

Comment

KNOW YOUR GENES. THE MARKETING OF DIRECT-TO-CONSUMER GENETIC TESTING

Genomics' problem of communication

Jenny Reardon

ABSTRACT: Since opening their doors in late 2006, personal genomics (PG) companies have faced skepticism and criticism from influential academic and government circles. While this has posed a clear problem of communication for these companies — one of effective promotion — I argue that the communication problem these companies face runs much deeper. It is a problem that lies at the heart of any genomics: the very understanding of communication and information around which genomics is built. While the value of genomic information for persons has been widely questioned, questions about the very notion of information that undergirds the production of genomic information rarely, if ever, has been broached. I suggest that making significant inroads into the vexing debates about PG would be greatly aided by addressing these more fundamental questions about the nature of information, and its genomic qualities.

In recent years, the production of genomic information has become fast and cheap, so fast and cheap that many now argue that it rightfully belongs in the hands and web browsers of the average person. Yet in practice, while widely celebrated in the press, the personal genome revolution has barely taken off. Reportedly, DNA Direct only ever had several hundred direct-to-consumer customers.¹ 23andMe, with all of its Google hipness and financial backing, musters 75,000-100,000 customers, a far cry from personal information giant Facebook's 750 million going on 1 billion users.² Founders of 23andMe cite entrenched practices coming out of human genetics that overly restrict the flow of genomic information, and position human geneticists as gatekeepers.³ Advocates for the Personal Genome Project fault a paternalistic government that casts people as too vulnerable to access powerful genetic information themselves.⁴ Both efforts figure personal genomics (PG) as a democratic cause: millions, even billions, might rightly access and manage the vast stores of personal information contained in their genomes if only the power of experts and the government could be broken, and rights of people recognized.

Such an argument taps into a deeply held belief in liberal democracies that information is a central good needed for rational decision-making. Perhaps nothing makes the importance of information in contemporary democratic (and economic!) orders more evident than Google. As the information behemoth states on its website, its goal "is to facilitate access to information for the entire world."⁵ Arguably, one reason for Google's stratospheric success has been its ability to position itself as providing this central democratic good, and its ability to cast its practices as "democratic" (i.e., "voting" with page clicks, open software development, etc.) It is to this democratic imaginary and corporate success of information that personal genomics harnessed itself when 23andMe opened its doors in the fall of 2007.

Financially and personally backed by Google, 23andMe would be the first PG company to open (along with deCODEme). Google co-founder and billionaire Sergey Brin and Google co-founder Linda Avey met through Brin's interest in Parkinson's. Avey sought faster progress in genomics. Brin also hoped for faster breakthroughs that might lead to treatments for Parkinson's disease, a disease associated with the LRRK2 genetic variant, a variant contained within his own genome. Powered by Google's money and algorithms, Avey and Brin's wife Anne Wojcicki created a company that sought to achieve these goals through turning human genetics on its head.⁶ Why assume that genetic information is dangerous and harmful, and that people should be protected from it? This, after all, led to a dreadful state of affairs in which geneticist overlords "protected" their research subjects through not giving them access to their genetic information, and not communicating with them.^{7,8} Why not assume instead, as Google did, that information is good; that information tailored to individual needs and wants is even better; that sharing this personal information is better still?

This narrative is compelling, but to date communicating it to a concerned Congress, key thought leaders in the social and life sciences, and persons with genomes more broadly has proved anything but easy. Despite the early fanfare around the sequencing of the human genome, and promises that it would reveal “the secrets of life,” in the last few years even prominent genome scientists have questioned the medical value of genomic information.^{9,10,11} As illustrated by last summer’s Congressional hearings on direct-to-consumer tests, many in Congress also fail to see the value of genomic information for consumers. Even worse, some believe that those attempting to sell it through personal genome services were the modern equivalent of “snake oil” salesman.¹² This resistance from influential academic and government circles clearly poses a significant communication challenge for PG companies: one of effectively communicating with and convincing important actors.

However, the communication problem personal genomics confronts runs much deeper than this one of promotion. Rather than a problem posed specifically by personal genomics, it is one that lies at the heart of any genomics: the very understanding of communication and information around which genomics is built. While the value of genomic information for persons has been widely questioned, questions about the very notion of information that undergirds the production of genomic information rarely, if ever, has been broached. What is information? Specifically, what notion of information informs/animates genomics? Is this the sort of information that is meaningful for the contemporary constitution of democratic societies and biological knowledge? In what sense is any of it “personal”? I suggest that making significant inroads into the vexing debates about PG would be greatly aided by addressing these more fundamental questions about the nature of information, and its genomic qualities.

To ask and answer these questions — to understand, as communication theorist Tiziana Terranova characterized it, “the informational quality that defines 21st century” — we require a new mode of reading that trains our attention not only on dramatic contemporary political and legal struggles over the meaning of information, but also the less spectacular specific material form information takes as it is imagined and enacted by physicists, mathematicians and engineers.¹³ Like Terranova, I begin by returning to Claude Shannon, an American mathematician and “father of information theory.” As Shannon explained in his code-breaking 1948 article, “A Mathematical Theory of Information”:

*The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. ... Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem.*¹⁴

In this formulation, information is a message (or signal). The problem of communication is how to relay this message in a manner that minimizes its distortion due to noise. Importantly, this problem does not address the problem of what the message means. The task is merely to encode the signal in such a way that secures its accurate reproduction. In this formulation, information is simply that which is not noise; meaning is irrelevant to the communication process, which is about preserving the integrity of the information.

Terranova argues that the “modern concept of information” has been subordinated to this “technical” understanding of information employed by “communication engineers”.¹⁵ My question is to what extent this technical approach to communication and information animates and shapes genomics?

Popular accounts of genomics offered by leaders in the field suggest an easy dismissal of any connection between genomics and this command and control approach to communication, an approach that divorces information from meaning. Most genome scientists state a strong interest in not just encoding, but decoding life. J. Craig Venter, head of the private effort to sequence the human genome, announces this interest in the title of his autobiography, *A Life Decoded*. Inside the covers he refers to “the human genome” as “the book of life” that reveals “the secrets of our inheritance”.¹⁶ Throughout, he uses a reading of his own genome to make sense of everything from why he has asthma to why he prefers to work at night.¹⁷ Francis Collins, head of the public effort to sequence the human genome, promises something arguably even more profound in his book, *The Language of God: A Scientist Presents Evidence for Belief*.¹⁸ Collins, paraphrasing former U.S. President Bill Clinton, describes the human genome as “God’s instruction book”.¹⁹

However, a closer read of these accounts of genomics, combined with an attention to genome scientists’ everyday practices, reveals a much more complicated picture of genomic information, and its relation to the realm of meaning. Consider that to render life loquacious — to make it speak its As, Cs, Gs and Ts — genome scientists inscribe it in silico. This was the secret to Craig Venter’s success. He abandoned biology’s analog approach to sequencing — the Sanger technique — for a digital one. The former entailed

what Collins described as the “arduous task” of pouring agarose gels between two glass plates, hand-preparing radioactive DNA samples, hand-loading the gel, laying the resulting gel on film, and reading the resulting X-ray induced images. Both Collins and Venter argue, and I concur based on my own experience pouring and reading these gels, that this process is prone to many problems, including that of human interpretation: “[M]any times the lanes would not run parallel to one another, so the farther one went down the gel, the less able one was to compare the black marks and gaps on each track to read the sequence”.²⁰ Is that a T or a C, and A or a G? Such questions often occupied the minds of those who spent hours pouring over films of gels, trying to record the correct sequence of As, Cs, Gs and Ts. Venter reported great frustration with these questions about what hash marks on a gel meant:

I found the room for interpretation particularly frustrating because I had such high hopes for molecular biology. Too many times I had seen science driven less by data and more by the force of a particular personality or the story on which a professor had built his career. I wanted the real, empirical facts of life, not those filtered through the eyes of anyone else (Ibid.).

For this, he sought to abandon humans for machines, the analog for the digital. He would be the first to successfully implement “automatic” sequencing on a large scale. Describing the ABI sequencing machines that made this automatism possible, Venter writes: “The four colors, representing the different nucleotides, provided a *direct readout* of the genetic code, transforming the analog world of biology into the digital world of the microchip” [italics added].²¹ There would be no more human-filtered readings of DNA. Machines would provide a direct readout of base pairs and encode them as clear 0s and 1s.²²

The public effort also quickly moved toward automation. As in other realms of human inquiry, this move took biology out of the data desert of the analog age and into the data deluge of the digital one. Production reduced from armies of humans pouring and processing large unwieldy gels to machines shining light beams on microcapillaries, and recording results on silicon chips. As the man credited with inventing the term *digerati*, George Gilder, argued, and Chris Anderson, editor of *Wired*, would repeat, the move to semiconductor machines greatly speeds up the production of information: “When matter plays so small a part in production [just the bit of sand needed to produce silicon], there is less material resistance to increased volume. Semiconductors represented the overthrow of material in the economy”.²³

We should, of course, resist Gilder’s too glib hyperbole — the move to semiconductors in no way represents an overthrow of material (this point will only become more important to recognize as the energy demands of data storage and processing become pressing institutional concerns). However, in many ways the move to semiconductors represented an overthrow of that fleshly material that we up until recently experienced as life. Increasingly one can play a central role in biological practice and never encounter the tissues of an organism.

For Venter, the digital also overthrew — or so he imagined — another kind of fleshly material: pesky persons, in particular the government bureaucrats who worked at the NIH, and who failed to support his research due to, on his account, petty politics. For him, genomics promised an escape from politics and deliverance into a space of pure science in which persons played an ever-decreasing role. Proudly, he describes how his automated machines “eliminated the need for thousands of workers”.²⁴ The few humans that remained worked in a “factory environment” where machines were predominant, and humans only ensured their continued function.

Leaving persons behind, Venter “rocket[ed] forward” into the genome age.²⁵ Soon, with the vast sequencing and computational facility he built at The Institute of Genome Research (TIGR) — what he described as “the beating heart of the organization” — he would be producing more genomic data than anyone else.²⁶ But once rocketed off mortal earth — plagued as it is by the slow and painful stuff of human interpretation — and into the stratosphere of data, did one also leave the land of meaning?

Throughout his memoir Venter repeatedly stresses his commitment to interpreting the meaning of genetic sequences. As a medic in Vietnam, and then a biochemist studying adrenaline receptors, Venter was interested in life and what made it tick. He describes this interest in interpretation as a key difference between him and then director of the Human Genome Project, James Watson:

Watson argued that our goal was to work out the sequence and let future generations of scientists worry about understanding it. I had always believed that interpretation was crucial to making the sequencing both efficient and worthwhile.²⁷

In practice, though, sequencing and speed would become paramount. On Venter's own account, he ultimately left the field of receptor biology to become a genome scientist, and while he kept his interest in the biological meaning of the genetic code, he realized that for the most part he "would have to fight that battle another day".²⁸

When others did try and work with Venter to fight that battle, it was never an easy task. The company created to make good on the investment in TIGR, Human Genome Sciences (HGS), faced significant problems making the data translate into something valuable for shareholders and patients.

*Paradoxically, the torrent of data generated by TIGR, which should have been an achievement to be celebrated, was the source of the problem: HGS was simply overwhelmed. ...Had I delivered them a single gene that was linked to a disease, they would have known how to mount a major discovery effort to turn the find into a test for new drugs. But I had given them thousands of genes over the course of a few months. ...HGS complained that to exploit the data 'was like trying to drink from a fire hose'.*²⁹

This problem of information overload is not unique to biology. Nor is this a new problem. In his 1981 book *The Political Unconscious*, the cultural theorist Frederick Jameson argued: "Unfortunately, no society has ever been quite so mystified in quite so many ways as our own, saturated as it is with messages and information, the very vehicle of mystification".³⁰ Jean Baudrillard, Mark Poster and other theorists of information have made similar observations, arguing that the increased speed and volume of informational language threatens the representational function of symbolic systems.^{31,32} As early as the mid-1990s, writers were coining terms such as data smog (which made it into the Oxford English Dictionary in 2004), terms invented to call attention to the problem of information overload, and the meaningless and potentially harmful environments this overload created.³³ These issues attracted renewed attention in late 2008 when *The Atlantic Monthly* published a piece by Nicholas Carr entitled, "Is Google Making Us Stupid?".³⁴

The question I pose is whether this crisis of meaning has now crept its way into biology, and if so, what we can do about it.

On some accounts, the outlook is bleak. As University of California, Berkeley professor Geoffrey Nunberg wrote in his *New York Times* book review of James Gleick's *The Information*: "There are no roads back from bits to meaning".³⁵ The technical units of information — 0s and 1s — just don't correspond to anything of social significance, or meaning. If this is the case, we might rightly begin to worry if Venter was right, and that it is now the machines and not humans who have the beating hearts. Put perhaps too starkly, might we have produced the life of information, and not the information of life?

While almost all seek the latter, signs that that the former might pertain increasingly are evident. Consider the claim 'Information wants to be free.' For Stuart Brand — who is attributed with coining the phrase in 1984 — the phrase referred to Moore's Law and the law-like manner in which the cost of producing information has to date decreased. In its contemporary usage, the phrase also carries strong moral overtones: not only is information nearly free to produce, it *ought to be* freely communicated. Today we expect information to be free in both senses of the term: without a price, and without control.

It is this understanding of information that animates genomics. Tales of the emergence of genomics frequently recount the bad old days in which human geneticists sought to control genetic data for their own good. In Venter's tale, human geneticists opposed automated sequencing and the free flow of genetic information it promised because they were "often more concerned that they win the race to discover a gene than that the race be finished as quickly as possible".³⁶ Avey and other genomic thought leaders describe the culture of human genetics as feudal, and genomics as a revolution that seeks to free up genetic information for the benefit of all.³⁷

Animated by this laudable and revolutionary goal, genome scientists today have become very good at freeing information, or at least at least information in the Claude Shannon sense. In an ever cheaper and more precise manner, they encode the As, Cs, Gs, and T's that inhere in the tissues of a mouse, human or worm and relay them so that they are faithfully reproduced on our computer screens. Through creating faster, cheaper sequencing machines, anonymizing data, and creating data commons, that information increasingly moves freely from computer to computer, cloud to cloud. As a result, genomic information is now a lively material reality. The genomics world is abuzz with discussion about where the data should live (and as ethereal as the cloud may sound, it is anything but; I prefer data hotel), how we can build computer powers big enough to process it, and how we will provide the electrical juice needed to keep all these data alive and well.³⁸ This is a long way away from what many had hoped we would be abuzz with at this point: medical breakthroughs; profound new understandings of life.

I am not saying that there has been nothing of biological significance, but only that the buzz is elsewhere — with the information.³⁹ I suggest that this is in no small measure because the notion of information at the heart of genomics does not address the problem of the meaning of life. Instead, the fundamental problem it addresses, and has made enormous strides towards solving, is noise — the problem of how to distinguish the signal of a nucleotide from a smear on a gel. This problem and this goal Venter and others have proven we can master. We can make information free. But that leaves us with a fundamental question: Does that free information have any value? As George Church, head of Harvard's Personal Genome Project, presciently observed a few years ago: "When costs come down to zero then people have to decide how to add value to it [the genome]."⁴⁰

Indeed. But who will those people be? After years of evacuating persons from the space of genomics, the idea of personal genomics poses a serious problem. What sort of person would want to live in the current meaning desert of genomics?⁴¹

Arguably, addressing this problem is what the field of personal genomics has been trying to do. While some may understandably feel anxiety about solving the problem through articulating biology to Google and Facebook cultures and practices, with or without these efforts, the problem remains: genomics spent decades addressing a communication problem that ignores the problem of meaning. As easy as it is to critique personal genomics for trivializing genomics through turning to flashy communication media, I suggest that in fact these companies are addressing a problem that is far from trivial. Genomics is built around a communication problem that provides few obvious routes back from bits to biology, from PCR machines to people. If genomics is to become "personal," this must change. In their devotion of funds to hire people to interpret large genetic datasets (in their lingo, to biocurate), their efforts to work together with lawmakers and regulatory agencies to decide when data should be considered accurate on the one hand, and meaningful for decision-making on the other, these companies have done more than probably anyone else to change this.

Rather than an indication that genomics is a space of the personal, perhaps the term personal genomics is better read as aspirational. After years of genomics privileging machines, it is undoubtedly true that if genomics is going to emerge as a human science, it must re-orient at this point to consider persons in all their meaning-loving fleshliness.

The first task, I suggest, is to adopt a different notion of information — one based in a theory of communication that presumes meaning, not noise, is central, and understanding, not freeing, is the goal.

Acknowledgments

I would like to thank Misha Angrist, Bob Cook-Deegan, Malia Fullerton and Kate O'Riordan for their thoughtful reading of this essay, and for their lively intellectual engagement. All errors and fuzziness of thought are mine alone.

Notes and references

- ¹ Partially because they recognized the direct-to-consumer genetics market to be large, the leadership at DNA Direct recently decided to cease taking on any new direct-to-consumer customers. Instead, the company has been bought by MedCo and is focusing on providing services for health providers and insurance companies.
- ² And of these 75,000 to 100,000 it is not clear how many purchased the test versus being offered the test by 23andMe for free.
- ³ L. Avey (2009), *The Life and Times of Lilly Mendel*, <http://lillymendel.blogspot.com/2009/11/inspiration-comes-in-many-forms-idea-to.html>, accessed April 2010.
- ⁴ M. Angrist (2010), Only Connect: Personal Genomics and the Future of American Medicine, *Mol Diagn Ther* **14**(2): 67-72.
- ⁵ <http://www.google.com/about/corporate/company/tenthings.html> (last accessed September 6, 2011).
- ⁶ *Wired* magazine's Thomas Goetz offers this description: "Brin proposes a different approach, one driven by computational muscle and staggeringly large data sets. It's a method that draws on his algorithmic sensibility—and Google's storied faith in computing power—with the aim of accelerating the pace and increasing the potential of scientific research. "Generally the pace of medical research is glacial compared to what I'm used to in the Internet," Brin says. "We could be looking lots of places and collecting lots of information. And if we see a pattern, that could lead somewhere." In other words, Brin is proposing to bypass centuries of scientific epistemology in favor of a more Googley kind of science. He wants to collect data first, then hypothesize, and then find the patterns that lead to answers. And he has the money and the algorithms to do it", http://www.wired.com/magazine/2010/06/ff_sergeys_search/.
- ⁷ Avey writes: "A seminal component of genetics research is, and has always been, the so-called human subject, a real, live person who contributes his or her highly valuable information... By establishing, and sometimes vociferously defending, the firewall

- between the subjects and themselves--with an intervening IRB-approved consent form typically guaranteeing no personal benefit back to the participants--researchers have been absolved of, or disallowed, as the case may be, any responsibility or ability to communicate directly with their subjects (sounds almost feudal, doesn't it?), as if post-study phenotypic details are irrelevant to the task at hand (Avey 2009).
- ⁸ See Avey (2009).
- ⁹ D. Goldstein (2009), *Common Genetic Variation and Human Traits*, *New England Journal of Medicine* **360**: 1696-98.
- ¹⁰ N. Wade (2009), *Genes Show Limited Value in Predicting Disease*, *New York Times*, April 15 2009.
- ¹¹ J.C. Venter (2010), *We Have Learned Nothing from the Human Genome: Spiegel Interview with Craig Venter*, <http://www.spiegel.de/international/world/0,1518,709174,00.html> (last accessed September 6, 2011).
- ¹² House of Representatives, Subcommittee on Oversight and Investigations, Committee on Energy and Commerce: Hearing Transcript, *Direct To Consumer Genetic Testing and the Consequences to the Public Health*, Washington, DC, U.S.A., July 22 2010, <http://energycommerce.house.gov/documents/20100722/transcript.07.22.2010.oi.pdf>, accessed August 2010.
- ¹³ T. Terranova (2004), *Network Culture: Politics of the Information Age*, Pluto Press, London U.K., page 6.
- ¹⁴ Shannon quoted in J. Gleick (2011), *The Information: A History, a Theory, a Flood*, Pantheon, New York U.S.A., page 221-222.
- ¹⁵ See Terranova (2004), page 12.
- ¹⁶ J.C. Venter (2007), *A Life Decoded: My Genome, My Life*, Viking, New York, NY U.S.A., Page 160, 190.
- ¹⁷ See Venter (2007), page 30, 79.
- ¹⁸ F. Collins (2006), *The Language of God: A Scientists Presents Evidence for Belief*, Free Press, New York, NY U.S.A.
- ¹⁹ See Collins (2006), page 90.
- ²⁰ See Venter (2007), page 98.
- ²¹ See Venter (2007), page 100.
- ²² See Terranova (2004) on the precision of digital over analog, page 32-33.
- ²³ Gilder quoted in C. Anderson (2009), *Free: The Future of a Radical Price*, Hyperion, New York, NY U.S.A., page 82.
- ²⁴ See Venter (2007), page 232.
- ²⁵ See Venter (2007), page 100.
- ²⁶ See Venter (2007), page 166.
- ²⁷ See Venter (2007), page 110.
- ²⁸ See Venter (2007), page 200.
- ²⁹ See Venter (2007), page 181.
- ³⁰ F. Jameson (1982), *The Political Unconscious: Narrative as a Socially Symbolic Act*, Cornell University Press, Ithaca, NY U.S.A., page 60-61.
- ³¹ J. Baudrillard (1988), *The Ecstasy of Communication*, Semiotext(e), New York, NY U.S.A.
- ³² M. Poster (1990), *The Mode of Information: Poststructuralism and Social Context*, University of Chicago Press, Chicago U.S.A.
- ³³ D. Shenk (1997), *Data Smog: Surviving the Information Glut*, Harper Collins, New York, NY U.S.A.
- ³⁴ N. Carr (2008), *Is Google Making Us Stupid*, *The Atlantic* July/August.
- ³⁵ G. Nunberg (2011), *Data Deluge*, *New York Times*, January 10-11.
- ³⁶ See Venter (2007), page 118.
- ³⁷ See Avey (2009).
- ³⁸ To power the computers needed to sequence the human genome, Venter reports that the local power company lacked adequate resources, and that Celera had to install a new transformer and power poles. See Venter (2007), page 251.
- ³⁹ As a sign of the buzz around information in broader cultural circles, note the attention drawn by James Gleick's new tome, *The Information*.
- ⁴⁰ This quote comes from George Church's discussion with Charlie Rose on the Charlie Rose Show on June 19, 2009. See <http://www.charlierose.com/guest/view/6321> (last accessed September 8, 2011).
- ⁴¹ This is not to say that there is no meaning attached to genomics. As Kate O'Riordan has argued, the interactive labor of engagement with personal genomes produces a person as part of a technologically engaged consumer public. Even if there is no or very little biological meaning being offered to this public, there is the symbolic power of the genome on offer, and the possibility of participating in a space with technological elites.
- K. O'Riordan (2010), *The Genome Incorporated* Surrey, Ashgate, U.K.

Author

Jenny Reardon is an Associate Professor of Sociology and Faculty Affiliate in the Center for Biomolecular Sciences at the University of California, Santa Cruz. She is convenes Science and Justice Working Group and the is the co-director of the Science and Justice Training Program at the University of California, Santa Cruz. Her first book, *Race to the Finish: Identity and Governance in an Age of Genomics*, was published with Princeton University Press in 2005. She is currently working on a second book manuscript entitled *The Post-Genomic Condition: Technoscience at the Limits of Liberal Democratic Imaginaries*. E-mail: reardon1@ucsc.edu.

HOW TO CITE: J. Reardon, *Genomics' Problem of Communication*, *Jcom* **10**(03) (2011) C06.