

Your genetic data

Eric Blair

7 June 2007

[Or, *The ethical implications of SQL.*]

Our paper on the genetic causes of bipolar disorder finally came out last week. The lead author has repeatedly said things like ‘we really couldn’t have done it without you,’ though, to tell ya the truth, I have only a limited grasp of the paper’s results, and have been unable to read it through, due to my lack of background in the world of genetics and biology in general. Fortunately, there have been press releases and a few articles to explain my paper to me.

The figure explains how this is all possible. It is what a genetics lab looks like. That’s a work bench, like the ones upon which thousands of pipettes have squirted millions of liters of fluid in the past. But you can see that it is now taken up by a big blue box, which hooks up to a PC. Some of these big boxes use a parallel port (like an old printer) and some run via USB (like your ventilator or toothbrush). The researcher puts processed genetic material in on the side facing you in the photo, onto a tray that was clearly a CD-ROM drive in a past life. Then the internal LASER scans the material and outputs about half a million genetic markers to a plain text file on the PC.

I know I’m not the first to point this out, but the study of human health is increasingly a data processing problem. My complete ignorance regarding all things biological wasn’t an issue, as long as I knew how to read a text file into a database and run statistical tests therefrom.

Implication one: Research methods We are in the midst of a jump in how research is done. Historically, the problem has been to find enough data to say something. One guy had to sail to the Galapagos Islands, others used to wait for somebody to die so they could do dissections, and endless clinical researchers today post ads on bulletin boards offering a few bucks if you’ll swallow the blue pill.

But now we have exactly the opposite problem: I’ve got 18 million data points, and the research consists of paring that down to one confident statement. In a decade or so, we went from grasping at straws to having a haystack to sift through.

As I understand it, the technology is not quite there yet. There’s a specific protocol for drawing blood that every nurse practitioner knows by heart, and another protocol for breaking that blood down to every little subpart. We have protocols for gathering genetic data, but don’t yet have reliable and standardized schemes for extracting information from it.

When we do have such a protocol—and it’s plausible that we soon will—that’s when the party starts.



Figure 1: The tools of the data processing field known as Biology

Implication two: Pathways If you remember as much high school biology as I do, then you know that a gene is translated in human cells into a set of proteins that then go off and do some specific something (sometimes several specific somethings).

So if you know that a certain gene is linked to a certain disorder, then you know that there is an entire pathway linked to that disorder, and you now have several points where you could potentially break the chain. [Or at least, that's how it'd work in theory. Again, there's no set protocol.] There are many ways to discover the mechanism of a disorder, but the genetic root is the big fat hint that can make it all come together right quick.

Then the drug companies go off and develop a chemical that breaks that chain, and perhaps make a few million per year in the process.

Implication three: Free will versus determinism One person I talked to about the search for genetic causes thought it was all a conspiracy. If there's a genetic cause for mental illness, then that means that it's not the sufferer's fault or responsibility. Instead of striving to improve themselves, they should just take a drug. And so, these genetic studies are elaborate drug-company advertising.

From my casual experience talking to folks about it, I find that this sort of attitude is especially common regarding psychological disorders. See, every organ in the human body is susceptible to misfiring and defects—*except* the brain, which is created in the image of ", and is always perfect.

Annoyed sarcasm aside, psychological disorders are hard to diagnose, and there's a history of truly appalling abuse, such as lobotomies for ill behavior, giving women hysterectomies to cure their hysteria, the sort of stories that made *One Flew over the Cuckoo's Nest* plausible, &c. Further, there are often people who have no physiological defect in their brains, but still suffer depression or other mood disorders. They get some sun, do some yoga, and everything works out for them.

But none of that means that the brain can not have defects, and that those defects

can not be treated.

The problem is that our ability to diagnose is falling behind our ability to cure. We know that certain depressives respond positively to lithium carbonate, Prozac, Lexapro, Wellbutrin, Ritalin, Synthroid, and I don't know today's chemical of the month. But we still don't have a system to determine which are the need-of-drugs depressives and which are the get-some-sun depressives.

Or to give a physical example, we don't know which obese individuals have problems because of genetic barriers and which just need to eat less and exercise. It's only harder because, like the brain, the metabolism is an adaptive system that can be conditioned for the better or for the worse, confounding diagnosis. Frequently, it's both behavior and genetics, albeit sometimes 90% behavior and sometimes 90% genes.

A genetic cause provides genetic tests. If we have a drug based on a genetic pathway, as opposed to a drug like Prozac that just seemed to perk people up, we can look for the presence or absence of that genetic configuration in a given individual. This ain't a silver bullet that will sort people perfectly (if that's possible at all), but having a partial test corresponding to each treatment is already well beyond the DSM checklists we're stuck with now.

Implication four: Eugenics We can test for genetics not only among adults and children, but even fetuses. On one small survey, five out of 76 British ethics committee members (6.6%) “thought that screening for red hair and freckles (with a view to termination) was acceptable.”¹

Foetal gene screens to determine Down syndrome or other life-changing conditions are common, and 92% of fetuses that return positive for the test for Down Syndrome are aborted [Mansfield et al.].

Biology has an embarrassing past in eugenics. And we're not just talking about the Nazis—the USA has a proud history of eugenics to go along with its proud history of hating immigrants (I mean recent immigrants, not the ones from fifty years ago, who are all swell). [My above-mentioned lead author refers me to this article on eugenics², and having read it I too recommend the first 80%.]

If I may resort to a dictionary definition, the OED tells us that eugenics is the science “pertaining or adapted to the production of fine offspring, esp. in the human race.” In the past, that meant killing parents who turned out badly in life or had big noses, but hi-tech now allows us to go straight to getting rid of the offspring before anybody has put in too heavy an investment.

Anyway, I won't go further with this, but to point out that what we'll do with all this foetal genetic info is an open question—and a loaded one, since the only choices with a foetus are basically carry to term or abort. The consensus seems to be that aborting due to Down syndrome is OK and aborting due to red hair is not, but there's a whole range in between. If you know your child has a near-certain chance of getting Alzheimer's 80 years after birth, would you abort? [This Congressional testimony³ approximately asks this question.]

¹<http://adc.bmj.com/cgi/content/full/88/7/607>

²<http://www.logosjournal.com/issue.6.1-2/jacobsen.htm>

³<http://www.hhs.gov/asl/testify/t960917c.html>

Implication five: the ethics of information aggregation This is also well-trodden turf, so I'll be brief:

- It is annoying and stupid that every time you show up at the doctor's office, the full-time paperwork person hands you a clipboard with eight papers, each of which asks your name, full address, and Social Security Number. By the seventh page, I sometimes write my address as "See previous pp" but they don't take kindly to that, because each page goes in a different filing cabinet.

You may recall Sebadoh's song on data and database management: "You can never be too pure/ or too connected." If all of your information is in one place, either on your magical RF-enabled telephone or somewhere in the amorphousness of the web, then that's less time everybody wastes filling in papers and then re-filling them in when the bureaucrat mis-keys everything. I have a FOAF whose immigration paperwork was delayed for a week or two because somebody spelled her name wrong on a form.

- Having all of your information in one place makes it easier for people to violate your privacy and security. As advertisers put it, it makes it easier to offer you goods and services better attuned to your lifestyle, which is the nice way of saying 'violate your privacy'. It means more things they can do to you on routine traffic stops.

The data consolidation=efficiency side is directly opposed to the data disaggregation=privacy side. There is no solution to this one, and both sides have their arguments. A prior entry discussed how information aggregation can lead to disaster⁴, but we should bear in mind that the same technology discussed there made the innocuous and essential U.S. Census possible. The current compromise is to consolidate more and put more locks on the data, but that doesn't work very well in practice, as one breach anywhere can ruin the privacy side of the system.

Back to genetics, when we have a few more snips of information about what all those genes do, your genetic info will certainly be in your medical records. This is a good thing because it means that those who need to will be able to diagnose you more quickly and efficiently; it is a bad thing because those who don't need to know may also find a way to find out personal information about you.

At the moment, you can rely on the anonymity of being a needle in a haystack, the way that some people who live at the top of high rise buildings are comfortable walking around naked and with the curtains open—who's gonna bother to look? But as the tools and filters and databases become more sophisticated, the haystack may provide less and less cover.

So we're going to have a haystack of data about you (and your foetus) right soon. Unfortunately, we don't quite yet know how to analyze, protect, or act on that haystack. I guess we'll work it out eventually.

References

Caroline Mansfield, Suellen Hopfer, and Theresa M Marteau. Termination rates after prenatal diagnosis of Down syndrome, spina bifida, anencephaly, and Turner and Klinefelter syndromes: A systematic literature review. *Prenatal diagnosis*, 19(9): 808–812.

⁴<http://fluff.info/blog/arch/00000199.htm>