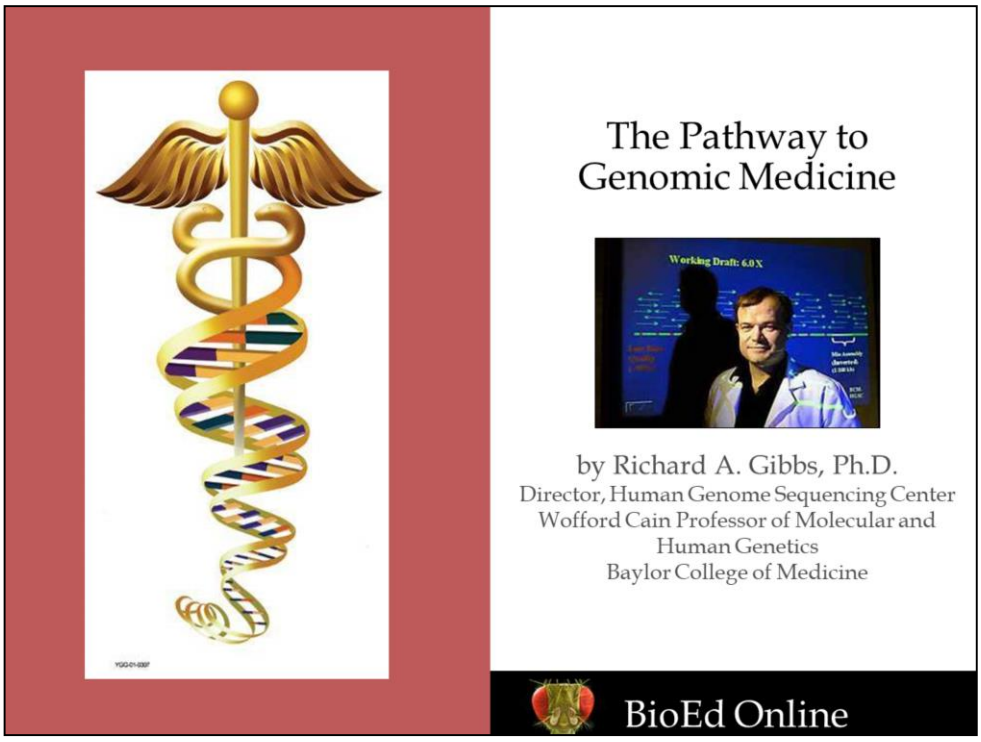


*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

We live in very different worlds—those of research and clinical practice. Genomic medicine, of course, infers that we're going to bring together genome information and genetics and clinical practice. I feel like much of the time, we know so little about each other's worlds [clinical and research], even though we're here [at Baylor College of Medicine] probably nestled together more closely and more intimately than in any other institution in the world. No one else can match our sheer mass of individuals on both sides of that fence.

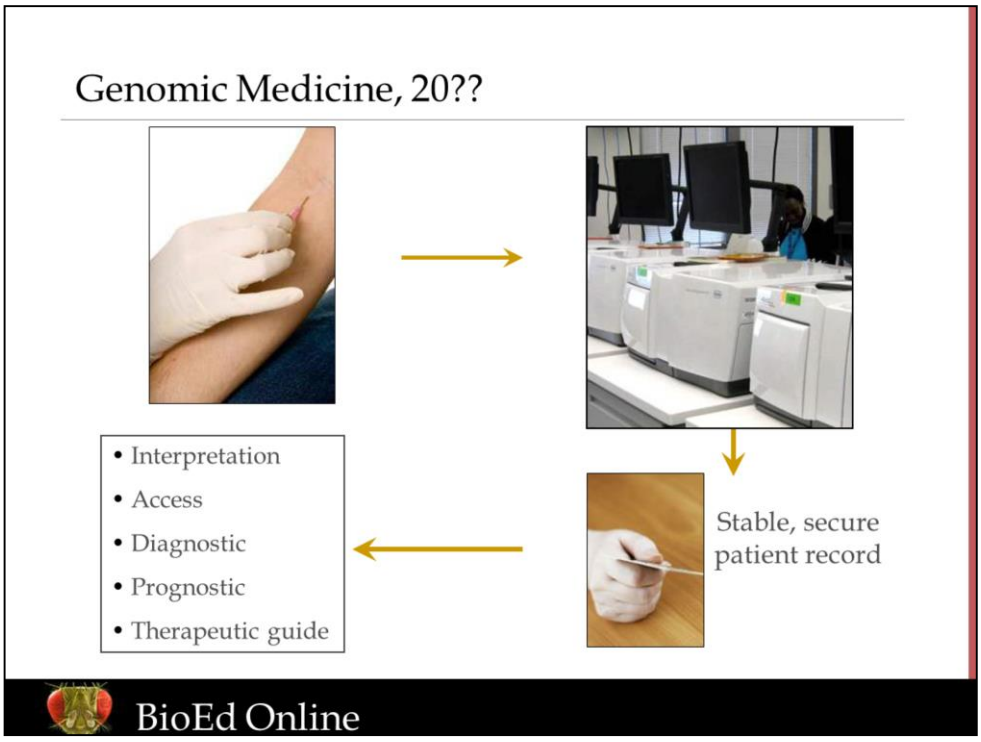
So I think we have this extraordinary opportunity, to overcome some of the enormous barriers to what we do. So that's why I jumped at the chance to come and talk here. I feel like every time we collaborate, we add one more little dot in the big mosaic that we're building. And maybe one day we'll all stand back and look at that mosaic, and we'll know exactly what genomic medicine is. In the meantime, I think we kind of have to struggle through to make our own definitions. So what I really want to do today is to give a bit of reality to some of the elements of the whole enterprise, and set the framework from the genomic point of view, as well as outlining how the study of genetics is unfolding in informing human disease.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image References:

Carduceus with DNA helix. (2001). 2001 U.S. DOE Primer: Genomics and Its Impact on Medicine and Society. U.S. Department of Energy Genome Programs. Retrieved 04-06-08, from

http://genomics.energy.gov/gallery/basic_genomics/detail.np/detail-01.html.
Richard Gibbs. (2007). The pathway to genomic medicine. Houston, Tx: Baylor College of Medicine.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

And so of course it's all directed to this magic model where somebody walks into your clinic, and I guess this is the near-term model, where you [the physician or researcher] send their sample over for a complete human genome sequence, and you use that to inform your decisions. The further-term model will be that they will actually walk in with that little device, or you'll look at it online, and they'll already have logged in their family genetic information, their own DNA sequence, and you'll just have to look it up to get it.

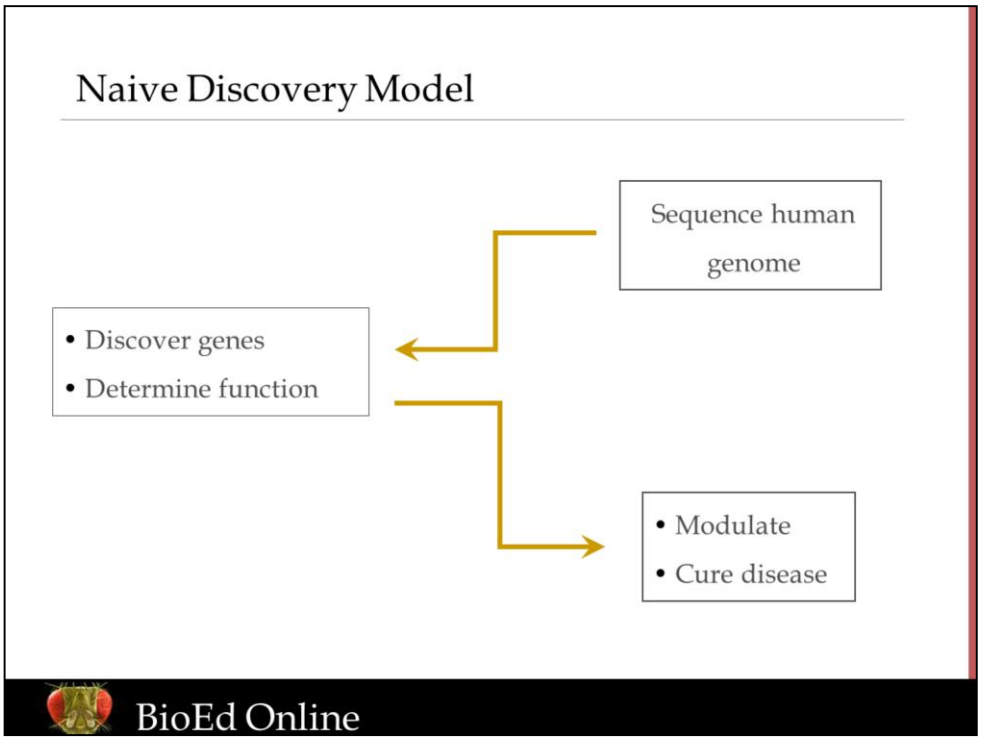
* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image References:

Hand with card, #926871. © 2005 - 2008 123RF Limited, and its licensors. All rights reserved.

Injection, #524330. © 2005 - 2008 123RF Limited, and its licensors. All rights reserved.

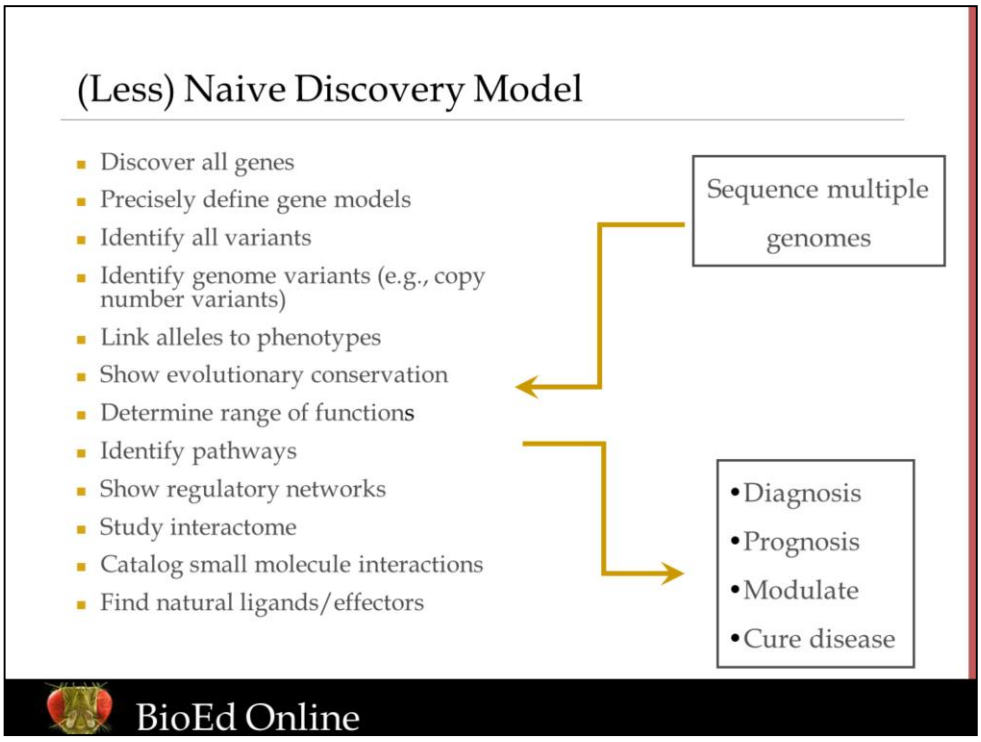
Scherer, S. (2008). *Human Genome Sequencing Center, Baylor College of Medicine*.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

If you subscribe to that model, which of course we all did at the beginning of the Human Genome Project, then you had to subscribe to this pretty simplistic model of how all this was going to work. That is, there would be a primary data discovery, the sequence of the human genome. This would easily lead to discovering genes and inferring their function, and we would just use that information to modulate and cure disease and do diagnostics and prognostics.

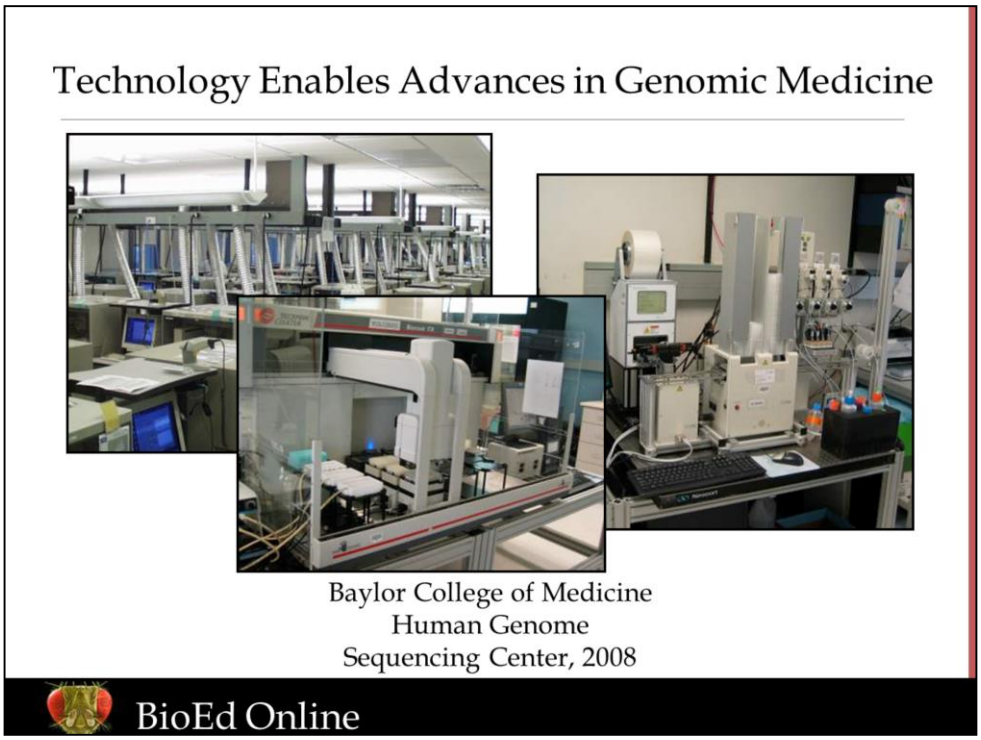
* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Of course, we all knew that that was a very naive model, and it had much more complexity to it. In fact, this big box on the left actually represents all of the biomedical research, and what everybody's been doing all along. I think what we've been doing since is really adding layers and layers, and smaller and smaller tics to this diagram, while acknowledging the importance of the integration of simple genomic approaches into the complexity of biomedical research and disease phenotypes.

* Notes in this slide presentation are adapted from the transcript of *"The Pathway to Genomic Medicine," a presentation by Richard* Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



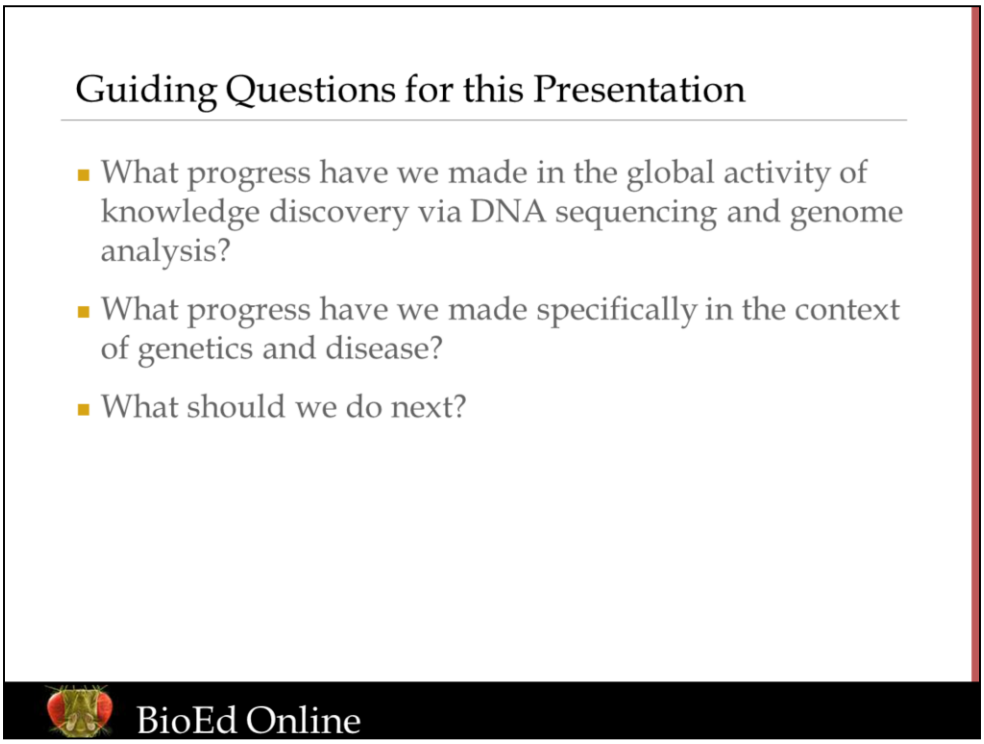
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Within the framework of genomic medicine, the technology has been critical. That is, every step in the development of the whole model has been enabled constantly by technology. That continues to be the theme. And at each point, the only real reality check we've had is to sit back and look and say, "Well, what have we achieved in terms of development of data sets and development of knowledge about individual disease?"

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image References:

Steve Scherer. (2008). *Human Genome Sequencing Center, Baylor College of Medicine.*



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

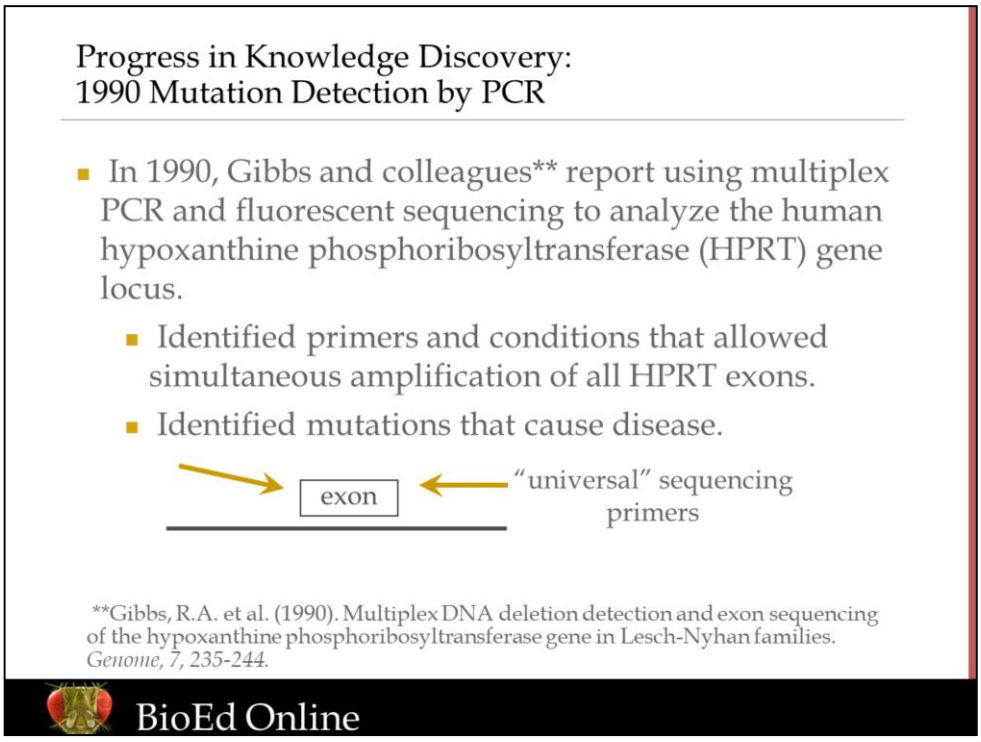
So that's the framework, as I already said, that I want to address today, in the context of these three questions.

How are we doing with this idea that a lot of DNA sequencing and analysis of genomes—and the follow up from that—can actually drive knowledge discovery?

In particular, how are we doing that in the context of genetics and disease?

What should we do next, particularly here at Baylor?

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



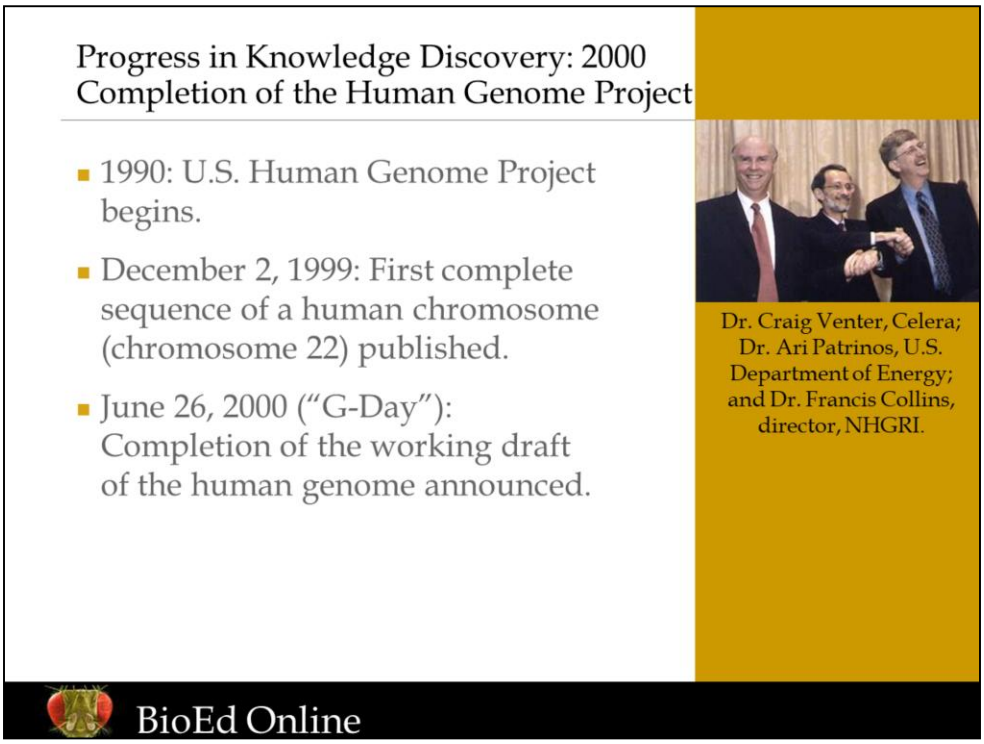
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

The first question is actually pretty easy to answer. We're really doing pretty well. This is my favorite slide from 20 years ago, when we did the first PCR direct analysis of a gene and found **heterozygous** mutations that caused a direct disease. It was a nice illustration of the first application of multiplex polymerase chain reaction, in combination with fluorescent sequencing, to find all the gene's structural mutations and to say that the rest of the gene wasn't affected, because we'd looked at every coding base in it.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Reference:

Gibbs, R.A. et al. (1990). Multiplex DNA deletion detection and exon sequencing of the hypoxanthine phosphoribosyltransferase gene in Lesch-Nyhan families. *Genome, 7, 235-244*.



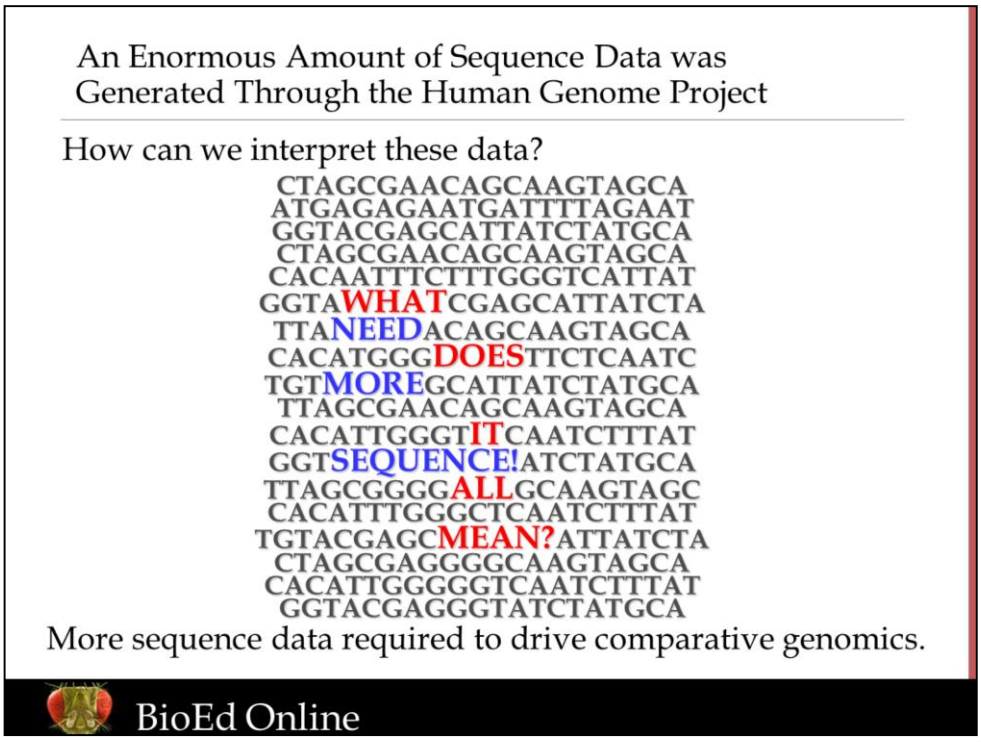
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

So we've gone from that now through sequencing the whole human genome, and here's the scene from the White House where everyone had a group hug, saying, "It's all done and the private and public sectors have all agreed to disagree about certain things, but the datasets speak for themselves." That's the shortest history of the Genome Project I've ever been able to deliver—down to about ten seconds?

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

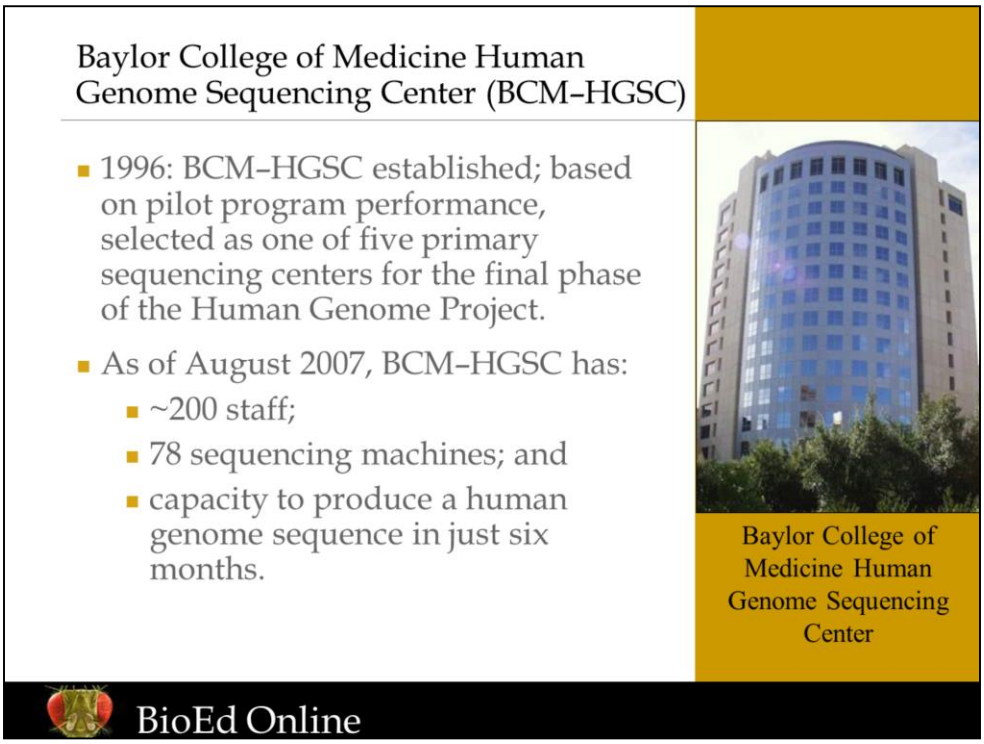
'Working draft' of human genome announced at White House. (2000). The NIH Record. Retrieved 04-07-2008, from http://www.nih.gov/news/NIH-Record/08_08_2000/story03.htm



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

And of course after that, we came home and figured that we had a lot of Ts, Cs, As, and Gs, and didn't really understand them, and asked ourselves the critical question, "How could we begin to interpret these data and eke out what was inside of them?" And we pretty much came to the conclusion that what we needed more than anything else was actually more sequence data, in order to drive comparative genomics and to generate some sources of genetic variation and so forth.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

So from that period, I guess it's now six years since that first draft of the human genome appeared, and four years since the finished genome, and we've continued to build the enterprise of being able to sequence DNA in the way that we learned to do it in the context of the Human Genome Project. So for those of you who've been upstairs—and if you haven't, you're certainly invited to come and look—we now have this big machine. It's three floors of this building, and it has 200 people, and it has a floor full in instruments and computers. And its job right now is to produce about a one times coverage of a mammalian genome every month [9/2007]. So in some contexts, that's remarkable.

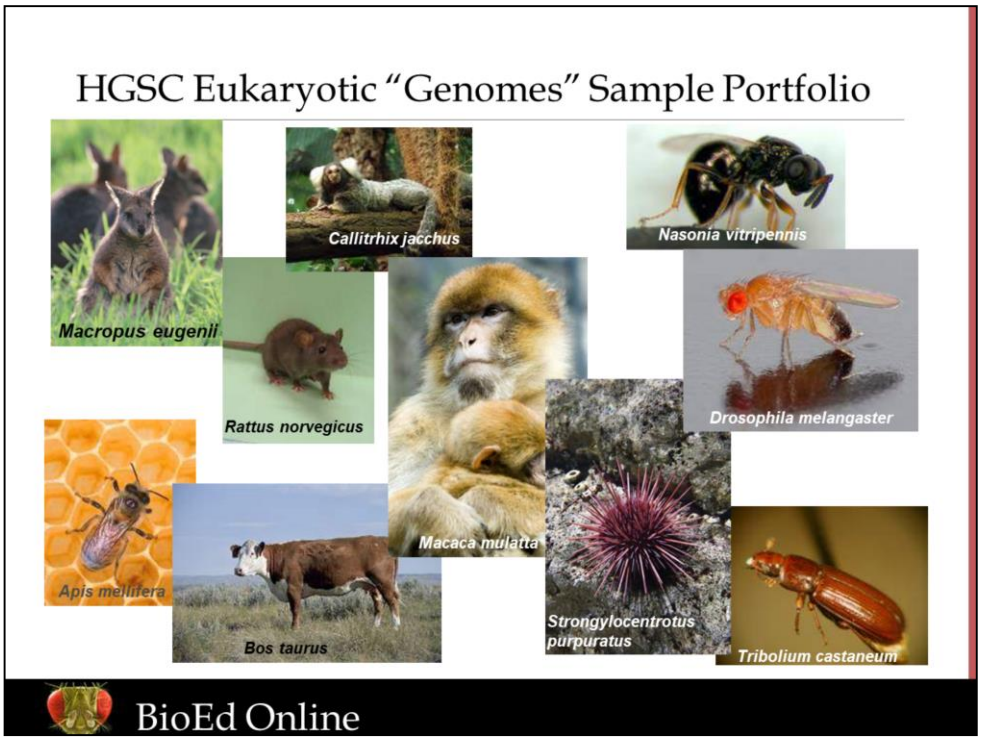
Because through the Genome Project, we started doing a very tiny fraction of that, and wondered how we could do this entire job. Now we can cover a whole human in one month. Normally, for reference, when we sequence a mammal, such as a human, we would do a six-fold coverage. In six months a facility like ours can actually push out a human genome sequence, equivalent to what the Human Genome Project offered, at least at its first phase. And so that's a number to keep in mind, please. We have a three billion—base genome. You have two copies of that. You have six billion bases; three billion from your mother and three billion from your father. And that's about the number of raw bases we can push out from the Genome Center.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

10

Gibbs, Richard. (2007). *The pathway to genomic medicine*. Houston, Tx: Baylor College of Medicine.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Now, since the finishing of the Human Genome Project, we've actually spent more time sequencing the genomes of other creatures to inform the function of human genes, than we have spent sequencing different humans in a disease context. This is a small part of the menu of all the creatures that have been pushed into this pipeline; and each one has its own story and own justification. Together, I think with the other centers and other parts of the international program, there's probably a hundred species possessing larger genomes, extending far up the tree of life, that are well advanced in having their genomes sequenced. And of course, many more hundreds if you count microbes and small eukaryotes.

So today I'm barely going to talk about any of these. But I do like to just touch on a couple of them to give you this flavor: on the genomic side of this genome medicine equation, we've still got a genomic enterprise that is about sequencing the genomes of animals to help inform our direction through biomedical research via comparative genomics.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image References:

Bauer, S. *Honey bee*. National Human Genome Research Institute. Retrieved 01-04-2008, from http://www.genome.gov/pressDisplay.cfm?photoID=76.

Day, E. *Red flour beetle*. National Human Genome Research Institute. Retrieved 01-04-2008, from http://www.genome.gov/pressDisplay.cfm?photoID=87.

MacNeil, M. *Bos tarus* . National Human Genome Research Institute. Retrieved 01-04-2008, from http://www.genome.gov/pressDisplay.cfm?photoID=67.

Macaque # 2034363. © 2005 - 2008 123RF Limited, and its licensors. All rights reserved.

Rattus norvegigus. National Human Genome Research Institute. Retrieved 01-04-2008, from http://www.genome.gov/pressDisplay.cfm?photoID=69.

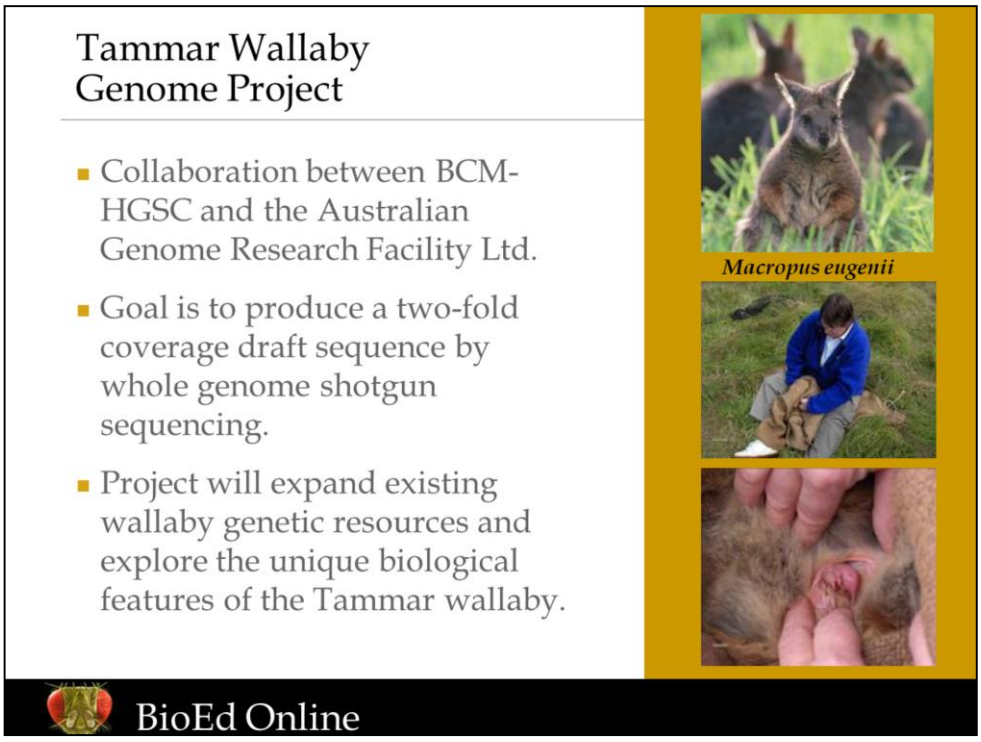
Sea Urchin # 2235255. © 2005 - 2008 123RF Limited, and its licensors. All rights reserved.

Spekking, R. *Marmoset*. National Human Genome Research Institute. Retrieved 01-04-2008, from http://www.genome.gov/pressDisplay.cfm?photoID=82.

Wallaby. National Human Genome Research Institute. Retrieved 01-04-2008, from http://www.genome.gov/pressDisplay.cfm?photoID=94.

Wikipedia. (2006). *Nasonia vitripennis* (M. Clark). Retrieved 12-28-2007, from http://en.wikipedia.org/wiki/Nasonia.

Wikipedia. (2006). *Drosophila melangaster* (M. Clark). Retrieved 12-28-2007, from http://en.wikipedia.org/wiki/Image:Drosophila_melanogaster_-_side_%28aka%29.jpg.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Example number one is this creature here, the Tammar Wallaby, which my friend and colleague, Marilyn Renfree, here is catching. Below, there's its embryo. And I told a remarkable story, I think at our meeting last week, of how all these offspring are born on the same day of the year—all the members of the species. So that's one piece of remarkable biology. But the other thing I didn't talk about is the antibiotic requirements for a creature that essentially does its development in the pouch, rather than in utero.

You know, these are not placental mammals, these are marsupials. So this thing develops most of its early life period outside or in the pouch. So at that time, it has a brother or sister who shares the pouch. So it [female parent] has two teats, one which produces a remarkable range of beginning-to-be-discovered antibiotics that can allow it [offspring] to survive in that period, and another which has other nutritional requirements. So what an opportunity for investigation in reproductive and early life biology.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

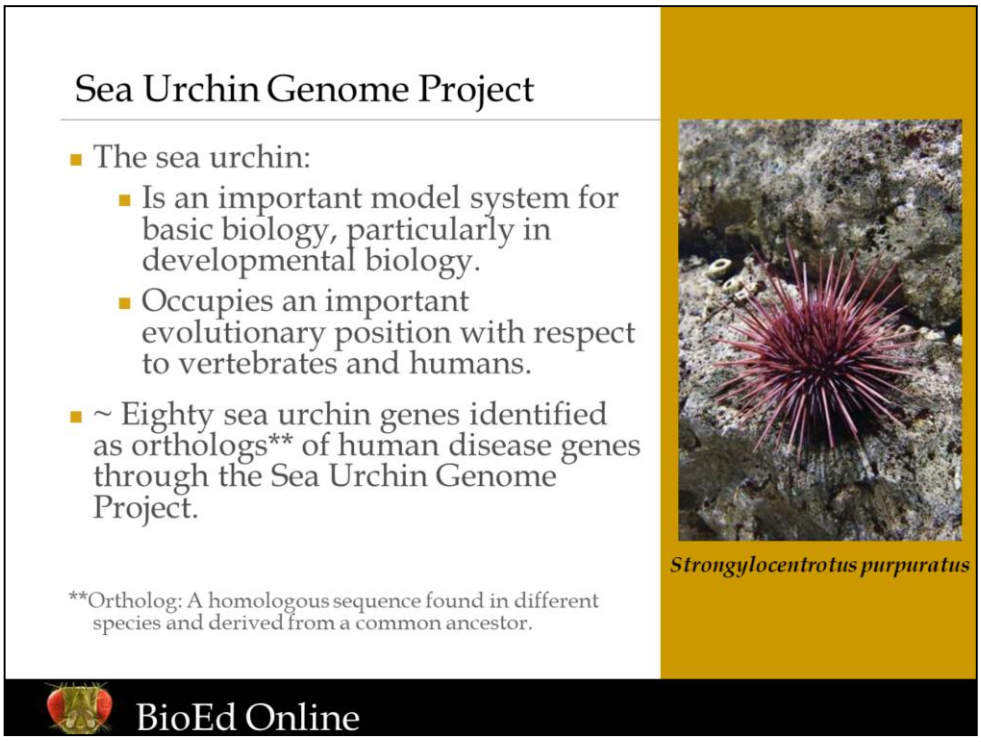
Reference:

Human Genome Sequencing Center at Baylor College of Medicine. (2005). Wallaby Genome Project. Retrieved 09-05-2007 from http://www.hgsc.bcm.tmc.edu/projects/wallaby/

Image References:

Adult Tammar Wallaby (Macropus eugenii). National Human Genome Research Institute. Retrieve 01-04-2008 from http://www.genome.gov/pressDisplay.cfm?photoID=94

Gibbs, Richard. (2007). *Marilyn Renfree with wallaby*. *The pathway to genomic medicine*. Houston, Tx: Baylor College of Medicine.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

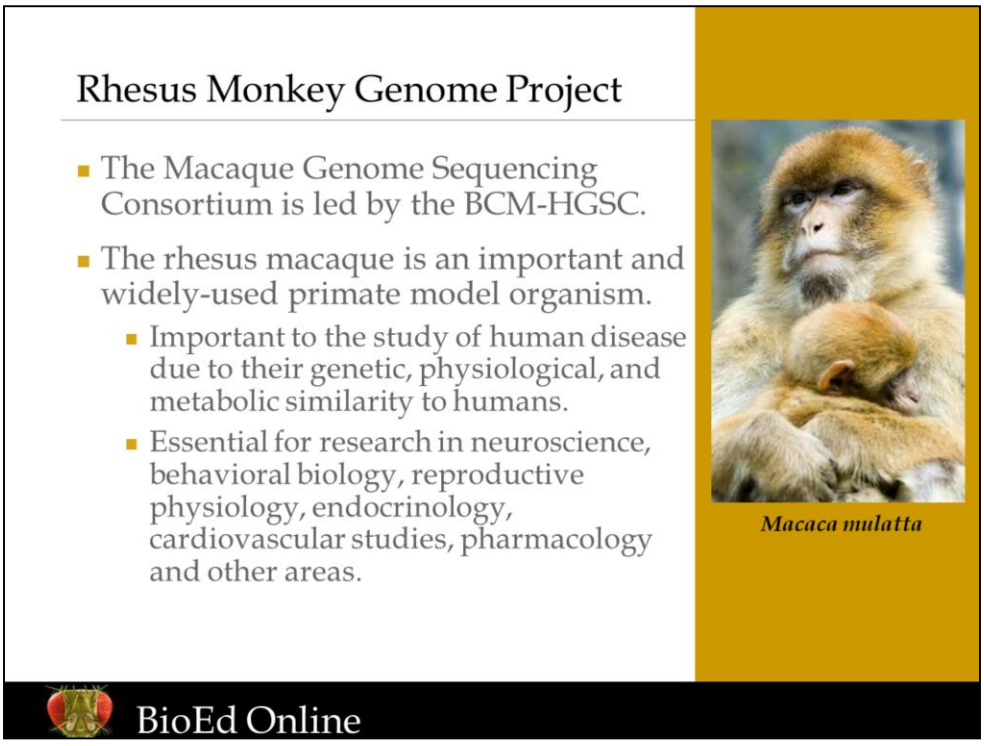
Here's another one of these creatures. This is the sea urchin, which probably some of you, or a few of you, might have worked on in some lab rotation at some stage. Of course, it's a model for development and another not too far distant eukaryote. Amongst the many things we discovered in the long, drawn out genome project was that there are eighty or so genes in this creature which really can now be used and studied as disease orthologs. If you're working on a particular human genetic disease and you know the gene, you can now consider the sea urchin one of the creatures that you might want to look at to investigate the biology of that gene.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Reference:

Human Genome Sequencing Center at Baylor College of Medicine. (2007). *Sea Urchin Genome Project*. Retrieved 09-05-2007, from http://www.hgsc.bcm.tmc.edu/projects/seaurchin/.

Image Reference:



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

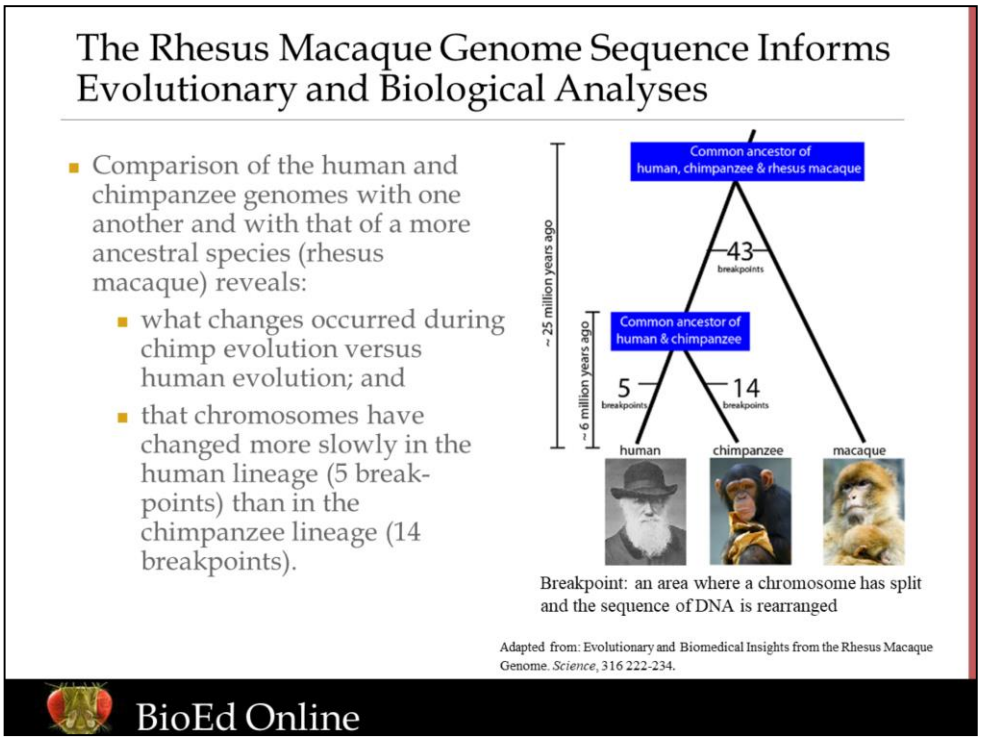
And the last example is the macaque, which we finished recently, as David mentioned. These primates, of course, are extremely popular and important laboratory models. They're used in neuroscience and vaccine studies, but also they're used in behavioral studies. What fabulous creatures, and how similar to us they are. Those of you who own dogs probably think that dogs are almost the perfect models for human behavior. It takes very little for a non-primatologist to see these primates in action and start to engage them as fellow creatures in the kingdom, for which we would love to understand much of their dynamics and how they work.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

References:

Human Genome Sequencing Center at Baylor College of Medicine. 2007. Rhesus macaque Genome Project. Retrieved 09-05-2007 from http://www.hgsc.bcm.tmc.edu/projects/rmacaque/

Image Reference:



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

So we studied the macaque, and along the way deciphered—using just its raw genome sequence and comparing it to the chimp and human genome sequence—features about the rate of primate evolution based on different chromosomal events. It turns out that as chromosomes have changed across the species tree, the changes have slowed down a little bit in the human part of the lineage, more so than in any other primate. This, together with other data, is very informative about the dynamics of these genomic changes across evolution. This is a long study; I'm not going to tell you much about it.

Another highlight was the few genes we were able to eke out of that genome by comparing it to chimps to humans, and asking, of the changes that we see, how many of those changes reflect the pattern of action of positive selection across the lineage? Because those are the genes, of course, which we suspect are most important in saying why we're humans and chimps are chimps, and macaques are macaques.

Some of those genes have even been duplicated as the genome has gone through its efforts to make that part of the genome more conservative, more important through the evolutionary process.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Figure adapted from original publication: Rhesus Macaque Genome Sequencing and Analysis Consortium. (2007). Evolutionary and Biomedical Insights from the Rhesus Macaque Genome.

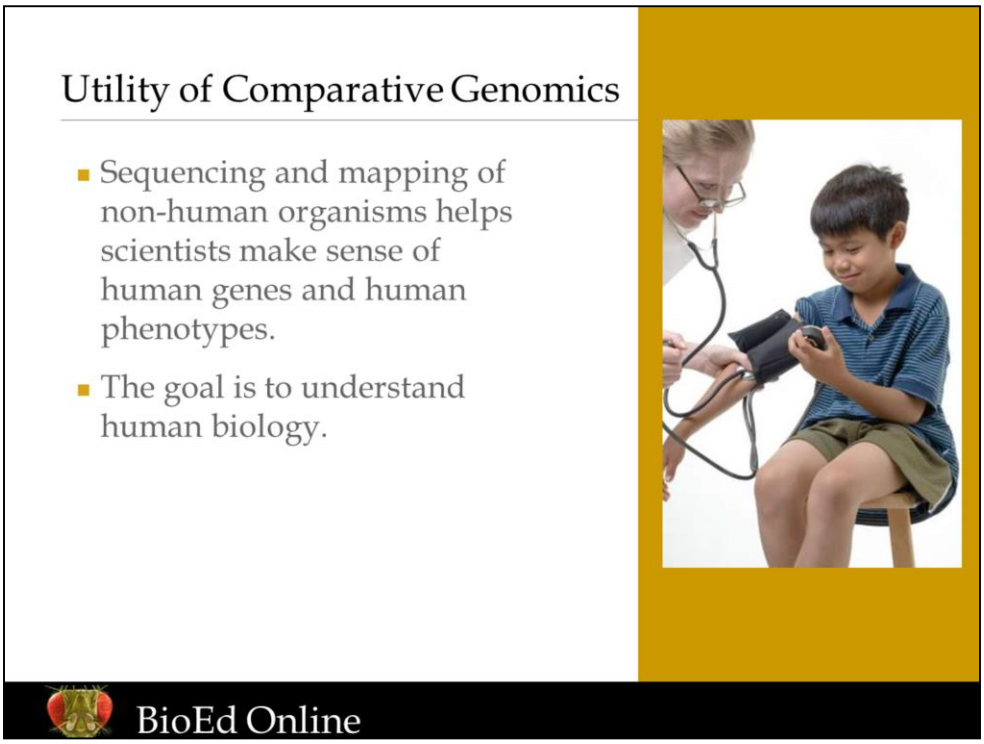
Science, 316 222-234.

Image References:

Macaque # 2034363. © 2005 - 2008 123RF Limited, and its licensors. All rights reserved.

Wikipedia. (2004). *Lightmatter Chimp*. Retrieved 01-04-08, from http://en.wikipedia.org/wiki/Image:Lightmatter_chimp.jpg.

Wikipedia. (2007). *Charles Darwin in 1880*. Retrieved 01-04-08, from http://en.wikipedia.org/wiki/Image:Charles_Darwin_1880.jpg.



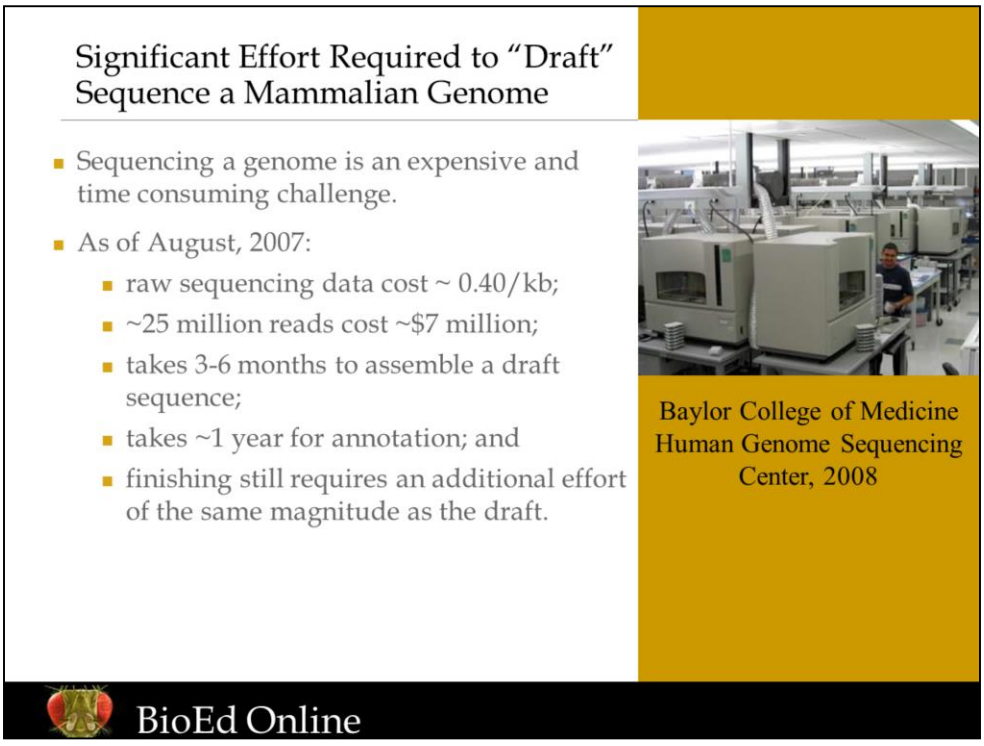
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

So we're continuing, of course, all these studies in these other creatures to drive our use of sequence and genome analysis to inform human biology, to make more sense of the human genes, to make more sense of the human phenotype, and to allow us to make more inference. When you talk about something to do with a human gene, what does it mean for the human biology?

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Boy with physician, # 439838.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

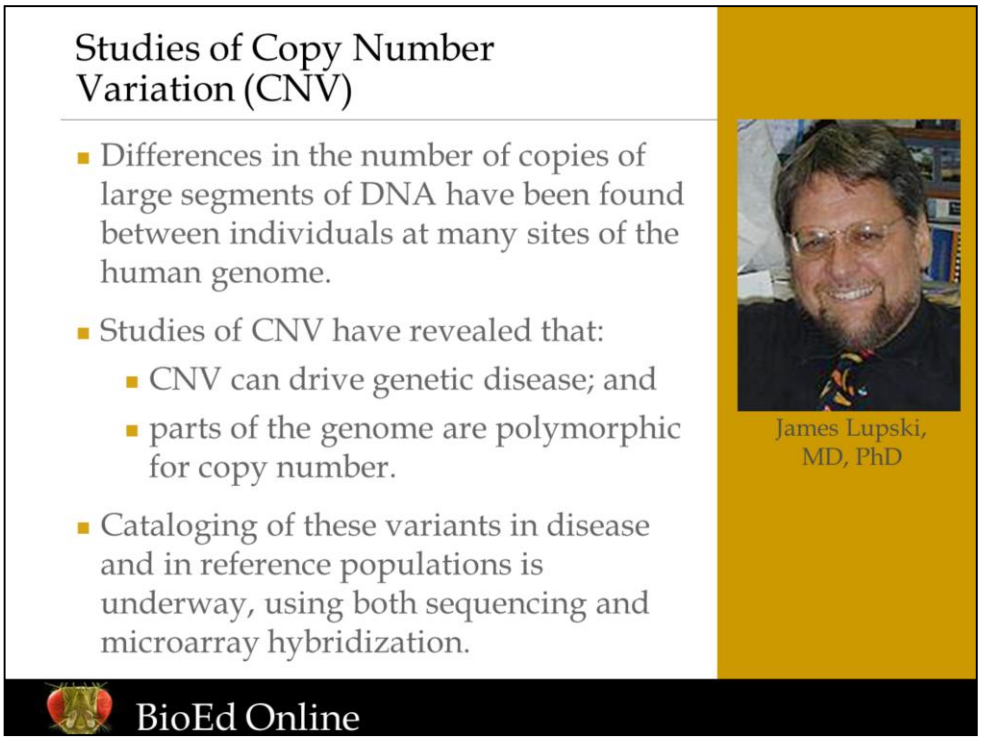
So as we're all geared to do in the genome world, we kind of make it sound easy, okay? But in fact, it's not. It's still an expensive and clumsy, growing science to produce whole genomes. And without going into the ugly details to prove that, to really exemplify it, I'll just tell you that to use those machines upstairs, it still costs about 10 million dollars to generate that raw data, and it still takes six months, which is remarkable, since it was a 15-year project using multiple centers the first time.

But it's still six months, and that's no diagnostic test, right? And that's just to generate the raw data. You've still got to get a team of people that you can barely find and train to do the assemblies and put the stuff together, and do the quality control and annotate the data, and then do all that stuff that makes a genome. And if you want a genome like the human, which is actually a very refined and high quality piece of scientific product, it's about the same effort again to get to that finished state. So the bottom line is, this is still a raw and underdeveloped science. There's still a lot of work to do just in this basic effort of genomics. Having said that, the programs have not been restrained in what we and they are trying to do with sequence as a basic discovery tool. And that's illustrated by a couple of things I'm going to just mention, but not talk about in any detail today.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Steve Scherer. (2008). *Human Genome Sequencing Center*, Baylor College of Medicine.



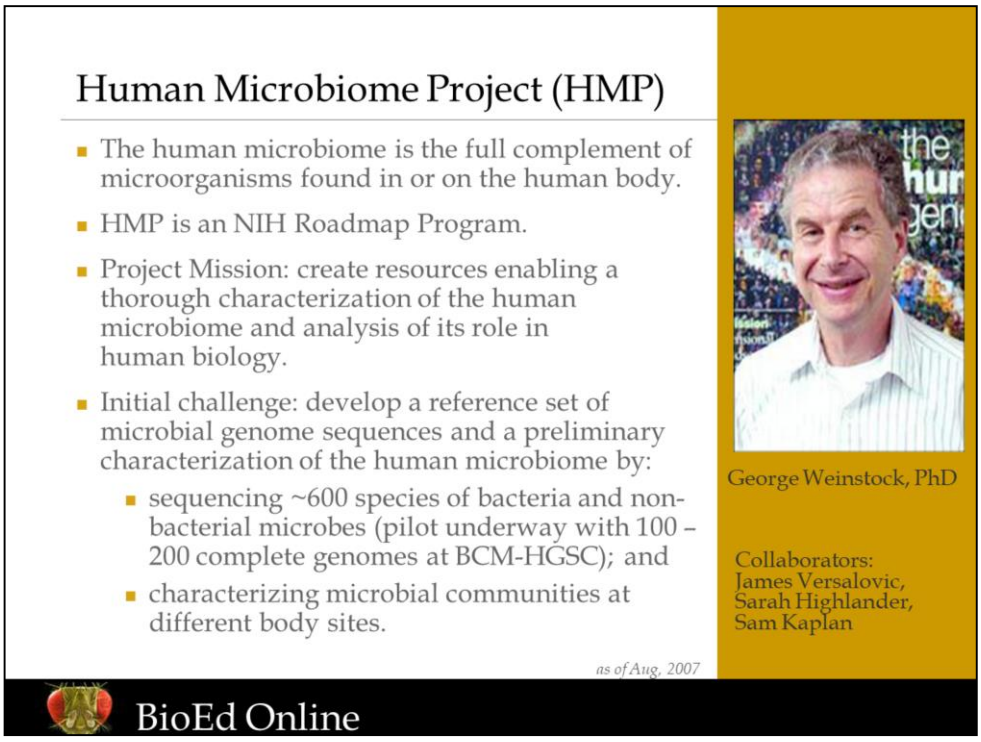
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

First, Copy Number Variation Studies - Jim Lupski will talk about this in this series. And what he will tell you is that following up his earliest discoveries are demonstrations that genome duplications indeed can drive genetic disease. That's quite controversial. Jim really nailed that maybe ten years ago. That, plus the observation by many others that parts of mammalian genomes are naturally polymorphic in their quantity. That is, you can look along the genome and find a part which is duplicated in one person maybe three times and you can go to another individual and it'll be duplicated five times. These can be large pieces of genome, with function, with genes contained within them. So the concept of copy number variation, CNV, plus this notion of the engagement in disease, has led to other programs which, in some instances, are using microarrays and other techniques. But in other cases, investigators are actually using sequence to drive the discovery of a full catalog of these events.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

James Lupski. Retrieved 04-07-2008, from http://www.bcm.edu/news/packages/lupski-bio.cfm.



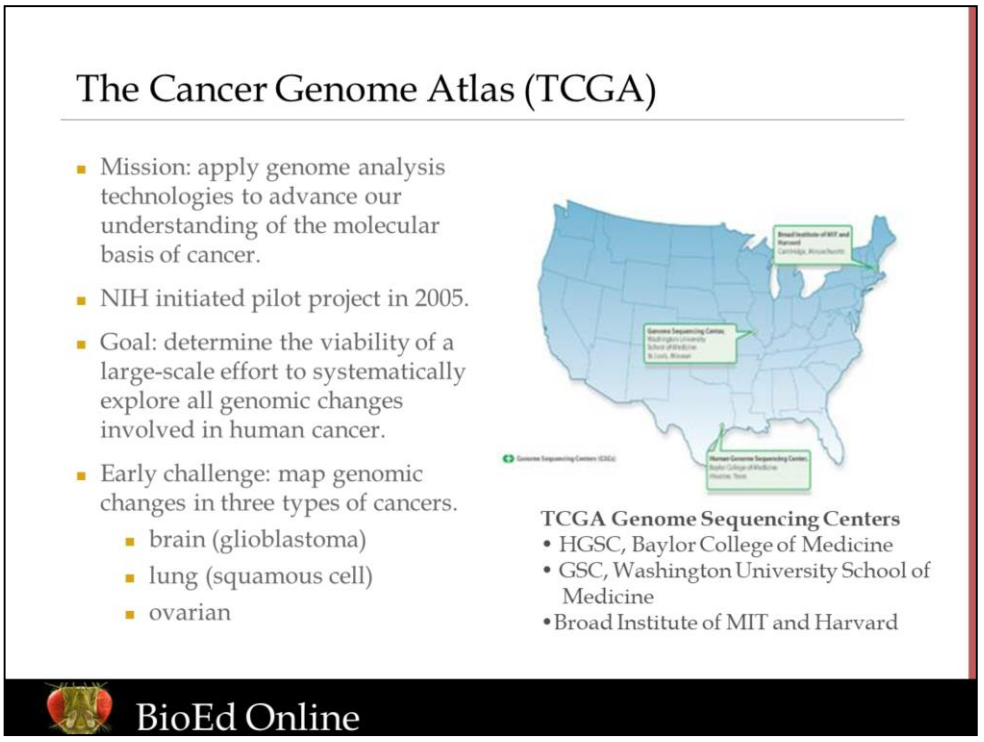
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Here's another program, called the Human Microbiome Project. We actually have a lot of this project anchored here [at Baylor College of Medicine] through George Weinstock's efforts, and he's going to talk to you, too. The idea is that since sequencing of bacteria is getting pretty routine, let's sequence them all. Now of course, you can't culture many of them, so that presents a big problem. So, the initial thrust is to sequence the 600 or so that are in the catalogs of all culturable and non-culturable bacteria, plus several non-bacterial microbes, and use this reference sequence, plus additional samplings of different body sites, different orifices and different groups of people, obtaining raw sequence to better understand the ecology of these positions in the body. This catalog is being built, and this is now an NIH [National Institutes of Health] Roadmap Program. This is going to be a big deal, and is going to impact on this whole concept of genomic medicine when we move past the human and the human genotype into other areas of sequence and those kinds of activities.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

George Weinstock. Retrieved 04-07-2008, from http://www.bcm.edu/genetics/facultyaz/weinstock.html.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

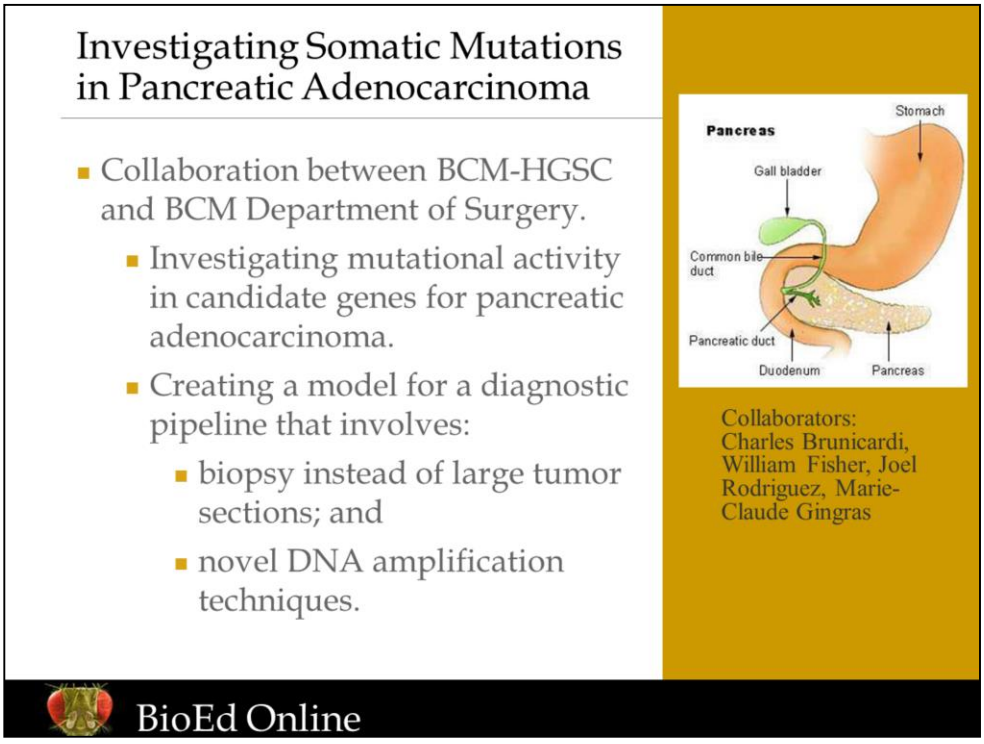
And of course, here's kind of the mother of all additional projects, the Cancer Genome Sequencing Project. This is another NIH [National Institutes of Health] initiative that was launched a couple of years ago. The idea is, "let's extend the sequencing model, and let's sequence a whole bunch of tumors. Let's extend our notion of characterizing different cancer diseases to a full and comprehensive understanding of genomic changes and the range of genomic changes that can affect those diseases. And let's use that to drive discovery of what loci might be important in those diseases when we haven't found those loci already."

This is controversial and well worth at least a one hour group discussion and meeting already. But suffice to say, it's off the ground. We have trials. Three diseases have been chosen by the larger program: glioblastoma (GBM), ovarian cancer, and squamous lung carcinoma. There's actually another whole component that was funded by a pilot before this, characterizing a lung adenocarcinoma. So this is all really rolling along.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

National Cancer Institute. *The cancer genome atlas*. Retrieved 04-08-2008 from http://cancergenome.nih.gov/components/gsc.asp



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

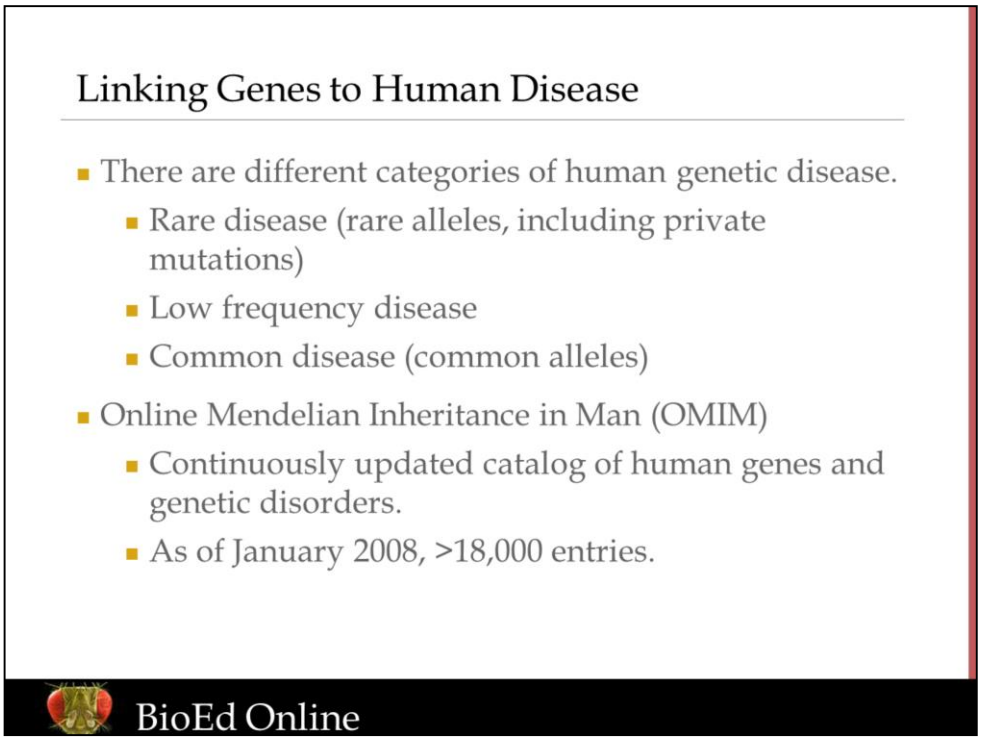
In addition to participating in those large multi-center programs, we have an initiative with the Department of Surgery [at Baylor College of Medicine] where we're looking at pancreatic adenocarcinoma. We're looking at both sampling some of these candidate genes to look for mutational activity in them, but also generating the pipeline where we can reliably and meaningfully move samples from the surgical suite to the lab to get these initial diagnostic profiles. So we're doing what I'd call a model of how one day we might have a true diagnostic pipeline that is meaningful, where you can use biopsy instead of large tumor sections, where you can use DNA amplification techniques that are novel, and have robust and meaningful results. But this is both a pretest of that model, as well as trying to inch pancreatic adenocarcinoma, and the group here of course, closer to being one of those 50 diseases at the top of the list for the national program. So those are the applications, in some of these big arenas, of using sequencing to drive the knowledge base.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Pancreas. Retrieved 04-08-2008, from http://en.wikipedia.org/wiki/Pancreas.

21



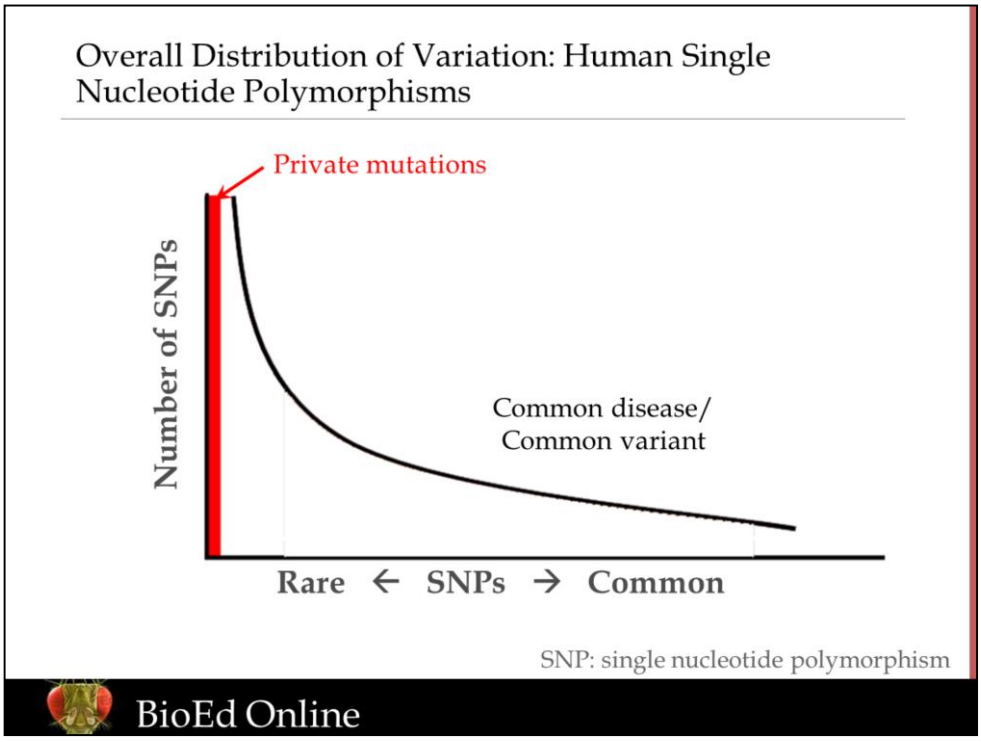
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

But what about human genetic disease? Because that is, remember, how we started here and what we're really focusing on today. It's most useful to think about human genetic disease in kind of three categories. And those categories are rare disease (rare alleles), common disease (common alleles), and the stuff in between. Now, all of you [in the audience] I know are familiar with OMIM [Online Mendelian Inheritance in Man], with the Catalog of Mendelian Disorders and Victor McKusick's list. There's something like 20,000 entries, which boil down to something less than 10,000 opportunities to find human genes where you might find a mutation and could link that mutation to a phenotype. And about 2,000 of those are actually solved now.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Reference:

National Center for Biotechnology Information. *OMIM: Online Mendelian inheritance in man.* Retrieved04-08-2008 from http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

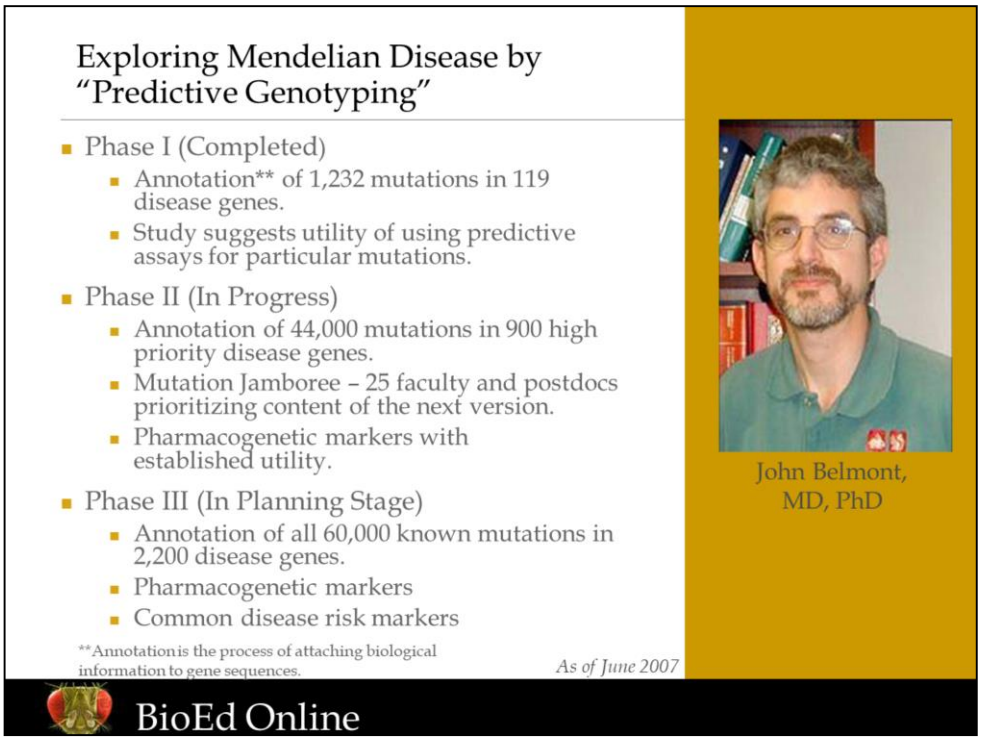
But all those disorders, or the vast majority of them, on this curve, belong over here [indicated by red bar]. So this curve is the fraction of SNPs (single nucleotide polymorphisms). On the bottom is the population frequency. So a rare mutation, one that is private to your family or your near family, would appear on the left-hand side of this graph. That's where Mendelian disorders are found, for the most part.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Gibbs, Richard. (2007). *The pathway to genomic medicine.* Houston, Tx: Baylor College of Medicine.

23



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Now in general, think about those [disorders] as a diagnostic challenge, okay? If you've got a Mendelian disorder that a family reports because a cousin had the disease, and you have no one else there, you pretty much have to do the full gene ascertainment for most of these disorders. A critical question in developing those diagnostics further is whether or not that will always be true. Will there always be a different mutation in these groups, or will there sometimes be a recurrence of certain ones?

John Belmont from our group has tackled this question. He did this because of what I just said, and also because we really don't know that much about these rare functional alleles and how they're distributed in the population, because we only ever see them when a family comes in with a disorder. Therefore, you get this real ascertainment bias of what the true frequency is. John has done two things, and he's going to report these to you in detail, but I'm going to give you a very high level view of it. First, if you look statistically across the larger databases at all these disorders, you see that although there is a preponderance of unique mutations in the databases, certain ones do occur more than once. And since we have this bias of ascertainment, you can infer that maybe those sites are either present in the population at a very low frequency for some natural reason, such as selection or population drift; or, they're actually mutation-prone sites because of the sequence context. So if you had a technology that was very efficient in screening for known mutations, you might want to include some of those in your repertoire of mutations that you screen.

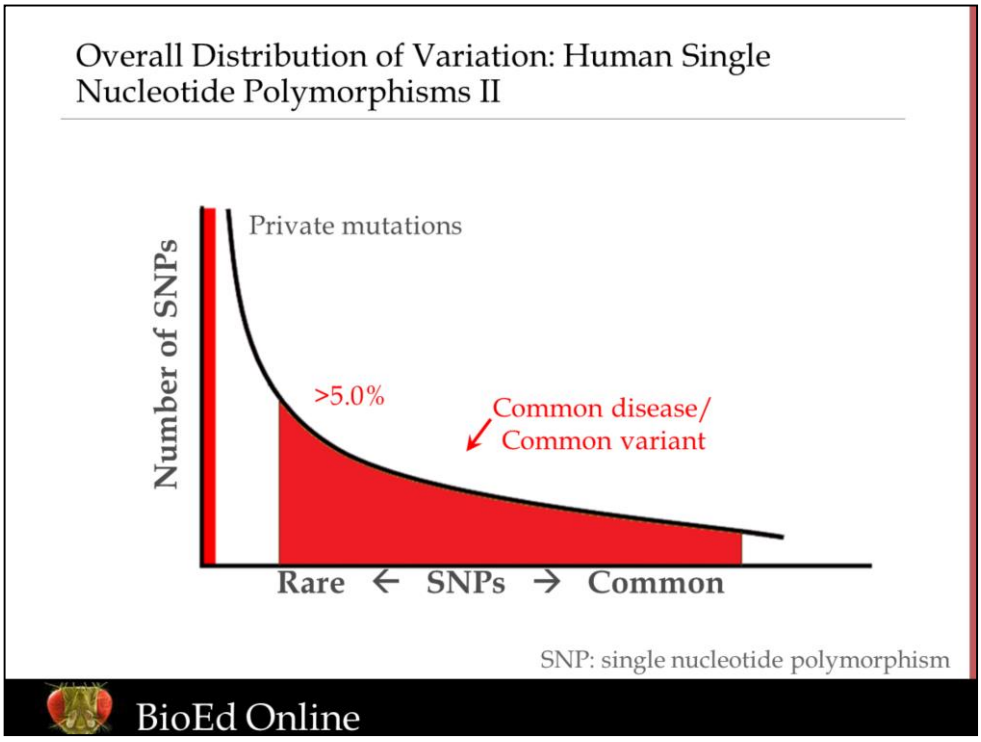
This is the frequency in certain disorders, and here's the test that we used in a large genotyping project that John adapted for this exact purpose. He took 1,232 distinct mutations from a database, mutations that were present in 119 disease genes currently screened in the Genetics

Diagnostics Clinic. But they're screened for mostly by sequencing. Generally, they just look for the de novo mutations. He made probes for those, and asked directly, were they present in different populations? And the story he's going to tell you is, he found some very interesting patterns. A subset of them [mutations], at least, were actually more frequent in different populations, and many of them turned up more than you would think. This finding is actually pretty profound, because it alters the way that you might want to generally approach screening. If you thought you always had to sequence to get a certain class of mutations, it turns out that in some cases, at least, you can perform an assay that is predictive for that particular mutation. That has tremendous implications for the implementation of this kind of work.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

John Belmont. Retrieved 04-08-2008, from http://www.bcm.edu/immuno/?pmid=1998.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

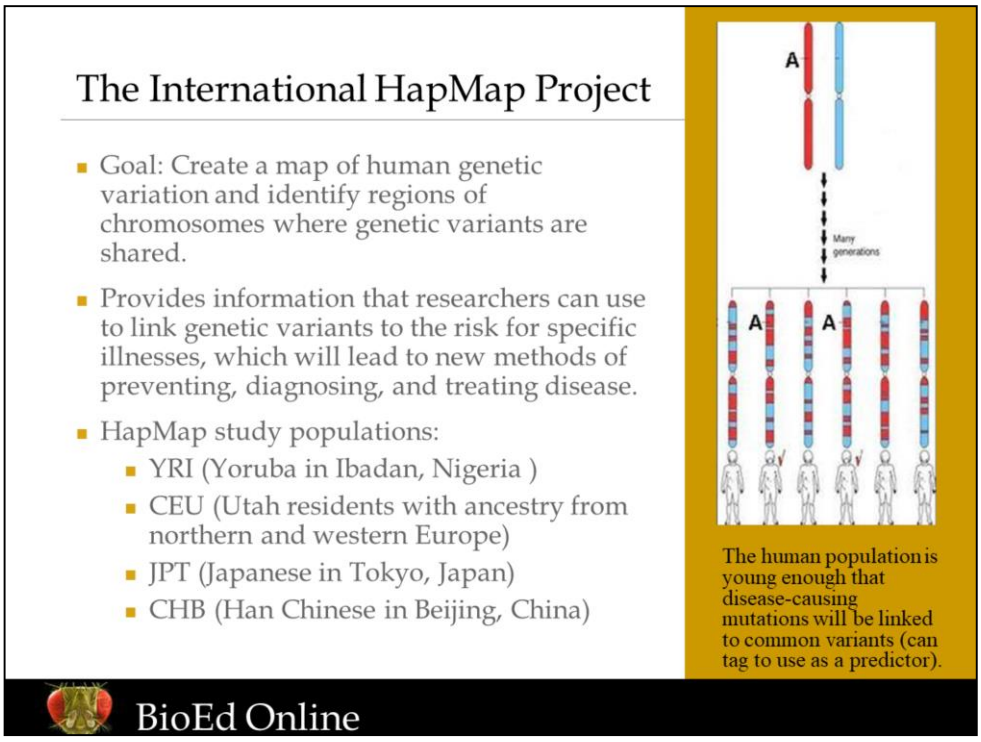
So much for that side of the curve and these rare mutations. What about the stuff on the other side, the common disease, common variant mutations? These are the ones, I think, that are driving all the excitement in this whole notion of genomic medicine, or at least genetic-genomic medicine in the first place. The idea is that you can get a predictor: go in at the age of 20, get a predictor saying that you're going to have a heart attack when you're 50, so you better do something different than if you didn't have the predictor. Now, faith in that predictor depends on this whole notion that there is a thing called a common disease, common variant hypothesis. What I just said is not quite true, but practically true, and you'll see why I qualified that in a moment. But it's the idea that there are actually alleles out there that are fairly common in the population, and if you happen to get a bad combination of them, they give you this propensity to have one of these disorders. These alleles, by their very nature, have to be common; otherwise the whole statistical model wouldn't fall together correctly.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Gibbs, Richard. (2007). *The pathway to genomic medicine.* Houston, Tx: Baylor College of Medicine.

25



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

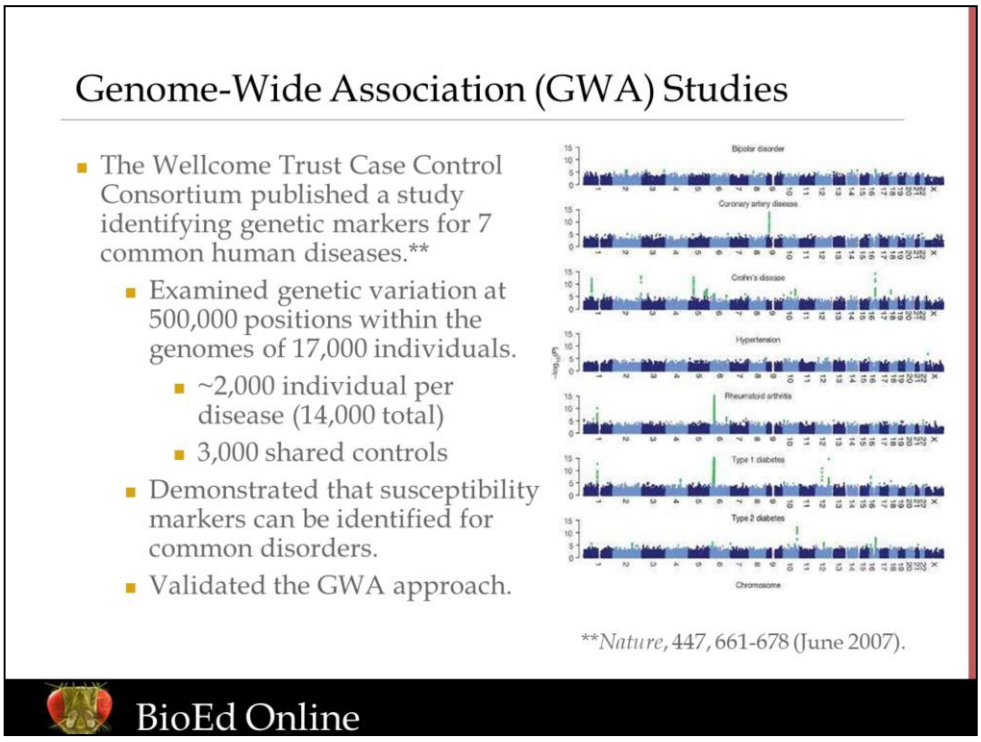
Another follow-up to the Genome Project was the HapMap Project. This project was designed essentially to allow us to test that hypothesis [see slide 25] and find out if there were indeed alleles out there that you could assign as predictive for these kinds of disorders.

The idea of HapMap was to take natural populations, reference populations, and find all the variation you could, given the available technology, then map that variation across the whole population and use it to tag different parts of the genome. So the underlying idea here is that humans aren't that old; we've only been out of Africa for 100,000 years or something: 10,000 generations, maybe less. And so our chromosomes haven't had that much time to recombine. If you picked two ancestral chromosomes, even through that many generations, in the current populations you can see large blocks that have been conserved over that whole period. This kind of genetic discovery and mapping can allow you to track those blocks in an efficient way. So that was HapMap.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

International HapMap Project. *The origin of haplotypes*. Retrieved 04-08-2008 from http://www.hapmap.org/originhaplotype.html.en with permission.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

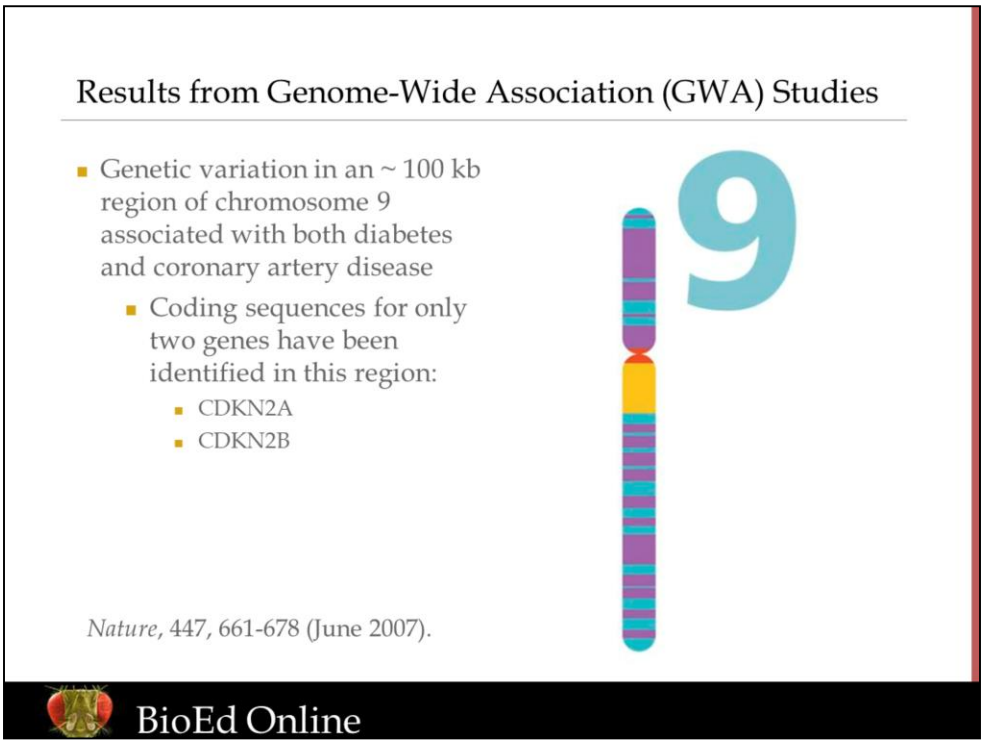
So the big experiment is this: sequence the genome, do the HapMap to find all these reference variants, and then take those variants and genotype them in critical populations—cohorts of diabetics versus matched controls, cohorts of schizophrenics versus matched controls, cohorts of Crohns' disease versus matched controls. Now that experiment has been done. We've been waiting with bated breath for this for at least as long as the Genome Project has been going. The Wellcome Trust funded the biggest study, and the one that's had the most publicity and the most success. And many of you might have seen this figure in Nature a month ago. We've all been going to meetings and celebrating tremendously. This is what happens if you do that experiment. You look for parts of the genome in these populations that have an unusual genetic conservation signal that reflects some increased propensity of that part of the genome to drive an increased risk for one of these diseases. You see those little green bars in the different disorders? Those are the success stories. Those are where there are actually signals from these disorders.

Now, once the champagne is finished and everybody sits back to think about it, it is striking, and it's stunning, and it's a fabulous result. But there's a lot of qualifiers here. You might ask, as a clinician, does that mean you're ready to run and use these markers and counsel people based upon these markers? Well, the odds ratios, that is the propensity for you to have the prediction versus not, if you have the allele or not, ranges from about 1.1 to 1.6, okay? So your frequency, your likelihood of having a heart attack, might go from 1 in 3,000 to 1.6 in 3,000, from these markers. They're really not of that much use when you distill things down to that level. What they are, of course, is incredibly useful for saying that there must be a gene under here and that gene must be involved in a pathway; other mutations in that gene might be more profound. They may be driving the disease. So this whole approach, looking for common disease, common disorders, is enjoying very good success right now, but it's still a tough game.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Reference:

The Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature,* 447, 661-678. Retrieved 04-08-2008, from http://www.nature.com/nature/journal/v447/n7145/full/nature05911.html by permission.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Here's one zeroing in on one region, this is one of the most exciting stories in the genome right now. One of those regions that was tagged was in Chromosome 9, and it's 50 to 100 kilobases. And it came up strikingly in the diabetes cohort and in the cardiac disease cohort. And it turns out, there's only two genes in the region, there's CDKN2A and CDKN2B. And they're well known for having mutations, somatic mutations, in pancreatic adenocarcinoma and some other cancers. They're the only genes in the region. So this has got everybody baffled. What's really going on here? So the hope is that either another gene, or maybe a gene controlling element, or some other part of the genomic biology here, will emerge and explain all these phenotypes and will have larger effects when there are different changes that are found in those same elements. But right now, this is the state of the study.

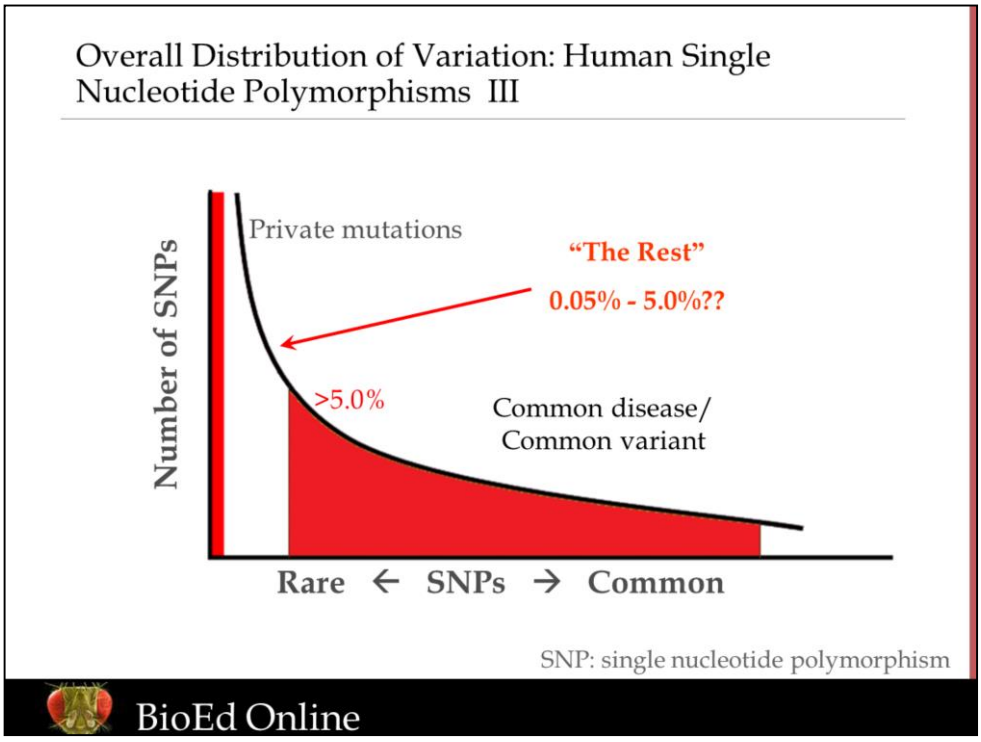
* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Reference:

The Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature,* 447, 661-678. Retrieved 04-08-2008, from http://www.nature.com/nature/journal/v447/n7145/full/nature05911.html by permission.

Image Reference:

U.S.Department of Energy. *Chromosome 9*. Genome Management Information System, Oak Ridge National Laboratory. Retrieved 04-15-2008 from http://genomics.energy.gov/gallery/chromosomes/detail.np/detail-18.html



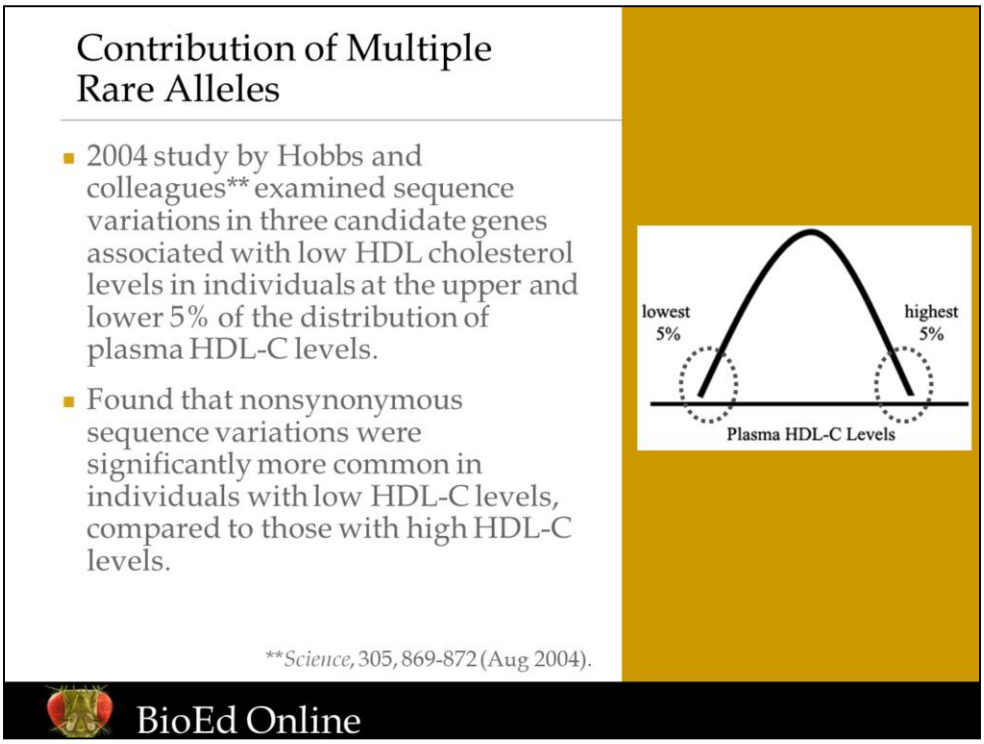
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Okay, so what about the rest of the whole spectrum? And this is where things really get kind of interesting. Because this actually turns out to be where most of the action is. Although there's been celebration of the success of the HapMap, and what I just told you, it's also abundantly clear that an awful lot of the variation is falling into this middle category. It's lower in frequency. As an individual event, it probably has more affect on phenotype, and it may therefore be more interesting, say, in the clinical context, but it's harder to find.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Gibbs, Richard. (2007). *The pathway to genomic medicine*. Houston, Tx: Baylor College of Medicine.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

A lot of excitement was generated about this class by our colleagues at UT-Southwestern. Helen Hobbs and Jonathan Cohen looked at various lipidemias by looking at various blood lipid levels. In their Dallas Heart Study cohort, they separated individuals out according to the distribution that you see. They just focused on candidate genes in pathways at the two ends of the distribution. And they actually showed, I think, maybe not for the first time, but most dramatically, that there were certain genes that simply had a lot of different mutations. Low frequency: 1% of the population, 0.1% of the population, had a lot of those at one end of the distribution, but not at the other. So if you're willing to go sequence a lot of these genes in people at the two ends of the distribution, you could start to bin them according to genes that had a higher overall mutation burden than others. And because they have the function and the biology to back this up, the story really makes sense.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

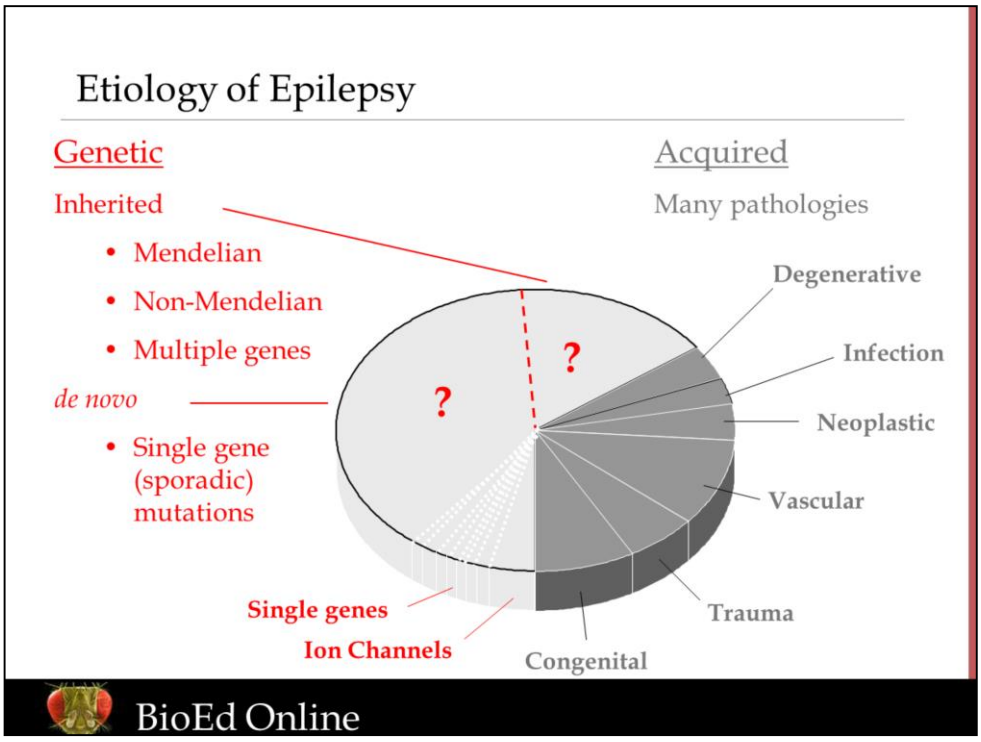
References:

Cohen, J. C., Kiss, R. S., Pertsemlidis, A., Marcel, Y. L., McPherson, R., & Hobbs, H. H. (2004). Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* 305: 869-72.

Cohen, J. C., Pertsemlidis, A., Fahmi, S., Esmail, S., Vega, G. l., Grundy, S. M., & Hobbs, H. H. (2006). Multiple rare variants in NPC1L1 associated with reduced sterol absorption an plasma low-density lipoprotein levels. *Proceedings of the National Academies of Science* 103: 1810-1815.

Image Reference:

Gibbs, Richard. (2007). *The pathway to genomic medicine*. Houston, Tx: Baylor College of Medicine.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

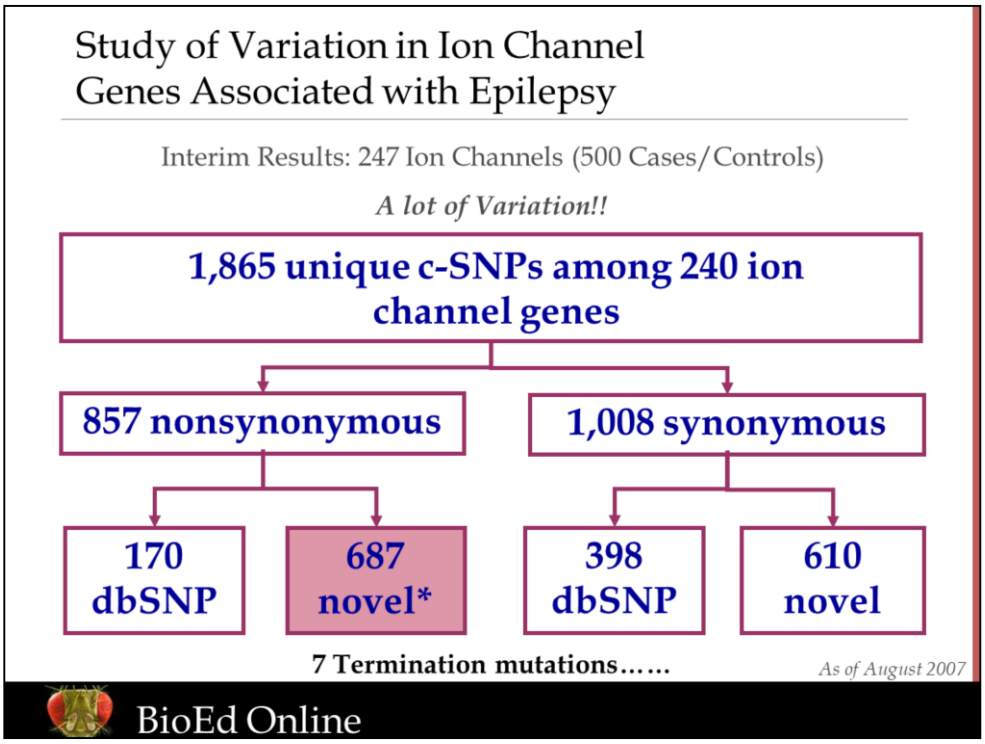
Many of us have been considering the other studies we're looking at, to try and fit them into this paradigm. Because we know that in these other studies we're not dealing with other diseases with common disorders, and we're not dealing with Mendelian forms; we're dealing with something in between. And here, we've worked with Jeff Noebels on epilepsy, and this just illustrates the range of possible causes, including some unknown genetic cause, with the biggest piece of the pie there.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Gibbs, Richard. (2007). *The pathway to genomic medicine*. Houston, Tx: Baylor College of Medicine.

31



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

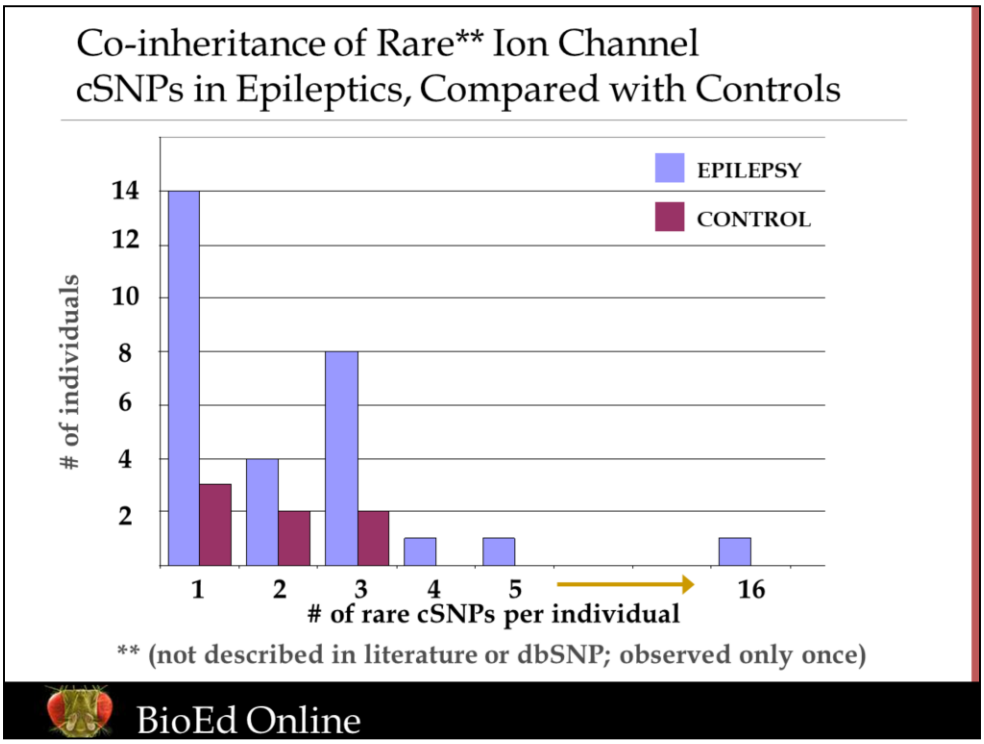
We've used the sequencing methods, sequenced about 200 people and about 240 genes that are in the candidate set of ion channels. And we have found nearly 1,000 changes in those people that might cause the disease. I said might, not that it does.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Gibbs, Richard. (2007). *The pathway to genomic medicine*. Houston, Tx: Baylor College of Medicine.

32



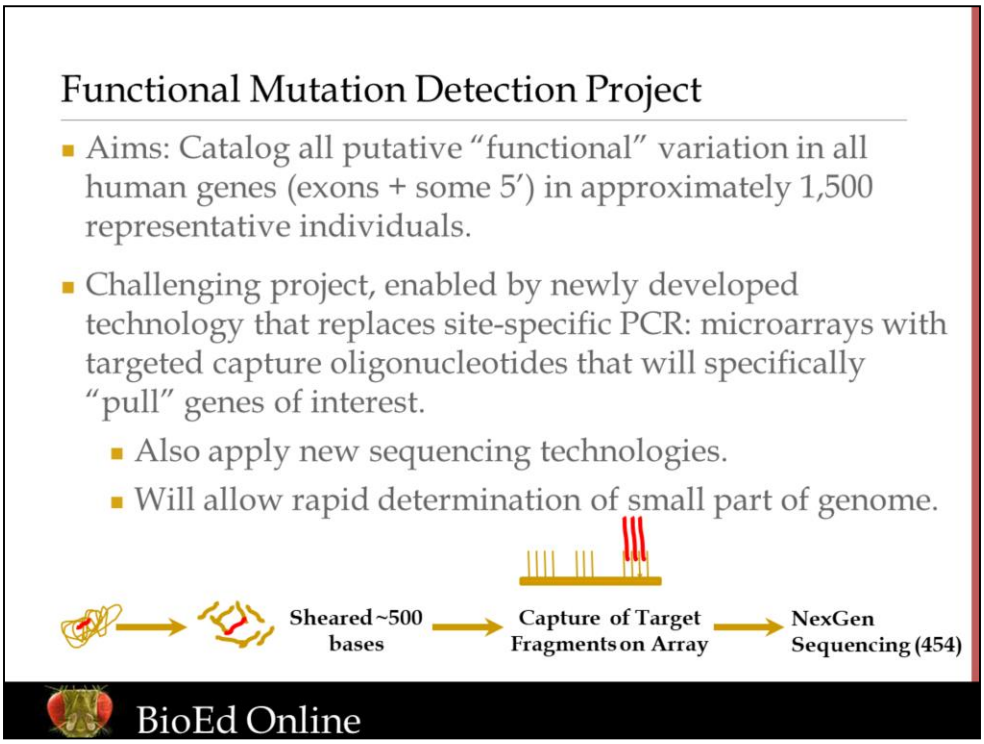
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

What we're doing now is trying to figure out, amongst those changes: Are they distributed statistically? Are there functional clues? Are these the knock-out mutations that kill the genes? Or [slide 33] are there other functional clues revealed by the number of individual mutations at those loci in the cases, versus the number of individual mutations at those loci in the controls? We're trying to distill from that whether or not we can really assign any of these genes as being part of the epilepsy syndrome that we're chasing. I want you to get the flavor of how difficult all this is. The combination of heterogeneous mutations in diseases with incomplete penetrants, with really uncertain genetic models, and trying to assign these mutations now with the quality of a diagnostic marker: that's a big gap, right? So we have a few stories like that, all these tough, unsolved genetic ones that we now have a lot more genetic variation data for.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Gibbs, Richard. (2007). *The pathway to genomic medicine*. Houston, Tx: Baylor College of Medicine.

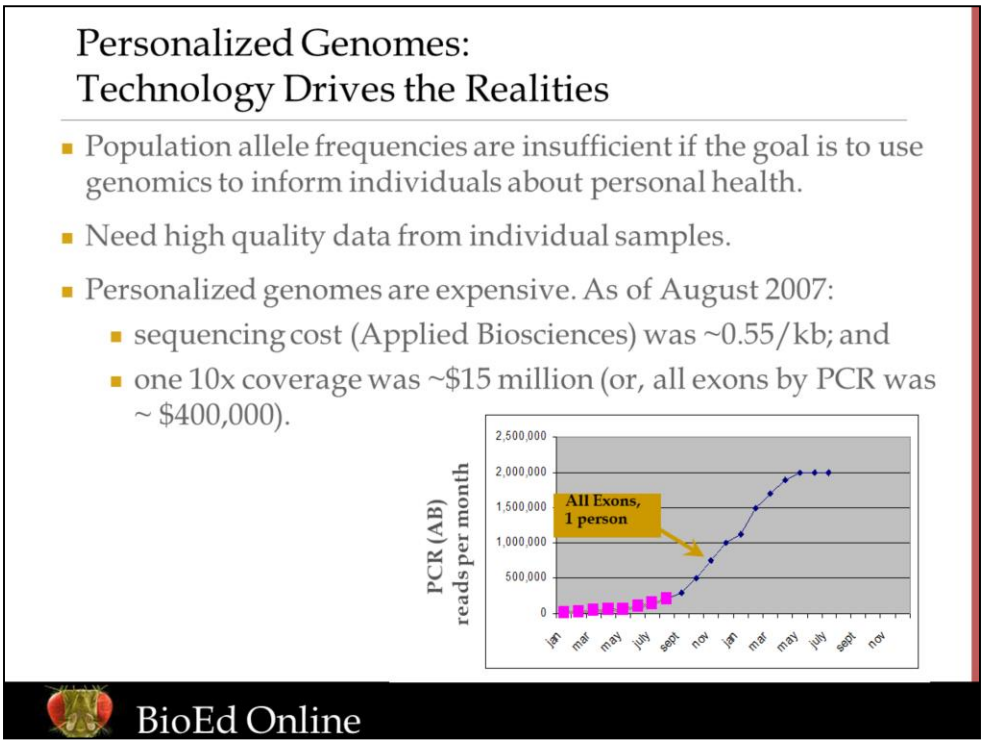


*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Now, there's some good news here, we think, which really comes from evolving technology rather than particular instances of solving the genetic paradigm. And they arise from the notion that maybe we don't have to be too obsessed with solving problems one at a time. Maybe we can actually go back to the reference populations and sequence all the exons in everybody. And then maybe, we're talking two years, maybe we can sequence all the genomes in everybody, and have this several thousand genomes-deep database that will really give us a map of these rare variants. This would make questions about what's really segregating with these diseases with very complex genetics more straightforward.

And there's even more good news. I don't want to get into technical nuances, but we've been slogging away at this with PCR, and with Sanger sequencing, just like I showed you in that slide from 1989, which is a long time ago. And only in the last six months, we have a real leap in the technology we think will change things. It's the idea we can use microarrays with targeted capture oligonucleotides on them that will pull out the genes of interest and allow us then to take these fragments off the array and put them into the new generation of sequencing technologies. Thus, giving us very rapid determination of the small part of the genome we think is the target for these groups. So that's the good news. We think we can build this big catalog.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



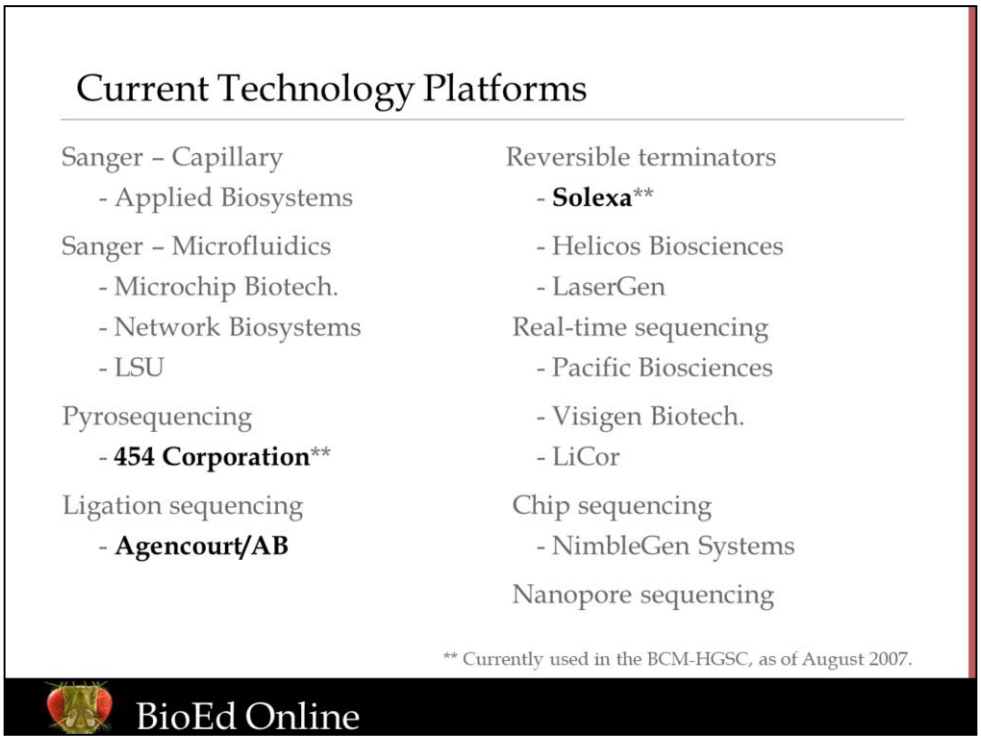
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Okay, so that's one, two, three. That's the rare alleles, the common alleles, and how we're tackling this group that's in between that's so important. But the thing about allele frequencies is that when it comes down to individuals, or you, or your patient, the frequencies are really either 1 or 0.5%, right? You're really not talking about a population question. You really have to do some kind of ascertainment that gives you very high quality data from individual samples. So within all this muddle of different activities, we have been steadfastly focusing on this notion: when can we deliver on this idea that an individual could come in and get a meaningful genome of his or her own? A snapshot of the cost of that today is probably—this slide's out of date—probably only $15,000,000 there, to do it the whole genome way. But to do it with the exons is about $400,000. Okay, if you were to get all your 300,000 exons—and we're missing a lot of the stuff people always say is important, the intragenic/regulatory stuff—by PCR and Sanger sequencing, about $400,000. And we could probably do it in about a month upstairs. So that's clearly too expensive.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

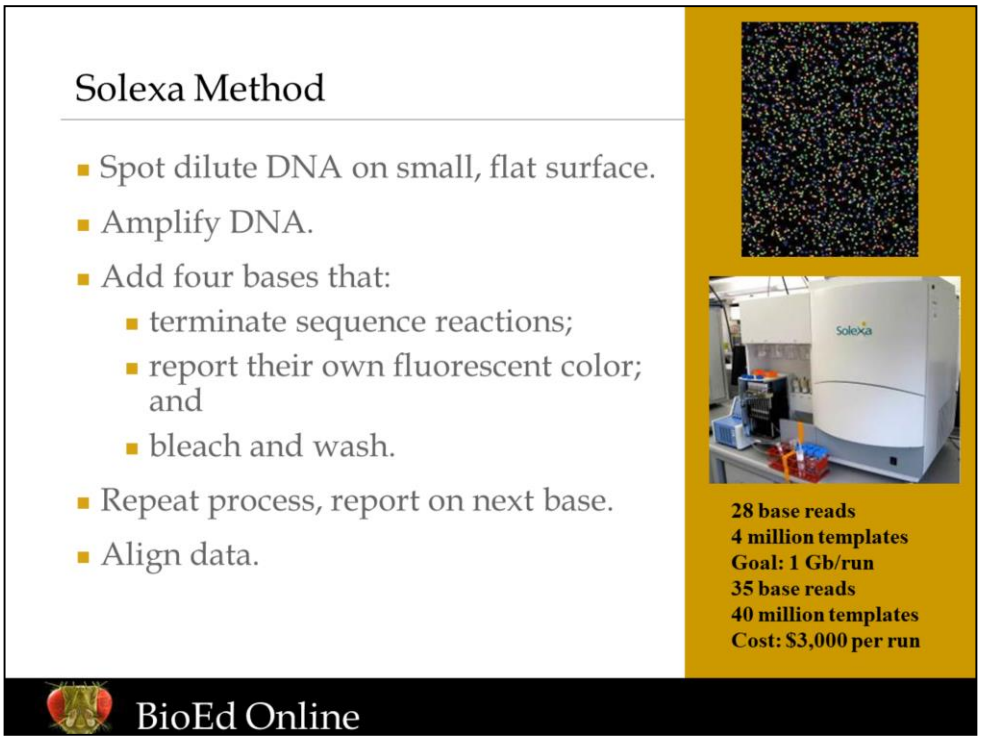
Gibbs, Richard. (2007). *The pathway to genomic medicine*. Houston, Tx: Baylor College of Medicine.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

The big question we have been facing is, where are these new technologies going to fit, and how will they enable that kind of effort [personalized genomes]? Here's the current repertoire of the new technologies. And there are only, well there are actually three now which are having some impact. The Agencourt/AB, at the bottom, is probably going to be a major player in the near future. But the two we have experience with now are called Solexa and 454 Corporation.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

The Solexa method is that—these ideas have been around for a long time, but they've just been enabled by clever chemistry and clever enzymes—instead of doing a Sanger reaction and using a gel or a capillary, you have a very small, flat surface, and you spot DNA on that surface in a very dilute fashion. And then by a surface-linked polymerase chain reaction, each of those molecules is amplified and grows to form a spot around the point it landed on. So then you have sequenceable molecules spaced through this little chip. Then you come along with a set of four nucleotides that, through very clever chemistry, terminate sequence reactions, report their own fluorescent color, and then by other treatment, can be made to bleach out the color and let the reaction proceed.

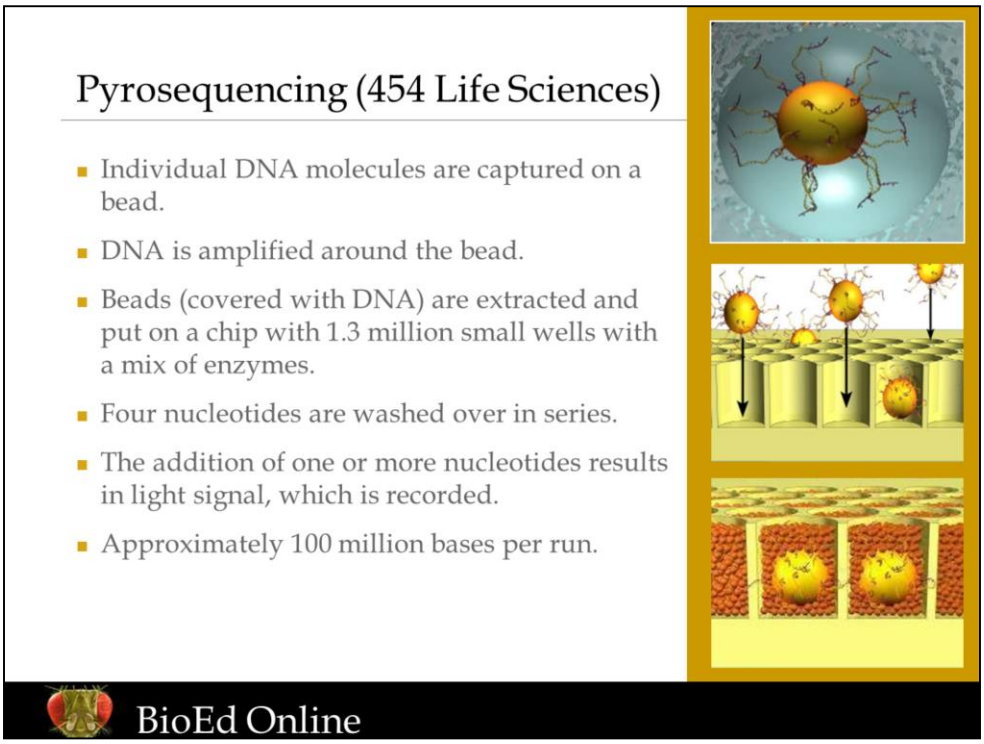
So you end up with this big field of molecules that you can treat with all four bases, which will be turned a particular color depending on what the next DNA sequence base is. You bleach it and wash it, then do the process again; the next base is reported in each molecule. It's quite amazing technology. Now, what are the problems? It takes three or four days to do a run on these machines, the error rate is really quite high right now, and there are a lot of general practical problems in implementation. The chemistry and enzymology is very difficult. But this gives you an idea of where things are going. This machine, when it works properly, could produce a billion bases in a run for about $3,000 [as of Aug., 2007].

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image References:

Steve Scherer. (2008). *Human Genome Sequencing Center*, Baylor College of Medicine.

Smith, Catlin. (2005). Genomics: Getting down to details (*Light fantastic: the first cycle in a round of sequencing by Solexa's method*). *Nature*, 435, 991-994. Retrieved 04-08-2008, from http://www.nature.com/nature/journal/v435/n7044/full/435991a.html with permission.



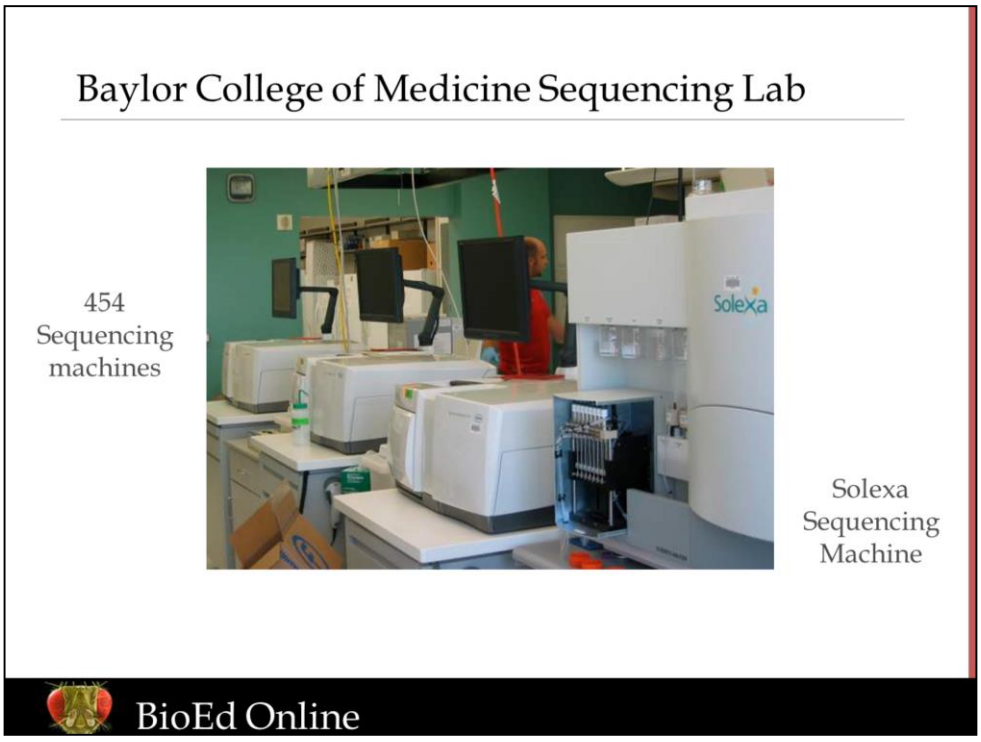
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

The other option that's really out there is the 454 machine, which uses a slightly different approach. Individual DNA molecules are captured on a bead, a very small bead, and then amplified around that bead in a test tube filled with oil and water. Little aqueous bubbles covered in oil form a little reaction chamber and allow that one molecule to amplify around that one bead. So you take your DNA, shear it, put it in a tube, extract the beads, and then put them into a flat chip that is actually 1.3 million wells with little etchings on it—it's glass fibers fused, shaved, polished, and etched; it's quite extraordinary. So those big beads in the bottom there are actually the beads with the DNA. The little beads are just packing beads. So that thing is pushed up against a camera, and the chemistry is done and you get to produce the DNA sequence. You get maybe 100 million bases per run. Those old machines from Applied Biosystems, 100,000 bases. So you see the growth in the technology.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

454 Life Sciences. *Sequencing*. Retrieved 04-08-2008, from http://www.454.com/enabling-technology/sequencing.asp.



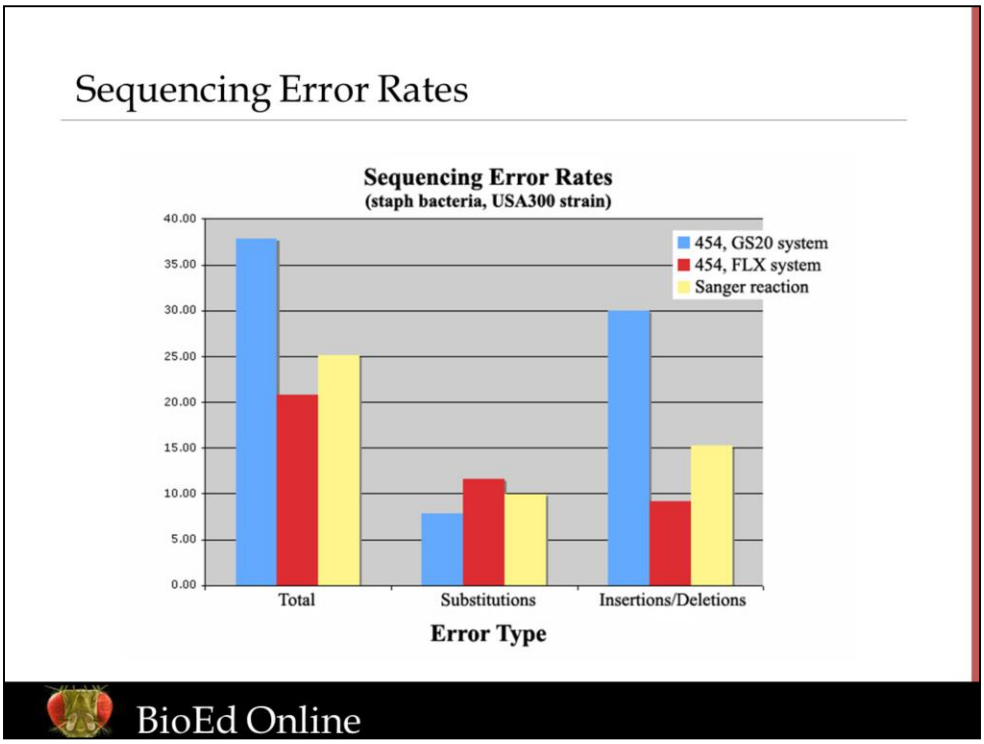
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

These are the 454 machines in the background, and the Solexa machine in the foreground. They're upstairs if you want to come and see them. And it works pretty well.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Steve Scherer. (2008). *Human Genome Sequencing Center*, Baylor College of Medicine.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

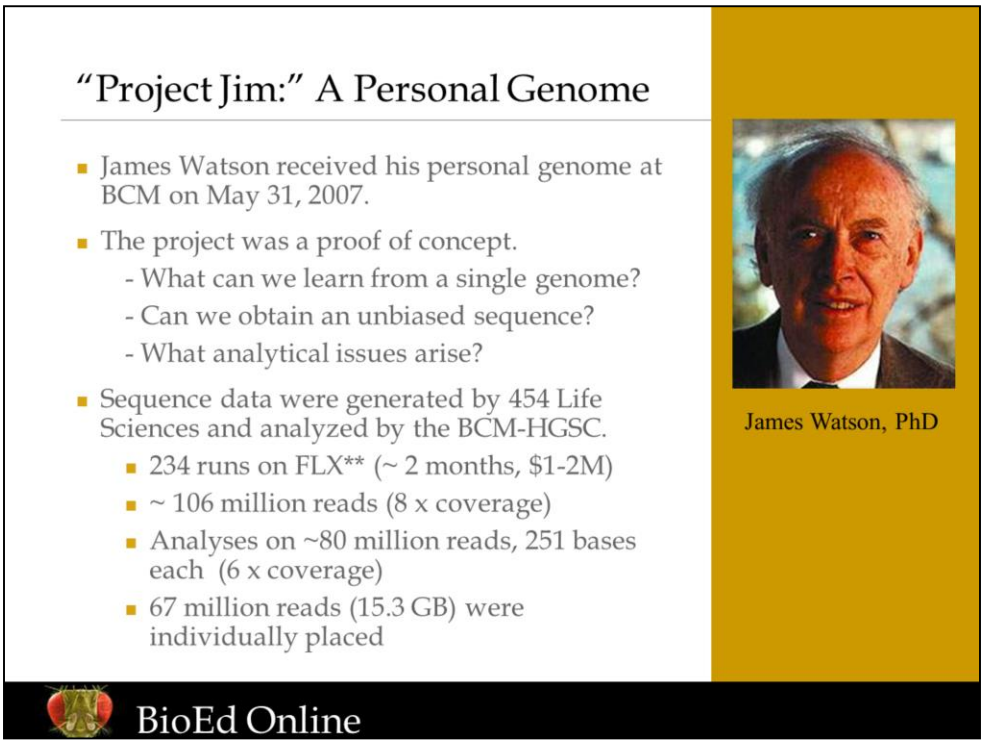
We have a lot of data on the 454. These graphs show that the error rates for the machines are down on the low end. Depending on the particular genome that's used, the 454 can give you a product that's what we're used to getting from the Sanger reaction.

454 = 454 Life Sciences

*Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Gibbs, Richard. (2007). *The pathway to genomic medicine.* Houston, Tx: Baylor College of Medicine.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

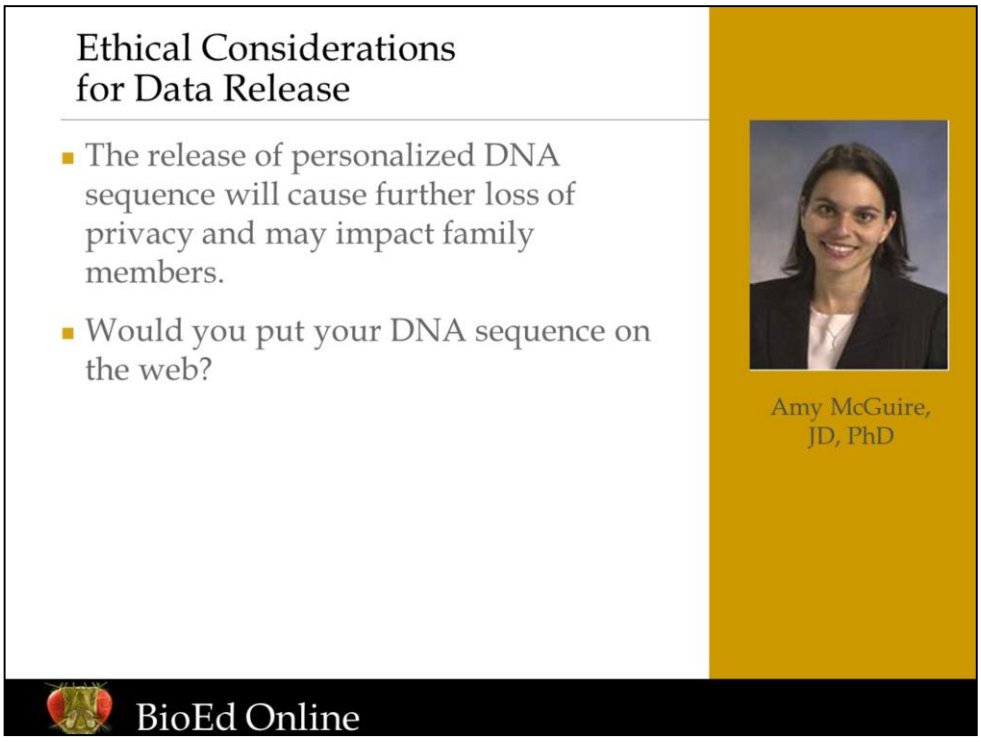
So this is what we used in the Watson Project. Our collaborators at 454 did the actual sequencing and paid for it. They (454) did the Watson Project as a stunt project to get some publicity, which sure worked. Whatever else you say about the project, that part was certainly successful. They generated all the data and allowed us to do the analysis.

**454 Life Sciences Corporation/Genome Sequencer FLX System (high throughput)

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

What a long, strange trip it's been…(Watson image). (2008). *Nature*, 409, 756-757.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

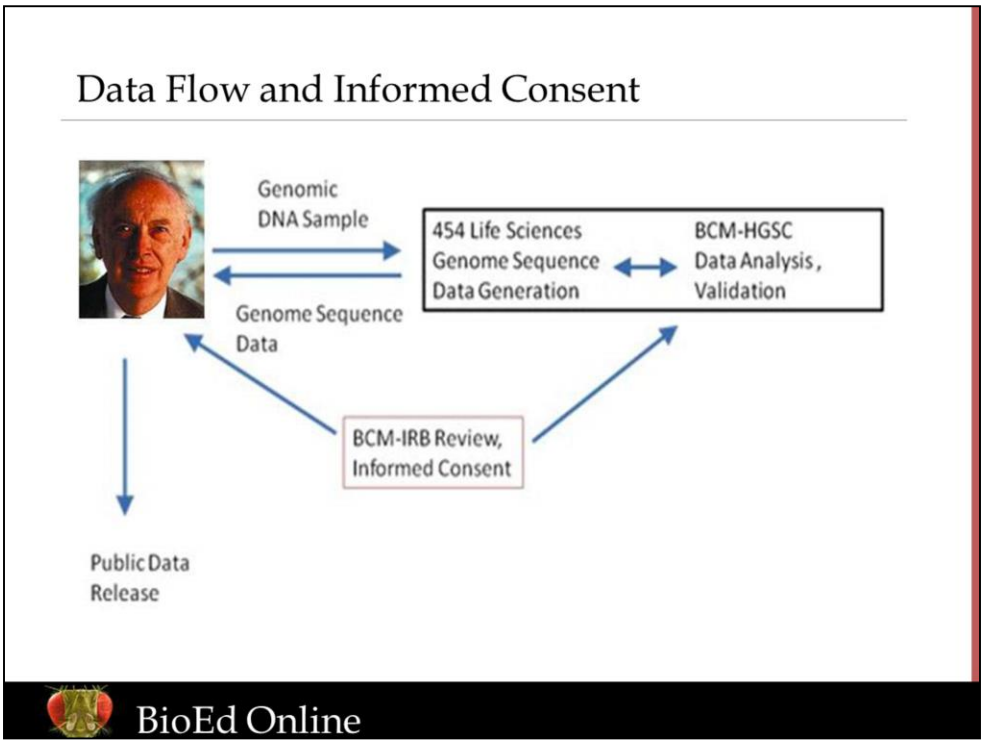
Now, Amy [McGuire, JD, PhD], who's in the audience, and who's going to talk to this group a little later, helped us manage the ethical concepts of data release with the [454] company. One of the premises of the project was that the participant would release his or her data—or would be inclined to release the data—I've got to choose my words very carefully, that wasn't an absolute requirement. But it should be somebody who would be inclined to release data to the public arena, just as we have been doing in the Genome Project all along. And Dr. Watson certainly was willing to do that. Through the course of the project, a lot of discussion emerged about really what that [releasing data] meant, and that it might not be as benign of a thing to do as you might think. So let me do the quick quiz: who here would put their data out there, complete DNA sequence on the Internet, without any qualms or thoughts tomorrow, if we offered to sequence you in the Genome Center? Are there any takers at all? You three? Okay, let's sign you up.

The risks of course, the model is that the risks aren't fully known. You may be incurring some problem for your family. Because remember, you share your DNA with your family. And you might be showing their propensity to do certain things that you might be completely proud of, but they might regard as something they don't want out there. So in any case, it's a complicated question.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Amy McGuire. Retrieved 45-08-2008, from http://www.bcm.edu/news/packages/mcguire-bio.cfm with permission.



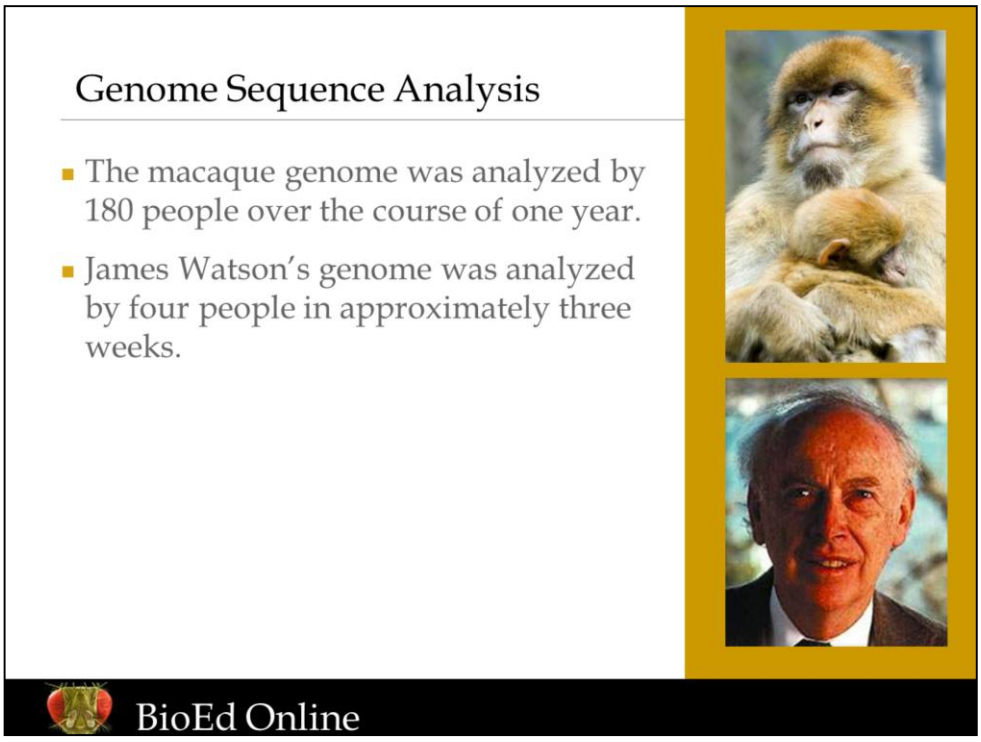
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

So what we are able to do, and Amy designed this data flow, was to make it clear that the submission to the public databases would be from the individual whose sequence was obtained. So here, we're able to do the analysis and validation of the data, and it was passed back to him [Dr. Watson] for his release to the public domain. Actually, someone sent me a link. You can actually download it right now if you want from the NCBI. It's very accessible.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

What a long, strange trip it's been…(*Watson image*). (2008). Nature, 409, 756-757.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Now, we did an analysis here in the Genome Sequencing Center at Baylor College of Medicine, and this slide's just to remind me to tell you that four people did it [the Watson genome analysis] in three weeks, as opposed to the macaque genome analysis, where 180 people actually did it over the course of a year (August, 2007).

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

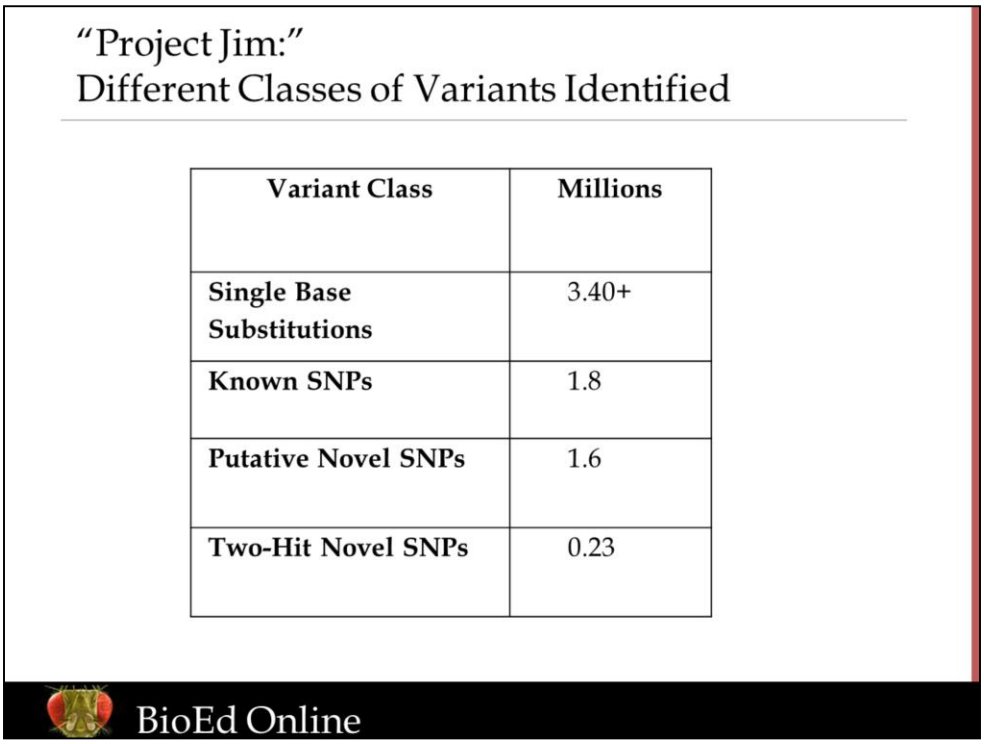
References:

Rhesus Macaque Genome Sequencing and Analysis Consortium. (2007). Evolutionary and Biomedical Insights from the Rhesus Macaque Genome. *Science*, 316 222-234.

Image References:

Macaque # 2034363. © 2005 - 2008 123RF Limited, and its licensors. All rights reserved.

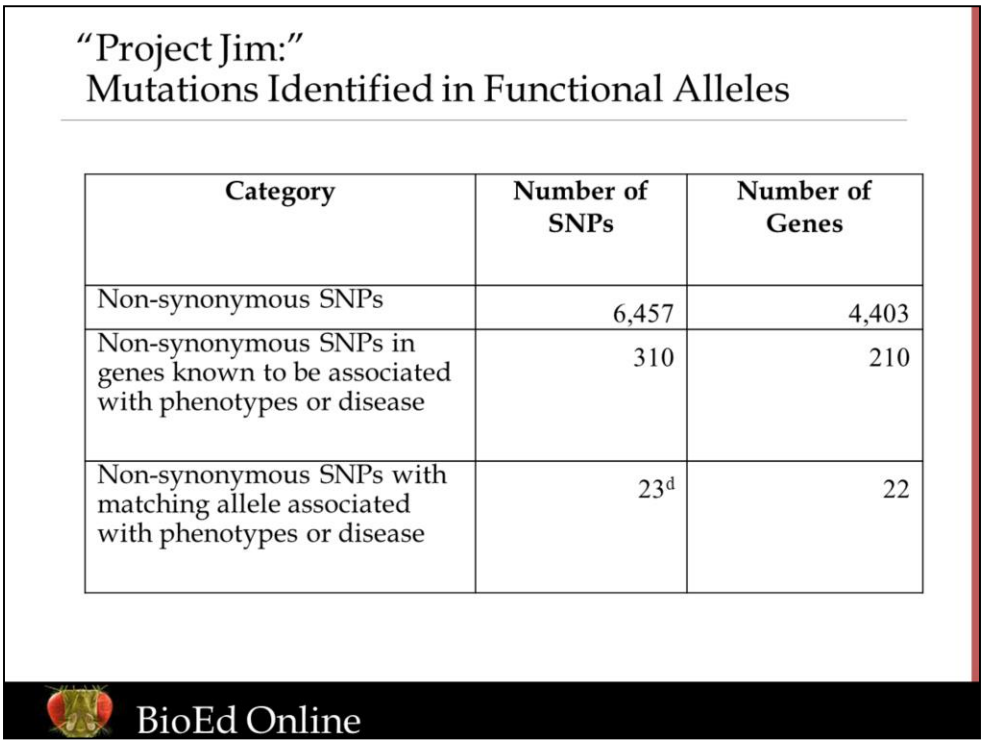
What a long, strange trip it's been…(*Watson image*). (2008). Nature, 409, 756-757.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

The analysis we got from Watson was less comprehensive in many ways than what we're able to do with the other major genome projects. Nevertheless, the fact that we were just aligning him to a reference sequence made many of the questions easier. But here's the very high points of what we found. We found a lot of individual based variants. We've actually redone these alignments, and the number's a little higher than 3.4 million in the complete data set. So we've got the expected number of overall single base variants.

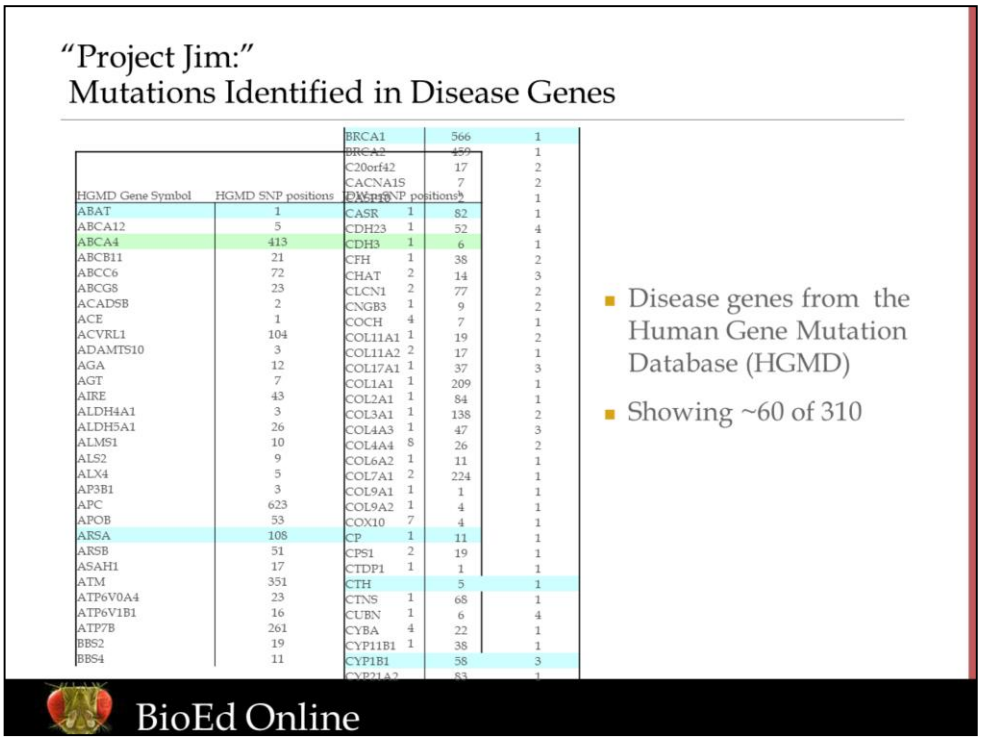
* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Amongst those [VARIANTS], though, about 6,000 clearly have the potential to alter the function of those genes. Those of you who think about mutation frequencies and mutational burden may or may not think that's a high number. Amongst those, 300 or so clearly were in genes that were in the disease list. And then, 23 of them were actual alleles that were known to have been reported before as bad alleles.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

What was really exciting about the project? Here are some of the genes. I bet if I leave that up, some of you will find genes that you already know quite well, and will wonder which mutations were in this individual. A really positive thing about the project was the taxing nature of it for the genetic counselors here, who are the best in the world, frankly, at single locus diagnostics. If you take a family with a disorder that someone in the genetics group here is familiar with and tell them about the mutational spectra, there are no better people to go to talk to about those mutations. But if you give them a dataset or a spreadsheet with a minimum of 300 genes and a maximum of maybe 6,400, where there are potential problems, that's a very different paradigm for them to deal with. But in fact, it's the one that we're leaning towards and moving towards as this work emerges.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

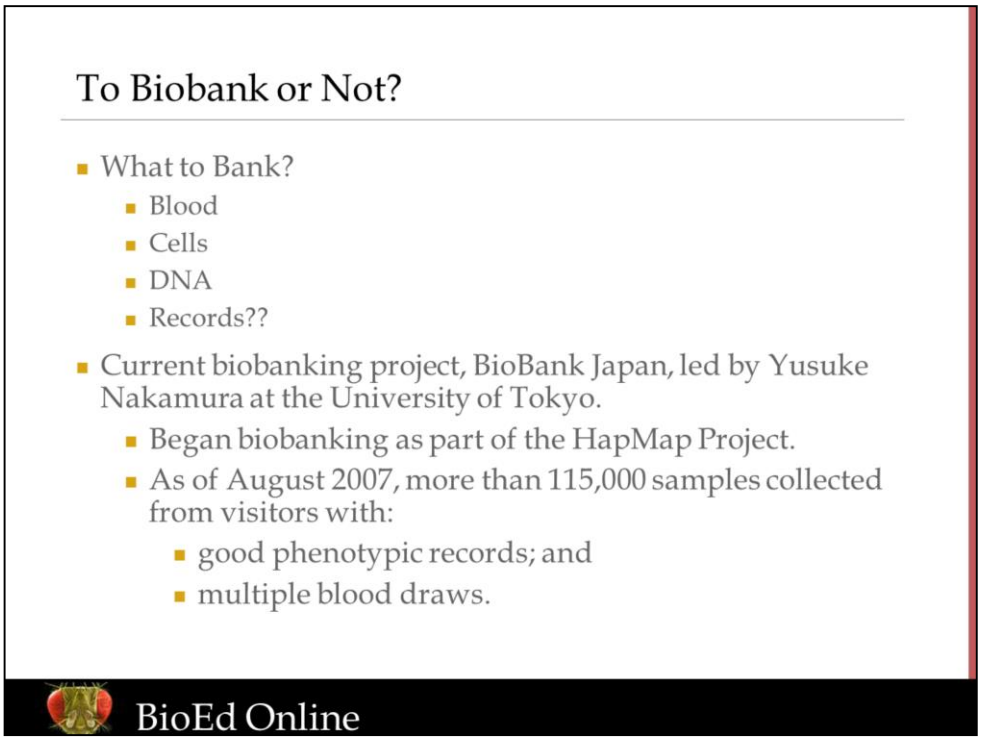
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

So here's our summary of the Watson Project that really has pressed the model: the concept of a genetic test, just how many mutations you can find if you look across a whole genome, and how indeed you simply handle that data—how you show it to the individual, how you share it with each other. So, it was a good demonstration project.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

Sorrelle, R. (2007). Nobel laureate James Watson receives personal genome in ceremony at Baylor College of Medicine (Image of *James Watson*). From the Labs. Retrieved 04-08-2008, from http://www.bcm.edu/fromthelab/vol06/is5/0607-1.html

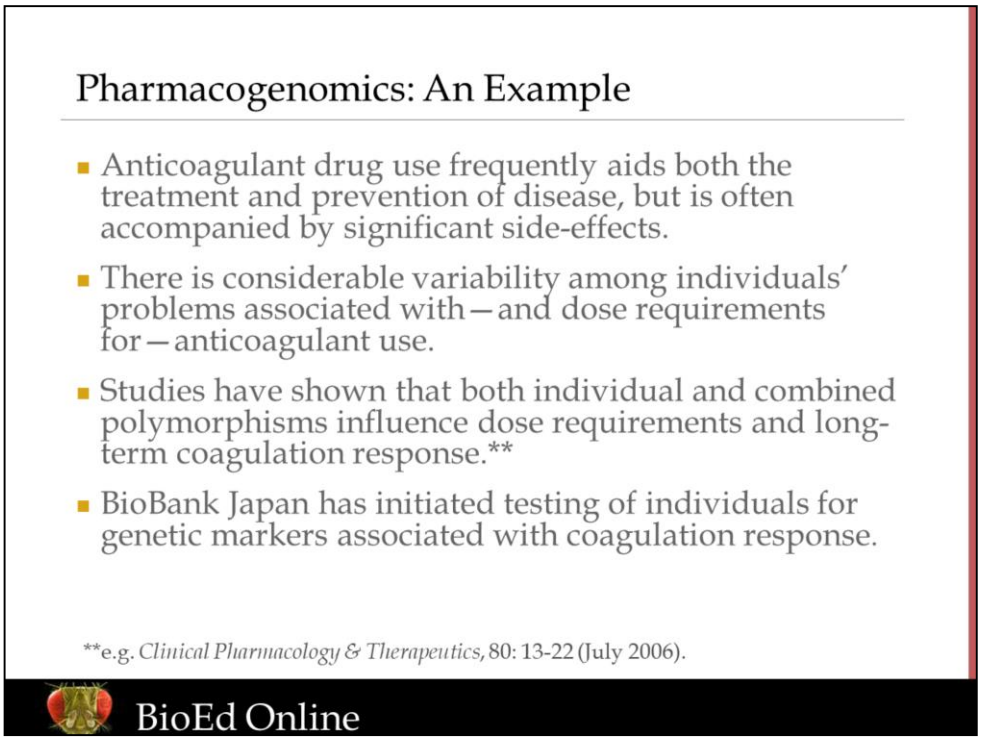


*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

I'm nearing the end of my presentation here. I wanted to bring it back though, for you to reflect on all of it. What should we do, then, to catch this wave that is coming, or crashing, depending on your perspective and how mature you think the data set is.

One of the most interesting sets of discussions that have been going on in the College [Baylor College of Medicine] for at least a couple of years has been the different efforts to collect samples in a more systematic way and to make them more available to investigators. Several of you in the room have been involved in those discussions or you have your own collections, and you have talked with others about how you might share them and how you might bank them. I wanted to encourage you to go and look at the University of Tokyo and BioBank Japan. Professor Yusuke Nakamura, as part of the HapMap Project, began biobanking in their clinic. He now has more than this—I just pulled that number down last night—115,000 samples from visitors and all with good phenotypic records, and all multiple blood draws. So they have 115,000 individuals with multiple draws from the same individual. Therefore, they have DNA and have multiple samplings over time. I don't know if they have individual tissues. But I do think these are a set of questions that we want to raise here for consideration. Maybe we should be collectively encouraging the College to put resources into this. [August, 2007]

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

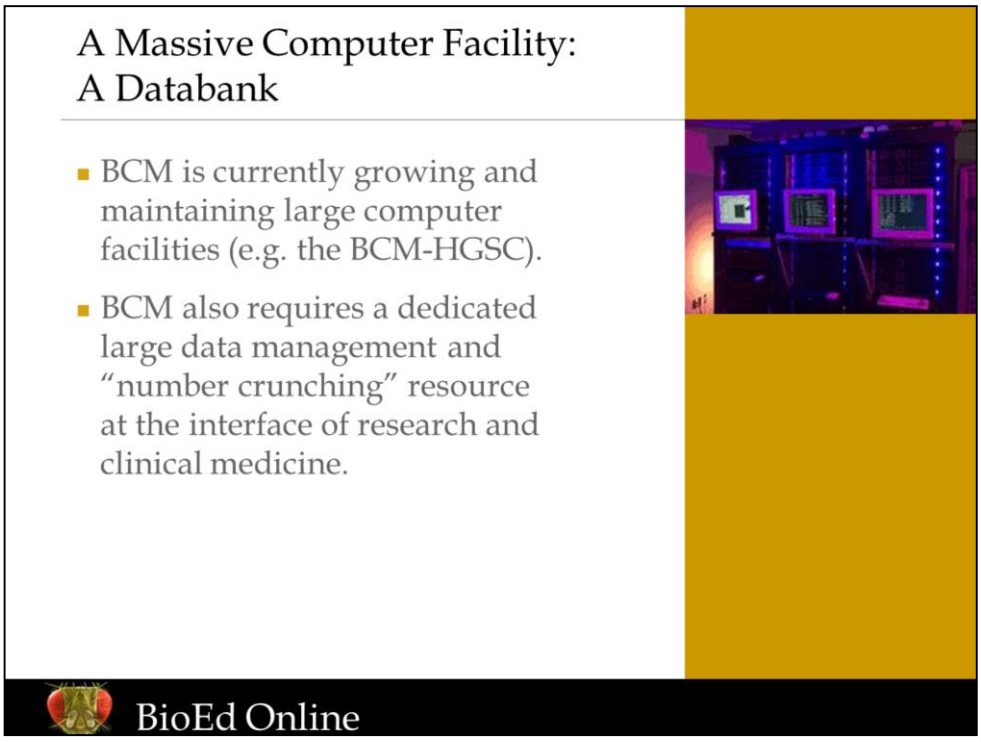


*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

I saw him [Professor Yusuke Nakamura] at the HUGO [Human Genome] Meeting in Canada last year, and he talked about warfarin sensitivity. And of course, I'm probably the least qualified person in the room to talk about this, but I discussed it briefly at a meeting last week, and I said I understood it was an empirical science to approximate the correct dosage or how much of this class of drugs you would give to particular patients. And the response was, "No, it's not a science. It's not even empirical." So I guess it's tough, right?

Now there are known genetic markers that influence, if not the direct response, then the long-term response to these drugs—the long-term coagulative response—and there's good literature on that. The point is that in Japan, they are certainly typing these individuals for those markers ahead of time, so that as part of their clinical record, they can anticipate what might influence their therapy. What a good idea!

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



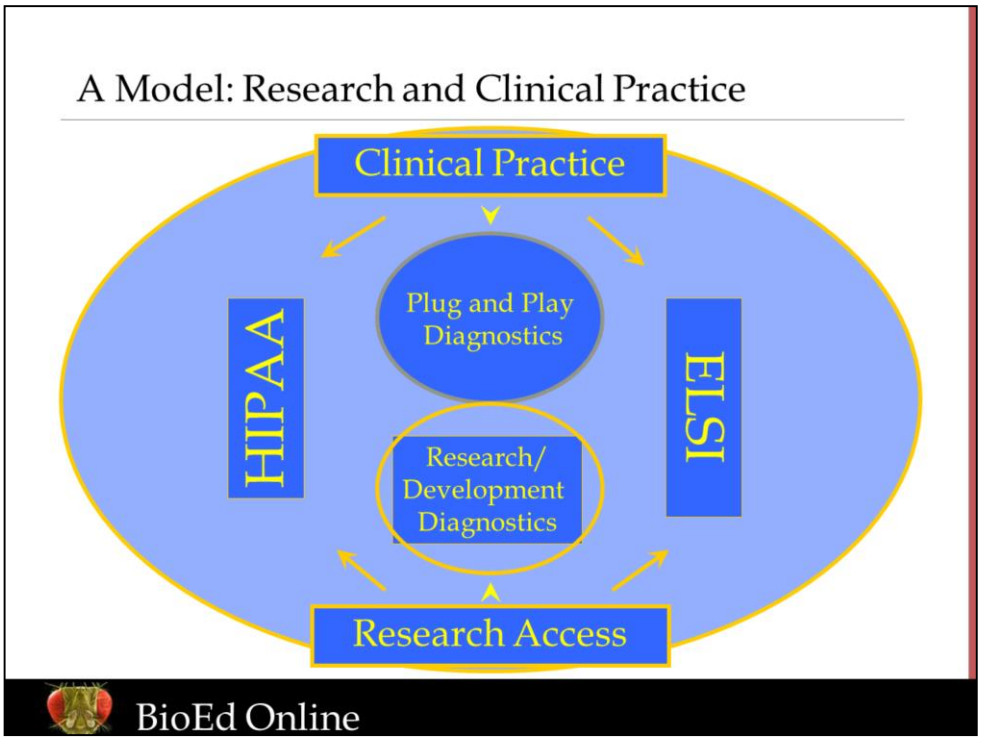
*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

The other thing is that we've had some good discussions (and I hope some of you have met Tony Elam—he's helped us a lot here too) about expanding our overall computer capacity and coordinating our data farming usage. There's no question that the Watson Project brought this home more than anything: data crunching and data handling is still a wall that we're always pressing up against. I think as part of our future, and as we move into the new [Baylor College of Medicine] clinic, it makes perfect sense that we should make some kind of long leap, between now and then, in our capabilities for doing that, just to give us the elbow room to do what we need to do as these kinds of data start to flow in.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.

Image Reference:

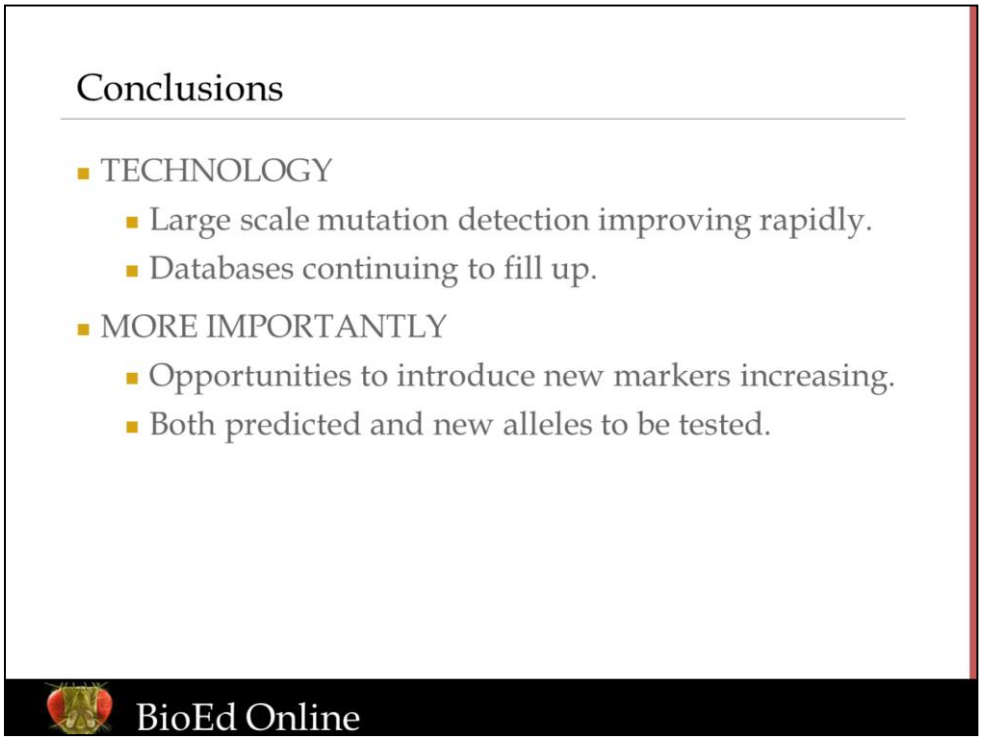
Informatics. Human Genome Sequencing Center. Houston, Tx: Baylor College of Medicine.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

Here's something of a model of how this might fit together. I'm sure many of you could add to the model. The idea is that if you were facilitated in your clinical and clinical science practices for sample archiving, and doing it in a way that was commensurate both with the patient record requirements and with the ELSI requirements, that would be a great asset both to you and to the research community here. The community presumably then could, through a proper but minimal development of protocols, access those sample collections in collaboration but without the onerous mechanics that are currently needed. They could access those samples and the accompanying data in order to do research activities that are not moving mountains, but are simply accessing this large body of data. I think that sounds like a pipe dream, but that, to me, is almost a prerequisite for pushing us into this next phase. Of course, the notion is that genetic diagnostics, I call it "plug and play diagnostics," and those assays that are already calibrated and working, that you can access, would be an intimate part of this. But really, we want to press that up against these developing techniques. I mean, you can't get a genome sequence now, but the minute you can, you'd want to have it somehow running through that "plug and play" arena there so it can be properly used by the clinician. So I think these two elements are really things that we all might want to press on.

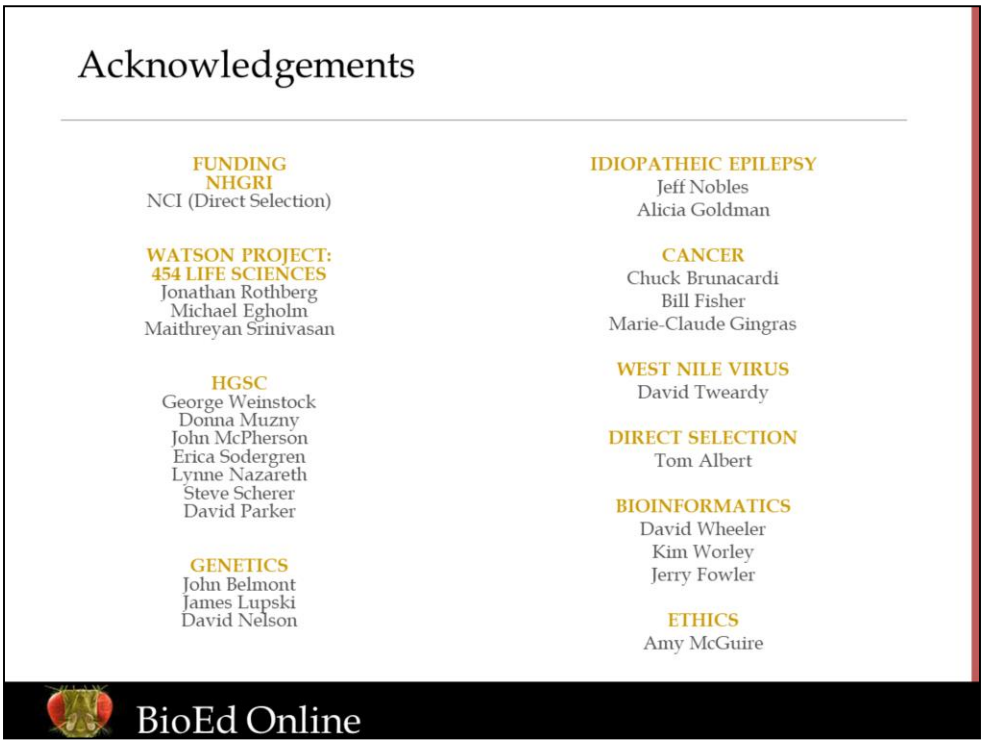
* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



*Edited Transcript from "The Pathway to Genomic Medicine," Richard Gibbs, PhD**

So in grand summary, then, I would conclude that things are moving pretty quickly. I don't think we have that many home runs to really point to in terms of alleles that were not known yesterday, but today everybody wants to have them screened. I think this groundswell is truly happening, and I think we can just point to the databases and the literature to show that that's true. We should brace ourselves. Now, thank you everybody, and of course there's many people here, and I hope that you know or might get a chance to meet a few who've worked on these different areas. This is the Genome Center Staff. Thanks a lot for listening.

* Notes in this slide presentation are adapted from the transcript of "The Pathway to Genomic Medicine," a presentation by Richard Gibbs, PhD, given in August 2007, as part of Baylor College of Medicine's Department of Medicine Grand Rounds Human Genetics Symposium.



"The Pathway to Genomic Medicine," Richard Gibbs, PhD