GENOME RESEARCH

Venter Wins Sequencing Race—Twice

The race to sequence the first complete genome of a free-living organism ended with a flourish last week when Craig Venter, director of The Institute for Genomic Research (TIGR), unveiled the complete sequences of not just one, but two bacteria: *Haemophilus influenzae* and *Mycoplasma genitalium*. Venter made his announcement at a 24 May session of the 95th General Meeting of the American Society of Microbiology (ASM), held in Washington, D.C.

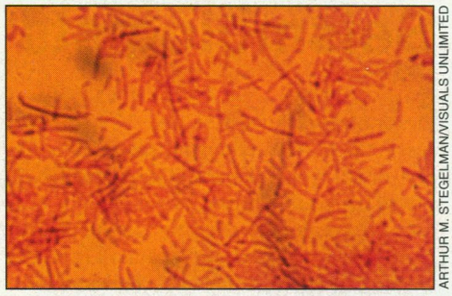
"A remarkable milestone" is how Francis Collins, director of the National Center for Human Genome Research at the National Institutes of Health (NIH), described the accomplishment. Jay Grimes, who heads the Department of Energy's microbial genome project, which funded the sequencing of *M. genitalium*, described himself as "ecstatic."

There is certainly reason for excitement. TIGR's accomplishment—the result of a collaboration with a team led by molecular geneticist Hamilton Smith of Johns Hopkins University in Baltimore who won the Nobel Prize for isolating restriction enzymes from *H. influenzae*—is a landmark for modern genetics. Although the tiny genomes of numerous viruses have already been sequenced, those organisms grow and multiply only with the help of the genetic information contained in the cells they infect. In contrast, the sequences of *H. influenzae* and *M. genitalium* are true genetic blueprints, containing all the information needed to sustain life (*Science*, 13 January, p. 172).

"Everything that should be there for independent life is there," said Venter, who told meeting attendees that a paper describing the effort to sequence *H. influenzae* is under review at *Science*. For example, among the 1749 genes packing the genome of *H. influenzae* (a bacterium that, despite its name, causes ear infection) are those encoding transcription enzymes that copy DNA into RNA and the enzymes that transmute nutrients into usable energy.

As more bacterial genomes are sequenced, researchers will be able to compare them and deduce which genes are responsible for bacterial features such as the ability to cause different diseases and adapt to different environments. And because there are many similarities in bacterial and human biochemistry, the sequences will also be useful for searching for human genes, including those implicated in diseases such as cancer. Bacterial gene sequences have already proved their value in that regard, having helped identify the "mismatch repair" genes that when defective trigger a certain type of colon cancer (*Science*, 10 December 1993, p. 1645).

Over the past few years, TIGR scientists have often been snubbed by public-sector geneticists because they helped generate a massive private database of human gene markers for Human Genome Sciences, the Gaithersburg company that funds TIGR. At the ASM meeting, Venter announced that the whole *H. influenzae* genome sequence will be deposited in public databases at the time of publication. He also took obvious pleasure in noting that NIH had refused to provide federal funds for the effort, because it feared, he said, that TIGR's approach to sequencing the whole genome would fail.



ARTHUR M. STEGELMAN/VISUALS UNLIMITED

Groundbreaking bug. *Haemophilus influenzae*, the first free-living organism to be sequenced.

Until now, the favored course for sequencing genomes has been to carefully break the DNA into ordered, overlapping segments—an extremely time-consuming task—and then sequence the pieces and reassemble them. But Venter and Smith took a more novel approach called whole-genome shotgun sequencing. They simply shattered *H. influenzae* genomes with ultrasonic waves to create DNA fragments whose order was unknown. They then sequenced the fragments and, using powerful computers, put them in order according to how their sequences overlap.

Even as they were sequencing *H. influenzae*, the TIGR-Hopkins team realized, said Venter, that "if we do it again, we can do it much better." After revving up the computer software so that it could process more information, the Venter-Smith collaboration went on to sequence the complete genome of *M. genitalium*, a simple bacterium associated with reproductive tract infections, in less than 4 months.

Despite the excitement generated by the completion of the first full genomic sequence of a free-living organism, such accomplishments may soon seem mundane—at least if TIGR has its way. The institute has the capacity to complete "10 or more microbial genomes per year with high efficiency," Venter told meeting attendees.

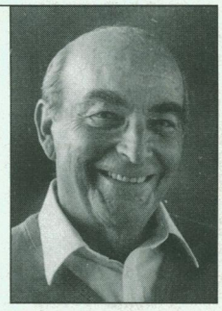
–Rachel Nowak

BIOTECH PATENTS

Scientists Named in PCR Suit

If your lab uses the polymerase chain reaction (PCR), your name may have come up in a San Francisco court. The Swiss pharmaceutical company Hoffmann-LaRoche recently gave the court a list of about 200 basic researchers who—so the company claims—are infringing on its patent for Taq polymerase, the enzyme used in PCR to amplify DNA. Those on the list haven't been charged with patent infringement or anything else. But they are being used by both sides as pawns in a legal battle between Roche and Promega Corp. of Madison, Wisconsin, over Promega's license to sell Taq.

The legal fireworks began after Roche bought the PCR-related patents from the now-defunct Cetus Corp. in late 1991. Cetus had, through those patents, licensed Promega and several other companies to sell Taq for non-PCR uses, such as DNA sequencing. Those companies agreed to refrain from promoting their Taq for PCR. But in October 1992, Roche sued Promega, claiming that the company's marketing and packaging of the enzyme encourages its use in PCR (*Science*, 4 December 1992, p. 1572). Promega denied wrongdoing and responded with its own lawsuit challenging the validity of the Taq patent (*Science*, 23 April 1993, p. 486). Both suits have slowly been working their way through federal courts in New Jersey and San Francisco, with no end in sight.



Opposes PCR patent. Arthur Kornberg.

Because "Roche accused us of contributing to and inducing infringement of their patents," says Promega President William Linton, "we simply asked the question, who are the direct infringers?" Roche was legally obliged to produce this list for the case to go forward. So, according to Kathy Ordoñez, president of Roche Molecular Systems in Branchburg, New Jersey, Roche scanned the literature for publications that mentioned PCR, searched the materials and methods sections for the source of the Taq polymerase, and listed authors who named Promega as their source of Taq. Roche submitted the list to the court 2 weeks ago.

Linton submitted it to assembled journalists at a San Francisco press conference a week later. He says basic researchers should be alarmed by the list, which he cites as evi-

dence of Roche's intent to limit their freedom to choose reagents. At the press conference, Linton had the passionate support of Stanford biochemist Arthur Kornberg, who won a Nobel Prize for his work on DNA polymerases, the class of enzymes to which Taq belongs. Kornberg argues that patents on PCR and Taq are invalid because the process and the enzyme were described in the scientific literature before their publication by Cetus scientists. Cetus's patents "were flagrant abuses of knowledge that has been in the public domain," Kornberg claims.

Ordoñez says Roche has no plans to prosecute the listed researchers. She accuses Promega of grandstanding—using the list in a publicity stunt to try to win the allegiance of the research community. "Promega asked for the list, took the list, and is waving it in the public eye," she says, "to create visibility and to cause anxiety among these institutions and scientists. If so, the strategy may be working.

Listed researcher Melvin Simon of the California Institute of Technology says it would be "outrageous" if Roche were to threaten individual investigators with legal action. He says his lab has not used Promega Taq for PCR in recent years because of Roche's "virulent stand" on the issue. The list, he predicts, will alienate researchers like himself who have "tried to go along with the system." Douglas Hanahan of the University of California, San Francisco, says he is not pleased with either Roche or Promega, noting that the legal sparring has become "pretty nasty" on both sides.

Stephen O'Gorman of the Salk Institute was one of several researchers who say their listed papers mention no source for Taq polymerase, and that they have never used the Promega enzyme. But O'Gorman says that in his case that wasn't due to concerns about license violations. "That's something we don't think about," he says. Basic scientists who use Promega Taq "aren't violating the patent for profit," O'Gorman says, "so I can't imagine that they would have any liability."

But Promega attorney Peter Carroll warns that "even if they [Roche] don't go after the researchers, no one should breathe a sigh of relief." If the courts accept that basic researchers can be considered infringers on the Roche patents, he says, that will overturn the general precedent that basic researchers are exempt from patent restrictions, leaving the door open for Roche to sue the researchers later. Says Ordoñez: "We have no intention to involve these or any other scientists in the litigation with Promega." But, she adds, "I wouldn't want to predict what action Roche would take relative to any patent ... in the future." Those carefully chosen words are unlikely to give comfort to researchers on the list.

–Marcia Barinaga

EARTH SCIENCE

At the Heart of Earthquakes, Seismologists Look for Pulse

Seismologists have it tough. While many researchers can delve right into their object of study—biologists manipulate DNA, physicists shuffle atoms about one by one, and meteorologists even fly through hurricanes—seismologists are separated from the source of earthquakes by kilometers of solid rock. What's more, the seismic waves that emanate from the depths of a fault during an earthquake only grudgingly reveal details about the processes that generate them. As a result, the deep parts of faults have largely remained terra incognita, leaving researchers unsure about why quakes strike where and when they do. Lately, however, seismologists have managed to glimpse the working heart of faults—and found persistent differences between the fault sections that generate earthquakes and those that don't.

Seismologists had long suspected that some fault property—the roughness of its faces, say, or the composition of its rock—varies along a fault and determines whether its sides slip harmlessly by each other or become stuck and eventually break loose in an earthquake. But until recently, researchers weren't sure if these "stuck spots," or asperities, really exist, or whether they persist for quake after quake, influencing their size, location, and timing. The stakes here are high. If the asperities do exist, they would help seismologists forecast when and where the next quake will strike, and how big it will be. If they don't, long-term earthquake prediction might be a vain hope. "It's a question of how orderly nature is," notes Stuart Nishenko of the U.S. Geological Survey (USGS) in Golden, Colorado.

"At the large scale or the small scale, there are physical features that control where rupture occurs and the dimensions of that rupture."
—David Schwartz

Two kinds of studies are now suggesting that nature is reasonably orderly—at least on some faults. New analyses of earthquake records by William Ellsworth of the USGS in Menlo Park, California, and by others have shown that smaller quakes often strike the same spot on a fault at more or less regular intervals, implying that the fault has a persistent stuck patch. Larger, more damaging earthquakes are generally too rare to reveal such clear regularities, but seismic tomography—a computer-aided technique for extracting images from seismic waves—is filling in by showing that quakes as large as magnitude 7.3 originate on patches of especially strong rock along the fault. "The bottom line," says geologist David Schwartz of the USGS in Menlo Park, "is that when you look at the large scale or the small scale, there are physical features that control where rupture occurs and the dimensions of that rupture."

Ellsworth suspected that was the case more than 20 years ago, when he and Robert Wesson of the USGS in Reston, Virginia, studied a 5-kilometer patch of the central San Andreas fault in Bear Valley, California, that had ruptured in a magnitude 5 earthquake in 1972. Based on details of the quake and its aftershocks, he and Wesson suggested that the site was an asperity that repeatedly locks up the fault until enough strain accumulates to break it in another nearly identical quake. That interpretation gained support when Ellsworth and Lynn Dietz, also at Menlo Park, compared the 1972 quake's seismogram with the reading from a 1951 quake. The seismograms matched like a pair of fingerprints, demonstrating that the two quakes had broken the same patch of fault.

Last month, Ellsworth and Wesson's interpretation looked even better when the same patch of fault ruptured again in an earthquake of about the same size. And the same pattern of repeating, nearly identical quakes is turning up elsewhere on the San Andreas, where some small and medium-size earthquakes recur at the same spots, sometimes with almost clocklike regularity.

Ellsworth and Dietz, for instance, have identified 10 series of repeating earthquakes of magnitude 4 to 5 along the central San Andreas. Each string of quakes includes as many as six events, which broke the same fault patch at intervals of from 10 to 20 years. Similarly, John Vidale of the USGS in Menlo Park, Ellsworth, and their colleagues recently showed that much smaller "microearthquakes" of magnitude 1.3 have broken the same 20-meter patch of the San Andreas just north of Parkfield nine times,