The near completion of the sequence of the *C. elegans* genome should provide researchers with a gold mine of information on topics ranging from evolution to gene control

Worming Secrets From the *C. elegans* Genome

In the early 1800s, American explorers Meriwether Lewis and William Clark traveled 8000 miles—from St. Louis, Missouri, to the Pacific Ocean and back—mapping newly acquired territory that expanded the size of the United States by more than a million square miles. The end of their 2-year journey marked the beginning of a century of expansion and growth for this new nation. Biology is now at a similar juncture, marking both an ending and a new beginning.

Lights on. The glow of green fluorescent protein reveals which of the nematode's nerve cells have a particular active developmental control gene.

This issue of *Science* signals the end of an 8-year effort to sequence the first animal genome, with the publication of the virtually complete sequence of the 97 million bases in the genome of a tiny nematode worm, *Caenorhabditis elegans*. But this milestone is also the beginning of a new era in biology.

For one, the worm-sequencing effort has helped pave the way for sequencing the 3 billion bases that make up the human genome, a project that will extend into the next century. The early successes of worm sequencing were instrumental in convincing researchers and funding agencies of the value and feasibility of large-scale sequencing projects (*Science*, 10 February 1995, p. 783; 2 June 1995, p. 1270). And now, the two groups who sequenced the worm genome, located at the Washington University Genome Sequencing Center in St. Louis and the Sanger Centre in Cambridge, U.K., expect to use the skills they've acquired to generate about half the human genome. "They cut their teeth and learned how to do high-level sequencing by practicing on the worm," says David Botstein, a geneticist at Stanford University.

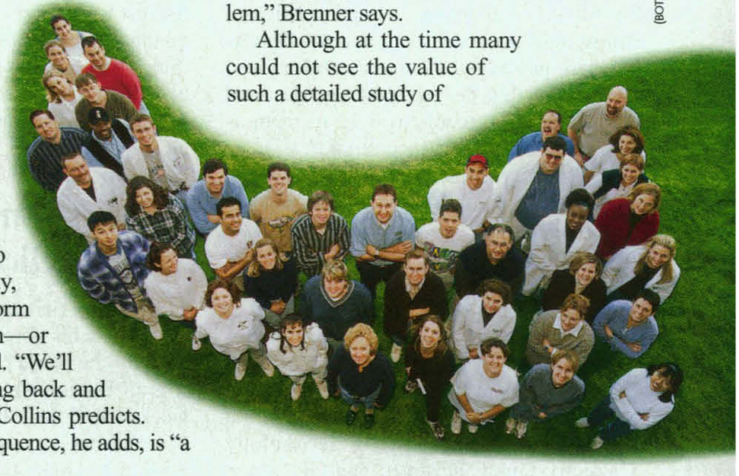
But beyond that, as the first sequence of a multicellular organism, the *C. elegans* genome should provide a cornucopia of biological information—and not just about the worm. As Gary Ruvkun, a developmental geneticist at Harvard Medical School in Boston, notes, "It's the first time we can see all the genes needed for an animal to function." As a result, says Francis Collins, director of the National Human Genome Research Institute in Bethesda, Maryland, countless other life scientists in addition to the 1200 or so who call themselves worm biologists will be tapping the nematode sequence for their studies.

He notes that studies of functions as diverse as muscle contraction, fear responses, digestion, and reproduction often lead researchers to some gene whose precise function is unknown. But because of the conservation between genomes, a matching gene can often be found in some form in the worm—even if the original organism of interest is only very distantly related, say a mammal such as the mouse or even a human being. And thanks to years of intensive study, the function of many worm genes is already known—or may soon be determined. "We'll be doing a lot of jumping back and forth between species," Collins predicts. The completion of the sequence, he adds, is "a significant milestone."

It should shed light not just on how existing multicellular organisms function but also how they came to be. Comparisons of the *C. elegans* genome with those of yeast and the other microbes that have been sequenced have revealed both similarities and differences that are sparking new thinking in disciplines from evolutionary biology to protein chemistry. For example, they can help answer the question of how genomes have expanded and changed to support multicellular life. The nematode sequence promises to be "a basic organizing principle" for all biologists, says Robert Waterston, who headed up the sequencing effort at Washington University. "It allows cross-communication [between] very different fields."

A humble beginning

The first person to sense that the worm might take on such a prominent role in biology was molecular biologist Sydney Brenner of the Medical Research Council (MRC) Laboratory of Molecular Biology in Cambridge. During the mid-'60s, he wanted to understand how the various parts of the nervous system get wired up correctly during development. In complicated species, like mice, humans, or even fruit flies, this problem seemed intractable. But *C. elegans* is both small—its 959 cells include only about 300 neurons—and transparent so that all the cells can be seen and followed during development. It "turned [this question] into a finite problem," Brenner says.



(BOTTOM) R. BOSTON

Although at the time many could not see the value of such a detailed study of a simple worm, Brenner did manage to recruit several young scientists to worm studies at his new lab. One was Waterston, an immunologist interested in muscle development who arrived in 1972; another was John Sulston, an organic chemist turned biologist, who by 1983 had traced the fate of each cell as the nematode transformed from a single cell to a full-sized worm.

The cell lineage map, described at the time as a "monumental achievement" by a Sulston group alumnus, Bob Horvitz of the Massachusetts Institute of Technology, laid the

CREDITS: (TOP) O. HOBERT AND G. RUVKUN/MGH

groundwork for determining just what influenced the development of the various cells. For example, researchers could destroy one cell to see what effect, if any, its absence had on the development of its neighbors. But to get to the underlying biochemical mechanisms that determine cell fates, researchers needed to track down the genes involved. So Sulston and Alan Coulson at the MRC lab decided to make a physical map of the worm genome, consisting of a set of landmarks, separated by known numbers of bases, along each chromosome. Such a map would enable them to home in on a gene's approximate location faster than they could before.

With help from a growing nematode research community, the two began the project by making a "library" of pieces of the entire worm genome, grown in bacteria. They then used a technique called fingerprinting to establish landmarks on each piece; by comparing the landmarks on the DNA pieces, they could determine which pieces overlapped and thus how to arrange all of the pieces and their landmarks into a map of the whole genome. As they struggled to link the pieces to cover entire chromosomes, Waterston realized that a technique developed by Maynard Olson, a colleague of Waterston's at Washington University, might be useful for filling the gaps. Olson was making yeast artificial chromosomes that contained pieces of human DNA; as David Burke in Olson's group later found, YACs turned out to be capable of expressing the missing bits of nematode DNA as well.

Transatlantic team. The Washington University (left) and Sanger Centre (right) groups line up to resemble their favorite experimental animal.

(BOTTOM) BARKER EVANS PHOTOGRAPHY

CREDITS: (TOP) MRC LABORATORY OF MOLECULAR BIOLOGY

In 1989, the nearly complete map took up an entire wall when displayed at the worm biology meeting at Cold Spring Harbor Laboratory in New York. Such a map was a prerequisite for sequencing a genome, as it divided each chromosome into smaller chunks whose makeup of bases could be determined and then fit into the proper place in the chromosome. Thus, to promoters of the Human Genome Project looking for a smaller genome to try first, the worm looked quite promising.

At that meeting, Waterston, Sulston, Coulson, and Horvitz mapped out a pilot project whose goal was to sequence 3 million bases—about 3% of the worm genome—by 1993. "The idea was that this was a dry run for the human genome," says Sulston, who is now at the Sanger Centre.

Getting support was not easy, however. "Close colleagues told me I was nuts," Waterston recalls. At the time, sequencing successes were measured in thousands of bases, and the nematode genome had millions. Also, many did not think it would be useful to spend millions of dollars "on something which didn't solve biological problems right off," says Sulston.

The worm researchers finally got initial grants from the MRC and the U.S. National Institutes of Health in 1990. But although the project was well on the way to meeting its first goal of sequencing 3 million bases in 1992, getting the rest of the money needed "was touch and go," Sulston recalls. Not until venture capitalists threatened to lure Waterston and Sulston to a private sequencing effort did the public support fall in place.

To go beyond the 3 million bases to the full genome, Sulston and Waterston first needed to make sequencing cheaper and more efficient. By 1992 they were doing a million bases per year, which meant that it would take nearly a century to finish the entire genome and would cost some $200 million. The researchers doubled, then quadrupled, the amount of DNA their automated sequencers could process at one time by adding more separation lanes to these machines and running the machines day and night. Both centers worked on streamlining the preparation of DNA and improving the data quality. Their labs mushroomed, and instead of managing a dozen researchers, Sulston and Waterston each eventually had 100 or more workers.

Wild about worms. Sydney Brenner (standing), John Sulston (left), and Alan Coulson (middle) pioneered nematode biology.

Unexpected biological hurdles slowed their progress, though. "There was a complexity [in the genome] that we weren't fully aware of," says Waterston. The researchers originally concentrated their sequencing efforts on the middle of the chromosomes, thinking that that's where most of the genes are. But there proved to be almost as many genes farther out on the arms, with the result that the number of genes turned out to be much higher—19,000—than the 15,000 they originally expected. That meant they couldn't relax their accuracy standards as they once thought they might be able to do in gene-poor regions. They also had to deal with hard-to-sequence repetitive regions throughout the genome.

Finally, because the researchers had done most of their sequencing on worm DNA pieces grown in bacteria, they had optimized the sequencing operations for the bacterial material. But the last 20% of the DNA was not contained in bacteria but in YACs grown in yeast, and that required figuring out how to revamp their procedures to deal with contaminating yeast sequences. (Even now, a few hard-to-sequence gaps remain to be finished.)

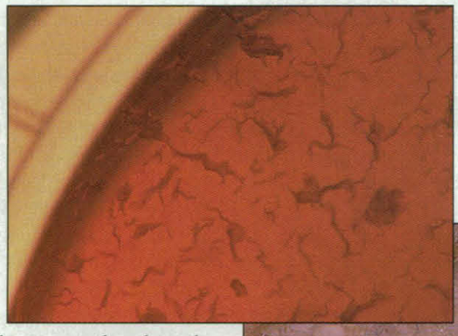
But by 1995, the tide had turned in favor of the sequencing effort. The two centers were producing several million bases of sequence per year. Many biologists had started to realize how useful the promised sequence would be for speeding up their own studies. "People were writing grant proposals around the idea that the genome was coming," says Sulston. "It was becoming clear that if you were serious about a research question, you couldn't address it without genomics." And that required having the full worm sequence in hand.

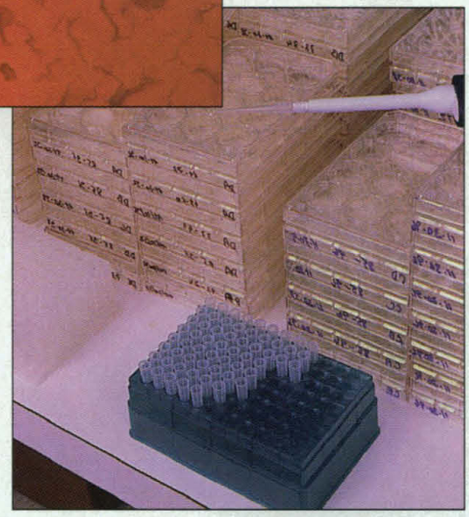
A new biology

The special section on the *C. elegans* genome, which begins on page 2011, hints at

the richness of this genome for neurobiologists, biochemists, developmental biologists, and researchers seeking to understand how genomes change through time. As these papers show, computer analyses comparing the worm's DNA with that of other organisms are already confirming the existence of a wealth of genes that have been conserved through evolution. For example, the genes needed for individual cells to function, such as those that code for proteins involved in DNA replication or protein synthesis, appear virtually the same in yeast and worm, and even in higher species. In addition, genes involved in development are often shared by the worm and other multicellular organisms. These similarities mean that "*C. elegans* becomes this radio beacon that you're triangulating all of biology from," says Ruvkun.

Stanford's Botstein agrees. He describes the similarities in the developmental genes and others as "an unbelievable boon to understanding what all these genes [do]." If computer gene-matching programs show that a target gene in another organism also exists in the worm, then knowledge about the worm gene—such as the identity of the cell regulatory pathway in which that gene's protein operates—can add to the understanding of its counterpart. "Suddenly you have not just your gene, but [the] context revealed," says Waterston. "You're looking at the forest, not just the tree."





Worm zoo. Stocks of mutant nematode worms, such as the ones shown at top, help researchers quickly home in on gene function.

And even if researchers don't know the function of a worm gene of interest—and the functions of 12,000 of the 19,000 *C. elegans* genes are still a mystery—it can be easy to find out. It is much simpler to evaluate gene function in the worm than in, say, a human.

For example, one 2-year-old technology known as RNAi (*Science*, 16 October, p. 430) has made it easier to knock out a worm gene and see how its absence affects the animal. RNAi involves simply injecting worm oocytes with a piece of double-stranded RNA that matches the gene and somehow blocks its expression. Experimenters then look for changes in the worms that develop from the injected eggs. There's even some hint that eventually researchers may be able to mix the RNA with a worm's meal and then study the effects of having the gene inactivated in the animal's offspring.

Moreover, because the worm is see-through, the gene under study can be linked with the gene for green fluorescent protein (GFP). Then, researchers can find out when and where in the worm the hybrid gene is expressed by monitoring GFP's glow.

To further speed the study of *C. elegans* gene function, Coulson is working with Robert Barstead at the Oklahoma Medical Research Foundation in Oklahoma City and Don Moerman at the University of British Columbia in Vancouver on a large-scale effort to generate large numbers of mutant worms. They have chemically induced mutations in about a half-million worms and are now screening them for interesting defects. Already, Coulson says, "we've had quite a lot of interest" from researchers studying human diseases who want to see if comparable defects crop up in the mutant worms.

Using the mutants, RNAi technology, and GFP proteins, these scientists can begin to learn more about how the gene at fault functions. "The worm is a highly tractable genetic organism," says Horvitz. "It's going to make a huge impact on mammalian genetics and human disease."

Researchers who want to understand how genes are regulated are also turning to the *C. elegans* genome. Take developmental biologist Stuart Kim at Stanford. He has teamed up with Yuji Kohara from the National Institute of Genetics in Mishima, Japan, to make a microarray, a glass plate with bits of worm DNA attached, which will make it possible to study when and where the worm's genes are activated by regulatory signals. By January, Kim says, the array will have some 12,500 genes represented on it, each one designed to produce a spot of fluorescence when exposed to a sample containing RNA messages from the corresponding gene. And in a few months more, he expects to have all 19,000 genes on the array. "Then we'll be ready to rock," he says.

Already some two dozen labs are gathering RNAs for testing on the array. And Kim is busy writing computer software that will let him discern global regulatory networks—clusters of genes whose activities are interconnected—from the microarray results. Thus he hopes to look at, for example, all the genes turned on by a particular regulatory protein, such as the cancer-promoting Ras, and then see whether some of those same genes are also activated by other proteins. In this way he hopes to discern the connections between various DNA regulatory pathways. This question "is hard to study one gene at a time," he points out.

And still other researchers are using the genome to address evolutionary issues. On page 2018, Neil Clarke and Jeremy Berg, biochemists from the Johns Hopkins University School of Medicine in Baltimore, compare genes for DNA regulatory proteins that interact with zinc in the worm, yeast, bacteria, and a separate group of microorganisms, called Archaea, best known for living in extreme environments. Only the yeast and nematode genomes code for large numbers of zinc-binding proteins.

But although many of the zinc-binding proteins in *C. elegans*—233 of them, in fact—serve as receptors for steroid hormones, relaying messages from the hormones into the cell, those receptors are completely missing in yeast. Most likely, the evolution and expansion of steroid hormone receptors "coincided with going from a single-cell to a multicellular organism," says Berg. Presumably, steroid hormones and their receptors were needed to help coordinate the activities of different cells throughout the body.

Other evolutionary studies have posed new puzzles. On page 2033, Ruvkun and Harvard colleague Oliver Hobert's first pass through the worm genome confirms that animals share not just genes but entire regulatory pathways, governing their development. So what, he asks, makes people look so different from worms? Is it that the genes and pathways shared by this menagerie are simply turned on at different rates or at different times during development? Or are there also sets of genes that are not shared, which account for why worms look one way and flies another? "The analysis of genes that work in specific organisms should answer this question," he predicts.

Just the possibility of asking such questions is what reveals the real promise of the completed genome, says the MRC's Brenner. "The sequence is not the end of the day," he emphasizes. "It's the beginning of the day." **–ELIZABETH PENNISI**

CREDITS (TOP TO BOTTOM): JANET DUERR; BRENT PURDY