

Leslie Biesecker
American Society of Law Medicine & Ethics
DNA Fingerprinting and Civil Liberties Project
RO1-HGH002836-02
Workshop 3 Presentation
May 3, 2005

Dr. Biesecker: Thank you very much for inviting me. It's a pleasure to come up to Harvard Square. Just this morning, I left Cold Spring Harbor Laboratory where we are planning the next phase of genomic research in human subjects. Some people might find that a little startling but we are plowing ahead and I'll mention that throughout the talk. I want to make two disclaimers at the outset. The first is that I'm not representing NIH policy in any regards here today. These are my personal opinions. And even though this meeting is sponsored by ELSI, ELSI cannot pay for my travel support today so I did that independently.

What is medical genomics? What are we trying to do? What are we trying to accomplish with these samples that we collect from our patients? The goals, very simply, are to diagnose and treat patients with disease. We do that in two ways. First, we have to understand the cause of diseases. It is important for people to recognize that we don't understand the spectrum of human genetic disease. There are thousands of disorders for which we have only the most rudimentary clinical and molecular understanding of disease. It is common for a third of the patients who come to see a physician or specialist such as myself to leave the clinic without a diagnosis. A third is a huge number of disorders that we don't yet even understand what the disease is, what the cause is and what the prognosis is.

The next challenge, after understanding the etiology and what the disease is, is figuring out how to treat the disease. You figure out how to treat it by dissecting the pathophysiology and then studying the therapeutic approaches.

Now this has to all to be put into a larger context of an understanding of genetics. That understanding began in the 19th century with our favorite monk, Gregor Mendel and proceeded through with the dissection and understanding of the structure of the DNA molecule.

In the 1960's, the first understandings of the components of human genetic disease were made and that was the chromosomal basis of disease. Human genetic material is contained within the nucleus, which is comprised of chromosomes. Nuclear DNA of human cells includes 46 chromosomes arranged in 23 pairs. The segregation of those 23 pairs of chromosomes is the cell's biological basis of heredity.

A special pair of those chromosomes, the X and Y chromosomes, is the biological basis for human genetic sex determination. And in the 60s researchers

determined that the abnormal segregation of those chromosomes caused a class of diseases in humans, most commonly exemplified by Down's Syndrome and Turner Syndrome.

Down's Syndrome is a disorder caused by the mal-segregation of chromosome pair number 21, where affected individuals have three of those chromosomes, instead of 2. That causes a number of problems including mental retardation, heart defects and other disorders.

Now from the 60's to the 80's, a very short span of time, scientists shifted their focus from the chromosomal basis of disease to the genetic basis of disease: the alterations in a single gene that cause human disease.

In the 1980's the first human disease gene was captured for Huntington's disease, a neurodegenerative inherited disease. That was followed soon after, in an increasingly frequent cascade of accomplishments, by muscular dystrophy in the mid 80's and Cystic Fibrosis in the late 80's. The dissection of single gene disorders is now called Positional Cloning. It is the approach to understanding the etiology of disorders caused by mutations in a single gene and inherited in regular patterns.

The key thing about Positional Cloning is that you can determine the identity of the gene that caused that disease by mapping it to a region of the chromosome even if you have no information or knowledge about the function of the disease. Prior to Positional Cloning you first had to figure out what the diseased gene was doing before you could figure out its identity. Positional Cloning allows you to jump directly to diseased gene identification with no understanding of gene function.

An example of this is shown here. This is the pedigree of a family affected by a disease. This couple presented in the clinic with a son who was affected by a particular malformation syndrome. And at the time they thought he was the first one in the family to have the disorder. But upon careful questioning it became clear that a number of his relatives had the same disorder. From that you can determine that all of these individuals who are highlighted must share the same altered gene. Furthermore, because you understand the inheritance pattern of that disorder, you can determine that these affected individuals are male. Therefore, you can logically infer that all of their mothers must also carry that altered gene.

Congruous conclusion can also be drawn that certain individuals in the family, males of women who carried the altered gene, who don't have the genotype must not carry the altered gene or the mutation. You can use those data to localize where the gene is because you can track how chromosomes segregate through the family. Then you can do a very simple one-to-one correlation: which

parts of which chromosomes correlate with the way the disease is segregated within the family. That's positional cloning.

Once you have a map location of a gene, you must then determine all of the known genes that are around that gene in that area and sequence all of those genes and figure out which one of them has the alteration that is causing that disorder. You can sequence many genes but only one of them is the cause.

Until the late 1980's, this process was extremely slow and inefficient. It took between 20 and 100 person years of work, mostly done by graduate students, to accomplish what I just described to you in two slides. At the end of the 1980's was what was called The Human Genome Project, which can be viewed both as an amazing project and a totally mundane project, made all of this faster and cheaper.

This project went through the 1990's rapidly setting up wide scale genetic maps to understand gene content, while at the same time facilitated more and more rapid elucidation of the genetic components of diseases like breast cancer. The first breast cancer gene was cloned in the 1990's and then we stopped sort of tabulating it because, frankly, so many of them happened so fast that you can't keep track. This is what we now call classical medical genetics; a group of disorders that are almost entirely attributed to chromosomal abnormalities we call aneusomy and aneuploidy abnormal chromosomes - Downs Syndrome, Turner Syndrome etcetera—mutations in a single gene that cause a genotype and affect at birth.

These conditions are certainly of great importance to affected individuals and I've spent my entire career taking care of patients who have stupendously rare genetic diseases. However, they do affect in aggregate a modest number of people and they are not a major case load in primary care.

As the Human Genome Project progressed, the tools became more sophisticated and more powerful. This allowed the geneticists to dissect more and more complex disorders. Remember the human genome, 46 chromosomes, 23 pairs, probably around 30,000 genes are made up of DNA and the DNA in a single genome is about 3 billion base pairs.

I should provide some explanation of my terminology - I talk sometimes about genetics and sometimes about genomics. Genetics is the study of genes one or two at a time. Genomics is the study of a genome as an aggregate, as a biological functional unit. It is a higher level assessment of inheritable material.

In April of 2003 the near-complete sequence of the human genome was determined, which markedly facilitated efforts at positional cloning and understanding the genetic components of human disease. Some people have suggested that because The Human Genome Project is done we are now in

what's called the "Post Genome Era." However, we emphatically reject that notion. This is not the Post Genome Era this is the Genome Era. Now that the genome has been sequenced, it is a tool that allows us to do more sophisticated analyses and do them faster and do them cheaper.

To qualify that our institute put out a vision paper describing what is hoped to be accomplished over the next 10 years. They likened this to the structure of a house where the Human Genome Project that sequencing of 3 billion base pairs which costs between 2 and 3 billion dollars is but a foundation – a scientific foundation -- to allow the accomplishment of other goals.

The goals are categorized in three groups; basic science or biology, health or medical research, and the social issues that surround those uses. I'll briefly go over those three areas. The genomic, the society level of the house if you will, goals are to develop policies for the medical and the non medical uses of genomics. The non-medical use of genomics is what funded this workshop today: understanding the relationship of genomics and race, understanding the consequences of uncovering the genomic contribution to human traits and human behaviors, and assessing how we should define the ethical boundaries for the uses of genomics. This is more commonly known as ELSI: Ethical, Legal & Social Implications.

Basic biology: Biologists have a very different role than physicians do in the genome project and there are tensions that result from that.

Their goals are listed here:

- To identify the structural and functional components of the genome, and dissect genetic networks and protein pathways—the notion that genes track with other genes. Genes create proteins and they all interact with each other to generate genotypes.
- Understanding the biology of heritable genomic variation, which is a huge challenge.
- Understanding evolutionary variations in the genome and the mechanisms that generate that variation.
- Policies that can be used to facilitate the diffusion of genetic knowledge out into other areas of research, so that not just geneticists are using the genome but all biologists, and also that it is being brought into the clinic.

More important and relevant here is translating genome based knowledge into health benefits in the clinic. What we hope to be able to do over the next ten years is to identify genetic contributions to disease and drug response.

I assume everyone here uses the healthcare system to some degree or another. And when you stop and think about how we use the healthcare system it's really appalling. The paradigm of how we use healthcare research now is a large randomized controlled trial that takes two groups of patients, divides them in half and treats one group one way and one group the other way. And if 63% of the first group responds favorably to a treatment and 52% of the second group responds unfavorably, then the first treatment is the winner. And that's what your doctor will do to you when you walk into the clinic, working under the assumption that you are in the 63% group A, which is ridiculous because you may not be.

What we are doing is cruelly applying treatments to people without knowing whether or not they'll benefit. In fact one of the hypotheses of the genome project is that individual genetic variations are what determine those responses and what we should be doing is assessing that variation and applying the appropriate treatment to the person who has the responsive genotype.

We also want to identify gene variations that contribute to health and disease resistance not just to disease. There are clear inheritable components to health and to longevity and there are genetic variants that are associated with those genotypes, and they can be mapped and they can be determined. They can be studied to see if we can manipulate human physiology to encourage health. We want to be able to develop genomic approaches for individuals to predict disease susceptibility and drug response, to allow early detection of disease and molecular taxonomy of disease.

Early detection of disease is a very exciting potential area. To wait until pathology manifests in a patient may be foolish. If in fact one can accurately predict the onset of a disease one may be able to interrupt that pathophysiology before the disease manifests with a much higher success rate than waiting until after the disease manifests. We want to be able to dissect genetic pathways in order to develop new therapeutic approaches.

We also need to think about how genetic risk information is conveyed in the clinical setting in order to improve health outcomes and improve costs. It is very well known in behavioral medicine that just telling people that they are at high risk of having some health outcome does not dramatically transform them into a different person; a lot of geneticists harbor that misconception. It just isn't true. We have to work very hard so that we can translate these predictive technologies into effective changes in behavior that will result in the health outcome that we all want.

We also want to be sure that the genomic based tools that are developed have wide spread benefits to all sectors of our society and our culture, which is not a straightforward thing to do.

To do all these things we need to have further quantum leaps in technology. We are currently able to genotype and sequence people probably a thousand-fold more efficiently than we could twenty years ago, yet we need probably another thousand-fold improvement in technology to do what we want.

We want to be able to genotype thousands of samples from large numbers of genetic markers for only thousands of dollars. Today this would cost billions. We want to be able to have a patient come into a clinic or a research center, submit a sample, and determine the entire sequence of their genome for a thousand dollars. That's the goal.

Current cost of the genome project as I mentioned was, I think, two or three billion dollars—the first one cost a lot. We can't afford two or three billion per patient. What we have to do is work on new technologies that will dramatically lower that cost.

Where do we want to go with molecular genetics research and healthcare research? We have to develop clinical genomics research technologies that will allow us to do large scale association studies. I talked about how we find genes that cause diseases that segregate in simple inheritance patterns. Well there are many diseases and disorders and malformations that don't segregate in simple inheritance patterns. They're much more complex. They clearly have a genetic component, but you can't see the inheritance by taking a family tree. This includes disorders like cleft palate, mental retardation, epilepsy, hypertension - I could go on and on. Those genotypes all have a strong genetic component but they are not inherited simply and we have to do more sophisticated studies with high throughput genotyping to find them.

The next phase that we're also moving into is to begin whole genome sequencing on individual subjects. I'm predicting that this will probably start within three years, which is amazing. Not only are we going to be learning your genotypes but selected markers around the genome of selected human subjects within the next several years will undergo high throughput sequencing where we will determine the entire genetic complement for that individual.

Why do we want to do this? We have to collect these data in order to dissect the pathophysiology of human disease. We have to understand what the disease is, how it's inherited; but not only how it's inherited but how the altered genes disrupt the normal pathophysiology of the body to cause the disease. We want to use these technologies to identify therapeutic targets. This is called pharmacogenomics. Where we can use altered pathophysiologic pathways to tell us what molecular targets are available and can be manipulated by pharmaceutical agents.

On an individual scale, we want to be able to take patients in the clinic, determine their pharmacologic responsive genetic profile then apply the proper drug to that

patient based on their individual profile, as well as develop individual disease risk profiles to allow for customized individual medical care.

It's not just molecular biology that has to undergo quantum improvements, it's bioinformatics. This is a huge challenge. These data that we generate in geometrically increasing volumes have to be stored and have to be analyzed and protected, and it's a huge challenge. We have to be able to put in genotypic data, gene expression data and phenotypic data, data about patients into computers in a way that is secure and allows us to do the correlations that we need to do to dissect these pathways. The statistical tools to do these kinds of analyses really just aren't available right now and need to be developed.

We also have to figure out how to cross a gulf – that gap between scientists and health care providers. These two groups of people don't think the same way about things, and we have to figure out a way to make that work better in order for these data to benefit clinical and health care.

The reality is that the Human Genome Project is done. We've got the sequence. You can just go on to a web browser – anyone can do it – and you can look at the human genome sequence. It's fascinating. There's lots of interesting things to look at but we really don't have a clue about what we're supposed to do with this or how to make it work.

DNA technology at its heart is a predictive tool but we can't fool ourselves into thinking that we are seers of patients' future. Science, by nature, is predictive. If you understand the variables that affect the system you can, if properly applied, predict the outcome of perturbing that system. In some ways that is predictive medicine. It's not the same as being a seer. I think we have to work hard to differentiate those two.