

**JOHNS HOPKINS UNIVERSITY**

**GENETICS AND PUBLIC POLICY CENTER**

**“GENEPOPS – SEQUENCING HUMAN HISTORY: THE  
GENETICS AND COMMERCE OF PERSONAL  
ANCESTRY.”**

**KATHY HUDSON,  
DIRECTOR, GENETICS AND PUBLIC POLICY CENTER**

**ARAVINDA CHAKRAVARTI,  
HENRY J. KNOTT PROFESSOR AND DIRECTOR OF THE  
MCKUSICK-NATHANS INSTITUTE OF GENETIC MEDICINE,  
JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE**

**SCOTT R. WOODWARD,  
PRESIDENT AND CHIEF SCIENTIFIC OFFICER, SORENSON  
MOLECULAR GENEALOGY FOUNDATION**

**SANDRA SOO-JIN LEE,  
SENIOR RESEARCH SCHOLAR, CENTER FOR BIOMEDICAL  
ETHICS, STANFORD UNIVERSITY MEDICAL SCHOOL**

**2:00 PM  
FRIDAY, MARCH 2, 2007**

*Transcript by:  
Federal News Service  
Washington, D.C.*

*(edited for clarity)*

KATHY HUDSON: Good afternoon. My name is Kathy Hudson. I'm the director of the Genetics and Public Policy Center at Johns Hopkins University and I'm pleased to welcome you to today's Genetics Perspectives on Policy Seminar, or GenePOPS for short. Today's topic is "Sequencing Human History: The Genetics and Commerce of Personal Ancestry."

The mission of the Genetics and Public Policy Center is to help policymakers and the public better understand and respond to the challenges and opportunities that are arising from advances in human genetics, and we hosted this GenePOPS series in order to explore a broad range of issues that are being touched by human genetics and foster discussion about their impact.

We're delighted today to have four experts joining us to talk about these issues. As we all learned from the human genome project, any two of us are 99.9 percent identical at the DNA level, but of course it's that 0.1 percent that makes each of us biologically unique and which occupies the time and energy of many scientists today.

These differences hold clues to our risks for future diseases and also to our pasts. Much of the work that's being done now really seeks to answer the ageless questions, where did I come from and how did I get here? Instead of using picks and shovels and examining bones and artifacts to unearth clues about our past, today's speakers use DNA to unearth clues about our past, genetic variation, and human migration. Today's panelists will discuss how DNA is being used to understand human history and human migrations over time and discuss how this work is being used by families and individuals to explore their pasts.

I'm going to introduce all four speakers before we begin. Each of them will speak for about seven to ten minutes, and then we'll open it up for a panel discussion and respond to your questions and comments.

One housekeeping order of business: if everybody could silence their electronic devices, cell phones, Blackberries, et cetera, I'd appreciate it. Also, Gina Paige, who had agreed to speak today, cancelled at the last minute and so I want to express special appreciation to Scott Woodward, who was kind enough to fill in for us at the last minute.

I'll introduce the speakers in the order in which they will be speaking. First will be Aravinda Chakravarti. He's the Henry J. Knott Professor at the Institute of Genetic Medicine at Johns Hopkins. Through what he describes as genetic dissection, he has uncovered the molecular basis of a number of common complex genetic diseases. He also played a prominent leadership role in the HapMap Project, which was an international project to map genetic variation. He received his Ph.D. from the University

of Texas, did a postdoc at the University of Washington, and today he's going to talk about the relationship between genetics, geographic origins, and medicine.

He'll be followed by Dr. Spencer Wells, who is the director of National Geographic's Genographic Project, which uses DNA to map the migration of people who inhabited the world tens of thousands of years ago. He's a scientist and author, a documentary filmmaker, and he's an explorer-in-residence at the National Geographic Society. He received his Ph.D. in genetics from Harvard, did his postdoc at Stanford and he'll share insights about the project and share a short clip from his documentary "Journey of Man."

He'll be followed by Scott Woodward, who is president and chief scientific officer for the Sorenson Molecular Genealogy Foundation and the principal investigator for the molecular genealogy research project. He's been involved in medical genetics research, genetic analysis of Egyptian mummies, and today he works mostly with DNA from living people to construct family trees. He received his Ph.D. at Utah State University, did his postdoc at the University of Utah, and he'll today discuss his foundation's collections and the work that they do with that, involving over 60,000 individuals.

And our clean-up batter is Dr. Sandra Soo-Jin Lee, who's a senior research scholar at the Center for Biomedical Ethics at Stanford University. She's an anthropologist by training with a Ph.D. in medical anthropology from the University of California, Berkeley. She studies concepts of race, ethnicity and culture in both science and medicine and today she'll talk about the social and scientific meanings of race and human genetic variation.

And with that I'll turn it over to Aravinda.

ARAVINDA CHAKRAVARTI: Thank you, Kathy. Thanks for inviting me with the others and I see my test is going to be to make sense of really a long period of human genetic history. Studying genes and their relation to geography and how it affects human traits is in fact a long passion of geneticists from its very, very early days in the 1920s or '30s.

So I wanted to start with this slide to try and give you a sense that genetic markers that define very well-understood genetic traits – here I'm showing you a specific one that describes the frequency of one kind of gene, if you will, which is the B-type of the AB of blood groups that all of us are aware of. Then this graph shows or this figure shows the distribution of the frequency of these genes, the dark colors representing a higher frequencies across the globe.

First let me tell you that these kinds of studies are not new, this study was done in fact in the 1950s and what this map will tell you are two basic features. There's a widespread variation of the frequency of any single genetic entity you pick up. Secondly, there is some clustering of the patterns of variation that you see, but perhaps not in the

usual ways in which you think of how individuals vary, say, by continental origin, at least by African descent, European descent, Asian descent and, much more recently, by American and Oceanic descent.

I was going to be simple and succinct and really point up three major themes that I'm sure that many of the others speakers will point to. We see ourselves as unique and very different, not only one from each other but also by various groups and populations, but if you look at the vast majority of human data, as well as data from other species and mammals, which we study fairly well, we really are not the most variable of species at the DNA level, or even among primates for that matter.

The second point is that most human variation is local. This is like politics and almost economics and everything else in that most of the variation is in fact local. And the third quite important point is that looking across our genomes, it's absolutely clear that human variation is largely continuous across the globe. It is because of these three points that you can bring it together to say that human variation can be titrated with geography. The closer we live to where our ancestors come from, the closer our genomes yet distinct; and the further our ancestors come from in space, the more distinct our genomes are.

So let me do the first part, which is quantifying human variation. Almost every human gene at the DNA level exists in multiple forms that we can study and easily identify. They have a specific name – alleles – but that's not what I want you to focus on. This variation is quantified by a specific term. I've highlighted it in yellow. We are, after all, scientists and quite geeky and scholars. "Heterozygosity" – but it's a very simple concept. The concept is if I pick two genes by chance – you remember that each of us have two sets of chromosomes, one from each parent – what's the chance that they are different?

As Kathy mentioned at the beginning, most of our DNA or our genomes are identical so if you pick two small stretches, they'll be identical and in fact, the difference is about 1 per 1000 per unit of DNA base for human beings. It's that 0.1 percent – so that is still a lot of DNA that's different between any two humans. To give you a sense, the variation is about three to five times larger if we study chimpanzees and gorillas, and that's simply because those ancestral populations were much, much larger than the ancestral population that led to all of humanity today.

The second point is most human variation is local, and there's considerable patterning of variation at the local level. We shouldn't believe that one continent is all of one type and a second continent is completely of a separate type. In fact, this map shows the distribution of a very well-known mutation, a sickle-cell mutation. It's a kind of frequency-gram – very much like a weather map except that it's a map of genes – a specific gene. And again by color what I'm showing is the darker colors are higher frequencies of this particular mutation, which by the way is not restricted to Africa but found throughout the Mediterranean region, throughout the Arabian Peninsula, and in

fact in the Indian Subcontinent, and the patterns tell you that there's considerable variation in the frequency.

But that's only one gene and I showed you the AB or blood group gene; if you take into account the hundreds of thousands of genes that we could study, in fact thousands of genes we already have, what you find is a pattern that shows continuous geographic variation. This is a false color map where similarity and dissimilarity has been color-coded in a way. This in fact is a classical study done by Luca Cavalli-Sforza – whom Spencer Wells has studied with – and what this tells you is if you picked individuals from close regions, then the colors are likely to be quite similar, but as you go across different region the colors tend to bleed out and become other colors, but there are seldom any sharp distinctive and sudden boundaries.

The other point on the bottom is that the differences we see in these patterns when quantified is largely within large regions, such as continents, and if you compared the individuals between two continents – say, Africa versus Europe or Africa versus Asia – then you will find it is of course larger, but it's about 13 percent larger. It's not 50 percent more, it's not 100 percent more, it's only about 13 percent more. And that's the sense in which I meant that most changes and variation are local.

This pattern is not a random pattern. This has come about because of the journeys that our ancestors took coming out of Africa and within the last 150,000 years, carrying with them their genomes that have now left their mark in each of our genomes, and this is the pattern that many geneticists and anthropologists still study and try to uncover, including some of the discussion that you're going to hear henceforth.

So it's in this sense that I mean that geography has been the major determinant of genetic differences and that we can't look at human variation only in terms of continental origins, although that is a big part of the difference that we see. And in order to exemplify this – I've been speaking generally about genes that in fact affect disease risk; I'd like to point out to you the way in which geography is in fact the major determinant of the kind of variation that we see.

For this I'm going to take you on a little journey – a journey that's not a random journey, but a journey that many of our ancestors took. And you'll notice from that little blue dot that starting in the southern part of East Africa, that if you went there and sampled individuals, this is fairly typical of the kind of individual that you might see. And as you go north – and this is the kind of journey that many individuals took – our ancestors came out of Africa and moved through the Arabian Peninsula and into Turkey and other places, going further and further north. You will find that there clearly are changes in the human phenotype, and this would be true for diseases as well, until you come to the most northern latitudes in the globe and you find individuals that largely look quite different than the individual that we've started off with or the genome that we've started off with.

Now, if I show you the whole pattern and you compared the young woman at the top with the young man at the bottom, they look quite different. This journey didn't happen in one year; it took several tens of thousands of years, but this difference is in fact the accumulation of small differences that arose as our ancestors went from one place to the other. This is the way that genes vary. This is the general lesson that I wanted to speak to you about, and this the way in which we have to begin to look at genes that impact human disease and affect a subject called genetic medicine – not aspects that are drastically different between populations, but in fact small genetic gradations that accumulate to fairly sizeable effects.

Thank you. (Applause.)

SPENCER WELLS: Thank you Aravinda. That was a great presentation, a good summary of the problem that we are trying to study which is really, as we say in the Genographic Project, trying to explain the patterns of human diversity we see when we look around the world. We all seem to be so different, but how different are we? And it turns out that DNA is carrying the answer to that question.

Now, I'm not going to talk to you about how we use DNA – the technical means that we use to study human diversity – in my lecture. I'd be happy to discuss that in the Q&A period. Rather, I'm going to tell what the Genographic Project is all about, and as someone who was described as, among other things, a filmmaker, I'm going to start up with a short film clip. So I'll tee that off.

(Begin film clip.)

NARRATOR: Around the globe. And how did your ancestors find their ways to the place where you live today? These and many other questions like these are the focus of a landmark study that will change the way we think about human history: the Genographic Project.

National Geographic and IBM are teaming up to collect and analyze DNA samples for human populations around the world, resulting in the most comprehensive picture of human genetic variation ever created.

WELLS: We're all part of one big family.

NARRATOR: Our partnership, in concert with field research support provided by the Waitt Family Foundation, is harnessing a worldwide team of scientists and institutions to sample human DNA on an unprecedented scale.

Our work will shed new light on our genetic and migratory history, helping us to better understand the connections and differences that make up the human species. We'll be able to answer key questions about migratory history and human diversity.

WELLS: The ancestors of the Native Americans made it through here about 50,000 years ago.

NARRATOR: The Genographic Project will also raise awareness of the threats facing indigenous populations and the world's disappearing cultures.

In the past few years, I've traveled to the four corners of world conducting ground breaking genetic research and sharing these results with the public.

WELLS: Your Y chromosome and his Y chromosome and his Y chromosome – they've been here for 40,000 years.

NARRATOR: As our DNA is passed on from one generation to the next, tiny changes occur from time to time. When these changes are inherited down through the generations, they provide us with the tool following lines of descent: genetic markers that never disappear.

WELLS: If you look at these particular position – see that change? – from C to T, that's the marker.

NARRATOR: These occasional changes or mutations are easily identified and accumulate in a particular order so that they act as a kind of clock. They provide us with a time machine, a way to follow family lines back to the deepest branches in our genealogy.

MAN: Why isn't it possible that the Africans actually come from us?

NARRATOR: Over the next five years with a team of colleagues from around the world, we'll sample DNA from members in indigenous communities to decipher the stories written in their genetic code.

WELLS: Given what you know about the people living in southern India, which population should we sample?

MAN: Yes, there are many –

NARRATOR: This research will be overseen by an advisory board of internationally recognized scientists, as well as representatives from the participating indigenous communities themselves.

MAN: That says that the Native American people are somehow connected to Central Europe?

WELLS: Yes, Central Asia.

NARRATOR: By sampling the blood of ethnically stable populations, like the San-Bushmen of Southern Africa or the Australian Aborigines, we'll be able to track the spread of genetic markers that arose hundreds or even thousands of generations ago.

WELLS: I think we've got about 16, 17 samples here one – (unintelligible) – so approaching some sort of statistical significance.

NARRATOR: Using cutting-edge scientific laboratories and IBM advanced research and technical resources, our team of genetic archaeologists will be able to flesh out the stories told by these ancient lineages, providing us with an unprecedented picture of the human family tree.

WELLS: This man is a direct descendant of a person who lived in Central Asia about 35 to 40,000 years ago, and his ancestor is also the ancestor of most Europeans and Native Americans.

NARRATOR: But not everyone is a member of an isolated indigenous group. As a truly global endeavor, the Genographic Project provides for, and in fact encourages participation by members of the public as well.

By purchasing a participation kit, swabbing your cheek and sending in your DNA sample, you'll discover details of your deep ancestral migratory history.

WELLS: So I am going to take these back to the lab and we will let you guys know the results as soon as they come in.

NARRATOR: Your anonymous test will only be used for non-medical purposes; all we're studying are markers that tell us about your ancestry. All results will be delivered privately and securely on a special Genographic Project Web site. Contributing your DNA will help with our exploration and mapping of the emerging genetic tree of all humanity. Furthermore, proceeds from the sale of the kits will be channeled back to assist local educational and cultural efforts among the participating indigenous communities.

WELLS: So what we're hoping to do with this study is to confirm your stories about being related to these Chinese sailors.

NARRATOR: The Genographic Project is one of the most important research efforts in the National Geographic hundred-plus years of explorations. It builds an IBM's deep legacy of supporting science, research and innovation.

MAN: Maybe after your study they can prove that the genetic evidence is better than archaeological evidence alone – probably, right?

NARRATOR: Access to Web site loaded with maps, new stories, a genetics primer, historical timelines, and the host of multimedia features on human origins,

migration, archaeology, climate, and linguistics, as well as information on indigenous peoples and population genetics.

WELLS: I'm really excited to be here. In a way you carry a secret in your blood.

NARRATOR: With your help, it will result in one of the largest collections of DNA samples ever assembled, culminating with the creation of a virtual museum of human history for all to access.

It's an ambitious attempt to answer fundamental questions about ourselves, to give us a better picture of our common past as an extended family: where we originated, how we came to populate the globe and we why, if we share a common ancestry, we look so different from each other? Out of our joint efforts, a story will emerge: the story of one family, the human family, seen in many phases.

(End of film clip.)

WELLS: So the Genographic Project – what is National Geographic doing getting involved in a genetic project? Well, it really stems back to the film you just saw some clips from, “The Journey of Man,” that I got involved in when I was actually at Oxford University. I was on the straight and narrow academic path – as you heard, Harvard Ph.D. and Stanford – and I was at Oxford running a laboratory there and as we did our research and published papers, people would often write to us and say we'd like to do an interview for a newspaper, or whatever it might be. And eventually someone contacted me and said we should be doing a film just based on this research and I said, okay, but I'm a scientist, I want to make sure that works for me in the right way. I want to be involved in it. And that film became “The Journey of Man” which aired on PBS in 2003 and was co-produced by National Geographic.

And coming at the end of that film, I started talking to the people at National Geographic about what this work was all about. They said, ‘we love the film as a film, but we're also really excited about this as science. In a way human origins research is in our DNA. We funded the Leaky family over the last half century, we have funded Jane Goodall, all the work that she has done out at Gombe. We really see the genetic data that you're collecting as being the kind of next wave in our understanding of where we all came from.’

And they asked a very faithful question. They said, ‘what do you want to do next? What's the next big step to take in your field?’ And I said, well, the information that we have collected that's led to a very broad-brush view of how we populated the world – the map, basically, that Aravinda showed you – was based on a few thousand people that have been studied for a handful of genetic markers, and that's not a very good sample size for 6.5 billion people spread around the globe. What we need, I said, is a much larger sample size. Basically what we're trying to do now is akin to an astronomer trying to assess the complexity of the universe using a pair of binoculars, and we need to build that high-power telescope. In my field – in our field – of population genetics, that

means more data. So instead of having a few thousand people, ideally we would have hundreds of thousands of people studied for more genetic markers.

And they said, okay, after their jaws were lifted up off the table, that sounds exciting, let's see if we can find a partner. We approached IBM mostly because of their experience in computational biology and setting up large databases, security, and so on, they were fascinated by this project – saw this as a real challenge for the computational group and they got involved and in 2005 we launched the project.

Now, the project has three core components. The main focus of the science is on sampling and working with indigenous people. What's an indigenous person? Well, many of us are living in Washington, D.C., or some other large city, and our ancestors come from all over the world. We're not necessarily indigenous or local to any particular location. I'm a mutt. My ancestors come from all over Northern and Western Europe and I live in Northwest D.C., but there are people who have lived in the same place for a very long period of time, and they're critical to doing this kind of work because they give us an insight into the genetic patterns of their ancestors. That is the core of what the science is focusing on.

Ten regional centers are set up around the world, each charged with sampling roughly 10,000 indigenous people from within that region. But when we were designing the projects, we figured there might be some interest on the part of people who wouldn't be described as indigenous in finding out about their ancestry, and so we wanted to build in a public participation component – an opportunity for people to send off for a kit, get their own DNA tested completely anonymously, get the results back; and if they wanted to, to actually be a part of the scientific analysis, to decide on the Web site, again totally anonymously: yes, I'd like to have my DNA analyzed with all the indigenous data, to actually be a part of a real-time scientific approach, which is pretty exciting, it doesn't happen very often.

Moreover, by purchasing the public participation kit, by going on and buying the kit on the Web site, you're helping to fund the research. You're helping to fund the field research we're doing with indigenous people around the world, but also you're helping to fund something we call the Legacy Fund, which is to give something tangible back to these indigenous communities, many of whom are the poorest of the poor in very poor parts of the world, and their way of life in many ways is endangered today. They're forced to leave behind their old ways, their ancient homelands and often move to a growing megacity. And when they do, within a few generations, they've lost touch with that ancient culture. They become part of the melting pot, possibly a good thing socially, it helps to break down barriers – traditional, ethnic or racial barriers – but it makes our jobs very, very difficult because we've lost that geographic context that the indigenous samples ordinarily give us. So the Legacy Fund is a way to get something back to these people whose way of life is endangered, and we want to raise awareness about these issues.

So the goal of the DNA sampling work, again, is to sample roughly 100,000 indigenous people over the course of the next few years. The project's going to wrapping up toward the end of 2010. This is being done in collaboration with 10 research centers – regional research centers that we've set up all over the world. All of it's overseen by institutional review boards both in local level and we have a supranational institutional review board at the University of Pennsylvania. We also have an advisory board of very well-qualified scientists and representatives of the indigenous communities, the goal being to engage with indigenous people and collaborate in an effort to understand our common history.

On the public participation side, as I said, you can go onto the Web site and order a kit. It's all totally anonymous: When you are sent the kit, it includes an alphanumeric code – a random number, in effect. We don't know which code goes in which kit. That is the only way that your sample is tracked from that point, the only way that you log in to get your results. So if you lose that number, you actually can't get your results back. And it tells you about the ancestral roots that your ancestors would have followed from an African homeland to wherever you live today. Now the Legacy Fund – I'm very pleased to say that we have raised about \$2.2 million for that so far. The net proceeds from the sale of the kits are being channeled into grants which, again, aim to give something back to indigenous and traditional groups. The grants are proposed by the indigenous peoples themselves and allow them to raise awareness about issues such as language loss, possibly do something to halt or slow some of the process of cultural change that's going on, but, again, it is a way to get something tangible back to these indigenous groups.

How are we doing? Well, on the research side, the indigenous sampling, we've sampled over 25,000 indigenous people around the world, so that is well underway. We're roughly on track to where we expected to be at this point. On the public participation side, we have far exceeded all of the expectations we had at the beginning of the project. We anticipated possibly selling 100,000 kits over the course of the project, over the five years, and not even two years in we're already up to nearly 200,000, which has raised a huge amount of money. This is a nonprofit enterprise, so all the money that we raise after recovering the cost of testing and producing the kits gets plowed back into the project, the majority going to the Legacy Fund. And as I said, over \$2 million has been raised for that. We're intending to give away around \$400,000 this year in grants.

What are the outcomes? Well, obviously to try and figure out some of these answers as to why we are so diverse as species, how we spread around the world; to educate the public through the Web site and through the media and National Geographic about these research results, about what it means to be part of this extended human family, about what it means to get your own DNA tested, what they might find out, what they can't find out; to give something back to the indigenous communities, as I said, in the form of the Legacy Fund; and ultimately the take-home message is this idea that we're all part of one extended family separated by no more than a couple of thousand generations. So that's an overview of the project and I will be more than happy to take questions when we get to that point.

Thank you. (Applause.)

SCOTT WOODWARD: We can look at Spencer's work and the things that he's doing as the foundation and the trunk of the tree for all of the large branches that come out of that that bring all of humanity together. The project that I'm going to describe focuses on the leaves of those trees and the veins within those leaves of the trees. We're looking at a more recent time period; in fact, the present, and moving back the last 10 or 12 or maybe 20 generations into the past and trying to use that information to bring families together – families in a traditional genealogical sense in that talking about grandparents and great-grandparents and great-great-grandparents.

The goal of the Sorenson Molecular Genealogy Foundation, of which I am the director, has that same feeling, that same kind of idea about bringing people together: that if people understand how they are in fact related to each other, that they will treat each other differently – that we really are, as Spencer said, members of one great human family. The Sorenson Foundation started collecting DNA samples six years ago this week. It began at Brigham Young University where I was professor. Within the last two years we have moved completely out of the university into our own facility, which I'll talk a little bit more about some of the numbers in just a minute.

The way that we would like to change the way people in the world think is by, first of all, creating a very large amount of data that combines genetic information with genealogical type information. Who are your father and mother, your grandparents, your great-grandparents, and what does our DNA say about the relationship within that family unit and between family units?

We talked earlier about 99.9 percent is the same and it's just that 0.1 percent or 0.2 percent that allows any kind of differentiation. And within that differentiation, the closer you are related to each other, the higher degree of sharing you have within that 0.2 percent.

As we create these databases of correlated genetic and genealogical information, we also provide the tools to be able to go into that data, understand what it means, how it works together, and what it means to be related genetically, and then to make that data available to the public. The goal has been and continues to be with our foundation – again, a nonprofit foundation – to provide this information back to the public free of charge. There isn't any charge or any cost to the public to participate with us in building this particular database.

Each of us contain with us our own genealogy. Some of you have members of your family who have gone to the court records and have written down all that information, the census records, who's my father, my parents, my grandparents, and so on, but we also have that type of record contained inside of our DNA. We have the information on how we are related. One of the goals of our project is to be able to take any two of you in this room and sit you down at the table and not just show that you are

related to each other, but specifically how you are related to each other and what is that common ancestor or, more realistically, who is that set of common ancestors that you share.

This genetic information that we're dealing with is maintained through generations. It allows us to link us together. There are specific regions of our DNA that can identify us as an individual unique from any other person that's ever lived on the earth, unique from any of the six billion that are living on the earth today, but it can also be used to link us together in a nuclear family – a very tight, father, grandfather and so on – extended family group and into other types of populations.

We define a population slightly differently, I think, from what we traditionally understand or think populations to be. We add both a geographic and a time element to that. I am a member of the Utah population in 2007, but in 1956 my genes inside of me were not in Utah, they were actually in Virginia. In 1930, my genes were not in me, right, they were in my parents, but where were my parents? My parents at that time were in Utah and so the genes walking around in me today in 2007 are actually representative of the Utah 1930 population in some way. And if we go back to 1900, each one of my genes has a 75 percent chance of being in that Utah population, but a 25 percent of being in Leed, Lincolnshire, England, because that's where my grandmother was from. And so if we continue that back, using the combination of genealogical information with genetic information, we can reconstruct pictures of these very recent populations in the past. That's the goal of our dataset is to be able to do that.

We're looking at three types of genetic information. I'm going to start at the bottom with mitochondrial DNA. That piece of DNA that we inherit from our mothers and that follows a matrilineal progression through the generations is one set of data. We also are looking at the Y chromosome information that follows a patrilineal inheritance pattern, and very useful bits of information that allow us to tie together with both immediate families, but also to find out where we fit in the world and where we came from, and the types of things that the Genographic Project is doing.

We would like to also look at our autosomal DNA, the pieces of DNA that we get essentially 50 percent from mom and 50 percent from dad and that in a lot of ways are a little more difficult to look at, but have a lot more information about who we are and how we are connected and our ability then to rebuild these genetic populations 100, 200, 300, 500 years ago.

We're looking at this very recent past as far as genetics and genealogy is concerned. As the Genographic Project gives us a foundation on where we came from, we would like to then move from the other direction in some place in the middle, we're going to come together – I hope. (Laughter.)

And so if you look at the traditional genealogy chart, you see at this generation you have 16 ancestors, a Y chromosome for the males can tell you something about one of those 16 ancestors. So if you have a Y haplotype, you know something about that

individual. Likewise, your mitochondrial haplotype tells you something about another one of those 16 individuals. But mitochondrial and Y by themselves doesn't tell you anything about the other 14 individuals that are in there. That lies in the power of the autosomes, a little more difficult to discern and to decipher and to sort out but it's doable, especially in the last 10, 12, 15 generations, something like that. That's a timeframe that is probably reachable with the type of data and information that we have right now.

If we look real quickly at the statistics associated with the Sorenson Molecular Genealogy Foundation database, you'll see that we've collected not nearly as many as Spencer has collected: we've only collected 75,000 people but these have been a very selective group of people. These are people who can also provide for us four or more complete generations of genealogical information, and so that's a pretty strict requirement for most people. I don't know how many of you in this room could give me four generations of your genealogy off the top of your head. Anybody? One, two, three, four, ten, twelve? That's good. Spencer too. Do we have your DNA in our database?

WELLS: Not yet. (Laughter.)

WOODWARD: That DNA then is linked in our dataset to four million people – four million ancestral records. The average size of a person's genealogy is about 151 individuals, our average depth of those genealogies is around 300 – 220 to 385 years, so you can see the time range that we're dealing with here. We ask for four generations; we often get more, and then our genealogists go out to work and find connections and links and extensions of those genealogies, so our genealogies now average about 11 generations into the past of the documented records.

Interesting thing: our generation time is actually real, measurable, and of all of that – and it depends on whether we're looking at a paternal line, a strictly paternal line, a strictly maternal line or we look at everything all together, it ranges somewhere between 20 and 35 years per generation and the number of terminal ancestors on each person's genealogy is around 73.

If we look at the genotypic information that we have in the database, we have typed only 19,000 of the 75,000 people on the Y chromosome and the reason for this is, first of all, only half of them are males, actually a little less than half. When we started this project, we collected DNA using blood and we found that we got about 3 to 1 females to males that participated. Since that time, we've moved to a mouth rinse and it's almost equal now. (Laughter.) We've also gone after individuals for whom we have the deepest genealogy, so we've selected the ones that we prioritized. Eventually all the male samples will have the Y chromosomes sequenced and both males and females will have their mitochondrial DNA sequenced, so we have 15,000 mitochondrial D-loop sequences. We have some specific information on the X chromosome that both males and females do have – a very interesting inheritance pattern that helps us to understand and answer some very local genealogical questions. And then if we look at our autosomes, we've now typed about 45,000 people at a number of different sites on the

autosomes. This has given us a little over 3 million genotypes associated with those individuals.

We have had some interesting success stories. What we want to do is bring people together; we want people to be able to find each other. It was built originally as, and continues to be a genealogy tool – another way to assist people in doing genealogy. One example here, the Cynthia Wilson story. A young woman who's a Seattle, Washington African-American very interested in who she is, where she's connected, runs into the roadblock. If any of you have done genealogy work, you will come to a point where you can't go any further based on the documented evidence. She was able to look at her mitochondrial DNA sequence; she found an exact match in our dataset, turns out that it matches with a person from Maui. We know that person's genealogy. We know what they have. They've now been able to make a connection, something that made a big difference in her life.

A nontraditional use of genealogy in South Africa, we have a group of 11 young boys off of the street who didn't know who they were or who they were connected to. We went in asking the question whether or not we could find their family. These were kids that have been essentially dropped off on the streets of Johannesburg and some of them had not been back to their family in over seven years. Some of them didn't even know where they came from, but by using some of the DNA typing that we did in our population sets in South Africa we were able to put these people back into their communities and actually found their kin.

However there are some difficulties associated – (laughter) – with assigning DNA to people and as we've seen just in the last week, some of the headlines, in fact one that I really like – this is the one. (Laughter.) “DNA Indicates Jesus Found” okay? (Laughter.) The headline – and I won't incriminate who wrote that... and you read the article and it had nothing to do about DNA, and in fact when it did say something about DNA. They got it all messed up because it said that the DNA they tested showed that that there was not a patrilineal relationship between the two DNAs, and they were looking at mitochondrial DNA – I mean, it was a mess. It's a disaster, okay? And so there really is a problem that we have to deal with here and it continues to be an educational problem.

And just as I end, it's better – it's better than it was five years ago, but there's a lot of work that still needs to happen for us to be able to bring the public into the loop so that they can understand what really is possible doing these types of DNA tests.

Thank you. (Applause.)

SANDRA SOO-JIN LEE: I wanted to start off by thanking Kathy Hudson and the Genetics and Public Policy Center. There is no doubt that there is this emerging genomic era and it's a changing landscape, and I think that that is characterized by this massive production of genomic data that is out there, and that what we'll see in the future is increasing research that associates this data with a broad spectrum of different medical

and non-medical conditions and traits. And so we have to think about how these technologies and the types of services that are provided to consumers will play out within this landscape.

Increasingly, there has been a focus on finding meaningful difference. Again, that's not yet defined, but it's where ancestry is believed to be important, both in the medical and the non-medical context. And as we heard and we know, increasingly products and services to determine answers have become readily available to consumers, so that the decision-making process when you're on the Internet and you want to buy one of these services is very quick, without perhaps a lot of time to think about what the broader implications might be.

Some of the issues that I thought we would think about – and there are many more, but here are three: what is the relationship between race and ancestry, how do genetic ancestry testing technologies inform our everyday ideas about how difference operates, and what race means for all of us.

The second is, how will commercially available genetic ancestry tests affect notions of personal identity and group membership? How will it inform how we think about populations? Will it challenge some of the sources of information that we draw upon when we think about who we're related to?

And then the third is, what responsibilities come with offering genetic ancestry testing?

Well, there is no doubt that the Human Genome Project, or the completion of the Human Genome Project, is a total advance in this emerging landscape of genomics. And one of the ubiquitous messages that we heard often with its completion is that we are all the same and again here it's been invoked I think already three times in the previous presentations, but here is a quote from President Clinton. When he talks about the completion of the Human Genome Project he states: "After all, I believe one of the great truths to emerge in this triumphant expedition inside the human genome is that in genetic terms, all human beings, regardless of race, are more than 99.9 percent the same. What that means is that modern science has confirmed what we first learned from ancient faiths: the most important fact of life on this earth is our common humanity."

Now, this quote dovetails nicely with what many social scientists already believed: that race is indeed not genetic, but somehow a product of sociohistorical processes, and changes depending on what social context you are in. However, what we've seen I think over the last few years is a pendulum shift in how we think about race and its relationship to genes.

Here this is an article from Nicholas Wade where he talks about some work done by Neal Risch, a geneticist now at UCSF, where he and his colleagues argue very strongly that race is very important in genetic research and that there are strong relationships between those two concepts. So we're back now to the question that

doesn't seem to go away: does race exist? And here we have the cover picture or the cover page of "Scientific American" that asked the question: "Does race exist? Science has the answer." This cover page refers to an article actually written by Steve Olson, who I know is in the audience, and Michael Bamshad, who is a geneticist.

I don't mean to pick on this article, but actually if you venture past the cover page, what you get is a very nuanced description about the history of human migration, much like what we heard in our first presentation. However, the choice that the journal made in publicizing this article is one where you have six women that are supposed to be emblematic of our concepts of race and if you look here you'll notice – what really is different between these individuals? Not much. Stripped of social historical location, cultural difference, we see six women about the same age, the only difference really is skin tone. And I think this reflects the dominant ideas around race: that even though we talk about the very nuanced history of human migration and that there are no sharp boundaries between populations, somehow the ways in which it gets translated to the public is through a prism of conventional ideas about skin color, phenotypes.

This is actually very similar to the first presentation in that it would be nice to be able to have a more full, nuanced idea about what difference is, but that doesn't seem to be the case. Now here, what do genes tell us about identity? I'm going to shift to the second question of what can these genetic ancestry testing technologies tell about our personal identity? And here we have Wayne Joseph, some of you might know the story. He is a high school principal in his 50s in Southern California who actually had gone through a lot of self reflection about his identity as a black American and has written quite a bit about race, but one day he read a story about DNAPrint Genomics, which is a company that offers ancestry testing, and he decided that he would send away for the test kit. He sent in his cheek swab and a few weeks later he got his results and he was surprised. This comment comes after learning that he is, according to DNAPrint Genomics, 57 percent European, 39 percent Native American, and 4 percent East Asian: 0 percent African.

He says here: "when you define yourself one way and then at 50 there are results that say you're something else, it does rock your whole world." I think this puts in stark relief how this information may challenge personal ideas about who and what we are, and even more so how it trumps other sources of information that one draws upon when thinking about what one's identity is.

So I think this points to the cultural significance of the gene. I think the reason why genetic information has become so popular to an extent is because it has this allure of specificity and technological precision. Individuals often turn to genetic information as a way of confirming or validating one's sense of self and predicting one's future, and that genetic information often trumps other sources of identity formation.

Now, how will commercially – this is the same question but it I think reinforces an idea that race or racial difference is genetically determined and it may be used in the future for different sociopolitical means. There was an act that was brought up by the

Vermont legislature to use genetic tests as a way of proving tribal membership, for example. It has been brought up as a possibility for when you apply for admissions and you are applying for certain programs, perhaps a genetic test would be a more scientific way of proving identity. It's been talked about. Also in terms of validating one's history with respect to slavery and for some campaigns around reparations. And one of my colleagues at Stanford has been asked repeatedly by immigration lawyers to use genetics as a way of validating refugee status, so literally cases where they want to know: did this person come on this side of the border or on the other side of the border? I think we need to think about what is being promised perhaps to the public or what the public is thinking about when they think about genetics and human identity.

Now, some of the responsibilities that come with genetic ancestry testing: privacy and confidentiality obviously are important issues. When one gives away one sample, how is that information protected? Disclosing the limitations of technology, so depending on which technologies are being used by various companies, it does privilege a certain aspect of one's history and I think that type of information needs to be given to consumers who are looking for this type of information.

And then the interpreting of results – and this dovetails with the last point of providing genetic counseling. Some have argued that in a medical context we always provide genetic counseling when one gives genetic information to individuals, not in the case yet of ancestry testing, but perhaps there may be a good argument to do so in that if this information has been given without proper interpretation or help with respect to what results mean, it may prove to be quite psychologically damaging or at least challenging in terms of understanding what one does with information that does not dovetail with how one sees oneself.

And then the very last point is the danger of selling ancestry testing as recreational genetics. Now, I heard this described as just recreational genetics. It's an extension of a hobby. And I think that we should be very careful when we think about it in those terms because what it does is it shifts responsibility from those providing information to the individual; this is something that's private and doesn't have implications beyond oneself and I think it's important to think about what the wider implications might be.

Thanks. (Applause.)

HUDSON: We're going to open up now for questions and comments from all of you. We are videotaping this session, which will be available on our Web site as streaming video in a month or so, so if you don't want to be immortalized on our Web site you might want to refrain from standing up. And we have roving microphones around, Rick has one and he's going over there now. When you ask your question, if you could say your name and the organization you're from.

Q: Hello, good afternoon, my name is Kevin Campbell, and I'm the Campbell surname administrator for the FTDA Surname Project. I guess I'd like to thank everyone

for coming today. I've been the administrator for four years and so I share your passion for using DNA for genealogical purposes, and we actually have a very exciting project with over 200 representatives, including some of the chiefly lines that go back 100 years.

I guess my question is for Spencer and that is: I heard you speak a couple of times in terms of what the project will do, but I saw on your slide that you anticipate publishing this year and so could you maybe tell us a little bit what some of your preliminary conclusions might be?

WELLS: I would rather take the fifth on that. (Laughter.) What I will tell you is that we have papers that have already started to go into the journals and that we will have other papers going in later this year, so we will be publishing the results and releasing the data into the public domain as we go along, but I would rather, as a scientist, let it pass peer review and then we can talk about it. Thanks.

Q: Hi, Jennifer Leib with Acumetrix. I have a question about the storage and use of samples I guess for both speakers who are affiliated with biobanks. I understand that the samples are anonymously used in the Genographic process. Do you store the samples? You destroy them once you genotyped them for this purpose? And if you do store them, for how long and at what point would you either use them for a different purpose, sell them to another entity for additional scientific or commercial interests? Who owns the samples, I guess?

WELLS: All these things we're being accused of. No, I mean, the samples completely anonymously. As I said, there is a code which goes into each kit. We don't actually know which code goes into which kit, and so that's the only way your sample is tracked. There is never a name attached to it. These samples are used only for testing markers that inform us about ancestry and history. We will never do anything that's medically relevant. And at the end of the project – so at the end of 2010 – we will be destroying all of the samples on the public participation side.

The indigenous samples are going to stay at the regional centers we've set up. Part of what we're doing in the project is building capacity, particularly in the developing world where a lot of these centers are located, and in the future people will be able to apply to the project – to the oversight committee we're going to set up – to receive these samples to do collaborative research, but again focused only on the ancestry and history. Those samples will remain under lock and key. We're not creating a biobank, we are simply doing the sampling. And, again, on the public side those samples actually will be destroyed in the end. They will certainly never be sold to anyone else.

WOODWARD: As far as the samples in our database in the Sorenson Molecular Genealogy Foundation, as we collect those samples, they are being stored and they are being consumed. The DNA is being consumed as we operate on these various tests. We expect that all of that DNA will be consumed over the life of the project, whether it is in the next five or 10 years, and that there will not be any more of those samples left on-site.

HUDSON: I have a question for Spencer and Sandra. Spencer, in your film you talked to somebody about being able to tell them or confirm for them their origin stories and then Sandra showed an example of somebody who received information about their origins that was surprising to them. So I'm wondering, in what way do you represent to people how they cope with genetic information that may conflict with their own community stories and cultural stories about where they came from and how they got to where they are?

WELLS: Very good question. I mean, what we do and what we explain very clearly when we're conducting outreach with indigenous groups, and the general public as well, is to explain that the stories that we're telling are simply one part of their ancestry. We're examining – Scott was talking about – on the Y chromosome and mitochondrial DNA less than 1 percent of their genome and there is a broader story out there.

Now, we may start to look at autosomal regions as part of the project. We certainly have IRB approval to do that. Again, nothing that's medically relevant, but potentially trying to bring in that other ancestry that's out there. But we explain that one of the risks is that some closely held personal belief, a traditional belief perhaps, may actually conflict with the story that we get from this particular piece of DNA that we're looking at. And that is a real risk and most people actually want to know the answer.

A great example: I was out in Chad, north-central Africa, back in October and November of 2005, and we were sampling a group called the Toubou who lived in the north in the Tibesti Mountains and they have very conflicting stories about their ancestry. Some people have told them that they came from Yemen, others that they came from Nigeria, some say that they are a lost tribe of Israel. They actually want to get the information back and that is something that we encounter all of the time working with indigenous people. They not only want to give us a sample, but they really want the information back so that they can incorporate it into their sense of self. It's not meant to replace a traditional belief; it's meant to accentuate.

SOO-JIN LEE: Right. Well, I think one of the difficulties is knowing what the motivations are for people who engage or who want to have these services done, and making sure that once they get the information they're able to process it in a way that makes sense. So I'm reminded here of the Havasupai Tribe where this tribe in Arizona signed up for research. They signed up to donate their DNA for research on diabetes and it turned out that the samples were then used for studies on schizophrenia, but also studies on history of human migration. And there was a lot of outcry because that's not what they signed up for. And so I think it's very important to make sure that the individuals who are signing up understand what the parameters are in terms for the uses for the DNA. And we live in a world where DNA does circulate and making sure that those that are donating understand what possibilities there might be in terms of the research that's going to be conducted on their samples.

WELLS: Yes, that's why these safeguards are so important, that's why informed consent is such an important issue. And believe me, we spent a long time thinking about the nature of the informed consent, group consent, individual consent, what does it mean to consent – all of these issues. And obviously people are consenting for a particular type of analysis to be done. DNA is not circulated, at least in our case; it is only used for that particular type of analysis.

CHAKRAVARTI: You're talking about several kinds of problems. One is to use genetic information, DNA information to try and find out the relationship of one group of individuals with another group of individuals – this is Genographic Project. There's a long history as well as an experience of that. In Sandra's example, which is individual ancestry, I think it's a young science. We should question how accurately we could do it and how would we know? There are many, many questions that are unknown at this point in time.

I think a number of times it has been mentioned that the fraction of humanity that we've sampled to try and say that this is where you come from – I'll just put to the fact that somebody mentioned that somebody was some X percent East Indian. I don't know whether that person was an Indian or a Pakistani or Bangladeshi. We have and they could have quite different histories culturally and socially, so I think the science of this really has to be probed quite deeply when we come to the individual and I think there's much more work to be done.

WELLS: And just to lead on from that point, I think it's always worth mentioning the little factoid that some of us like to throw out in lectures, namely that there can be more genetic variation between two African individual samples from the same village than in the rest of the world combined, so we know so little about the patterns.

HUDSON: Thank you. Do you have a question Charmaine?

Q: Charmaine Royal from Howard University. My question is for Aravinda. Aravinda, in your last slide, I really liked that graphic where you went from the person in Africa migrating to different parts of the world. You talked about the differences being due to small changes – small genetic changes – but you didn't talk specifically about the interactions of these genes with the environment. And I just wanted to know if that's implicit in your discussion of how the person at the top differed from the person at the bottom, because you talked only about genetic changes that occurred.

CHAKRAVARTI: Yes, my God, I only had somewhere between seven and 10 minutes – (laughter) – but it is implicit. By the way, I would say one of the large foci in genetics has been looking at DNA in a very – what should I say? – as a dynamic piece of information, but it changes only in really great historical time. So it's not only implicit – very deeply implicit that some of these differences clearly are due to environmental changes that have occurred over time and how they sort of shift our genomes, but our DNA responds in function also to many immediate environments. That leads to patterns

of disease, and it could be diabetes, it could be hypertension, could be anything. And so there are many, many other things that DNA does. Exactly how it goes about its business is really still not understood, but from the functional view point the differences between those two individuals or even groups of individuals are very many.

I was just pointing out that even though DNA differences have accumulated over a long period of time and that's our story and it is not to be believed that something happened and suddenly we created two groups of people that are completely different from one another.

WELLS: One of the big unanswered questions in the physical anthropology – and that is basically what we're talking about here – is if we have so little genetic variation, if we're 99.9 percent identical, why do we have so much morphological variation, more than any other species of large primate? And we don't know. The environment, probably adaptation's played a large role in that, probably what Darwin called sexual selection – choosing people to mate with on the basis of what you find attractive, and that varies around the world – probably played a role, probably random genetic drift – small changes as these populations moved around. But what we do know is that there is a lot of variation on the surface and selection has certainly played a role in that.

Q: Hi. I apologize in advance for my lack of technical expertise, but I wanted to ask in terms of assisted reproductive technologies and other scientific advances, do you see ART providing a challenge for genealogical analysis, for example, mitochondrial DNA in terms of surrogacy and different sort of IVF techniques?

HUDSON: And can you tell us who you are?

Q: Again with the caveat of not having scientific expertise, I'm from the American Association for the Advancement of Science. (Laughter.) And my name is Anita Williams.

HUDSON: Okay. (Laughs.) Thanks, Anita.

WOODWARD: What we're looking at are biological lineages, so we would be looking at the donor of that biological material and the lines associated with that and as we've seen, as a result of that – maybe an individual that grows up in an environment much different than where the donors of the biological material are. It may make them a completely different type of individual. That's the sociology question I think that we have to deal with: is it your genes, is it the culture, is it the environment that made you who you are – that determines what race or what group of people you belong to? And so what we're dealing with are actually the biological lines and that may or may not have direct correspondence to a person's personal understanding of who they are and how they are connected to the world.

HUDSON: I have a related question having to do with misattributed paternities, so there's sort of lore in the genetics community that misattributed paternity accounts for – what percent, Aravinda?

CHAKRAVARTI: I'd say in our studies, anywhere from 5 to 10 percent.

HUDSON: Five to 10 percent, and it's rumored that for people who have sons named "Junior" it's about twice that. (Laughter.)

CHAKRAVARTI: Is that actually on the record? (Laughter.)

HUDSON: So here's my question: in your case, how much do you account for mistaken family relationships. When you're going back your four generations when you don't really know that your grandfather wasn't your grandfather, for example, what –

MR. WOODWARD: We have a way to test that in an interesting way in that we have collected samples not just from the general population, but from people who claim to have some known genealogical information, so we can actually test whether or not the genealogy that they provide to us on the paternal line is in fact what they think that it is, whether it's expected. When we do that and look at our dataset, we find something much different than what you see out in the general public, but we're less than a half of a percent of that mispaternity.

This could probably be attributed to a number of things. One is that we are going to people, we're telling them that we're doing biological tests that can in fact determine this type of information, so if you're a little suspect don't participate, okay? There's probably that part of it also, but these are datasets that we can actually look and get a real hard number as to how many non-paternities we might have in that dataset.

HUDSON: And what do you do with misattributed paternities?

WOODWARD: That information is not returned – as you noticed, we don't send information back to the individual participants, so people would never know in our dataset.

HUDSON: I think we're over here.

Q: Hi, my name is Judy DeWitt. I'm waving to all my DNA cousins who have sent me here, in effect. My question is, I'm trying to work my way through Philadelphia to find somebody who will actually look at my mutations. I have recently done the full sequence mtDNA. My computer has entry genes and mital map and I think of that an awful lot, but I really need somebody who knows that next chemical level. I need somebody who can point to one of my mutations or two and say, if I were you I really wouldn't be taking that vitamin pill with all that vitamin C or iron or copper or something because my metabolism is – well, I'm an H, the vanilla pudding of Europe, I have refined my mitochondrial sub clay down so far at this point, it's scary, but I'm excited about it. I

want to know more. I know my family history. I want somebody to offer that counseling. I'm not afraid to hear what you've got to say. There should be somebody for people like me. Thank you.

HUDSON: You're interested in finding out information that can help you make decisions about how you live your life, so health-related information.

Q: Absolutely, do I buy long-term care insurance today or have I got a little bit more time. (Laughter.) Really, these are things that are important for us and if we're going to do all this testing, I really need to know more. I need to know what this means.

CHAKRAVARTI: I'll give you two really kinds of questions – or answers rather, depending on the kind of studies that I've done, specifically for the mitochondrial DNA. The mitochondrial haplotypes, or types that are largely determined by the tests that we have, really tell us very little so far as to whether the bearers of that mitochondrial type have even, say, twofold or fivefold or tenfold increased risk to disease X or disease Y.

WOODWARD: Yes, but some of those just only change the risk by a little bit. But there clearly are mutations within the mitochondrial sequence that may be very strong predictors for some very rare types of hereditary disorders which usually are also evident in the family history, but not always. So if you're concerned about any of those kinds of disorders of phenotypes or traits, then almost any medical geneticist would be able to sit down with the genetic test, provided it's been done the right way, to tell you whether you are at particular risk or not. But most of our mitochondrial types are much better markers of where we've come from than saying, at this point in time, that we are at much greater risk or protected from, say, metabolic disorders or any other disorders.

Q: A follow up on that, too. My problem is I – (off mike). Outside of Philadelphia – again, many calls, genetic counselors, university programs for genetic counselors who have people on staff from all the Philadelphia universities. I called Penn and the gatekeeper's secretary person told me that we don't do mitochondrial here at Penn. I thought, okay, who do I get to talk to? I even went to dinner with a neurologist and his genetic counselor wife. I knew more than they did that night.

MR. CHAKRAVARTI: Let me tell you the following: we can, I can put you into the right sources but really my main thing is you should have come to Hopkins and – (laughter) – that –

Q: I have flexible time. I can stay.

HUDSON: We're going to move on to the next question. We have a question here?

Q: Thank you my name is Assya Pascalev of Howard University and my question is for Mr. Wells. I would like to learn a little bit more about the protection of the

confidentiality and the privacy of your subjects from the general public. You mentioned that this is done through coding and also it's an anonymous database or databank –

MR. WELLS: That's right.

Q: But of course coding doesn't imply necessarily anonymity and obviously the fact that something is anonymous only if it's not traceable to the source. You also mentioned that the only way for the public to purchase the kit is through going on the Web site, which I assume also implies you provide credit card information. And although the purchaser doesn't necessarily have to be the person whose DNA is sent, chances are that would be very high overlap there. So how do you get around this issue? If I'm a very suspicious yet curious member of the public, what would you say to kind of...

WELLS: Yes, that's an excellent question; it's one we thought about quite a bit when we were launching the project. You do have to provide a credit card to order from the Web site, but that's not linked in any way with the code that goes into your kit. Again, we don't know which code goes into which kit. It's done totally anonymously, unknown to us. There's a central repository where we put the stickers on and put them into the kit, it goes out, another group of people puts them in the outlets. There is no way to connect a name to that number and that's the only way to track the sample. It's also the only way to get your results back. If you lose that number, you have to order another kit. You're SOL, so to speak. (Laughter.)

Q: Can I ask you a follow-up question? You're talking about the cheek swab part of it. So for the indigenous part of it where the samples are actually being stored and then presumably will be available for researchers, how are you –

WELLS: Those are assigned the same sort of number that we give to the general public, so anyone from the indigenous community who does have Web access can then take their number and anonymously go on and get the results back. That's the only way the samples are tracked but of course, in the case of the indigenous groups, they are coded and the names are actually kept under lock and key at the regional center.

HUDSON: Questions back here?

Q: Good afternoon, my name is Lois Tully. I'm from the U.S. Department of Justice, Office of Science and Technology, and my question is about the individual who was surprised when he got the results from DNAPrint Genomics, it rocked his world, but is there any concern about the accuracy of the testing that was done in that case or is there any way to kind of monitor organizations that do this type of service for the general public and how accurate their protocols and their markers, et cetera, are?

WOODWARD: I don't think we have any representatives of any of those commercial companies here. I think that's a big concern and I think that needs to be looked at very carefully, and I think that's part of what we need to do with the public is to

make sure that they know what they are getting into when they look for a genetic test of any type, particularly on the Internet, and send money and DNA, that they know what it is that they're getting back.

I know in our case we do double blinds as we process our information, we use a laboratory that is AABB certified and CLIA certified and take those kinds of safeguards to make sure that we're performing to the highest standards in the laboratory, but I think the question you raise is an important question and that we need to make sure that people do have access to reputable testing companies.

MR. WELLS: And I would follow on by saying that is part of the reason we're not doing that sort of testing in Genographic. I suppose there are some individuals who would like to know what percentage – and basically these are racial groups; they may call them continental groups, but they're effectively old-fashioned racial groups – what percentage they are of African and European and East Asian.

In terms of the history of our species and the migratory patterns that we are studying, that really doesn't tell us very much. I mean, because there's so much diversity within Africa and within Europe and within Asia. Moreover, I do know of cases where people from the same families have sent out the samples and gotten fairly different results from those tests. It has to do with chromosomal segregation and recombination and things like that, so I would argue that the accuracy isn't quite as good as the companies might be saying that it is.

MS. SOO-JIN LEE: And I'd add that there is healthy deference on the part of the public for these technologies. In fact in the case of Wayne Joseph, he not only didn't contest the results, but he actually said that in the next census he's going to mark off Native American. (Laughter.)

MS. HUDSON: We have time for one more question over here.

Q: I'm Maggie Fox with Reuters and I just want to follow up on this because given what you guys just presented and given the little bit that I know about this, what that guy in Southern California got was – I don't know if they tested his Y chromosome or his mitochondrial DNA, but he just got one line of one ancestor, is that correct? So how did that work and how do people interpret those?

CHAKRAVARTI: It's always very difficult when also a member of the public tries to think of what the company really did, but I think it's fairly well advertised that DNAPrint, if that's the company – many of these companies use a number of markers that are called ancestry informative markers: markers that are essentially all or none for the great, at least, continental groups of humans. Again, very small groups of individuals have been sampled so far to say that – so what we do is we study group X, find the marker that's only existent in that group and not in the others, and we say you have a certain fraction of belonging to that group. But we, of course, have no idea that there aren't 20 other such groups in the world that only have that marker as well, and so I'm

not saying that it's the principle of the approach that is incorrect, but we have far too little data to be sure that this is working at this stage.

And secondly, it's like any other test in medicine or otherwise – that the data have to be published and argued over and squabbled over before we really get enough faith in it to using in a wide-scale way.

WELLS: And then, of course, there is the whole issue of what is race and to what extent is it a social construct and what does it mean to be 37 percent European and 46 percent African and so on. These are big issues and I would argue that those are actually very real issues for those types of tests and I would say in the grand scheme of things, those are the ones that worry me the most.

HUDSON: And on that provocative note, I'd like to ask you all to join me in thanking our speakers, and I'd like to thank you for coming. (Applause.)

(END)