SCOTT WEISS: So this is an outline of what I'm going to talk about, and we're going to begin by of getting at this question of why complex trait human genetics is so difficult. And then go through each of the steps that you would do if you were actually doing this work. The first question you would get asked on an NIH grant application is your phenotype heritable? And so some evidence of heritability, or doing a study to determine heritability, is often the first step in a genetic study.

And so we're going to talk a little bit about heritability and how you define that, and then identifying disease phenotypes-- so that's a critical issue-- the difference between a subphenotype and an intermediate phenotype, how you want to look at those things, then developing your study design.

And the paper that I just was telling you about, this paper that's coming out in science, these people looked at two relative genetic isolates-- the population of Finland and the Saguenay-Lac-Saint-Jean population in Northeast Quebec, French Canadians. So these are populations where they had a limited number of founders, and so the thought is that they're more genetically homogeneous, and it might be easier to find genes in these populations. But what's the big concern, if you found a gene in the Finns? What would be your biggest concern about that?

AUDIENCE: That it would be replicable?

SCOTT WEISS: Exactly. Maybe that gene might not replicate or be a significant gene in an outbred population in a country such as the United States, where there's a lot of ethnic variation and diversity. So it may be easier to find genes in genetic isolates, but it may be more difficult in terms of their generalizability.

So give me some other examples of relative genetic isolates around the world-- populations that would be considered relative genetic isolates?

AUDIENCE: [INAUDIBLE]

AUDIENCE: Iceland.

SCOTT WEISS: Iceland. Yes, absolutely, Iceland. So that's where Decode is working-- Decode is our number one competitor in COPD research. They're the only company that's actually doing COPD stuff. What else? Where else could you go?

AUDIENCE: South America. [INAUDIBLE] generally from Iceland [INAUDIBLE] in South America. I guess tribal populations in South America and [INAUDIBLE].

SCOTT WEISS: Maybe-- what might be the disadvantage of a tribal population, or disadvantage of a small island population like Tristan da Cunha, which is where they first went to do genetic isolate work in asthma?

AUDIENCE: The small number of people.

SCOTT WEISS: Yeah, small number of people. So you've got a limited number of meiosis in a population like that. You're not going to get too many combinations. So you just run out of gas because you don't have a big enough sample size.

AUDIENCE: Swiss. Swiss families up in the Alps somewhere.

SCOTT WEISS: Possibly. The Finns, the Swiss, the northern Netherlands, the Ashkenazi Jews, Costa Rica.

AUDIENCE: Why? Why Costa Rica?

SCOTT WEISS: Why Costa Rica? Because Costa Rica is surrounded by volcanoes. It was the one place in Central America that the Spanish-- there's not a huge Spanish influx because there was no gold there. So there's a limited number of Spanish founders in the 15th century. They pushed the Indians to the periphery, and they settled the Central Valley of Costa Rica. So you've got this 200 founders in the 14th century, very little intermarriage, and perfect church records-- very large pedigrees.

It's probably the closest, next to the Saguenay--Lac-Saint-Jean population in Quebec, genetic isolates in the Western hemisphere. And we're actually doing a big study there. We actually have six big pedigrees with over 120 people in each pedigree. We've just finished the collection of these pedigrees. We're about to do a genome scan that Tom is actually doing for us.

AUDIENCE: [INAUDIBLE] most of the people are isolated by geography, as you've mentioned with the exception of Ashkenazi Jews. So to broaden, I guess, the definition of genetic isolate to include things like economic-- like, things that usually [? happen to ?] people, like royal families back in the old days. They only intermarried because of social status [INAUDIBLE]. But now there's things where people with certain economic status [INAUDIBLE] in the coming years [INAUDIBLE] see a certain set of people where they marry and--

SCOTT WEISS: Yeah, it turns out, I think, the geography is actually a much better--

[INTERPOSING VOICES]

AUDIENCE: --happening to royal families where people have a chance to procreate [INAUDIBLE] proximity as well?

SCOTT WEISS: I think the geography is a lot better. Historically, it's a much more reliable guide to a genetic isolate than any social convention. I think, exactly for the reasons that Zach--

AUDIENCE: Aren't groups like the Ashkenazi Jews

AUDIENCE: There's so strong social pressure.

AUDIENCE: Yeah, like they [INAUDIBLE].

SCOTT WEISS: They may be the exception to the rule, but I think populations like Iceland, Costa Rica, Tristan da Cunha, Finland--

AUDIENCE: How about the Amish?

SCOTT WEISS: The Amish would fall into that group. The Mormons-- maybe a little bit less so. The big advantage of the Mormons is not so much that they're a relative genetic isolate, but that they have very large families and they have very good church records. So those are other characteristics that are helpful. But the idea behind the genetic isolate is that you've got a relatively homogeneous set of alleles that's circulating in the population. Mormons do intermarry, and so there may be-- I don't know that you'd consider them really a genetic isolate, whereas I think that the Amish and the Hutterites and the Ashkenazi-- these are people where there is some set of code social conventions that those people are much more likely to intermarry.

AUDIENCE: The reason I ask is because it seems like the trend might be, especially nowadays with transportation the way it is, that finding geographic isolated populations is going to decline and [INAUDIBLE].

SCOTT WEISS: This isn't the only way to do this. And I think it is important to make the distinction between linkage and fine mapping, and those two things may be somewhat different.

The advantages of an outbred population is just that the degree of linkage disequilibrium will be relatively narrow. In some of these genetic isolated populations, the degree of linkage disequilibrium can be very large. That means you can do the linkage part pretty effectively, but the association part is more difficult because you've got these big LB blocks that you've got to work with, and you may not be able to get to the gene.

So anyway, we'll go through this stuff. These are all of the steps that you would do in a typical study, and we'll just go through each of these. So I'm going to use asthma as my example because this is the disease I know the best, and the point here is that asthma prevalence in Western developed countries has gone up a lot. So do you think that this is a genetic thing or do you think it's something else? So over the 20-year period, '80 to 2000, we've had more than a doubling in the number of cases. So over a 20-year period, doubling in the number of cases?

AUDIENCE: Genetic?

SCOTT WEISS: Genetic or environment?

AUDIENCE: Environment.

SCOTT WEISS: Why?

AUDIENCE: Environment [INAUDIBLE].

SCOTT WEISS: So there are three potential-- what are the three population genetic mechanisms that something like this could occur-- a genetic explanation? So I'll give you at least half credit, but you're definitely not 100% right. The first genetic mechanism would be spontaneous mutation. You had some spontaneous mutation, and it caused this epidemic of diseases. Is that possible?

Well, you already said no, it's not possible. And you're right, it's not, but you have to know what the spontaneous mutation rate is, which is about 1 times 10 to the minus 8 base pairs per generation. So it's pretty low. We're spontaneously mutating all of the time, but we're not spontaneously mutating fast enough to double the number of cases in a 20-year period of time.

So what's the second possible genetic mechanism? Natural selection, right? So is natural selection going to do--

AUDIENCE: [INAUDIBLE].

SCOTT WEISS: Well, particularly with the disease like asthma, where there's no selection pressure and no reproductive advantages and disadvantages-- you all know plenty of people with asthma, and they're able to reproduce just like everybody else. So couldn't be natural selection.

And what's the third population genetic mechanism that could--

AUDIENCE: [INAUDIBLE] the third one.

SCOTT WEISS: Genetic drift. So you had some mutant asthmatic that came into the American population, and over 20 years of time, they intermarried with all these other people and doubled the asthma rate-- plausible or implausible?

AUDIENCE: Implausible.

SCOTT WEISS: No, it can't happen. So you're right, it can't be genetic. So I give you half credit. I mean, most geneticists don't think like this, but they should think like this. The reality is that all of these genes operate in a developmental and an environmental context. All of your genes do. So the true underlying model for disease causation is gene by environment interaction. So it could very well be that there was some dramatic change in the environment, and now that's interacting with some other genes that it wasn't-- they weren't interacting before with, and now you've got this marked explosion in a number of cases.

And that's almost-- certainly is the most comprehensive explanation, but it would have to devolve from some of environmental change, rather than some of primary change in the genes. But it could easily be that there is interaction between whatever the environmental exposure is and some of the underlying polymorphisms that may be disease related that are different now than they were back here when the disease rate was a lot lower.

So big health problem. I don't want to dwell on this, because it's not the purpose of this course, but all this means is that people will give you money to study this, and they weren't so keen on doing that 20 or 30 years ago.

The other important point is this disease is a disease of children. So you're going to think about-- I usually tell people, this is data looking at the age of onset in a closed population in Olmsted County. What's the famous medical center in Olmsted County, Minnesota?

AUDIENCE: Mayo?

SCOTT WEISS: You weren't supposed to answer that. They're supposed to answer that.

AUDIENCE: [INAUDIBLE].

SCOTT WEISS: Mayo Clinic. So Mayo Clinic is-- everybody in Olmsted County goes to the Mayo Clinic. Now Saudi princes and sheikhs and famous people from all over the world and who's the King of Jordan-- King Hussein? He went there for his-- so Zach's mother, another famous person, went to the Mayo Clinic. So they get a lot of people from outside, but this data is based on the people who live in Olmsted County.

Now if you live in Olmsted County, you don't go anywhere. You just sit right there and you stay there. So this is a fairly stable population. And they were able to capture all of the incident asthma cases and document them because they were all going to the Mayo Medical Center, and they had their chart records.

So 90% of all of the people who were diagnosed as asthmatic in Olmsted County were diagnosed before the age of six. So this is a very, very important point. So this is the opposite of Alzheimer's disease because if you think about genetics, this is great for geneticists because I only have to wait six years from the time the kid is born, and I'm going to know whether they've got the disease phenotype or not. If I was waiting for Alzheimer's cases, it would be *Waiting for Godot.* I'd be waiting a long time before I'd get my cases.

Now there's ways around that for the old people right. And what did geneticists do? How did they find the BRCA1 gene? What did they do to enhance the probability that you would find-- if you're looking at older people, how do you enrich for a genetic cause of a disease? What do you do?

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: Education doesn't have anything to do with it, I'm afraid. What characteristic of the cases would make you think it's more likely to be genetic?

AUDIENCE: Family history.

SCOTT WEISS: What?

AUDIENCE: Family history.

SCOTT WEISS: Family history, but what specifically-- the family history-- age of onset. You're looking for genetic causes of heart attacks, you're going to take the people that have heart attacks when they're age 50. So Ed Silverman, who's the world leader in COPD genetics, in my laboratory, is looking for early-onset COPD cases. So he gets cases where the age of onset is younger than the age of 52. So that's young.

So if you were looking at Alzheimer's cases, you'd say, well, we want all of the cases of Alzheimer's people before the age of 60. And this is how Mary-Claire King found BRCA1. She looked at all of the early-onset breast cancer cases, people who got breast cancer in their 20s, their 30s, their 40s, instead of looking at older postmenopausal women, which is almost certainly another disease.

So if you're looking at old people, one of the clever ways that geneticists enrich for genetic susceptibility is by looking at early age of onset.

AUDIENCE: So they're selecting, potentially, a special case in a particular disease?

SCOTT WEISS: Absolutely right. It's a little bit like, in some perverse kind of way, it's a little bit like the geniculate isolate. You might find a gene that is specific for that particular type of early onset disease. So you find a gene for early onset Alzheimer's, but not for garden variety old age Alzheimer's that occurs in virtually everybody by the time they're 90. So you're right.

But we're still in, I think, what geneticists would say the early stages of this. And because we're still in the early stages, most of us would be happy if we found any gene. So you're going to be in science, if you find that early-onset gene, and nobody's going to be criticizing you because it's not the gene for all breast cancer or all Alzheimer's.

AUDIENCE: [? So yet ?] another example of low-hanging fruit.

SCOTT WEISS: Exactly. So a little bit more about the disease-- most of the kids are allergic. Allergy is probably the big reason why the asthma epidemic occurred, and that means that they have this particular type of an inflammatory process where antigen presented to dendritic cells in the airways activate these CD4 positive T lymphocytes, which then elaborate this series of inflammatory cytokines, which go to these inflammatory cells, which then infiltrate the airways and set up an inflammatory reaction with coughing, wheezing, airways responsiveness, et cetera.

This is all well-known, but it does suggest a whole host of other potential phenotypes that you could potentially look at. And it also gets at this concept of ontogeny of the immune system, where T null cells, at some point, differentiate into these Th1 and Th2 cells, which are determined-- their phenotype is determined by which cytokines they actually elaborate.

And I've got a question mark here, but, actually, this particular step, which is the crosstalk and interaction between these two types of cells that are controlled by two specific genes that elaborate cytokines-- IL-10 and TGF beta, and we genotyped both of those genes in asthma and COPD, and they're important in both diseases.

Now it's important for you to understand that I skewed things a little bit because I told you that asthma is a Th2 disease, and Th2 diseases have increased. There's this increase in allergic rhinitis, food allergy, asthma, et cetera. This novel gene that I was telling you about just a few minutes ago-- it's going to come out in *Science* on Friday-- that gene is expressed in the skin and in gut epithelium and in airway epithelium, suggesting that may be important in all these different types of allergic diseases which, again, has heightened people's interest in the gene and its potential importance.

But it's also important to recognize that Th1 diseases have also increased. So give me some examples of some Th1 diseases. So the epidemiology-- and the reason I'm bringing this up is that most of the immunology community is focused on-- this is why it's important if you're going to be a good geneticist, you've got to really know your disease.

You can't just wave your hand at it and say-- and I think that age of the generalist geneticist, geneticists that sort of, oh, you know, I'm going to study this disease and I'm to study that disease-- not with complex traits. That's not going to work. You're going to have to really know your disease because you have to know the environment, you're going to have to know the natural history, you're going to have to know the intermediate phenotypes, and you have to really understand the biology as well.

The point here is these Th1 diseases-- give me an example of a Th1 disease.

AUDIENCE: Well, I'm thinking about that-- could you tell me the autoimmune diseases, like inflammatory bowel disease, what are they [INAUDIBLE]?

SCOTT WEISS: They're Th1. So Crohn's disease-- Th1 disease. Juvenile rheumatoid arthritis-- Th1 disease. Psoriasis-- Th1 disease. Juvenile diabetes-- Th1 disease. And the reason-- so if the prevalence of these has gone up, and the prevalence of these has gone up, people are thinking that there's something going on further up here that has to do with Treg cells, cells that regulate T cells in terms of their differentiation, because it can't be just at this level that the immunologic defect is.

So it raises the possibility that there are genes-- a Foxp3 getter, T-bet, a whole bunch of other genes that are proximal to the Th1, Th2, C4 lymphocytes that may be important in all of these autoimmune diseases.

And people are just now starting to look at that. And, obviously, the environmental and genetic factors that influence the differentiation of the immune system, or how do people actually tolerize the foreign antigen-- that's the kind of really simple complicated question that, if you could figure out an answer to that, you'd win a Nobel Prize. So that's what my laboratory is starting to work on.

So this is just to show you, again, what I've already told you, that there are a bunch of factors, mostly bacteria and viruses and parasites that influence this Th1/Th2 differentiation. And environmental factors that influence those things are presumed to be important. And one would want to know both the genes and the environmental factors that are involved in this particular disease. There happened to be a whole host of environmental factors that are correlates of those sorts of changes. And I've listed a bunch of them here.

We get very interested in this one. We went to China in 1996 to do an asthma genetics study, and I noticed how different the environment was there. And this left-hand category would summarize what you would see if you were standing in rural China and what you would see in terms of the environmental exposures. Very, very low asthma rates in rural China-- it's about 1%. And it is a progressive increase in gradient in terms of disease prevalence as you march towards Beijing or Shanghai-- much, much higher rates.

AUDIENCE: Let me ask you some questions. Is it not factually wrong that Chinese populations-- I thought the families were [INAUDIBLE]?

SCOTT WEISS: No. See, that's again-- a little bit of knowledge is a bad thing. If you get into rural China, actually, where farming is what everybody does-- although the central government would say there's a two-child policy or a one-child policy, in rural China, they just they have as many children as they-- they may not register them with Social Security, but if they need three kids to run the farm or four kids, they have as many kids as they want. So we found a lot of families with four or five, six, eight kids.

AUDIENCE: All right, here we go, [INAUDIBLE].

SCOTT WEISS: So why has it been presumed to be difficult to do this kind of work? What's the reason that it's difficult? And I think these are some of the reasons. And some of them relate to the issues of study design-- the things that we were talking about. One is this whole idea of genetic heterogeneity. Particularly, if the underlying model here is gene by environment interaction, presumably you could get the same phenotype-- and these phenotypes are determined by multiple genes.

You could get the same phenotype, either high IgE or airways responsiveness in population A with a very different constellation of genes and environmental exposures, and you can get the same thing in population B with different genes. So this is the geniculate heterogeneity thing is a reason for focusing on a genetic isolate. But then you have to worry about the generalizability question.

So, in fact, in asthma there are four positionally cloned genes, counting the paper that's going to come out on Friday in *Science,* and of those four, the first is the only one that's people have really attempted to replicate, and it's gotten mixed results. There are some people that have replicated it and some people that haven't. So it's one of those genes that probably falls into this category of, well, it's not a major gene. It's a minor gene. It's one of the 200 genes that determine asthma, but it's not one of the top 10 in every population.

AUDIENCE: What's your guess about this new gene is going to be [INAUDIBLE].

SCOTT WEISS: My guess my guess about this new gene is that it's a major player. But having said that, the point that I made to the *Science* writer who was doing this is it that's what science is all about is replicating this, seeing how important it really is, and seeing what actually happens.

I mean I think you can get a clue as to whether you've got hold of an area where there's a potential major locus or not, by looking at the replicability of the linkage peaks in a particular region for a complex trait-- in other words, if you've got a region where there's a linkage peak, and there are 10 different studies and 10 different populations, and there's always a peak in the same region, then the chances are that there's a major gene in there that's probably going to apply to a bunch of different populations.

Well, this is-- going back to the *Science* article-- this is a region where there have been a number of people have found a peak there.

The other problem here is that unlike single gene disorders, where there's a known mode of inheritance, you get everything under the kitchen sink here. So you get some of these genes that are autosomal recessive, and some are autosomal dominance, and some are-- so you're getting a whole bunch of things jumbled up in one phenotype, which makes it very difficult.

And then there's this problem of phenocopies. So what's a phenocopy?

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: Give me an example from your own clinical experience of somebody who's-- a phenocopy who's not due to a-- well, just like can get these diseases from genes, you can get them from exposures in the environment. So what if you've got some guy who smokes four packs a day and he's 50 years old and he has a whopping big heart attack? Well, maybe when you quiz him, and he has no family history, but you smoke four packs a day.

Well, you can get a heart attack from smoking four packs a day, and you don't need to have any genes at all for heart attack. You can just-- so that's a phenocopy. He's going to look like somebody who's a genetic susceptible because had a heart attack at 50 years old, but it's all due to an environmental exposure.

And complete penetrance. So this is a problem even in single gene disorders because they're clearly examples-- hemochromatosis, cystic fibrosis-- very different spectrum of diseases in these-- we know that the CFTR gene causes cystic fibrosis. You've got some people who have completely normal lung function, no lung disease at all, and all they've got is mild pancreatic insufficiency, and you've got other people who are totally debilitated from it. So part of that can be penetrance. Part of it can be environmental exposure. But incomplete penetrance is important. Then you've got this problem of multiple genes.

People have very-- the lay public has a very delusional-- they think the genes are immutable. If you've got those genes, that's it. It can't be changed. And they also-- they're monolithic. They're really big, whereas the reality is that anyone-- you've got 33,000 genes in the genome, take a disease like asthma, which isn't very complicated-- maybe there are 200-250-- I don't know, a lot-- that are probably important. Maybe 10 to play a role in most every population.

And a lot of environmental things going on, and it makes it very complicated, and that's why guys like this guy are going to make big bucks because they're going to be able to model all of the different pathways and all the different genes together in some more realistic model of systems biology or some actual way of looking at this.

But the point here basically is look it's complicated to do this stuff. But again, going back to what I said earlier, it's getting a lot easier. So 2002-- one position we cloned gene for asthma. 2003-- two positionally cloned gene from asthma. 2004-- first paper is already out, and there's probably going to be four or five more. So it's going to be four or five this year, year after that, they're going to be probably 10. And all of a sudden now, you've got 20 genes identified for the disease by positional cloning.

And that is the history of complex trait genetics, and it's going to be-- it's happening right now. Right this very moment, all across the world, labs like mine are right in the middle of the fray doing this stuff. This is, simply put, the single most exciting time to be doing human genetics. And it's going to go on for a while, but who knows for how long?

So then there's this other problem of pleiotropy, which can have one gene and it can do a lot of different things. You got the CFTR gene that gives you lung disease, pancreatic insufficiency, and fertility. It all has to do with mucosa and epithelia and different organ systems where this particular genus expressed.

So one of the genes we're looking at-- did Jeff talk to you about CRHR1? Did he show you the data about CRHR1 last week? So that gene is expressed-- is that gene expressed in the lung? Yes or no?

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: No. Not expressed in the lung. It's the receptor for CRF or CRH, and it's expressed in the brain. So what other disease might that gene potentially be important in? He's an endocrinologist. He's forbidden from answering.

AUDIENCE: [INAUDIBLE]

AUDIENCE: What?

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: It comes from the hypothalamus, actually.

AUDIENCE: What clinical disease? Well, I just-- [INAUDIBLE] hypertension.

SCOTT WEISS: Hypertension?

AUDIENCE: It's not--

SCOTT WEISS: There are endocrine causes of hypertension.

AUDIENCE: [INAUDIBLE] course yet?

SCOTT WEISS: So anybody had physiology?

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: HPA?

AUDIENCE: OK, so [INAUDIBLE]?

AUDIENCE: No, think common diseases, man.

SCOTT WEISS: Common disease. [INAUDIBLE] tell you. It's depression. It's been studied a huge amount in the section brains of people who committed suicide, and this--

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: I mean, all kinds of things show that CRF and CRHR1, which is the ligand and the receptor are important in affective disorders.

AUDIENCE: Is there any link to [INAUDIBLE]?

SCOTT WEISS: Well, there's an association between our [INAUDIBLE] type and depression, and Julia [INAUDIBLE] Mexican-American--

AUDIENCE: Really?

SCOTT WEISS: Yeah.

AUDIENCE: Very cool.

SCOTT WEISS: Yeah, so that's pleiotropy. And then, obviously, you've got this problem of penetrance, which is-- individuals with a genotype who actually express the trait, and IgE genes can be important in hay fever. They can be important in asthma. And there are some people who don't have high IgE at all, even though they've got the genes. And these are some other examples of things like--

The basic point I'm trying to make here is that these are reasons that have been given for why doing this stuff is hard. But I'll tell you something. Really, the hard part has been developing the bioinformatics infrastructure, the tools, the bioinformatics tools, and cheap, reliable genotyping. Those have really been things that have been important. And just in the little bit of time that I've been doing this, my genotyping costs have gone from $1.20 a SNP genotype, down to next year, I'll be downed at about $0.15-$0.20 a SNP genotype. And there are three million SNPs minimum-- three to five million-- in the human genome.

Now I'm not going to type at three million, but I've got to type-- in any one experiment, I've got to be able to type 1,000 over a 1020 megabits region-- one of these linkage peaks. So I got to do a lot of genotyping and a lot of people, and it's expensive. The very first position we cloned gene for asthma took six years and $15.6 million. We could do that experiment today for $2 million and a regular NIH grant, and that has totally changed the field. That's the kind of thing that's really making this possible.

So I already said that if you're going to think like a geneticist, everybody has to know a little bit of population genetics. So you have to understand the concepts of linkage disequilibrium, drifts, natural selection, et cetera. Remind you of the fact that somebody asked the President of the United States whether he believed in evolution, and his answer was, the jury's still out.

AUDIENCE: Really?

SCOTT WEISS: Yeah, that's what he said.

AUDIENCE: This is your former classmate, right?

SCOTT WEISS: Right. I went to high school with a president. And so this is the first question that you're going to get asked if you're writing a grant. First thing that you have to address is the disease or the phenotype that you're interested in-- is it heritable? So there's lots of different ways to measure this. You can calculate a heritability estimates. You can do twin studies. You can develop this concept of risk to relatives, which is you look at the risk in the pro bands divided by-- the relatives divided by the risk in the population at large, or you can look at familial aggregation.

But the point is, you've got to gather the evidence, and if you don't know that your phenotype is heritable, you're going to have to demonstrate that it's inheritable before anybody is going to give you a grant to study it because that's what geneticists say. They say they want to know that. They want to know the answer to that question.

So I think it's-- asthma doesn't necessarily have a high heritability, but it clearly is a heritable disease. This is data from one twin study from Danish Twin Registry that looked at the concordance of asthma in identical and fraternal twins. Identical twins share 100% of their genotype. Fraternal twins share 50% of their alleles. Everybody knows the twins also share the environment. So that's another factor that's at issue here. But the reality is that there clearly is evidence of heritability of a disease. You get very different--

The problem here is that heritability estimates are always dependent on environmental exposures as well because the true underlying model-- and disease prevalence-- so the true underlying model for all of these diseases is clearly still going to be a gene by environment interaction.

So after you've decided that the phenotypes that you're interested in are heritable, then you've got to go out and you've got to say, OK, I've got these phenotypes and I'm going to genotype them in a population. You can either look at disease phenotypes-- the advantages of this is that people want to look at asthma. They want asthma genes. They want to find, quote, the gene for asthma, unquote, which we already know is probably a false concept.

But the problem with a lot of disease phenotypes is that even though there may be binary clinically, there may be real problems in terms of making that diagnosis in a way that would be useful for a research study.

The problem with asthma is it's a syndrome. I mean, there is no one way of diagnosing asthma so that you can say, you take this test, and I can guarantee you that everybody that takes this test is going to have the disease and everybody who has a negative test doesn't have the disease.

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: Yeah, but the FEV1 doesn't tell you whether somebody's got asthma or not. I can show you people who have reduced FEV1 and have cystic fibrosis or have interstitial lung disease or have COPD. I mean, they can have a lot of different things. So it lacks sensitivity and specificity the FEV1. And that's true for every single test. I mean, elevated IgE-- well, you could have elevated IgE from parasitic disease or from eosinophilic pneumonia or from 20 other different things.

So there is no single test, and the same may be true for most complex traits. There may be some phenotypes that are a little easier to measure like, say, well, I want to study obesity. Well, how fat is fat? Or is people who are fat like this different from people who are fat like this? I mean, there's all sorts of different ways of looking fat or being fat. So any one of these phenotypes has complications.

And I can tell you this from-- when I first got into this, all I knew was phenotype. I was a world-class phenotyper. I knew all of the nuances of phenotype and everything there is to know about phenotype. And that tends to be what happens when you talk to clinicians because they understand that more.

So this stuff is really, really important, but it's not going to get you very far if you don't know all the other stuff. You've got to know all the other things. I think the point is you do have to know this, and, again it's a--

AUDIENCE: The current problem is a lot of genomicists that don't understand.

SCOTT WEISS: Exactly. Well, it goes back to the point that I was making earlier, which I think is that genetics is moving from a field where genetics were generalists to a field where geneticists are specialists. You get people who specialize in respiratory genetics, cardiovascular genetics, obesity genetics, diabetes genetics. The days of the person that can roam around and do all of these things-- no, I don't think that's going to happen. In five years, six years, you're going to have to be able to go in there and focus on a specific disease because it's going to be too complicated for you to be able to do otherwise.

Then you've got this other type of phenotypes where you can say, well, OK, we want to look at asthma, but what about looking at intermediate phenotypes? So give me some examples of an intermediate phenotype related to my disease of interest? What would be an intermediate phenotype?

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: Absolutely, FEV1. What else? I had it up on a number of slides. So IgE level, right? It's a measure of allergy. Skin test reactivity, airways responsiveness, symptom score, sputum production, exhaled and out. I mean, the list goes on and on and on. You can create 100s.

So for obesity, it could be looking at body mass index is the primary phenotype to define obesity, but then you could look at absolute fat mass or percent body fat, or waist to hip ratio, or insulin resistance, or do CT scans of somebody's abdominal fat deposition. I mean, there's a million different ways of potentially going at this.

The advantage here is that sometimes these are more objective than a subjective, oh, it's asthma, it's not asthma. And it may be closer to the gene in the sense that you've got somebody's IgE level, you have some idea of genes that determine that. And it can be quantitative. You can do a different approach statistically to quantitative traits then you can use if you're looking at binary traits.

AUDIENCE: [INAUDIBLE], like if someone comes in and you look at them and say, OK, this person has a symptom of asthma. They have difficulty breathing. [INAUDIBLE] then you start delving into looking at

SCOTT WEISS: All these intermediate phenotypes.

AUDIENCE: And that correlates then with narrowing the diagnosis from, OK, you don't really have [INAUDIBLE], you have this or you have this type of asthma or is that--

SCOTT WEISS: Well, it's the way I prefer to think about it. And I think it's probably a better way for you to think about it is that you've got to get away-- this is where thinking like a doctor and a clinician is bad. In the world of clinical medicine, it's just like religion. You either have the disease or you don't. There's no such thing as being a little bit pregnant. You're pregnant or you're not pregnant. You have to have bypass surgery or you don't. Clinicians live in a binary world.

Real scientists live in the world of continuous distributions. So you can have-- when are you fat? Are you fat with a body mass index of 23, 24, 25 26. When do you have high blood pressure when it's 130 over 80? Or 140 over 90? Or when is that?

And the other thing is that way to think about these is kind of like overlapping Venn diagrams. The clinical phenotype is actually a composite of these overlapping Venn diagrams that all have separate genetic determinants and things that contribute to them, separate genetic-- it's like dissecting a layer and peeling an onion, where you've got all these different things.

But I think, in many ways, being a clinician can help you as a research scientist. But in some ways, it can also hurt because you start to think in these absolute terms.

So I think the better way to think about it is that these intermediate phenotypes overlap to create clinical phenotypes. And yes, what you're trying to do is stratify in some way or classify people in some way so you're creating homogeneity so that you can actually identify the genetic determinants of a disease or an intermediate phenotype. So you want to go in that direction. But most of these things lack sufficient sensitivity and specificity to really be terribly helpful.

So this is just a list of some of the phenotypes that people have looked at in asthma. And I've studied some of the ones that people have focused on in terms of linkage peaks that have actually been identified. But this is interesting because there's clearly a bias in the literature because there's a whole bunch of these other phenotypes where you could just as-- and I could create a list of 30 more of these where people haven't looked at.

So this just gets to the point that there's plenty of work here for anybody who wants to do this stuff because you can go out and I've got a junior person in my lab, he's got a bunch of phenotypes that he's really interested in, and he's going to go out and he's going to determine their heritability, and then he's going to write another grant, and he's going to map the genes for them and so on and so forth because he wants to have his own little area to work on.

So then the next thing-- so now we're kind of at the point where you've got to I've got to move a little faster or we're not going to make our way through this. But you've got to have a study design, and there's a bunch of different ways of doing this. You can do linkage. You can do association. And amongst the linkage studies, you can do a little sharing methods which are distribution free, or you can do continuous distributions and focus on that.

There are two types of genetic association studies-- the family-based and the case control. Important point here is that they're very different. Here, you have to genotype three people. Here, you have to genotype only two people. Different hypotheses-- here, you're looking at the alleles or the genotypes in the cases relative to controls. It's the frequency of the genotype, frequency in the cases, versus the controls here.

You're looking at a transmitted alleles from a heterozygous parent to an affected offspring. So very different hypotheses, different study designs, and important thing to recognize is that in any association study, the association between a variant and a phenotype can be due to a causal relationship. It can be the linkage disequilibrium.

Or it can be due to population admixture, which means that, usually in the context of the case control study, not a family-based study, you've got different allele frequencies segregating in the cases in the controls because you've got different population histories, evolutionary histories, that have determined those allele frequencies.

So most extreme example would be I had 1,000 Italian cases of asthma, and I'm comparing it to 1,000 Swiss controls who don't have asthma. And even though these two groups are predominantly Caucasian, their evolutionary history may be different, and the allele frequencies may be different as a result of that. So even within an ethnic group, you can get these different allele frequencies. And this is because ethnicity or self-designated ethnicity is only a weak predictor of evolutionary history.

AUDIENCE: What was your example or [INAUDIBLE] example. So you compare [INAUDIBLE] assoiciations [INAUDIBLE] between Germans and Italians, those two populations. Sure enough, we found a linkage association between pasta eating and some piece of genome because, in fact, what you'd be looking for is linkage to the fact that you're an Italian, just by the fact that Italians have a distinct [INAUDIBLE] polymorphisms then the Germans [INAUDIBLE] creates this [? past ?] association when, in fact, [INAUDIBLE] looking at different populations.

SCOTT WEISS: Some of the guys in my lab wrote an article demonstrating all of the potential problems in the case control type of genetic association study. And one of the things that's really impressive about this paper in *Science* is that we all use genetic association as part of the fine mapping process to map a linkage peak.

But-- this is very important-- but because even if you can get rid of the population admixture problem, linkage disequilibrium is always an issue. And so you're never going to know for sure if you're at the gene or you're just close by to it. And so you're going to have to have something else to show that you've actually found the gene. You're not getting into science just with genetic association.

And so the people in the paper that's coming out this week, they have expressed the gene in bronchial tissue. They've done immunohistochemistry to show that the gene is expressed in epithelium. They replicated their results in a different population, et cetera.

So the thing about the [INAUDIBLE] about the case control studies and about even family-based association is that these studies are really easy to do. And so there's lots of them in the literature. So it's really important for you to know, going back to this slide, it's really important for you to know these potential problems because you want to be able to read this literature and say, yeah, these guys really found something or maybe they didn't.

So the advantages of this candidate gene thing is that it's cheap and easy to compare-- remember, I said that now, four positionally cloned genes that have used this type of genome screen approach, four that have been identified since the human genome was mapped in 1996. Well, that's seven years. That's not even one gene a year. That's pretty meek, or weak. And that's because this is very expensive, technologically intensive.

But the thing that's great about this is come up with a novel gene at the end of the time. So it's not dependent on what anybody knows about pathobiology. And so you can go this way, and you can say, look, I know that IgE is important in asthma. So I know that we ought to be screening IL-13 IL-4, IL-4 alpha receptor, CTLA-4, all of those genes in the pathway that determines IgE makes sense-- screen those genes because we've already said that people with asthma have high IgE. Well, you check those genes and yeah, in fact, they are-- most of those genes are asthma or allergy genes.

It's not real exciting, though. I mean, it's not like everybody's going to jump up and say, oh my god, IL-13 is an asthma gene. Well, molecular biologist says, yeah, well, we knew that 10 years ago. What's new? What's great about that?

I mean, there are interesting things about it because you actually can get to the level-- it's going to change molecular biology too because you're actually going to get the level where you say, well, it's these three variants in the promoter, it's this variant in exon 1, and it's this particular haplotype that's determining the effect on IgE level.

So molecular biology is going to change because people aren't going to just be-- aren't going to get away with knocking out a gene or looking at a whole gene effect. They're going to actually have to go in there and determine the particular variants that are important in terms of the molecular mechanisms. So I don't want to denigrate this because this is-- we do all of us-- do a lot of this stuff to keep ourselves busy while we're trying to do these really big experiments that are very expensive and take a long time.

Skip that. So let's talk a little bit about linkage. Linkage is this idea of take these microsatellite markers all the way across the genome. It's a property of families. It's not a property of individuals. And you're looking to see if there's a particular region of the genome contains a gene that's related to the phenotype of interest that's segregating in these families, using identity by descent.

So what you do is, when you have some extended pedigree like this, what you could do is you could do segregation analysis to develop a model to see how the disease is actually segregating in this population. But that's pretty difficult for complex traits. It's not easy to do.

You could also use this approach, the allele-sharing approach, which assumes no mode of inheritance. It just says we collected a whole bunch of sib pairs who are affected, and we're going to test whether these affected relatives have inherited a region of the genome identity by descent more often than expected under a random Mendelian segregation.

And the nice thing about this is that it's easy, but it's not very powerful. I mean, the problem is you need a lot of sib pairs and even then even with over 300 sib pairs, you don't get such great power using this approach. So power goes up if the disease is more heritable, and you can do with less sib pairs, but the reality is that even with a huge number of sib pairs, you may not have a lot of power if the lambda is down here, which it is for asthma, probably.

So I think that this is why people have focused on extended pedigrees in these relative genetic isolates, and that's why we're so excited about Costa Rica. The Finns are clearly excited about Finland. And Decode is doing what it's doing in Iceland. Whether we're going to be successful or not, I don't know.

But the basic approach is that whether you're using an outbred population or a genetic isolate, and whether you're using sib pairs or pedigrees, is you've got these usually di- and trinucleotide repeat STR microsatellite markers, most of the genome services use about 400 of these markers equally randomly spaced across the genome, and what you do is do just do a form of logistic regression basically where you do a LOD score, log of the odds ratio calculation between relating phenotype in the family to these markers, and what you do is get a linkage peak that is the log score for that relationship between the markers and the phenotype.

And what that says is there's a gene or multiple genes in this particular region on a chromosome that's associated with a particular phenotype. And then you have to then go in down and put more markers-- first more STR markers and then SNPs-- and gradually map that region until you've actually got it down to a very small region of a particular 1,000 base pairs or whatever. We can say it's a gene or one or two genes in this relatively large region. So that takes a lot of genotyping and a lot of work.

So our experiments, now, over the next year, we have all these linkage peaks in asthma and COPD. Each experiment is going to be about $200,000. There's going to be 1,500-1,600 SNPs in each of these regions, and we're going to fine map three or four regions over the course of the next year. And hopefully, we will be in *Science.*

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: SNPs 400 base pairs kind of thing, on the order of one per 1,000-- about one per 1,000 bases. That's about what we're shooting for.

So this is just a summary of all of the genome screens that have been done in asthma just to show you that most of them have been sib pair studies. Most of them have been relatively small. But we do get a substantial amount of replication. These are regions across the genome-- this one right here, that's the gene that was just mapped. Several populations, including the Finns, showed a peak in this region. And they got this gene, and then they went to the Canadians and they said, can we replicate it in your population?

The interesting thing is it was asthma in the Finns, but it's high IgE in the Canadians. So it shows you that this problem of phenotypic heterogeneity and genetic heterogeneity is a big issue here. So it isn't a perfect replication at the phenotype level between these two populations, but they've got all this other stuff-- the expression and everything else-- that proves that they've really got the gene.

But the one we're working on is actually not on here. I didn't leave it off intentionally, but it's 12q, and it's one of the ones that's the most replicable.

Now here it is here. It's in this slide right here. So this is a very good region, and there's an [INAUDIBLE] here, but it's also got a very low p-value. So that's one of the better ones.

Now you can already see from this, each one of these-- this region has five or six different genes in this region. It's the cytokine clusters here, beta 2 adrenergic receptors here, IL-13 is here, CD-14 is here. So there's a whole bunch of small genes in here. Nobody knows whether there's a big gene or not. And it may be that linkage peak is just being given by the fact that there's a whole bunch of small genes in that region.

This one, the one we're working in, this is 30 mega bases. That's huge-- huge region. But you can see from just looking at this that-- one, two, three, four, five, six, seven, eight, nine-- and there's another-- I mean, these are 20 regions, each of them about 20 to 40 mega bases. There could be five or six genes in each one of these regions, and at least two of the positionally clonal genes-- there were two genes in the region, and you couldn't tell from the articles. In fact, this Finnish article that's about to come out-- there's a second gene identified, and they don't have the molecular biology on that in the paper. And they're not sure what that gene is doing.

AUDIENCE: So you're actually going to do a [INAUDIBLE] genome.

SCOTT WEISS: Probably.

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: So these are some of the issues in doing the type of linkage studies that I talked about-- multiple markers, multiple phenotypes, multiple comparisons. Phenotypes are correlated, markers not independent-- you know you've got to do.

So there's a lot of statistical issues-- this work is really exciting, I think, because it combines genetics, clinical medicine, molecular genetics, statistics, evolutionary-- all of this stuff is all mixed together. So a lot of important statistical issues in doing these genome screens.

So then you got to genotype the people. We've already said that SNPs are the primary genetic variation in the human genome. But we found INDELs, we found repeats, we found SNPs and INDELs together. I mean, there's all kinds of stuff. In general, SNPs occur about between 1 and 1,000 and 1,000 and 2,000 base pairs. They're approximately three-- maybe three to five million in the human genome. And it's using these as the primary source of genetic variation that we're actually going at trying to map these genes.

There's a whole host of questions about how do you pick SNPs? We wrote a paper together, Zach and I, with some of our colleagues, about haplotype tagging SNPs. There's other approaches to using linkage disequilibrium to define the SNPs that you want to genotype. So lots of issues there, where bioinformatics is interfacing with human genetics.

And no one really knows-- this is probably not 30 million. This is probably three. But no one really knows how many of these SNPs are actually coding, and I think everybody does know that there are more than coding SNPs that are important. Our motor SNPs are important, our coding SNPs are important, SNPs in the 3 prime UTR are important because they're going to change transcription factor binding and potentially change message level-- whole host of different-- and any one SNP is probably, in and of itself, isn't going to change function in a gene all that dramatically.

So people are going towards this idea of analyzing data at the molecular level by looking at relevant functional haplotypes. If you've got a couple of SNPs in the promoter and another that's a nonsynonymous C SNP and an exon, another that's at a splice site, another that's in 3 prime UTR that's determining message level of stability-- you combine all of those SNPs to try to get an effect across that whole gene in terms of looking at that gene and its impact on phenotype.

So this is just a little bit about data analysis. You can either look at continuous quantitative traits or qualitative traits. There are parametric and nonparametric approaches to this. Then you use all this stuff you actually find the gene. I think the people are not doing-- the initial work was done with the [INAUDIBLE] clones, but now we're past the idea of doing that because there's enough markers with a HapMap project across the genome that we can go into almost any region now in the genome, and we can come up with validated SNPs across that region so that we can actually pick SNPs and genotype them and go directly-- and this is what's accelerating the pace of positional cloning at the moment.

So these are some of the things that-- I haven't really talked about this. This is introductory, this lecture. But you really get into this-- how do you do haplotype analysis, ancestral haplotype analysis or linkage disequilibrium mapping, molecular methods or tissue expression-- all of these things can potentially be helpful in the fine mapping process.

We've been very interested and have a project with Zach, where we wanted to use mouse expression and mouse QTL analysis to help us with human positional cloning. We're not sure if our project is going to be funded so we don't know if we're actually going to get a chance to do that.

At the end of the day, you want to be able to look at the impact of polymorphic variation in the gene that you found and see whether that polymorphic-- how much of the phenotypic variance is explained by that polymorphism. And that gets back to this question of, well, you found a gene by positional cloning. How do it's really a significant gene? Well, does it replicate across different populations in different conditions? Is it important in different kinds of asthma? Does it seem to be explaining a significant amount of the variation?

So this is one example. It's a poor example because it's not a really good one. This is a gene, CD-14, that we genotyped in the program in genomic applications. This gene is the gene that binds LPS, or lipopolysaccharide, to the membrane of the monocyte, and then transduces that signal to the T cell to produce Th1 cytokines. So we found a polymorphism in this gene as part of the program in genomic applications. It's a CDP polymorphism. So here's the T variant. Here's the heterozygote. And here's the C.

And you can see that if you look at a dominant model, where the C's are together, anybody that has a C genotype actually is likely to have more positive skin tests than those who are T, T. And that genetic variation is associated with variation in soluble CD-14 levels in peripheral blood. So there's a relationship between genotype and intermediate phenotype, and a relationship to allergy, ultimately.

AUDIENCE: I'm sorry, is that supposed to show a difference between the two?

SCOTT WEISS: It's small, but it was significant. Well, I think the point here is that this is one SNP. This gets back to the point that it's not even a haplotype in this gene, and still there is-- and these are modest numbers. They're not huge. But there was clearly a difference, probably the level of difference you'd expect if it was just a single SNP. I mean, none of these effects are going to be very large at the level of an individual variant. At the level of a gene with a haplotype, with a really significant gene, maybe so, but certainly not one SNP.

So these are some of the skills that if you guys want to do this work, if you were going to come to my laboratory, I would want you to know something about it. You'd want to be want to know something about this. And how to genotype and apply this to the disease, study design, statistical methodology, phenotyping, environmental exposures, and I probably ought to add to this list bioinformatics because without good bioinformatics skills, you're going to be lost.

And it's hard to know exactly where on the spectrum people want to-- if could do this and never have anything to do with the phenotyping and just focus on the functional variation from genes that these guys are actually finding, or you might situate yourself somewhere in the middle, I've got people in my lab that are doing just this and very few people that are doing just this, but I have some that are sitting in the middle.

So where is this going in the future? I mean, I think that what's driving the field is high-throughput sequencing and high-throughput genotyping combined with bioinformatics in the presence of having lots of populations to do this kind of work. That's what's really necessary is you've got to have well phenotype populations. In my lab, these are all the different populations that we have for asthma. We've got these extended pedigrees. We've got affected sib pairs. We've got trios, and we've got individual cases and controls so that we can test the genes in multiple different populations and under different conditions.

So why don't I stop there, and I'd be glad to answer any questions that people have about any of the things that I said.

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: Structure of proteins?

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: Well, I mean I think that means is that you actually getting into-- once you've got a relationship with a gene, what you have to do is really get down and figure out what are the variants in that gene and what are they doing, and that can proceed. Human genetics can contribute to that at the level of genetic association.

So, for example, Lori Glimcher, who's an immunologist at the School of Public Health, identified a gene that controls T-cell differentiation. It's TBX-21, or T-bet is the name of the gene. And she created a knockout mouse, and when you knock this gene out in the mouse, you get tremendous airways responsiveness and allergic inflammation. It looks like an asthma gene in the mouse.

And so we sequenced that gene, and then we started to look at-- we found a variant in the gene that's in the coding region. It's a nonsynonymous C SNP in the coding region. It's very rare. It's only occurs in about 3% of people. But it turns out that coding region variant determines which patients who get inhaled steroids get better. The people that have that variant and get inhaled steroids have their airways responsiveness completely return to normal.

AUDIENCE: Has this been published?

SCOTT WEISS: We're about to submit it to *The Lancet.* We're actually working with Lori to-- it's pretty exciting [INAUDIBLE]. And it's exciting because it's an example of how you can actually-- you don't have to even go to the animal model. And so she, then, has created her mouse model. She started to do some experiments with steroids, and steroids are probably important in controlling T-bet expression. And she didn't know that. So that's an example of structure-function relationships, where you're trying to figure out what a gene actually does.

And it is important to recognize, there are some genes that have been around for a while, and people still don't know there's a relationship to a disease phenotype, but we don't know how they work. So figuring out that structure-function stuff can take a long time, potentially. And doing the genetic association and the fine mapping may actually now proceed at a faster pace and not take as much time. But I think you can actually do a lot of structure function stuff.

Usually, what we do is when we get an association, we will type every damn-- we'll sequence that gene, we'll type every damn variant we can find in that gene in population and look at everything that could be related to an interesting phenotype because we're searching for clues to how to help our molecular biology colleagues in trying to help them figure out what the gene is actually doing.

AUDIENCE: [INAUDIBLE] of that group [INAUDIBLE] see how [INAUDIBLE] changes [INAUDIBLE].

SCOTT WEISS: Well, you could do that. I mean, we're trying to work with this guy. He's got people in his lab who have ideas about how to get clues. So like the stuff that [INAUDIBLE] showed you last week, that gene CRHR1-- we know a relationship to steroid treatment response, but we don't know what the variant is in the gene. So we sequenced the gene completely, and now we've got two INDELs in that gene that sit right at intron/exon junctions.

So the presumed-- what we're thinking is that those insertion deletion polymorphisms may be changing alternative splice sites. So we're going to have to try to prove that. That's one of the hypotheses that we're going to investigate in the renewal of the grant is trying to look at that. So you have to let the gene tell you where its variation is and how it might be contributing to phenotype.

And so the first thing, usually, is to sequence the gene completely. Second thing would be to then do a very careful analysis of the new variants and the resequenced variants that you found in relationship to the phenotype of interest or phenotypes of interest, and see if you can find either haplotypes or individual SNPs or insertion-deletion polymorphisms or transcription factor binding sites or things that could potentially explain the genetic association. So then you can do that, and then you have to go into an animal model and test those in a more rigorous way, usually.

AUDIENCE: How far off do you think is the day when a commission [INAUDIBLE] will be able to come to you and say, I have a disease. I have 500 [INAUDIBLE] cases and 500 controls. I think [INAUDIBLE] on the long arm of a chromosome. And I want to [INAUDIBLE].

SCOTT WEISS: So you're really asking, the question is-- I think the question you're asking is how far away is whole genome association?

AUDIENCE: How far is whole genome association where it's within the reach of significant but not impossible clinical studies.

SCOTT WEISS: Max three years. Max.

AUDIENCE: All right.

SCOTT WEISS: I mean, George Church-- it's all about the genotyping costs, Zach.

AUDIENCE: [INAUDIBLE]

SCOTT WEISS: I mean, he-- listen, he thinks he's close to the $1,000 genome. So if he's really close to the $1,000 genome and SNP genotyping costs really drop, continue to drop as dramatically as they've dropped over the last three years, I would see whole genome association being within the range of a reasonable budget in a two or three-year period of time.

AUDIENCE: All right, on that note, thank you very much, Scott. And--