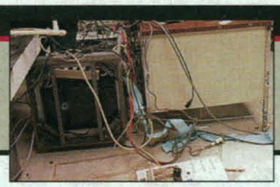


PAGE 410
Animal rightists trash Univ. of Minnesota labs

415
The world's largest known bacterium

GENOMICS

Drug Firms to Create Public Database of Genetic Mutations

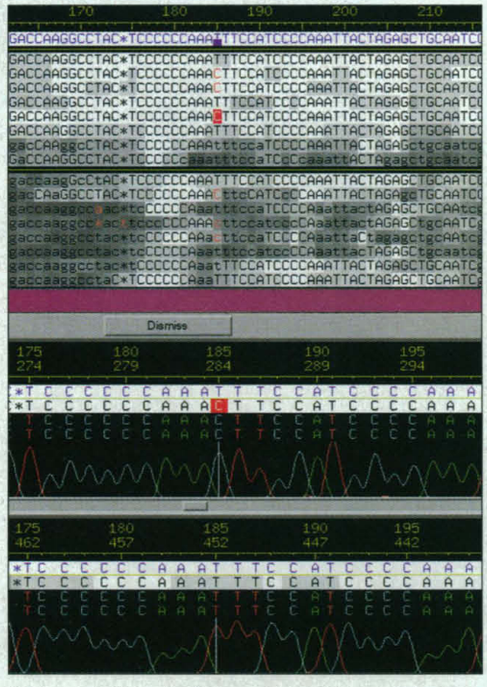
Private investment in genome research has changed the way biology is done in the 1990s and created some odd partnerships. But a remarkable venture, announced this week, could set a new standard for high-impact science by an unlikely set of collaborators: Ten large pharmaceutical companies and a British charity will spend $45 million to create an archive of human genetic variation—and give it away. These fiercely competitive companies will team up to bankroll work by a network of academic labs, and the collaboration will release proprietary data free of charge to all comers. In a field known for cutthroat competition and secrecy, this arrangement is, to say the least, an anomaly.

The project's goal is straightforward: It will identify variable points in the human genetic code—known as single nucleotide polymorphisms, or SNPs—as quickly as possible. SNPs are single-base variations that can serve as physical landmarks along the 3 billion bases of the human genome. SNPs will be used as analytical tools, making it easier to trace inherited disease risks and abnormal responses to drugs. The nonprofit SNP Consortium, or TSC (its official name), will publish data every quarter on the Internet, organizers say, no strings attached.

TSC isn't altruistic, though. The companies backing the enterprise—a Who's Who of the drug industry—expect that SNPs will enable them to develop and sell drugs more effectively. And by creating a public database, they will avoid having to buy multiple, private data collections from the half-dozen or so biotech firms that have been collecting SNPs since 1997, hoping to stake a proprietary claim on the data. In addition to the Wellcome Trust philanthropy of Britain, which is contributing $14 million to the project, the sponsors include Bayer Group AG, Bristol-Myers Squibb Co., Glaxo Wellcome PLC, Hoechst Marion Roussel AG, Monsanto Co., Novartis AG, Pfizer Inc., Roche Holding Ltd., SmithKline Beecham PLC, and Zeneca Group PLC. Each of the 10 companies put in $3 million apiece.

Arthur Holden, a former biotech executive who heads TSC, says the nonprofit labs that agreed to do the research have specific marching orders: They are to find 300,000 SNPs in 2 years. SNPs are believed to be evenly distributed along the human genome at a frequency of about one per 1000 bases. Once found, the SNPs will be tracked to positions on the genome. The goal is to have 150,000 mapped in this way by mid-2001.

Few SNPs are likely to be directly involved in disease, but a database of several hundred thousand, researchers hope, will make it easier to track smaller segments of the genome and identify patterns of inheritance that affect health. The SNP map, they anticipate, will make it possible to diagnose illnesses earlier and avoid giving drugs to patients likely to experience side effects, including drugs already in use. This last possibility is why the companies are eager to get SNPs as soon as possible.



SNP spotting. Software identifies a cytosine-thymine substitution in this sample.

Some scientists are so excited about the open nature of this effort that they're calling it "a new model" of public-private collaboration. "This is absolutely unique," says Eric Lander, director of the genome center at the Massachusetts Institute of Technology's Whitehead Institute for Biomedical Research in Cambridge, Massachusetts, a partner in TSC. The companies have recognized, he says, that the SNP map ought to be "common infrastructure" and that "locking it up as private intellectual property didn't advantage anybody." The private sector "wanted all these SNPs," Lander says, but company executives "realized it didn't make sense to compete on them." Instead, these companies will compete on developing drugs. Lander hopes this will inspire other companies to donate research tools to biology.

Francis Collins, director of the National Human Genome Research Institute (NHGRI), also sees it as a "groundbreaking" event, adding that "I know of no previous example of this kind of collaboration" between companies and basic biologists. However, in its technical objectives, the project echoes a government-funded effort to create a database of 60,000 to 160,000 SNPs, which Collins and NHGRI launched last year (*Science*, 19 December 1997, p. 2046). Collins says the two efforts are "nicely complementary." TSC will be collecting random data across the entire genome, he explains, while half of the government-funded investigators will be focusing SNP hunts on specific genes involved in disease. In any case, Collins adds, the more SNPs people find, the better: The new consortium "in no way implies that we should diminish our efforts. …"

The TSC's effort is divided into several stages: hunting for SNPs, mapping and archiving them, and releasing them to a public database. The hunt will be carried out by three large human genome sequencing teams: Lander's group at the Whitehead, a group under David Bentley at the Sanger Centre near Cambridge, U.K., and a third team under Elaine Mardis at the Washington University genome center in St. Louis. According to staffers, all three will use newly developed biological tricks and software to identify SNP variations in randomly sequenced human DNA.

After candidate SNPs have been found, they will be forwarded to an archive at the Cold Spring Harbor Laboratory in New York.

CREDIT: WASHINGTON UNIVERSITY

There, bioinformatics leader Lincoln Stein will double-check the data and run a computerized scan against previously banked human genome sequence data, looking for matches. Wherever the computers find a "hit," they will note the genomic location—a process Stein calls "mapping in silico."

At present, only a fraction of the human genome is available for this kind of SNP mapping. As genome centers release more DNA sequence over the next year, Stein says, in silico mapping will become more powerful. Meanwhile, as insurance, researchers at Stanford University and the Sanger Centre will use a "radiation hybrid" marker system based on cloned fragments of the genome to fix the approximate location of SNPs until they can be located more precisely. Holden, TSC's chief executive, says the mapped SNPs will then be sent to a law firm, which will file patent applications but convert them to simple registrations of invention to prevent others from claiming priority. Every quarter, starting on 15 July, Cold Spring Harbor Lab will release the data on a public Web site. No one—not even key sponsors like Glaxo Wellcome (GW)—will get an early peek.

Allen Roses, GW's worldwide director of genetic research, says the organizers didn't set out to create this public resource; it just grew naturally. Glaxo got interested in SNPs about 18 months ago, Roses says, when an in-house experiment demonstrated that they could be used to speed up the search for disease susceptibility genes. Glaxo suggested to other companies that they jointly fund a collection of SNPs and a map of the human genome, dubbed "Atlas." The cost estimate was high, though: $150 million. "We didn't think that anybody's board would go along if it was a public sort of thing," Roses says, so Glaxo considered creating a private venture that would not release all the data. But when the partners insisted on public release, all agreed.

Last summer, according to Roses, there came "a critical point that made it seem like [a public consortium] was going to work" after all. The academic centers leading the human genome sequencing project expressed a willingness to collaborate. The Wellcome Trust helped in negotiations with academic centers. At the same time, Celera Genomics Inc. of Rockville, Maryland, and NHGRI announced that they were going to speed up the pace at which the human genome will be sequenced (*Science*, 18 September 1998, p. 1774). If Celera and NHGRI produced a whole genome, it would be easier to conduct in silico mapping of SNPs, the consortium leaders realized. The cost estimate for the SNP project dropped sharply—to about $40 million. By the end of 1998, the goals were set. And in January 1999, the Wellcome Trust funded SNP-hunting pilot projects at the three big sequencing centers. Soon, the project was on its way.

Will it really be possible to find 300,000 SNPs and map half of them in 2 years? Perhaps so, if indications from the pilot projects hold up. "They have come in with more SNPs than we had anticipated finding," says Lander. Bentley and Mardis echo his optimism. As new sequencing technology gets installed, Roses concludes, "I think we're going to exceed the goals." –ELIOT MARSHALL

EPIDEMIOLOGY

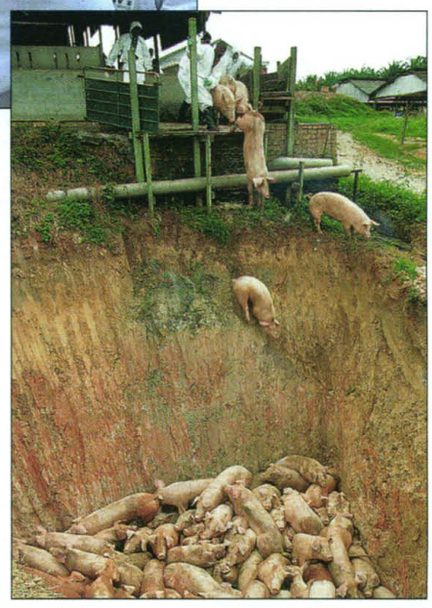
New Virus Fingered in Malaysian Epidemic

Scientists have unmasked a killer responsible for the deaths of at least 95 people in Malaysia in the last 6 months, most of them pig farm workers. The culprit, named the Nipah virus for the small town from whence the strain was first identified, is a previously unknown virus that replicates in pigs and seems to be easily transmitted to humans. It is closely related to another notorious agent, the Hendra virus, which surfaced in Australia in 1994 and killed two people and more than a dozen horses. But the new virus spreads much more rapidly, making it "an emerging virus of grave concern," says John Mackenzie, head of the department of microbiology and parasitology at the University of Queensland in Brisbane, Australia.

The country's health authorities initially assumed they were dealing with an outbreak of Japanese encephalitis (JE), which causes similar symptoms, and some critics have accused the Malaysian government of being slow to consider alternative causes. Indeed, even now authorities insist that the country is battling a "dual epidemic" of JE and the new disease. But last week the U.S. Centers for Disease Control and Prevention (CDC) in Atlanta said that the Nipah virus is the main culprit, and that the JE virus has played at best a marginal role in the continuing tragedy. The epidemic peaked around the middle of March and seems to be waning now. Early this week, the Malaysian Ministry of Health put the total number of cases at 251. Since early March, Malaysian health authorities have killed over 800,000 pigs to halt the spread of the virus.

The first cases of the disease occurred in late September near the city of Ipoh, in the northern state of Perak. The victims, all of whom worked in the pig industry, came down with high fever and encephalitis (an inflammation of the brain), and some died. Officials concluded that they had succumbed to JE, which is transmitted by *Culex* mosquitoes and known to replicate in pigs.

Circumstantial evidence supported that theory. A few dozen people contract JE in Malaysia yearly, and the numbers usually rise at the end of the year. In addition, tests at the University of Malaya in Kuala Lumpur and the Institute of Tropical Medicine in Nagasaki, Japan, confirmed that blood and cerebrospinal fluid of some patients contained antibodies against JE. To contain the outbreak,



Deadly business. Over 800,000 pigs were killed to halt the spread of the Nipah virus, first isolated by Lam Kai Sit (*top right*) and Chua Kaw Bing.

CREDITS: (LEFT TO RIGHT) UNIVERSITY OF MALAYA; ANDY WONG/AP