

## **The legacy of the human genome project: lessons and implications**

The historic announcement of the draft sequence of the human genome calls attention once again to a project that has been ongoing for almost fifteen years. What are the implications of this project for the average scientist and research program? Has it been worth cost?

There are many ways that this, the first national and international project in biology and medicine, has had an effect on our science, but some are more subtle than others. We believe that even with the overheated media coverage of the past month, it is difficult to overestimate the ultimate effect. The announcement symbolically marked nothing less than the transition of biology from its adolescence to a mature science--a science that will provide the foundation for much of 21<sup>st</sup> century medicine. The scientific community knows well that this raw information provides a beginning to the science, not the solution to the problem of understanding how it all works. It knows as well that the genetic information is not the single determinative factor in organisms like ourselves, but the community has an obligation to bring this understanding to the public. Public statements from scientists seeking to deflate the importance of the genome sequence in the interest of countering ideas of genetic determinism serve us and the public poorly .

Aside from its scientific importance, the Human Genome Project (HGP) itself, as a social enterprise, holds important lessons for science policy. There are decisions that will soon be thrust upon us in the conduct of affairs between nations, especially between the wealthy and the impoverished, that derive in part from the HGP. Indeed, the way we respond to the choices will tell us more about ourselves than the long stretch of chemical units that underlie our individual biology.

The Human Genome Project has changed the culture and conduct of biomedical research in the United States. Early attempts by a few prominent leaders in the biological community--notably Robert Sinsheimer, Leroy Hood and Walter Gilbert--to initiate a genome project were frustrated largely because genomics could not exist within the biological research paradigm of the time. Biomedical research was overwhelmingly hypothesis driven and reductionistic in a narrow sense, with scientists studying primarily one protein or one gene at a time. The genome project, on the other hand, was to focus on the development of large-scale resources--enabling technologies, including maps of the human genome, that would allow information about thousands of genes to be accumulated simultaneously. This shift away from a very narrow view of what was worthwhile in biology, was driven by the realization that biology, for all its successes was floundering in ignorance of the basic elements of its science, the encoded sequences of its structural, enzymatic and control elements in the genome. What science could mature without this kind of fundamental information resource? Along with the basic information of the genome sequence the project has delivered a substantial enrichment of the technological tools for the biological sciences. We are only beginning, for example, to realize the benefits that came from the funding of a wide range of projects developing

computational methods for the manipulation and analysis of biological information. It was arguably the early struggles of the genome project to find ways to deal with the anticipated map and sequence data that seeded an explosion of activity in bioinformatics and computational biology that is now beginning to be harvested. Ten years ago few of us would have predicted, for example, that computational algorithms could find genes in stretches of genomic sequence. Databases and the attendant software, structural annotation of coding sequences, and a wide range of other needs for the genome project led to support for work that has stimulated and enriched the field immensely.

The technical fallout of the genome project is difficult to assess because it was certainly not the only force, or the only funding source, for technological development in biology, and the full impact of these technologies has not yet been felt. The difference the project made, however, was evident in the study sections and review panels over the past decade. Proposals for innovative technology development turned again and again to the genome program for support it could find nowhere else. Among the areas strongly affected was that of automation of biological laboratory techniques, an area essential to the genome centers, and rather poorly addressed by the technology of the 1980's. New methods and applications were developed, but major contributions came from the genome-driven adaptation of automation methods from other fields.

The practical needs of the Project eventually elevated the development of technology for research to the level of a respectable endeavor in itself, which was a very new attitude in the era of the origin of the project, the early 1980's. Because the Project had to develop outside the accepted biomedical paradigm, it is not surprising, at least in retrospect, that the idea found a fallow field at the NIH, but relatively fertile soil at the Department of Energy (DOE). DOE is devoted primarily to funding physics and engineering. It supports, for example, more than 90% of the high-energy physics research in the US. In high-energy physics, collaboration between engineers and scientists with varying backgrounds to build large, complex and necessarily expensive enabling resources has been common for decades. The history and culture of the science programs of the Department were reflected from the very beginning not just in receptivity to the idea of producing an enabling resource that could have an unprecedented effect on biology, but in the style of the earliest human genome proposal. While news media were reporting scientific conversations about full-scale sequencing, DOE formulated a plan in 1986 to sequence the genome as a two-phase, 15-year project beginning with the development of instrumentation, robotics and mapping. Only afterward would large-scale sequencing begin. Eventually this was the strategy carried forth by NIH and DOE and their international collaborators. This history illustrates the central importance of diversity among funding agencies. Just as vigor increases with diversity in the world of biological ecology, so it appears to do the same in the world of organizational ecology.

The Genome Project from its beginning has been a model for inter-agency and international cooperation. Early discussions with Senator Domenici were crucial to developing a consensus between the Administration and the Congress on the importance of an initiative that initially appeared to be too costly and too far beyond the boundaries of mainstream biomedicine to sustain. The discussions nevertheless encouraged the

evolution of a plan and a Project that has been a spectacular and virtually unprecedented success in the scope of collaboration between scientists in different nations, and between federal funding agencies. Success required that the biomedical leadership adapt to a new culture; a new style of approaching science. The dexterity of their response, and alacrity with which the computer science community-- which was fully foreign to biological science prior to 1985-- responded to the information and mathematical challenges, is one of the most remarkable stories in the sociology of science.

Nor could the Project have been completed as efficiently as it was without the rapid adaptation of the National Institutes of Health, and especially the leadership provided by James Watson and Francis Collins, and the managerial and diplomatic skills of Ari Patrinos, the current Director of DOE's biological and environmental research programs. The history here illustrates that cooperation in complex social enterprises, even between organizations with very different agendas and cultures, can take hold if the goals are important enough. It underlines the importance of the multi-agency approach to funding science. It also illustrates, somewhat surprisingly, the rapidity with which paradigm shifts can occur.

One of the Department of Energy's goals from the very outset of the Project in 1986 was stimulation of the national economy, which we expected to occur with the growth of a fledgling biotechnology sector. This growth appears to be happening at rate exceeding even our most optimistic expectations. The emergence of Craig Venter and other talented scientist/entrepreneurs, and the enormous contribution of Celera to accelerating the final phase of the Project, is a remarkable example of two sectors of the research community -- private and public--reinforcing one another for the common good in spite of competition and some conflicts. The private sector will now continue to play a major role in the realization of the benefits of the genome project.

The Genome Project has been important in another way, in recognizing that the effect of a major new technological shift on society is complex; that it irreversibly changes the social fabric, and that the change must be controlled if the technology is to fulfill its promise. For this reason DOE initiated a policy in 1987 to set aside 3% of the genome budget to study the social ramifications of genomics, including civil liberties, issues of genetic privacy and discrimination, and national and international challenges posed by genetic engineering. The funding level was subsequently raised to 5% by the NIH and the DOE followed suit. Although many of the ethical dilemmas raised by modern biological technology have, in one form or another, been with us for millennia, modern technology has quickened the tick of the clock, making solutions extremely urgent. It is society at large--the scientific community, our political leaders, and the lay public -- which must now act to bring about the changes necessary to solve these dilemmas.

The completion of the Genome Project is biology's graduation. It signifies the completion of an important enterprise, while symbolizing the beginning of a new era. It marks the beginning of the time when our understanding of biology will increase at a rate unimagined a mere 15 years ago. The new knowledge that we acquire will provide the

foundation for a medical revolution rivaling in importance any technological revolution in the history of the human race.

Although this view of things to come is surely not hyperbolic, it is nevertheless important to add that when we make such strong predictions, we must not end with just the promise of a new genesis. Rather we must recognize that with the current appalling economic stratification of the human race, the vast majority of humanity is constrained to remain outside the reach of the benefits of this kind of scientific project. It need not be so. We must recognize that the greatest challenge of the 21<sup>st</sup> Century is not a scientific or technological challenge: it is the social challenge of bringing to those who need it most the fruits of the inevitable technological revolution. The intricacy of this problem is much greater than the intricacy of the genome, and its solution will require far more cooperation than did the Human Genome Project. But the success of the genome project offers us hope in this field too. It provides an example of the rapidity with which deep-rooted assumptions can change, and the success of an enterprise that many first-rate scientists thought was preposterous 15 years ago. It is not too early for national leaders to bring together the best minds from all disciplines to begin formulating a plan to assure that genomics makes real contact with humanity. We believe a new and extraordinary chapter in our history is just beyond the horizon, stimulated by modern technology, and symbolized by biology's graduation. The childhood of the human race is about to end.

\*\*\*\*\*

Note on the authors:

**Charles DeLisi** is Arthur G. B. Metcalf Professor of Science and Engineering at Boston University and Dean of the College of Engineering. **David Galas** is Vice President and Dean of the Faculty at the Keck Graduate Institute in Claremont, CA, a new college dedicated to the application of the life sciences. Dr DeLisi was Director of the DOE biological and environmental research programs from 1985-1987; Dr Galas was Director from 1990-1993.