

**Policy Perspectives from Members of the Scientific Community**  
*Gerald R. Fink, Ph.D.*

---

DR. WILLARD: So, Dr. Fink, welcome.

DR. FINK: I know you can see me. Is there some way we could just pan around so I can see you? You were a blur.

Good morning. I was director of the Whitehead Institute from 1990 to 2001, and I should lay my cards on the table. I'm not, in principle, for or against big science projects or the kind of large population study which you're being asked to evaluate. In fact, I was the director of the institute and responsible for managing a portion of the Human Genome Project that one of my faculty, Dr. Eric Lander, spearheaded here at the Whitehead. In fact, it was a common joke that I was Eric Lander's boss.

I think that the scientific community now is in general agreement, but in retrospect that the Human Genome Project was successful big science, and that Francis Collins was a wonderful leader in this effort. But, you know it wasn't always that way. In fact, if you go back and read about the Human Genome Project at its inception, the scientific community was not completely behind it.

For this reason, I think it's worthwhile to use the beginnings of the Human Genome Project and its ontogeny as a guide to how a new project of the magnitude of the one you're considering might be successful.

So I'm just going to list some things that struck me when I considered this. The Human Genome Project was very focused.

(Videoconference connection lost.)

DR. FINK: Hello? Are we back on?

PARTICIPANTS: Yes.

DR. FINK: Can you hear me now?

PARTICIPANT: Yes, you're live. You're back on. Unmute it. We can't hear you.

DR. FINK: I muted it by accident when we were cut off.

So we need defined benchmarks, and in fact the yeast genome was the first genome. You carried out the genome sequence, but the human genome began with a map of the genome, and then there were successive increases tied to various benchmarks. There was a defined endpoint, namely the sequence of the entire human genome. There was a defined cost, and I think the interesting thing is that the cost of the study kept going down. That is, the price for every base sequence kept going down rather than going up.

I think this affected what I would say was the scientific community's trust in the project. That is to say, the benchmarks were met, the endpoints were reached, the costs came down rather than going up. So there was originally very great skepticism about this project both in the scientific media and in the public domain, and there was even skepticism I would say about the science and

the fate of the R01s, basic research grants. Would this take away from other important science was a question that was booted about by many scientists. But as each tier was completed and the promise realized, the basic scientists became the greatest supporters of the project because it actually added value to basic research. I think that was an extremely important feature, the feedback from the Human Genome Project into the basic research effort.

Where would skepticism arise in the current project? For this, I have to take a slight diversion into some science, but I think it ends up being a policy issue because the science of quantitative trait loci, QTLs, has a long history, and I need to just give you a sense of where I see some concerns about the scientific issues. This is not to say that this is an unsolvable problem, but this is not untrod territory.

What do I mean? Many of the model organisms have been used to try to map complex traits. Perhaps the most well studied area is in the plant field, because the agricultural scientists in this country have tried to breed plants for traits that deal with yield per acre, very complex traits that have to do with productivity and so on, over the years. So this is a very well developed field in the plant field. In the fruit fly field it's also well developed.

I'm going to mention a study in an area that I know best, and that is yeast, and I should say in all the organisms in which these quantitative trait loci have been looked at, and at great resolution, they're all sequenced, all the data is computerized, and there are no ethical issues.

So I'm going to talk about just briefly quantitative traits, associating disease phenotypes with genotypes, because that's ultimately what such a study would like to do. I want to point out that this has been extremely difficult, even in model organisms.

Yeast, as I mentioned, the genome has been sequenced. It only has roughly 6,000 genes. It's possible to have all of the polymorphisms between two strains on a single chip, and there are no ethical concerns about backcrosses. Brothers can be crossed by sisters, brothers by mothers, by grandmothers, et cetera. This is not a big deal in this organism. Those crosses turn out to be extremely important for the resolution not only of simple Mendelian traits but of quantitative traits.

The trait that was looked at very carefully by a terrific scientist was the ability to grow at high temperature. So what you see on the left is a Petri dish, and you can see a Petri dish in which, on the left, a strain that doesn't grow at high temperature and two strains that do. So here the phenotype is very clear, very easy to identify and easy to identify amongst the cohort that one is studying. These research workers wanted to know how many genes were involved in the difference between growth at high temperature and growth at low temperature, and they did a cross and did all sorts of, I must say, extraordinary genotyping at a level that dwarfs anything that could involve a human population because, of course, it's possible to grow billions of yeast, not just millions, and it's possible to look at this.

The best they could do was to map it to 32 kilobases of a chromosome. They tested this by 3,400 markers, and they localized this there, but there weren't just simple differences between these, so they had to resort even to more exotic breeding experiments and genotyping experiments to try to localize actual differences in the genetic code that could account for the heat resistance, and their conclusions of the best study in yeast was that they couldn't find a single difference that was necessary or sufficient for high-temperature growth. They could not find any marker trait association. Furthermore, they concluded that it must have required combinations of common

and rare variants to underlie the quantitative traits, and the number of genes that controlled this were far greater than expected.

I went through this brief discussion because I think this system and many others that are model organisms point out how difficult it has been to associate these quantitative trait phenotypes with genotypes. So I don't want to minimize -- it is a policy issue, which is what kind of data will we be able to get, and how can we maximize the possibility of identifying the key genes that are involved in a multigenic disease?

So my sense is that one should follow the experience of the Human Genome Project. I think a pilot study would be the equivalent of the early benchmarks in the Human Genome Project. Experience -- I don't know who said it, but experience lets you recognize the mistake, especially when you make it again. So like the genome project, it seems to me that picking some heritable multigenic disease with a defined benchmark, a target, would gain the confidence that one is going to get statistically significant data. I think this is especially important considering that less than 5 percent of medical records are computerized, whereas all the data that I showed you for yeast and for all these model organisms is easily computerized. The phenotypes are very clear. One would have a defined endpoint, then. One could know what a segment of the project could cost, because I think again, for the scientific community, cost is a big issue.

Finally, what would the government do with the information where the particular variant genes increased the risk of disease a few percent? These are just questions. I certainly don't have answers. Does that mean that that would then be a place to look for cures, collaborations with pharmaceutical companies? Finally, what would be the consequences for the R01s? Would this take away from investigator-initiated research?

I think the project, like the Human Genome Project, would initially be viewed with some alarm because of the scientific issues that I discussed and, of course, the risk to funding. The community would be much more supportive, I think, and reassured if there were proof of principle.

Furthermore, I think a crisp definition of the question. I think the goal as listed in the information I got of understanding the relationships of genes, health and common complex diseases, I don't think that actually works. It seems too general certainly for the scientific community, and I think that a large population study to identify risk factors for a specific disease would gain further trust in the scientific community.

Just in discussing this idea, which is, of course, an idea that was generated from the onset of the Human Genome Project, having discussed this idea with many colleagues, questions that have come up and which I don't have time to discuss, and you obviously have much more time and the expertise to discuss, is the NIH really the right organization? Others have suggested the CDC, pharmaceutical companies. I think this is an interesting question.

Finally, ethical issues. Clearly, I am not an expert in this area, and many of you are, but my experience is that one can never anticipate the ramifications of genetics studies that inevitably evoke race, gender and age, buzzwords for unanticipated responses. This reinforces my sense that without a successful pilot program, one could, by indirection, create antipathy towards an otherwise laudable goal.

That's the end of my remarks.

SACGHS Meeting Transcript  
October 19-20, 2005

DR. WILLARD: Thank you very much for that, Dr. Fink.

Can you hear me, Gerry?

DR. FINK: Slightly. Can you hear me?

DR. WILLARD: More than slightly. That's fine.