

A Deal for the Rice Genome

For the second time in just over a year, *Science* is at the center of a debate over public access to the data behind a major genome paper it is publishing. The issue: Should journals refuse to publish any DNA sequence paper unless the authors make the data freely available through a public database such as GenBank?

On page 92, a team from the Switzerland-based agricultural biotechnology giant Syngenta describes a draft sequence of the *japonica* subspecies of rice. Under an agreement reached with *Science*, the company is making the data publicly available through its own Web site (tmri.org) or on a CD-ROM, rather than through GenBank. Scientists can use the partially assembled raw genome sequence without strings for research, and Syngenta will permit researchers to publish papers and have Syngenta deposit a gene's worth of DNA data in GenBank without negotiation. (The raw data include minimal notes, an official says, such as labels on DNA likely to be "nonrice in origin.") Larger amounts will require a specific agreement. The company seeks no "reach-through" intellectual property rights, but scientists doing commercial work must negotiate their own data-access agreements.

Last year, *Science* touched off a furor when it struck a similar deal with Celera Genomics of Rockville, Maryland, as a condition of publishing Celera's draft of the human genome (*Science*, 16 February 2001, p. 1304). Celera gives noncommercial researchers free access to raw DNA sequence but charges a fee for access to its annotated gene database. Criticism in a more muted form surfaced again several weeks ago when word of a possible Syngenta agreement with *Science* began to spread in the genomics community. A score of leading researchers—including Michael Ashburner of Cambridge University, U.K., David Botstein of Stanford University, and Maynard Olson of the University of Washington, Seattle—circulated a letter arguing that failure to insist that the sequence be deposited in GenBank constituted a "very serious threat" to genomics research.

"We understand that concern," says *Science* Editor-in-Chief Donald Kennedy, noting that it would be ideal to have "one-stop shopping" for all genomic data at GenBank. But, Kennedy said at a press briefing last week, the company would have been unwilling

to publish its raw data if it had been required to deposit the sequence in GenBank. "We think that the public benefit of bringing this important science out of trade secret status greatly outweighs" the cost of granting an exception, Kennedy said.

The arrangement has not so far prompted the intense reaction that greeted the Celera agreement. One reason is that Syngenta has promised to work closely with publicly funded groups to produce more complete drafts of the rice genome (see Letters, p. 45). Monsanto of St. Louis, Missouri, which produced its own draft of the *japonica* sequence 2 years ago but hasn't published it, is also cooperating in this endeavor. Members of the public consortium working with Monsanto say that 30% of the data they have released to GenBank originated from the company.

The Syngenta sequence will be useful in refining draft sequences. "Thanks to Syngenta, I don't think it will be so hard" to close gaps between the more than 100,000 fragments in the draft sequence of the *indica* subspecies—also being published this week (p. 79)—says Wong Gan Ka-Shu, a leader of the research team that sequenced *indica*. (The team's draft sequence has been deposited in GenBank.)

As a result, much of the Syngenta sequence is likely to end up in GenBank over the next "12 to 18 months," mingled with data the public groups will be depositing, says Steven Briggs, head of Syngenta's Torrey Mesa Research Institute in San Diego, California, which oversaw the company's sequencing project. Asked why Syngenta is not prepared to deposit its sequence in GenBank now, Briggs said last week that Syngenta believes it has "a significant commercial advantage" and isn't ready to permit unrestricted use of its data by its competitors.

Susan McCouch, a rice genome researcher at Cornell University, is disappointed that Syngenta's data are not going directly to GenBank. This would have made whole-genome comparisons "easy," she says, enabling more rapid discovery of gene function. Despite the decision not to deposit data in GenBank, Rod Wing of Clemson University in South Carolina has concluded that the new data-sharing terms look "very good," particularly because there are "no reach-through terms" seeking to patent scientists' discoveries.

—ELIOT MARSHALL

ment, RGP was in the strongest position, receiving \$10 million in 1998 for the first year of a projected 10-year effort. But no government sequencing funds were forthcoming in the United States until 1999, when the U.S. Department of Agriculture (USDA), the National Science Foundation (NSF), and the Department of Energy finally came up with \$12.5 million to fund two groups to work on chromosomes 10 and 3. "The thing that slowed us down the most was the late entry of the U.S. funding agencies," says Burr. Other countries fared even worse. The U.K. and Canadian groups never won funding. And Thai researchers joined the consortium, contributed a small amount of sequence data, but then withdrew to concentrate their scarce funding on gene discovery.

Japan, however, picked up some of the slack, and France, Taiwan, South Korea, India, and Brazil stepped in to share the burden. Finally, by the beginning of 2000, IRGSP seemed on its way toward its goal of completing the *japonica* rice genome se-

quence by 2008, possibly earlier. But the international consortium soon had company.

The tortoise and the hare

In April 2000, Monsanto announced that it had sequenced the *japonica* genome. Working with researchers at the University of Washington and the Institute for Systems Biology in Seattle, Monsanto had produced an incomplete, but very informative, version that the company promised to share with individual academic researchers and with

IRGSP. The news shocked, then worried, but ultimately delighted the community, because the data promised to speed up the IRGSP effort.

"The Monsanto data has been very helpful and very valuable," says Machi Dilworth, who oversees plant genomic programs at NSF. The French group sequencing chromosome 12 turned to Monsanto for 75 of the 109 clones it is now sequencing. And the Shanghai group set aside its *indica* project, picked up the Monsanto materials, and finished a draft of chromosome 4. The Monsanto material "is now proving to be of value in speeding up the IRGSP sequencing and the cost effectiveness of the overall project," says Ed Kaleikau, a plant biologist at USDA's Cooperative State Research, Education, and Extension Service.

The Monsanto boost was not decisive, it turns out. In January 2001, Syngenta reported that it, too, had sequenced *japonica*. It had contracted Myriad Genetics in Salt Lake City, Utah, to work on rice

GENOME SIZES	
Species	No. of bases
<i>Arabidopsis</i>	125 million
<i>Drosophila</i>	180 million
Rice	430 million
Maize	3000 million
Human	3000 million
Wheat	16,000 million