obtaining broad patents and controlling access to the information will not be happy. Other firms might rejoice in a policy that obviates their having to negotiate licenses in order to develop gene-based products. In a discussion earlier this week, Craig Venter, CEO of Celera Genomics, told me that he didn't think that patents ought to be granted on raw DNA sequences without accompanying information on its usefulness (See attached background paper).

Process for Deciding Which Gene Patents will be Approved

On December 21, the Patent and Trademark Office published draft guidelines to help clarify requirements for obtaining gene patents. This issue has been debated for the past few years, with some friction apparent between PTO and the National Institutes of Health. In the new guidelines, PTO indicated that the agency's policy was moving in a direction that makes NIH and the scientific community happy, but the situation is still not resolved. There is general agreement that raw DNA sequence without accompanying evidence of function of the gene or its protein product is not patentable. On the other end of the spectrum, it is also clear that an applicant will be granted a patent when he/she knows the sequence of a specific gene, what protein the gene codes for, and what the protein does in the body. This information would satisfy the utility criterion in patent law. I thought you made both of these points very clearly in your interview. The problem is the grey area in between.

PTO's draft guidelines state that an applicant must disclose a "specific and substantial utility that is credible." My staff (with OVP, DPC and OMB) met with Todd Dickinson, Patent Commissioner, on Jan. 7, to discuss the utility standard. At that meeting, PTO promised to issue examples of how this statement would be interpreted, but have not yet done so. These examples are critical to understanding the somewhat vague guidelines. I will communicate to Todd Dickinson the need to issue the examples in short order and extend the comment period, if necessary.

The PTO is a service organization serving the needs of the scientific and technological communities, in industry and academia. I think your interview was particularly useful in letting all parties know your position on the value of making genetic information public.

Public/Private Cooperation in the Human Genome Project

On a final note, I believe negotiations have stalled between the private and public sectors engaging in human genome sequencing. Both Craig Venter and Francis Collins (NIH) profess continued interest, however. I think some form of collaboration would demonstrate the complementarity of the two activities and accelerate scientific progress. This is a noble goal, and I will pursue it but have no illusions as to whether or not it can be achieved.

Attachment

cc: John Podesta
Chris Jennings
David Beier
Eddie Correia
Tom Kalil

GENE PATENTING

This paper describes the current state-of-play with respect to gene patenting. The handful of companies that have pinned their hopes on patenting and licensing human genetic information are enjoying a meteoric rise in stock prices. Their fortunes depend almost entirely on where the U.S. Patent and Trademark Office draws the line between patentable and unpatentable subject matter. Other players including pharmaceutical and biotechnology companies, and academic biomedical researchers, are also concerned about the potentially broad scope of gene patents. The uncertainty of what subject matter PTO may declare patentable forces "defensive patenting" to ensure a patent position in the absence of a definitive policy.

The Constitution gives Congress the broad mandate "...to promote the Progress of Science and useful Arts." By statute, the U.S. Patent and Trademark Office grants patents on inventions that have utility, that have been adequately described, are novel and non-obvious. Patents provide inventors with rights to exclude others from making, using or selling their invention and are, therefore, a significant incentive for inventors to disclose the invention. Public disclosure of an invention affords others the opportunity to develop improvements under appropriate licensing agreements with the patent holder; and to avoid duplicative efforts.

Gene sequencing used to be a very laborious process that was undertaken *after* isolating a gene thought to be associated with a disease or normal biological function. With the advent of automated sequencing in the late 1980's, it was possible to ascertain the order of letters making up the human genome *before* one knew the function of a particular gene, or even where one gene ended or another began. This made it conceivable that with gene recognition software, patent applications might well be filed without any knowledge of what a gene or its protein product actually did in the body. In fact, the first notable set of gene patent applications contained not full gene sequences but only pieces of genes.[1] Since then, according to PTO, the level of sophistication of patent applications has risen, and applications are more likely to contain the full gene sequence and some degree of information regarding the function of its protein product. It is now possible, once the DNA has been loaded into a sequencing machine, to ascertain its sequence, compare it to other known sequences, and generate and file a patent application, without any human intervention. Some critics view this possibility as a failure in the patent system.

Patent law requires an applicant to explain how the invention might be used before a patent can be granted. This criterion might be met, for example, by claiming that a disclosed gene encodes a protein that is a key component in the immune system and might be useful in the treatment of leukemia. At issue is the degree of certainty the PTO will demand to satisfy the utility criterion. It appears that PTO will allow an applicant to obtain a patent based on a computer-generated function or utility, a position that attracts substantial concern. It is hoped that PTO's forthcoming examples of patentable matter will help clarify the boundaries of patentability left unresolved by the Dec. 21, 1999 guidelines.

---

[1] In 1991, NIH filed three patent applications claiming over 6,000 partial gene sequences obtained in the laboratory of Dr. J. Craig Venter.

THE WHITE HOUSE

WASHINGTON

February 9, 2000

MEMORANDUM FOR JAKE SIEWERT

FROM: NEAL LANE

SUBJECT: 2/10 L.A.TIMES INTERVIEW

Peter Gosselin consulted with my staff in preparing his Feb. 7 article on gene patents. This morning he shared his possible focus for tomorrow's interview. He is likely to ask these two questions:

Question: Should the public have access to the human genetic code?

Answer : Yes – absolutely. Genetic discoveries are going to have profound impacts on American life. I want to make sure that the public reaps the maximum benefit from these new advances and is protected from any possible abuses. I am working to see that science and technology can proceed within the bounds of our most deeply held human values. One of those values is to seek ways to reduce human suffering.

The best way to ensure that scientific progress is made as quickly as possible is to get the results out there as quickly as possible, free from any encumbrances. That is the rationale behind the policy adopted by the supporters of the Human Genome Project—the National Institutes of Health and the Department of Energy—as well as our British partner, the Wellcome Trust. They require that their grantees make public raw human DNA sequence data-- that is, the four letters of the genetic code in their proper order—within one day of deciphering the code. I fully endorse the policy of providing rapid, unencumbered public access to human genome sequence data, and applaud scientists who adopt this stance, in this country or abroad, in academia or industry.

Question: If so, do you have concerns that the rush to patent genes will get in the way of public access?

Answer: I don't believe that good patents will get in the way of rapid dissemination of the genetic code. Patents have an important role in encouraging companies to invest in product development. Intellectual property protection—the grant of some rights over the markets created by one's own inventions—is a vital incentive to getting the genetic code translated into the next generation of gene-based health care products. That's the central tenet of our patent system, as designed by Thomas Jefferson and defined in our Constitution, to promote the "Progress of Science and the Useful Arts." So in order to stimulate innovation, the government has to worry about what happens all along the pipeline, from the most basic research, which we conduct in our labs or sponsor in universities, to the more applied end of the continuum, which requires a smooth handoff to the private sector. One of the things the government can do at this end of the

innovation cycle is to provide adequate incentives to industry in the form of intellectual property protection. Without intellectual property protection and private sector risk-taking, we will not get maximum economic or medical value out of all of our scientific efforts.

The issue of gene patents came up when the First Lady and I hosted a Millenium evening on October 12. Eric Lander, one of the leaders in genetic research and director of one of the five principal human genome sequencing centers, expressed his concerns about overly broad gene patents. I gather that there is some uncertainty in the scientific community, both in industry and academia, as to where the Patent and Trademark Office draws the line on patentability of the results of genome research.

Right before Christmas, the PTO issued draft guidelines that are intended to clear up this situation. These guidelines are still out for public comment but I think there is general agreement that the Patent Office is going in the right direction. This Office has the job of upholding patent law and determining whether or not the claims in a patent application meet the criteria in the law. But within these bounds, Commissioner Dickinson also exercises his judgement and I think he is sensitive to the demands of a rapidly moving technology like genomics. The Commissioner has indicated that PTO will not issue gene patents without a specific, substantial, and credible utility. We need good, strong patents in order to provide adequate incentives for industry to pursue. It is through good patents that the public will get a return on its investment in fundamental genome research. This return will come about by way of the revolution in medical treatment based on discoveries from the human genome and biotechnology inventions of all sorts, many of which we can't even imagine right now.

Attachment:

Feb. 9 note regarding draft U.S.-U.K. Joint Statement to Ensure that Discoveries from the Human Genome are Used to Advance Human Health

cc: John Podesta
Joe Lockhart
Karen Tramontano
Chris Jennings
David Beier

February 9, 2000

Note from Neal Lane re: Draft U.S.-U.K. Joint Statement to Ensure that Discoveries from the Human Genome are Used to Advance Human Health

I have been working for several months with my British counterpart, Lord David Sainsbury, to construct a joint statement to be issued by our respective heads-of-sate. The draft statement (attached) would endorse the policy of rapid, unencumbered access to human genome sequence data that is currently in operation within the publicly-funded, international Human Genome Project. The statement has been cleared by relevant EOP offices and agencies including the Office of the Vice President, Domestic Policy Counsel, White House Counsel, National Security Counsel, National Economic Counsel and the Department of Health and Human Services. There is one point left to resolve with the Patent and Trademark Office.

When our draft is cleared, I intend to send it to Lord Sainsbury, who has communicated informally that Prime Minister Blair is receptive to issuing such a statement (see attached note to Lord Sainsbury from Jonathan Powell, the Prime Minister's Chief of Staff). However, it would be premature to mention this statement in public before the U.K has had the opportunity to vet it.

Francis Collins and others in the scientific community are eager to have this statement made public as soon as possible.

Attachments

Neal

10 DOWNING STREET
LONDON SW1A 2AA

From the Prime Minister's Chief of Staff

4 February 2000

For information
Helen Fl... 8/2

Dear David,

US/UK DECLARATION ON THE HUMAN GENOME PROJECT

Thank you for your letter of 18 January proposing that the UK should indicate that it is willing to participate in a joint US/UK declaration on welcoming the publication on the Internet of the Human Genome Project.

The Prime Minister would be willing to make such a declaration if Bill Clinton agreed. We will need to find a suitable peg. I should be grateful if you could discuss this with the White House on this basis.

Yours ever,

JONATHAN POWELL

The Lord Sainsbury

JOINT STATEMENT TO ENSURE THAT DISCOVERIES FROM THE HUMAN GENOME ARE USED TO ADVANCE HUMAN HEALTH

In the last decade of the twentieth century, scientists from around the world initiated one of the most significant scientific projects of all time: to determine the DNA sequence of the entire human genome, the human genetic blueprint. Progressing ahead of schedule, human genome research is rapidly advancing our understanding of the causes of human disease and will serve as the foundation for development of a new generation of effective treatments, preventions, and cures. Intellectual property protection, for fundamental genome sequence information that has been isolated from its natural state and shown to be useful, will play an important role in stimulating the development of important new gene-based health care products.

To realize the full promise of these discoveries, fundamental data on the human genome, including the human DNA sequence and its variations, should be made freely and broadly accessible to scientists everywhere. Unencumbered access to this information will promote discoveries that will reduce the burden of disease and improve health around the world. Discoveries related to the human genome will improve the quality of life for all humankind. We applaud the decision by scientists working on the Human Genome Project to release information about the human DNA sequence and its variants rapidly into the public domain, and we commend all scientists worldwide who adopt this policy.

224
1006