Lennon at the Lawrence Livermore National Laboratory in Livermore, California, distributes the related clones, which can be used to search for detailed biological information.

Like many other academics, Rubin views the Merck Gene Index as a godsend. Had the pharmaceutical firms kept all the human data to themselves, "it would have been a disaster," Rubin asserts. "I am extremely grateful to companies like Merck that have made available their precompetitive information. ... It furthers my research and that of many people."

In November 1995, the Wellcome Trust announced another gift to public databases: It pledged to give the Sanger Centre $75 million over 7 years to begin sequencing the complete human genome. In February 1996, the Howard Hughes Medical Institute in Bethesda, Maryland, awarded a $2.3 million grant to Waterston's group to create a complete gene index for the mouse, a valuable tool for gene-comparison studies. And the National Center for Human Genome Research (NCHGR), part of the NIH, followed Wellcome's move in April 1996, with $22 million in support for five U.S. pilot projects that have now begun sequencing the human genome at an accelerated pace. In another effort still awaiting final approval, Wellcome is expected to announce that it will offer $25 million in grants for the sequencing of microbes. In all cases, sponsors have insisted that the data be made public rapidly.

To reinforce this ethic, several research sponsors have adopted a series of increasingly pointed guidelines for grantees. In 1992, NCHGR and DOE jointly announced a policy novel to biomedical research: It asked grant applicants who were likely to generate "significant amounts of genome data or materials" to specify exactly "how and when" they would make the results available to the public. The policy also says grantees should not retain work for more than 6 months "from the time the data or materials are generated," whether or not they were part of a published study.

The Wellcome Trust and the Sanger Centre, joined by NCHGR's director Francis Collins, built on these principles in February 1996. At a meeting in Bermuda of newly funded sequencing teams, Sanger Centre director John Sulston proposed that everyone agree to release raw data on a daily basis, or "as soon as possible," without seeking patents on the raw data. There was no audible dissent, according to geneticist David Bentley of the Sanger Centre, who was present. Bentley has defended the policy, in a Policy Forum in *Science* (25 October 1996, p. 533), as a way to limit duplication, stifle "inappropriate" attempts to garner early patents, and avoid giving any group preferential access to data.

Collins endorsed the policy again when he announced the NCHGR's grant awards in April. And he added a new touch, asking grantees not to seek patents on "raw genomic sequence" data, because this "could have a chilling effect" on future investments in gene technology.

Still, it is not yet clear just how these principles will translate into action. First, the new rules have not met with universal praise. Venter and his TIGR colleague Mark Adams, for example, recently attacked their underlying assumptions in print, arguing that the rules would encourage sloppiness and discourage researchers from trying to publish journal articles (*Science*, 25 October 1996, p. 534). They argue instead for release "as soon as ... data have passed a series of rigorous quality control checks and have been annotated." Also, the anti-patenting rule clashes with a 1980 federal law, called the Bayh-Dole Act, that encourages federal grantees to patent their discoveries.

And the issue of when and how to share sequence data is especially complicated when it comes to labs that take both private and public funds. TIGR's allegiance to HGS already has caused many headaches over data release (see sidebar). GTC also exists uneasily in two worlds: In addition to its private income, it received at least $37 million in grants from the U.S. government between 1990 and 1995. Yet it has released only random genomic data from parts of the microbial genomes it set out to sequence. Vovis explains that the federal grant was mainly a "technology demonstration" project, one that was never meant to yield complete genomes. But as the company notes in its annual report, the grants helped defray the company's overhead research costs.

"I am extremely grateful to companies like Merck. ... It furthers my research and that of many people."

Gerald Rubin

The debate over who should control DNA data, which has been going strong for at least 5 years, could easily continue for as many more. It is hard to predict whether the campaign for daily release of genomic data will prevail, or the patent seekers will come out ahead in the end. But one thing is clear: The amount of genomic sequence available in public databases is growing at a breathtaking pace. Venter, for one, fondly wishes that, as a result, the "whole argument" about who owns the genes "will just go away." But nobody is betting yet that it will go quietly.

–Eliot Marshall

INTELLECTUAL PROPERTY

Companies Rush to Patent DNA

Getting rich on human genes has become a fantasy for many investors in the 1990s. Big, savvy pharmaceutical companies and brash biotech start-ups are spending huge sums of money in the hope of gaining exclusive property rights to uncharted areas of the human genome. But who ends up getting rich may have more to do with their skill at navigating patent law—and with the unpredictable decisions of federal judges—than the importance of the biology they have discovered.

Although agencies around the world have been awarding patents based on DNA for more than 15 years, it's still not entirely clear which discoveries are patentable and which are not. One major unresolved issue is just how much biological data on the function of a DNA sequence is needed to win a patent. Applications based on whole genes whose function is well known stand the best chance of being awarded patents. But some less-than-complete gene sequences also have been patented in the past, when their commercial use was well defined.

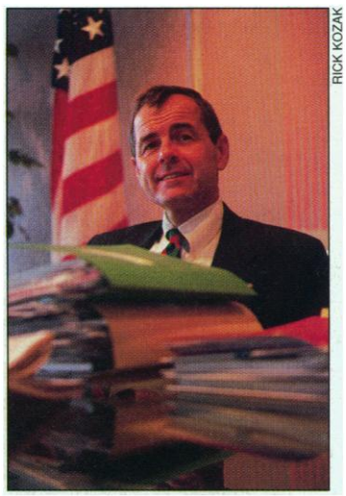
This question has been brought to the fore by a mass of recent patent applications that try to lay claim to thousands of genes by patenting DNA fragments that can be used to reconstruct whole gene sequences. Even if these fragments, called "expressed sequence tags," or ESTs, are ultimately deemed unpatentable—and many experts now believe they will be—the filings could still cloud the legal picture for many years. The reason: These applications will create a priority date for the discovery of many genes, making it hard for later gene hunters to argue that they have made a truly novel discovery. This uncertainty about who can claim priority has been deepened recently by moves to place vast amounts of sequence data in public databases (see p. 777).

Opening the floodgates. While the policy on gene fragments may be in a muddle, the notion that a whole gene can be privately owned was firmly established in 1980; when the U.S. Supreme Court ruled that Ananda Chakrabarty, a molecular biologist then working for General Electric, could patent a genetically engineered organism. Chakrabarty had spliced a gene for an oil-dissolving enzyme into a microbe, creating a bug that could clean up oil spills. The U.S. Patent and Trademark Office (PTO) initially rejected the application on the grounds that life

couldn't be patented. But the court ruled that what Chakrabarty had described—although living—was an artificial substance, and that Chakrabarty had a right to patent it.

Later decisions made it clear that even "normal" DNA sequences are considered artificial products and therefore patentable. John Doll, the PTO's biotech section chief, explains: While "nobody 'owns' the gene in your body, inventors can own the right to exploit it commercially. ... You can't turn over a rock and find a gene."

Since 1980, Doll says the PTO has received more than 5000 patent applications based on whole genes. And it has granted more than 1500 patents on them. This estimate generally tracks the results of a study published in *Nature* last year by a science policy group at the University of Sussex in Britain, led by S. M. Thomas. Between 1981 and 1995, the Thomas group found, the patent offices of the United States, Europe, and Japan issued 1175 patents on human DNA sequences.



RICK KOZAK

Swamped. PTO Commissioner Lehman, overwhelmed by DNA fragments.

The genes covered by these claims range from DNA coding for human interleukin and interferon—immune-system regulating proteins—to genes for bone and brain tissue. Most inventions are aimed at treating medical problems, and in the United States and Europe, more than half of the patents are held by public sector institutions. The single most valuable human DNA patent, however, may be one covering the human erythropoietin gene, which is used to produce a hormone needed by kidney disease patients. In 1991, the U.S. Supreme Court affirmed the validity of this 1987 patent, which now earns its owner, Amgen Inc. of Thousand Oaks, California, more than $1 billion a year.

A fragmented picture. Just as the legal picture seemed to be clearing, with patent offices and the courts upholding claims based on whole genes, it was thrown into turmoil again in 1991 by the U.S. National Institutes of Health. NIH filed a set of applications for patents on thousands of EST gene fragments. Private companies have since staked their own claims on DNA fragments covering most of the genes in the human body.

Between 1992 and 1996, for example, Human Genome Sciences Inc. (HGS) of Rockville, Maryland, together with its nonprofit partner The Institute for Genomic Research (TIGR), also in Rockville, launched a factorylike effort to sequence gene fragments. HGS, which owns the commercial rights in this enterprise, applied for scores of patents on ESTs. Incyte Pharmaceuticals Inc., of Palo Alto, California, set up a similar project. Today, at least 350 patent applications, covering more than 500,000 gene tags, are pending at the PTO. The largest single application contains 18,500 sequences.

The ultimate fate of these monster claims is far from certain. In 1993, the PTO rejected NIH's application in a preliminary ruling, largely because NIH had not explained how the gene fragments, whose biological function was unknown, would be used commercially. Harold Varmus, who became director of NIH in 1993—and who had come under pressure to abandon the claim—decided not to appeal. But companies with big investments in gene fragments—particularly HGS and its main corporate funder, SmithKline Beecham (SB)—are continuing to argue the case for EST patents.

SB's chief of research George Poste asserted in *Nature* a year ago that patenting ESTs "is no different from the patenting of other biomarkers such as the *BRCA1* breast cancer gene" whose functions are unknown. And both Poste and HGS CEO William Haseltine have tried to persuade skeptics that ESTs are patentable as research tools. But these arguments have drawn derisory responses from others in the biotech community hoping to make money from sequence data. Mark Hoffer, counsel to Genzyme Corp. of Cambridge, Massachusetts, a developer of genetic tests and medical products, for example, likens them to "filing a claim on miscellaneous bolts" and arguing "they could be used to make a car." Even PTO Commissioner Bruce Lehman has said, "A lot of this stuff is just data." And as he points out, data alone aren't patentable.

So far, the PTO hasn't taken legal action on EST filings other than those from NIH. Nor has the PTO appeals board or any court touched this issue, because NIH never appealed the rejection of its filings. But Lehman is trying a clever tactic to reduce the backlog of EST patent claims, which he calculates would take his entire biotech staff a year to sort through if it did nothing else. Last October, after consulting with industry, Lehman issued a ruling that no application may contain more than 10 DNA sequences. As a result, companies would have to file thousands of new applications—at $400 to $800 per shot plus legal fees—to maintain all their current claims. "I think our policy will cause the companies to focus on what is the real innovation they're coming up with here," says Lehman dryly.

But even if these applications do not become patents, they could still have a long-term impact on the biotech business. Companies such as HGS and Incyte, for example, hope to split off subsidiary claims covering specific genes after they have investigated them more thoroughly, using the initial EST filing date to establish that they were the first discoverer of the genes. Doll, for example, says there is some concern that a company with thousands of ESTs in its portfolio could create "submarine patents" that vanish today but resurface later, when the company decides to take advantage of an early discovery date. But Doll also notes that a new international agreement limiting the life of a patent to 20 years from the date of filing will put a lid on such submarines.

A public threat? Academic scientists, funding agencies, and at least one major pharmaceutical company have launched a counteroffensive to undermine large proprietary claims on the human genome by encouraging researchers to deposit sequence data in public databases. Some agencies are making quick release of data a condition of their grant awards. These moves may have undermined the value of private EST databases amassed by HGS, Incyte, and others. And some analysts worry that the rush to make public genetic sequences may also undermine the patentability of whole genes in the future.

Poste argued in his article defending gene patents that by publishing DNA data before filing for patents, researchers may render genes "obvious" under patent law, making them unpatentable. And investment analyst Matthew Murray of Lehman Brothers in New York expressed a similar concern in a paper on genomics last September. "The opportunity for genome companies to capitalize on gene discoveries is somewhat limited," Murray wrote, by the "rapid progress being made" by projects that release DNA data quickly. "Obtaining gene patents will become more problematic once the entire human genome sequence is in the public domain."

The impact of public databases and EST applications are just a couple of the uncertainties that hang over the world of gene patents. "It makes a lot of people nervous," says David Galas of Darwin Molecular Inc. in Seattle, but "there's no way of answering these questions until you see what happens" in court. Says attorney Reid Adler of the Morrison & Foerster law firm in Washington, D.C., who filed the first of these EST patent claims for NIH: "I'm sure most of the genes will be identified before these issues are resolved in the courts."

–Eliot Marshall