

Table of Contents

Point-and-Click Biology (Why Programming is the Future of Biotech) 1

Point-and-Click Biology (Why Programming is the Future of Biotech)

By Eugene Thacker

Should we play God?" "Is Dolly a monster or a miracle?" ask popular discussions of biotechnology. But what these banal formulations fail to understand is that the current phase of biotech research is characterised by a massive move towards computerisation, with the boundary between wetware and software becoming blurred. Eugene Thacker calls for a more accurate framing of the whole debate

[IMAGE]

>> Human Genome Project

THE HIGH SPEED DATA DUMP

The race to map the human genome — some estimated 80,000 genes — will quite probably go down in histories of science as the moment when science and commerce, infected by the magnetic force of information technology, could no longer stand to remain apart. The word 'biotechnology' has always been associated with start-up companies. And this association has been fraught from the very outset, from at least the late 1970s, when the first recombinant DNA patents were filed and the first biotech start-ups started up. The science community has been polarised ever since. Some wished to preserve the more altruistic, public face of medical research, and viewed biotech companies as a degrading intrusion into the domains of scientific knowledge. Others not only found in the corporate model a means of speeding up the timetable for clinical trials and concrete results, but also argued that generating the billions of dollars required for macro-scale projects takes a hefty amount of finance capital. The latter group, needless to say, have won, and as a result the genome project has become a fascinating, monstrous hybrid of new lab technologies, flows of currency, mergers between biotech companies, push-pull relationships with governments and wildly dystopic media reports topped only by the extreme science fiction visions of scientific journals and press releases. What the scientific community once thought of as a discipline of pure scientific inquiry and medical application has mutated into a high-speed database upload/download.

A key moment in this shift occurred in 1998, when Craig Venter, a former National Institute of Health (NIH) researcher and self-styled 'maverick scientist', announced that his new company, Celera Genomics, would sequence the entire human genome for less money and in less time than the government-funded Human Genome Project (HGP). Venter's statement sent a shockwave through the scientific community, horrifying some (like James Watson, co-discoverer of the structure of DNA, who resigned from his post at the HGP) and exciting others (such as Perkin-Elmer, who stood to gain a great deal by selling its proprietary gene sequencing computers to Celera). Celera's challenge to the HGP started a high-tech race whose stakes were nothing less than the data-mapping and informatic colonialism of our own 'human software'. It also begged big questions: how far should the government intervene in commercial ventures? What if those commercial ventures are also medical ventures? Where do researchers, CEOs and government officials draw the line between pure economic interest and pure medical/health interests?

The relationship between the HGP and Celera can best be described as like that between a responsible but sluggish parent and its brilliant but delinquent child. Celera is, of course, a business — a company in the business of genetic mapping, a field known as genomics. In this it manages private databases of genome information (it has done the *Drosophila* fly genome and is working on others), which are only available to paying subscribers. With its array of automated, around-the-clock sequencing computers,

it has developed a controversial method known as ‘shotgun sequencing’, in which the entire genome is randomly chopped up into small pieces, sequenced, then reassembled. It raises money not through government institutions, but through investors willing to risk putting down capital on (yet another) biotech company promising the key to the future. Celera must also, like many biotech companies, ensure that it covers its costs by filing patents (on gene-related products) and providing for a range of services (such as selling subscriptions to its genomic database). And finally, Celera also forms alliances with businesses providing technology development, as well as strategic business agreements with pharmaceutical corporations (or ‘Big Pharma’) interested in genetic drug development. It is this outright interest in corporate-economic models, along with the adoption of high-speed, cost-effective, product-oriented technologies, that has made Celera the *enfant terrible* of modern genetics.

While this is going on, public debate shows a lack of understanding of the key issues involved. The common, banal discussions about the ethics of the human genome project end up painting themselves into a corner: are you pro-science or pro-nature? Should genetic information be made ‘public’ or ‘private’? Do scientists have the right to ‘play God’? What many of these positions miss are the ways in which a given argument gets structured the way it does: that is, not the ‘pro’ or ‘contra’ but rather how the entire debate — the ‘vs.’ — gets articulated in the first place. We all know that multiple-choice surveys are formatted a certain way to elicit a certain range of responses. If we are unable to question the very structure of the discourses and practices — what Bruno Latour calls ‘actor networks’ — then effective discussion and intervention will never come about. Despite the involvement of government agencies such as the U.S. NBAC (National Bioethics Advisory Commission), the terms of the debate on biotech are largely being set by the biotech companies, in their tensions with the government-funded researchers.

Genetics itself is about mutation, permutations within sequences of coding that can occur ‘naturally’ or through artificial means. A similar set of permutations, I believe, are called for within the practise and discussion of the field itself: the letters of the genetic code, the human source code, need to be first read, then analysed, and then re-assembled and re-inserted into the discourses and practices of the biotech industry. Below is an attempt at beginning to do the difficult work of critically thinking about and acting on the issues presented to us by biotech and genetics research.

TRADE SECRETS

We should start by stating some of the untold secrets of biotech and genetics.

Firstly, out of all of the information in the human DNA in each cell, researchers estimate that only 3% of it is ‘useful’. That’s right, folks: according to researchers in genomics, only that much of your total DNA is ‘coding’ genetic material that corresponds to the production of some amino acid or other biochemical molecule. That means that roughly 97% of human DNA is what is called ‘junk DNA’. When maps of the human genome are sequenced, they are, of course, focusing on that 3%, with the remaining 97% as a kind of afterthought. In an amazingly complex molecular network, intricately enough designed to correlate biochemical events that enable, say, metabolism or muscle development or regeneration or memory to occur, it may seem difficult to accept that this network simply has untold amounts of trash which has never been taken out. Several research teams have, however, offered alternative views, by exploring the idea that the junk DNA is involved in controlling gene function and expression (not the sequence but what the sequence does), or that junk DNA comprises a vast resource of residual evolutionary DNA (which is why researchers often use the mouse genome as a comparison, due to its large genetic similarities to ours).

Secondly, despite the seductive metaphors of the ‘Book of Life’ and the promised ability to be able to ‘read’ an individual through their code, the completed human genome map may very well turn out to mean nothing at all. As William Haseltine, CEO of Human Genome Sciences, has stated, the biggest secret of the human genome project is that the sequence itself is of no value; it’s what the sequence *does* when operating in the living cell that counts (which is why Human Genome Sciences has an impressive number of gene-related patents in its arsenal). The popular notion that an incredibly long iterative sequence of four bio-molecules will provide existential answers is way wide of the mark. However, the notion that this permutation-machine is in some way involved in a complex network of biochemical processes is more plausible, and more articulate. It seems naive for us to look to DNA for religious, philosophical, or ethical answers, though the fact that this is often the case (consider the recent debates around the genes for homosexuality, for criminality, for obesity) is more indicative of how much we desire DNA to be all those things, and for science to play the role of mediator, exegete, high priest.

Thirdly, and most importantly, the current phase of biotech research is characterised by a massive move towards computerisation. Areas such as bioinformatics and biological computing (the use of DNA to accomplish computational tasks) are now governed by information management practises, and the biotech lab looks more and more like a computer lab, with emphasis being placed on such tools as DNA chips, microfluidics stations and genetic analysis software. But while this shift may bump biotech up a level on the info-tech ladder, what goes unmentioned is the increasing fusion of the biological and the technological, genetic data and computer data. Is DNA like a computer? Does DNA operate like a computer? Or are our attempts to model ‘Artificial Life’, neural networks or genetic algorithms like DNA? When we speak about a genetic ‘code’, do we actually mean something that can be uploaded into an online database? Celera seems to think so, as does the HGP, which updates its public database daily. A further question arises here: when computer code touches genetic code, when does that touch the individual patient?

SAVE AS... HUMAN SOURCE CODE

This last, vital development is not altogether new. At least since the discovery of the structure of DNA by James Watson and Francis Crick in 1953, modern genetics has been saturated within a rhetoric emphasising textuality and information — ‘the book of life’, ‘the code of codes’ and so forth. In the late 1940s, physicist Erwin Schrodinger, in a lecture entitled ‘What is Life?’, suggested that the development of genetic science had made possible a new understanding of biological life in terms of a ‘code script,’ in which the whole of the individual might be found encapsulated in the nucleus of the cell. This reference to information was no accident: the complex roles which mainframe computers played in the war, as well as their increasing dissemination in business and industry, made for a fertile discourse on computerisation and society. It was a few years later, in the post-war era, that mathematicians such as Norbert Wiener and Claude Shannon separately began to develop theories of information. Wiener was primarily concerned with information as related to systemic regulation (through feedback loops), as well as the larger social implications of cybernetics and the grouping of humans and machines under the categories of information and negative feedback. Shannon, working at Bell Labs, was primarily concerned with a technical question: how to transmit a message at point A to point B with the greatest accuracy and lowest probability of error. More than one historian of technology has suggested that Shannon’s research significantly contributed to telecommunications advances and even the Internet itself. A key component to these theories of information was that the technical question (a quantity of information) took precedence over signification (a quality or ‘meaning’ of information). As long as the data arrived intact, that was all that mattered.

The implications of this classical version of information theory — as well as for molecular genetics — is that, as the author N. Katherine Hayles points out, "information begins to lose its body". As Hayles suggests, it is when information is equated with disembodiment that a critical intervention is needed in order to point out the embodied contingencies of information and the subjects immersed in information. As a fellow thinker on the subject, Richard Lewontin, has stated, "the final catalogue of 'the' human DNA sequence will be a mosaic of some hypothetical average person corresponding to no one."

It is within genetics that this perspective on information is translated into a form of genetic reductivism, whereby two assumptions are made. The first is that the organism, the body, the subject, is essentially its genetic code. That is, researchers — and, assumedly, physicians in medical genetics — will first and foremost concentrate on the code of the individual/patient as a means of treating disease. The second is that, with an array of new computer and networking technologies, genetic code is treated as computer code. This is increasingly true as developments in computer software and simulation programs begin to blur the line between wetware and software. And when the individual is equated with a disembodied, abstract patterning system, genetic reductivism effaces all extraneous elements in favour of an economy of the database. This is already occurring in the selected uses of genetic screening, disease profiling and various gene therapy techniques: medical genetics will become a matter of sending Java applets or cookies embedded into chromosomes to monitor trace routes.

DATA IS DATA, BUT NOT ALL DATA IS EQUAL

Given this propensity to approach biology through info-tech — and as info-tech — we can now consider several ways in which human genome projects act as a form of database management.

The problem of 'genetic difference' has been an issue with the HGP since its beginnings. Human genome mapping projects currently derive a single genetic map by way of statistical averaging (Celera has utilised the DNA from an unnamed male individual, combined with six other individuals of varying ethnic backgrounds). Such tactics determine which 'letter' or which gene will be entered into the database at a given locus. Given the claims by genetics researchers that each individual differs from others by less than 1% of its DNA, such an averaging seems to be a sensible, practical way of working. However, given that there are at least some 80-100,000 genes and over 3 billion base pairs ('letters') in the human genome, and that a number of known genetic disorders (such as Parkinson's or Sickle Cell) are in part triggered by a single base pair mutation (a single letter reads wrong), such homogenising practices take on a different tone.

At first glance, this massive homogenisation (and genetic simulation) of the human genome seemed to be corrected by the Human Genome Diversity Project (HGDP) when it was initiated in the early 1990s. The HGDP's contribution to the human genome mapping project would be to consider the ethnically-based genetic differences from a range of 'other' cultures, especially those cultures without dense histories of trans-ethnic genetic migration. Its job would basically be to account for the excess genetic material not within the euro-centric field of consideration of the HGP: the HGP aims above all for universality, while the HGDP concentrates on the local. Recently, though, the methods and practices of the HGDP have come under fire, primarily in the course of debates surrounding the patenting of genes and cell lines from indigenous cultures. Certain practices of the HGDP thus not only constitute literal examples of bio-colonialism (many of the cell lines representing indigenous cultures are collected without proper consent and intended to be archived — as a type of genetic museum — within government health organisations), but they also raise questions about the ways in which ethnic-genetic marginalisation is translated into a dominant system involving science, politics and governmental-economic structures.

As recent debates over Intellectual Property Rights and the patenting of genes and cell lines indicate, genetics and genomic mapping so easily slip away from bioethics simply because the modernist, humanist subject is nowhere to be found in their discourses. Genetics and biotech do not speak the language of the 'possessive individual', whose inalienable corporeal rights have been violated by patents on biological material. Rather, embodied and culturally-located subjects are reinterpreted as particular informational patterns within a statistical field which precedes the subject (e.g. the genomic map of a New Guinea tribe, of 'the' African-American, of 'the' Asian, of 'the' homosexual etc.). Genomic profiling of a population can dangerously slide into cultural-ethnic profiling as well.

Epistemologically, the HGDP ends up reduplicating the universalist, statistical assumptions of the HGP. While it considers cultural and ethnic specificity, it is too tightly bound to genetic determinism to be anything but a messy conflation of the discourse of nature and the biology of ethnicity. Here genetic difference doesn't simply account for cultural-ethnic specificity; it demands it. The HGDP, which has taken a lot of flack in the past, has been relegated to a more low-key position in its labs at Stanford University. But this does not mean the problem has faded away; the very disappearance of the HGDP in the public scene only means that its techniques and approaches to bio-colonialism are spread throughout biotech generally, whether in the external form of genetically isolated populations, or in the internal form of genome diversity and population health informatics. What biotech adds is a product-driven and service-oriented interface to the data-gathering begun by the HGDP. In producing a series of ethnic gene maps, projects such as the HGDP intimately forge a bond between genetic difference and ethnic-cultural difference. When considered alongside the genomic mapping projects, one sees not only the development of a genetic technology of normativity, but also the situating of ethnic-cultural excess in relation to the human genome: that is, mapped out as 'otherness.'

ERROR 404: NOT FOUND

Now there is a tendency, of course, to view biotech's increasing engagement with computer science and info-tech as yet another dystopian scenario, in which the mega-corporation descends upon the victimised 'people' and through some technology zaps the humanity out of them — so we'd become something like genetic zombies. Certainly there is some truth to such narratives, especially when we consider the veritable monopoly which the private sector, and mostly U.S. and European governments, all have on the genome that is supposed to represent the entire human race.

But a critical perspective needs to also steer away from the fatal mistake of demonising technology in the process. To be sure, all technologies are enframed by social, political, and economic elements, but by the same token critiquing biotech corporations need not mean doing away with information technology altogether. Biotech's investment in info-tech is not about disembodiment or the liquidation of the body, the 'human'. It is more subtle, and more perverse, precisely because those in the biotech industry do not notice it. Biotech is all about a series of boundary-blurrings, which are dictated not by intention, but by the combined network effects of info-tech, the economy, distributed research projects and the intersections of governments, corporations and laboratories. We should not fear disembodiment, but rather a particular *type* of embodiment. Biotech is not about turning the human into pure data or a stream of media images; biotech is about compelling individuals to become a genome, and through that, being able to monitor the population through database management. Genomics is a powerful discipline, because it enables biotech to have it both ways: you can approach the individual on the basis of information and code, and you can also use that code to literally synthesise and modify the biological body.

What is at stake here is not the human, not the body, and not the threat of technology. What is at stake is the negotiations over a future vision of the body, medicine, and definitions of 'life' itself..

Eugene Thacker <maldoror AT eden.rutgers.edu>

Biotechnology Industry Organisation (BIO): <http://www.bio.org>.

Celera Genomics Corporation: <http://www.celera.com>

Human Genome Diversity Project: <http://www.stanford.edu/group/morrinst>

Human Genome Project (DoE Division): http://www.er.doe.gov/production/ober/hug_top.html Human

Genome Project (NIH Division): http://www.ornl.gov/TechResources/Human_Genome/home.html

National Bioethics Advisory Commission (U.S.): <http://bioethics.gov>

The Biotech Bandwagon:

who got on when, and where they sit.

In the early Fall of 1999, Celera Genomic Corporation — the first privatised biotech company to initiate its own human genome project — announced that it had delivered its first major database instalment of human genome data to its pharmaceutical subscribers.

In December of 1999, representatives from the Human Genome Project — now funded by the US National Institute of Health (NIH), Department of Energy (DoE), and Britain's Wellcome Trust — announced that they had completed a complete map of the first total human chromosome, number 22.

In a State of the Union address on 'the responsibilities of science and technology,' U.S. President Clinton announced his public support of biotechnology, and inaugurated the new millennium by declaring January 'National Biotechnology Month'.

In anxious expectation of the completion of the human genome map, biotech stocks suddenly experience an unprecedented increase during the early Spring of 2000; financial media sources from CBS Marketwatch to Bloomberg christen biotechnology as 'the new dotcom'.

At around the same time, however, President Clinton and Prime Minister Tony Blair, in a joint statement, highlight the importance of human genome data being "publicly available to all". In a matter of nanoseconds, biotech stocks fall sharply, hitting even biotech companies not directly involved in genomics.

At the biotech convention BIO 2000, organised by the Biotechnology Industry Organization (BIO), groups of public protestors gather outside the convention building, holding grotesque 'Frankenfood' sculptures and donning monster-masks. Protestors consider dressing up as genetic mutants and chasing the convention attendees on their morning jogs, but were unable to circumvent security.

Just this past April, Celera announced that it had completed the sequencing of one entire human genome, and that it expected to be finished with the assembly phase at the end of the summer. In reply, spokespeople for the government-funded Human Genome Project expressed both scepticism at Celera's methods, and urged caution in the use of language suggesting anything 'finished'.