SEQUENCE THE HUMAN GENOME

"Know thyself." That was chiseled in rock in the temple of the Oracle of Delphi—presumably because the priestess's murky utterances revealed only the supplicant's inner thoughts—a sort of verbal Rorschach blot.

Knowing thyself is good policy—assuming that thyself is somebody worth knowing and that self-study is not the only course in the curriculum. And that, in a sense, is the problem facing us on the molecular level in the debate over whether—or when—to sequence the human genome.

The question arose, again, during *Nature*'s recent conference, "Exploring the Human Genome."

Harvard's Walter Gilbert argued at length that sequencing should begin *now*. With current technology adding new sequence information at the rate of 10,000–100,000 bases per research team per year, Gilbert estimated it would take 100 team-years to map the genome, 3,000–10,000 team-years to sequence it, followed by about a million years of communal cogitation to interpret the results.

Gilbert predicted that new instruments and new techniques could raise the new-data flux to a million bases per team per year. At that rate, 150 researchers could sequence all 3 billion bases of the human genome in twenty years. With support staff, that force grows to 300 people—the size of an institute that might cost $30 million a year to operate, about one half of one percent of the National Institutes of Health's current budget.

Still more advanced equipment (see *Bio/Technology* **4:**890, Oct. '86), perhaps using Gilbert's own technique of genomic sequencing, could boost that rate as high as 30,000 bases per researcher-day—roughly 10 million bases per research team each year. At that rate, the whole genome is within reach.

There are other barriers, of course—technical, financial, and philosophical. Otherwise we would have not a debate but a bandwagon. Accuracy is still a problem. The most recent reports on automated gene sequencers claim error rates of about 1 percent, about 10,000-fold too high to develop meaningful data. And there are, as Gilbert pointed out, unclonable sequences and filling in those gaps could be difficult.

Some have asked, "Whose genome?" It doesn't matter, as long as the donor provides a normal karyotype, viable cells, and a large supply of starting material. Others have objected to the project on the grounds that the great mass of non-coding DNA is of no scientific interest. Do we really know that?

If the project is begun, there could be unpleasant consequences, of course. The worst would be a repeat of what NASA's shuttle program did to funding for other space research—dried it up. It is possible, though, that funds for human genomic sequencing could come from new federal sources, where the impact, if any, would be felt by researchers outside the life sciences. (How is that for parochial?)

As the known sequences grow, wet-lab researchers could well wind up spending far more time at keyboards, constructing probes and expressing proteins *in electro*. And it might conceivably get harder for researchers to win grants for independent sequencing projects—even though the final consolidated sequence should properly be used only as a reference.

Still, the potential benefits outweigh the drawbacks. The sequencing project should start now. Begin with amassing human DNA libraries (*Bio/Technology* **4:**537, June '86), making maps, and developing the instruments for fast, accurate, automatic sequencing. In a few years (5–10, says CalTech's Leroy Hood), begin the sequencing in earnest.

Some three millennia after the Delphic Oracle first admonished pilgrims to know themselves, Alexander Pope took up the theme again in his *Essay on Man*. "Know then thyself, presume not God to scan, / The proper study of Mankind is Man, / ... The glory, jest, and riddle of the world." The glory and jest will have to remain. Some of the riddle can be unraveled.

—Douglas McCormick