Work in Progress

George M. Church, Ph.D.

Professor of Genetics, Harvard Medical School, and Director of the Center for Computational Genetics, Boston, MA

I have been interested in math, biology, and computers since I was very young, although for a while I couldn't figure out how they related to one another. One of the first connections that worked for me was in X-ray crystallography of the first folded nucleic acid, transfer RNA (tRNA). I wrote some software 30 years ago that's still in use today. Ever since then, I've been trying to automate, computerize, and integrate different subfields of biology.

From my first year in graduate school, 1977, I was already obsessed with trying to automate and multiplex DNA sequencing of everybody's genomes and make it affordable. Sung-Hou Kim was my first mentor. I was already excited about computational chemistry and crystallography, but he got me thinking about it in a more professional way. Wally Gilbert encouraged my breadth in science. Gilbert's was the ideal lab because he was so entrepreneurial and was a sequencing pioneer (with the lac operator). We published a paper on genomic sequencing in 1984, which was also the year I participated in a meeting in Alta, UT, where a few of us suggested the genome project.

The genome project was not dedicated to technological leaps, but to incremental improvement and good business practices. It's been a back and forth dance between the forces of technological change and the forces of production. Now everybody's getting very excited about the technological leaps, having spent the first 15 years of the project mainly on production. We are co-developing so-called next generation sequencing technology using various polony (polymerase-colony) flavors, including in situ gel PCR, emulsion PCR beads, and isothermal no-gels-no-beads, and also focusing on exons, each of which may improve sequencing by a factor of 100 in production and price. A third component is coupling polonies and exons to what we call the personal genome project (PGP). In PGP, a large number of participants are willing to contribute their DNA sequence and a lot of their phenotypic information in a way that can be integrated to establish comprehensive associations.

That squarely faces the ethical, legal, and social aspects of privacy. The current practice is to promise anonymity. The PGP instead says there are many mechanisms by which data can become public, that is, viewed by many eyes, including many strangers, and you may be identifiable because

DNA is identifiable and your phenotypes are identifiable. Many of the ethicists and I feel that it's more honest and straightforward than saying, "Trust us, we'll keep your DNA secret and you will be unidentified." The data may also be scientifically more manageable in an identified form. We're recruiting people who understand the risk, and we will try to keep it as private as we can. I give talks to people who are not scientists and who want to volunteer for the PGP about these ethical issues.

As this moves out of the lab, privacy may turn out be easier to maintain for individuals who want to sequence their own genome without involving their employer or insurance company, as the price becomes more affordable. With the technology I'm talking about, \$3000 may get you 95% of the information you can interpret today (from 1% of the genome). Software companies are going to spring up to satisfy all these early adopters who have a genome all dressed up and nowhere to go. Similar to what you'd do with a home pregnancy test, you would use it to decide when to go to the doctor. People are used to having software advise them. You will have software that tells you in the privacy of your home what diagnostics you should ask your doctor to explain earlier, which ones later, and which lifestyle changes you should make. Certainly the software that we're writing ourselves, and there will be others, will tell you in advance, before it tells you anything, the expected percentages of false positives and negatives.

My advice to people in my lab is to dream big and follow those dreams, because you'll spend almost the same amount of time following up something that's of little consequence as something that has a lot of impact. I encourage technology as a viewpoint, because it allows you to think the slightly bigger thoughts. If you accept current technology, you will think more incrementally. I also encourage people to be nice to their colleagues, think about the environment and the impact of what they do, and not just say, "It's not my department because I'm a scientist." When they ask, "Should I cooperate with this lab or should I hide the data?" I say, let them have it, don't worry about it. Sometimes all it takes is a little nudge to do the right thing, and create virtuous cycles.

It's turned out that a noticeable chunk of the world has gravitated toward the combination of computing and biology I had sought. I happened to be at the right place at the right time. I happen to have an obsession or addiction that society thinks is good. I'm glad it worked out so far, but this is still a work in progress.

-As told to Lynne Lederman, a freelance medical writer based in Mamaroneck, NY.