MICROBIOME
Larry Thompson
National Human Genome Research Institute
06/13/12
9:00 am ET

Operator: Good morning and welcome to the Human Microbiome Project Press Conference hosted by the National Institute of Health. This press conference will last for 60 minutes. There will be five principal speakers who will provide brief remarks and then members of the media will be invited to ask questions. To ask questions, you can press the star key and the number one on your touchtone phone to enter the queue. You leave the queue by pressing the number or hash key. This call will be recorded, transcribed, and available soon after 1:00 PM Eastern today on the website of the National Human Genome Research Institute, www.genome.gov. Now, I will turn the program over to moderator, Larry Thompson, Chief of Communications at the National Human Genome Research Institute.

Larry Thompson: Good morning, everybody. I'm Larry Thompson at NHGRI, one of the 27 institutes and centers at the NIH. I want to welcome everyone to the call and I think will be pretty interesting. I want to remind everyone on the call that the contents of this briefing may not be made public until 1:00 Eastern time today in keeping with Nature magazine's embargo policy but it is otherwise all on the record. If you need a copy of the press release, do call the Communications office at NHGRI and of course, we'll be posting all these materials on genome.gov at 1:00 today. Today's briefing focuses on a series of scientific reports from the Human Microbiome Project which is usually referred to as HMP. The papers are being published in a coordinated way with the journal, Nature, and several journals in the Public Library of Science or P-L-O-S or PLoS. These papers represent the work of many, many people. Since we can't have everybody on the call,

their work will be represented by five speakers this morning. Dr. Eric Green is Director of the National Human Genome Research Institute which manages separate for NIH. He will talk about why this study was done. Dr. James Anderson is the Director of the NIH division of Program, Coordination, Planning and Strategic Initiatives and which oversees the NIH Common Fund and will talk about how NIH organizes these kinds of efforts. Dr. Bruce Birren is the Director of Genome Sequencing Center for Infectious Diseases and co-director of the Genome Sequencing and Analysis Program at the Broad Institute of MIT in Harvard and will describe how the study was done in some of the principal findings. Dr. Phillip Tarr, the Melvin E. Carnahan Professor of Pediatrics and Director of the Division of Pediatric Gastroenterology and Nutrition at Washington University School of Medicine in St. Louis will talk about how the HMP data will be used in medical research. Dr. Tarr's running a little bit late so we're hoping he'll be joining us and he's now on the phone I'm told. Okay, so that's a good thing, and Dr. Amy McGuire, the Associate Professor of Medicine and Medical Ethics and Associate Director of Research for the Center of Medical Ethics and Health Policy at Baylor College of Medicine in Houston who will provide an overview of HMP-funded research in the area that's usually called ELSI or the Ethical, Legal, and Societal Implications research program. That was started way back in the days of Human Genome Project. I've asked each speaker to give out three to five opening minutes of remarks. In addition, Dr. Lita Proctor from the Genome Institute who has served as the Program Manager for the Human Microbiome is here in the room with us and will be available answering the questions. So I'd like to go – we'll start with opening remarks then we'll go to questions and we'll proceed from there, so Dr. Green?

Dr. Eric Green: Thank you, Larry. Good morning, everyone. Thank you for joining. Physicians and researchers alike have long known that human share their bodies with trillions of microorganisms. For example, previous estimates suggest that there are 10 bacterial cells for every human cell in our bodies but because of their small size, microbial cells only make up about 1% to 3% of our body's mass. Still in a 200-pound adult, that is 2 to 6 pounds of bacteria, a rather remarkable amount. Now, some people talk about the human body as a super organism composed of human cells and the community of microbes that live in us and on us, that is, our microbiome. Most of the time, we live in harmony with these microbes and we are aware of evolutionary forces that have shaped those relationships but sometimes that harmony breaks down resulting in disease. To prevent illnesses or to restore health once illnesses occur, we need to understand better what the normal microbiome is like and what happens to it when it changes to cause or to influence to disease. This requires studying and understanding the interaction of communities of microbes in our bodies, not just single microbes one at a time. The challenge is that microorganisms are hard to isolate from the body and grow in the laboratory to study them in detail. In fact, even today we are unable to grow in the laboratory the vast majority of the bacteria that live in us and on us. Further to date, the fields of microbiology and pathology have mostly focused on microbes that appear to cause disease and as such, only a few hundred species have been isolated from the human body and studied in the laboratory. Thousands and thousands of additional microbes remain unstudied. As such, we know very little about the overall composition of the normal human microbiome until now. Recent advances in genomics now provide the opportunity to open our eyes to the human microbiome in an unprecedented way. Specifically, the powerful new DNA sequencing technologies can be harnessed to inventory the

microbes in our microbiomes without first having to grow those microbes in the laboratory. We can simply sample a body site, purify all the DNA in that sample, sequence that mixture of DNA and then use computational methods to analyze those sequences and identify what microbes are present and at what amounts. The beauty of this approach is that it identifies everything that is there giving us complete views of the microbiome at a given body site like an explorer mapping the coastline of a newly-discovered continent for the very first time. Because of the massive number of samples involved, an effort like the Human Microbiome Project could not have been considered if the cost of DNA sequencing had not plummeted in recent years. When we started the Human Microbiome Project in 2007, we suspected that the new DNA sequencing machines that were just arriving into genomic laboratories would profoundly reduce the cost of DNA sequencing to the point of making large-scale microbiome research feasible. Indeed, over the last five years, the cost has dropped from hundreds of dollars per megabase of DNA sequence to fractions of a penny. Moreover, the lessons we learned from the Human Genome Project is that we needed a reference database of what is normal before we could start asking questions about what is abnormal. The Human Genome Project gave us our first complete reference, human genome sequence and that has proven to be invaluable from wide ranges of disease studies. To do the same for the Human Microbiome, we need to know what normal looks like. What microbes live in what parts of the body? What do they do when they're there? How do they interact with each other and with healthy normal cells? Well, to learn this, NIH set out to create a data set of genomic information that define normal and that is what the Human Microbiome Project consortium is presented in the papers we're discussing today. Finally, I want to emphasize remarkable scale, the work we describe in these papers. This

has been a massive project starting with the recruitment and sampling of some 300 healthy volunteers from which samples were derived and then fed into the DNA sequencing pipelines. Because the microbiome is important for so many different disorders, it is the kind of research that many NIH institutes and centers are interested in and that is why this has been pursued as an NIH Common Fund initiative. I'd now like to ask my colleague, Jim Anderson, to talk about why this became a Common Fund project and how the work was organized. Jim?

Dr. James Anderson: Thank you, Eric, and good morning, everyone. The NIH Common Fund supports research on broad challenges and opportunities in biomedical research that require coordinated approach and have the potential to fundamentally change or transform an existing field of research or actually to create a new one. These programs are typically quite large in scale and involve multiple complementary initiatives that are led by teams of NIH institute and center staff. New programs are identified by a strategic planning process but first, yes, there's broad input from the scientific and lay communities about emerging opportunities and challenges in biomedical research and then through an iterative process of analysis of the NIH research portfolio complemented by expert input, the ideas are refined into specific research initiatives designed to have the greatest impact to advance the field. The HMP was launched in 2007 in response to growing public interest about the role of microbes in health and disease and as Eric pointed out, very importantly, the availability of new emerging approaches to genome-wide sequencing and analysis. The centerpiece of the program being highlighted today was designed specifically to enable the community at large by defining the microbial makeup of healthy individuals and to serve as a reference for disease-specific studies. The program exemplifies the power of community engagement and science. It

took more than 200 researchers at 80 research institutions and five years of work to define what's normal. Other initiatives in the HMP support complementary projects to synergize microbial researches both nationally and internationally. The HMP Data Analysis and Coordinating Center or DACC for example is providing rapid public access to microbiome data generated through the program. Other HMP initiatives are fostering the development of new approaches to overcome obstacles to isolating and sequencing the difficult-to-culture organisms and testing new computational approaches to make sense of the vast amounts of sequenced data that are being generated by the program. To begin to apply this vast amount of sequenced data to ask biological questions, a series of demonstration projects are investigating the relationship between the human microbiome and different diseases and I anticipate that Dr. Phil Tarr will be highlighting some of these findings later in the tele-briefing. Recognizing the importance of Ethical, Legal, and Social Implications or ELSI for human microbiome research, an initiative devoted to ELSI is a key component of the program and Amy McGuire will be speaking on this topic later. The Common Fund provides a strategic approach to supporting large-scale complex science that no single NIH institute is likely to take on, yet all institutes serve to benefit from. The Common Fund investment has been substantial to date at about $153 million plus several NIH institutes have invested about $20 million through co-funding arrangements making the total NIH investment in the HMP about $173 million. Overall, I'd say the investment is paying off by establishing a new field of microbiome research and catalyzing dozens of disease-specific studies. So Dr. Birren, I believe you're next.

Dr. Bruce Birren: Thanks, Jim. I'll be describing work of the two HMP clinical centers, the four sequencing centers, the data analysis and coordination center as well

as the many members of the research community who contributed to this project. Together we worked closely to develop methods needed to identify healthy human subjects and eliminate obvious environmental influences that might alter the microbes we needed to measure. We developed methods for sampling microbes from different body parts and for extracting their DNA and both laboratory and analytical methods needed to sequence and identify the organisms present in these samples as well as the genes they carry. Although metagenomic studies have been carried out before, nothing of this scale or complexity had been attempted and we carefully validated our laboratory and computational protocols to ensure they were both reproducible and accurate. Every step from sampling to analysis has been documented in detail and this information has been shared. HMP participants created numerous working groups to tackle the key challenges in data production and analysis and kept in close contact through regular meetings and weekly phone calls. This close coordination allowed us to take advantage of two complete shifts in sequencing technology that took place over the course of the project which in turn allowed us to generate the largest metagenomic data set yet produced. Our sampling used subjects in two US cities with each subject receiving medical and dental exams prior to sampling. In addition to the large number of healthy subjects and the many different body sites we examined, this careful clinical characterization of each subject makes these data unique. Our sampling focused on five main body areas and within these, we looked at 15 sites in the male subjects and 18 sites in females. Subjects were sampled up to three different times. We used two different DNA sequencing methods to examine the microbiome. First, we measured the bacterial content of thousands of samples by sequencing a specific gene known as the 16S gene. It is found in all bacteria but not in humans. The 16S gene varies just enough between different bacteria that

the 100 million 16S sequences reproduced served as barcodes allowing us to count and identify the groups of bacteria in each sample. In addition, from hundreds of these samples, we sequenced all the DNA we isolated producing over 3.5 terabases of microbial sequence to identify the many different genes present within these communities. Simultaneously, we sequenced more than 800 reference genomes from the human microbiome to help us interpret metagenomic data. Our findings include the fact that healthy humans carry a remarkable diversity of organisms. We found over 10,000 species of microbes living in and on our subjects and while the human genome contains roughly 22,000 genes that specify the human organism, these microbes contribute more than 8 million different genes, many of which play critical roles in our own development, nutrition, and health. We found that different body habitats have different signature organisms, thus, the microbes on skin are easily distinguished from the microbes in the gut but the diversity and the abundance of these signature organisms vary widely between people. Apparently, there are many different ways to be healthy when it comes to our microbes. Despite this extensive variation in the microbes in a given body site between different people, it is striking that the metabolic functions the microbes carry out are relatively constant, that is, although different people may have different collections of organisms on, for example, their tongue, each person's tongue microbiome has roughly the same pathways for breaking down different energy sources like simple sugars. Apparently, it's like a potluck dinner and that the same person doesn't always need to bring plates or forks as long as someone brings them, everyone will eat. Overall, our sampling encountered between 81% to 99% of the genera, the enzyme families, and the community configurations occupied by the healthy western microbiome. Our future work must go on to look at additional populations and also study how our microbiome is established

in infants and is maintained throughout life and is modified by our diet, lifestyle, environment and genetics as well as by disease. Dr. Tarr will describe the value of these data and the approaches taken in the clinic.

Dr. Phillip Tarr: Okay. This is Phil Tarr from Washington University and I'm going to provide some perspective in three different realms. I'm going to talk as a participant in this project who is now trying to focus on the data that have emerged and how they can be used to benefit human health. The areas I'm going to discuss are – try to impress upon the audience are this is really a new vista in biology. It opens up many, many new opportunities to improve the health of our populations and finally, I'd like to make a few observations on the value of this incredible effort in the Human Microbiome Project. So first, this is a new vista on biology. You've just heard the technological tour de force that has gotten us to this point in the Human Microbiome Project. We've now been introduced to this biomass in each and every one of us. These organisms, these bacteria are not passengers. They're metabolically active. As a community, we now have to reckon with them much like we have to reckon with the ecosystem in a forest or a body of water. We are going to be moving out of the old paradigm of one germ, one disease, one person and more into the paradigm of community affects both health and for disease. So this is a whole new way of looking at human biology and human disease and it's awe-inspiring and it's also incredibly – offers incredible new opportunities. The second area I'd like to talk about are these new opportunities. So since the time of Louis Pasteur and Robert Koch, we have thought about again, one germ causes one disease and we know how to handle that that very focused reductionist approach to problems caused by infectious agents. Now, we have to say okay, pathogens happen but they might happen because the other organisms, the other germs in the

body are not a healthy ecosystem. One disorder which we're starting to really focus on this is an infection called C. difficile. This is a germ that frequently attacks hospitalized patients. It can be quite deadly and we don't have a good handle on how to prevent it and we have only fair tools for how to treat it. However, it's becoming apparent that C. difficile does not just happen. It probably happens because there are perturbed ecosystems in the gut in these patients. If we could somehow modulate that, if we can somehow get a handle on that, we should be able to make incredible inroads in treating and ideally preventing this infection which is a major challenge to healthcare systems. Another area in which we predict medical advances will be in the metabolic exploitation of these germs. These are both contributors to health by the products they make. They may also be contributors to weight gain and in children, for example, in resource-poor environments, they may be contributors to stunting and poor growth so we now need to identify what the organisms are. We're at that point now where we can economically, efficiently, and accurately do this by virtue of the HMP project to date. We need to know what differs in people who have metabolic problems possibly related to bacteria and what those germs are making so that we can now come up with novel ways to move forward. Finally, I'd like to make an observation on this fantastic example of [theme] science. This was an outstanding collaboration between federal support – industry. The machines were not invented by the NIH. They were invented by industry and there was a fantastic interaction between advancing technology and federally-funded science and this is also a terrific example of open data. As soon as these sequences were generated, they were made public. Now, they're being amalgamated into these fantastic publications and as a physician who's interested in this area, this is going to be a whole new ballgame and that's been our perspective. Thank you.

Larry Thompson: Okay. Let's move on to Dr. McGuire, please.

Dr. Amy McGuire: Okay, thank you, Phil and good morning, everybody. As both Eric and Jim mentioned Ethical, Legal, and Social Implications research has really been an integral part of the Human Microbiome Project since its beginning. Ethics and legal scholars were included in discussions about study design and implementation and the NIH Common Fund funded six independent research projects that were exploring the ethical, legal, and social implications of this area of research. Outright, we had one of those independently-funded projects and then our project involved interviews with investigators and NIH project leaders as well as individuals who are recruited to participate in the Human Microbiome Project at Baylor College of Medicine. We asked them questions designed to elicit their concerns and hopes for the project focusing mostly on the ethical issues that the scientists were dealing with as they conducted the research. What we found was that for many of the ethical concerns that were raised like informed consent and participant privacy, these are issues that we've been dealing with for many years as they relate to human genetics research and so there was a general feeling of comfort and confidence with the ability of the research teams to address these issues. So one example is that the microbiome project as we heard is a community resource project and it was designed to make data generated from the project widely available to the research community. So as in the Human Genome Project, the genomic information that related to the microbes are being studied, put into a publicly-available database on the internet and as Jim mentioned, this is done largely through the HMP Data Analysis and Coordinating Center or the DACC. However, in order to protect the privacy of the individuals who contributed biological samples, project leaders decided to

put any human DNA that was analyzed into a control database so that only approved researchers could access it. This is what is commonly done now in the field of human genetics. However, it was discovered part way through the project that some of the human DNA was present in the microbial DNA that was being publicly released. So the research team needed to develop new filtering methods to deal with this human contamination. This is done very quickly and although there was some recognition that as technology advances, there may arise new risk to privacy. Most felt comfortable with the protections that were put into place with these new filtering approaches. Perhaps even more interesting are some of the more philosophical, legal, and social issues that this project raises. For example, as Eric mentioned, the primary goal of the HMP was to create a data set of genomic information that defined normal. What it means to be normal and how that relates to what we think about it means to be healthy really turned out to be quite controversial in this project. There are also very interesting questions about whether the fact that we have more microbial DNA in and on our bodies than human DNA changes how we think about what it means to be human. There is a lot of talk throughout the project and in the media about us being super organisms and how does that influence how we think about ourselves, our health and the causes of disease. Questions have also been raised about who owns the microbiome that co-inhabit our bodies and what legal and social implications that might have and of particular interest I think are questions about how different prebiotic or probiotic products will be regulated in the future. So right now, many of the probiotic products like different yoghurts or supplements that are out there on the market are regulated as foods and not as drugs but as we develop more scientific evidence about their health benefits, will they be treated as a class of drugs and what kind of evidence will be required for companies to make health

claims about their product. One of the other funded ELSI research projects being conducted by Diane Hoffmann at the University of Maryland Law School is specifically looking at this important legal issue. The other ethics projects that were funded through this consortium focused on issues related to the study of human microbiome research among indigenous communities and this is done by Paul Spicer and his colleagues at the University of Oklahoma. Patient perceptions about therapeutic probiotics which was conducted by Rich Sharp at the Cleveland Clinic, an analysis of how risks and benefits are conceptualized in human microbiome research and this research was conducted by Mildred Cho at Stanford University and Pamela Sankar at the University of Pennsylvania and a philosophical exploration of the ethical, legal, and social implications of human microbiome research conducted by Rosamond Rhodes at Mount Sinai Medical School. So I'll now turn it back over to Larry who will open it up for Q&A.

Larry Thompson: Thank you, Dr. McGuire. So let's open this up for questions and I would ask the operator to take the first question and let us know who the questioner is and I would ask my colleagues here to re-identify yourselves as you're speaking. Not everybody has memorized all your voices yet and we'll know who's answering the questions. So first question please?

Operator: The first question comes from Mark Johnson with Milwaukee Journal Sentinel. Please go ahead, your line is open.

Mark Johnson: Thanks very much for the presentation and for taking our questions. Two-part question, I wanted to just check what cities the subjects came from. I'm also very interested in learning how you go about assembling a reference microbiome. Are you looking at what's normal in terms of the

difference – the variation of different microbes or are you also looking at what's normal in terms of the specific genes or the specific DNA in each microbe?

Larry Thompson: Dr. Birren, you want to give that a shot?

Dr. Bruce Birren: I would. This is Bruce Birren. The clinical centers were in St. Louis and Houston and when we talk about assembling our reference microbiome, what we mean both the organisms that are present and their abundance. Both of those can vary between people and we want our reference collection to capture that wide range of variability across a healthy population. Separately, we sorted through the metagenomicro – or shotgun data to look at finer scale variation just because we think a bacterium is present, there can be subtle differences in strains that might be in one person or another and our data captured that as well.

Larry Thompson: Anybody else like to add anything? Okay? Mark, does that meet your need?

Mark Johnson: Yes, it does. Thank you.

Larry Thompson: Great. Want to move on to the next question please?

Operator: Next we have Matthew Herper with Forbes. Your line is open.

Matthew Herper: Hey, I have a couple of questions I think mostly for Bruce. So you've spoken about a change in about two shifts in technology – two complete shifts. Did you mean the switches to 454 and Illumina?

Dr. Bruce Birren: Yes, I did.

Matthew Herper: Okay. When we talk about being able – how do we understand what the metabolic pathway is when we're talking about the gut or the tongue or the skin, your potluck dinner analogy? How do you – and can you give us a little more on what the diversity within and – or can anyone give us a little more about the diversity within a person between and within [sizes]? I mean there were some nice graphics in the paper but I'm having trouble figuring out how to digest them into a sentence.

Dr. Bruce Birren: Alright. Well, let me start on the metabolic functions and how we recognize those. We actually used a couple of different methods to wade through this ocean of data we produced. The first was to take the reads and ask what bacteria do we know about are they most similar to and what are the genes in those bacteria and then simply by counting those genes and knowing what functions they carry out, we could make lists of what we would expect that community to – what metabolic functions would take place. We also took all of those data and assembled them into a metagenomic assembly that would also help us recognize genes in them and similarly asked what do we know about those genes and other environments that tell us what they do and that's what's allowed us to recreate these pathways that for instance would tell us about the microbes that could digest fats as a source of energy or that might digest complex carbohydrates. So in terms of diversity – oh, the term "diversity" conveys a lot. We mean that for instance, the organisms we would see in any given body sites could be highly variable between people. If you look at the figure two in the paper, the very colorful figure two that describes how this varies, you could see that in the gut for instance, in certain people up to over 90% of the organisms present would be just from a single grouping

where other people had virtually none of those organisms. So that would vary between people and people. Other aspects of diversity are the number of different organisms present in the site. For example, the oral cavity and the gut microbiome are tremendously complex, made up of many, many different organisms whereas the microbiome in the vagina was relatively simple and that typically we found fewer organisms present there.

Larry Thompson: Okay, Matthew?

Matthew Herper: Yes. I think that answers those.

Larry Thompson: Great. Alright, let's move on to the next questioner please, John? Operator?

Operator: Next is John Lauerman with Bloomberg News.

John Lauerman: Hi. Good morning. This is really interesting stuff and I have I guess three questions which I can give all at one time or – and ask them in – you can reply to them however you want to do it.

Larry Thompson: Go ahead. Why don't you ask them, John?

John Lauerman: Ask them all three? Okay.

Larry Thompson: Sure.

John Lauerman: My first question I guess is I don't think I've heard the word "symbiotic" and I'm wondering if it's appropriate to use that word or not and if you

could talk about whether these are symbiotic relationships or are they approach – are they approximate that or not?

Dr. Phillip Tarr: I can potentially take that.

John Lauerman: Who is it? Who's talking?

Dr. Phillip Tarr: Sorry. This is Phil Tarr.

John Lauerman: Okay, Tarr.

Dr. Phillip Tarr: The answer is sure, the microbial population is providing some benefit to the human host and in return, the human host is harboring those bacteria enabling them to replicate and be stable and perhaps even thrive then I think that would meet the definition of symbiosis. So we think they are not only what we find pathogenic or harmful populations but we're going to be able to find symbiotic and good populations and this is again a new opportunity to manipulate these ecosystems but the first step here is the definition which we've now been given by the HMP and then comparison of health versus disease states.

John Lauerman: Okay. That's great. Second question is the human DNA. Were you surprised to find that in the samples or was that sort of a [unintelligible], et cetera and why were you concerned about that and what are the potential risks of releasing human DNA sequences in this public database?

Larry Thompson: So let me ask Lita Proctor to take a first shot at it and then Dr. McGuire to also to respond to that.

Dr. Lita Proctor: Hi. Good morning, Lita Proctor. No, we weren't surprised at all to find human DNA. The samples, remember, these were sampled from humans.

John Lauerman: Right. Okay.

Dr. Lita Proctor: So you do a swab of your skin or a swab of your tongue, you're going to collect both the microbes that live there as well as the human material and when you're sequencing as Bruce mentioned, whole genome shotgun sequencing or shotgun sequencing, you're sequencing all the nucleic acid from that sample. So it isn't a surprise. What we found was there's a terrific range of human so-called contamination in samples, everything from less than 1% in the stool which acted as a proxy to the gut or the gut tract all the way up to over 90% for vaginal microbiome samples or skin microbiome samples. So the range was tremendous but the presence was not a surprise.

John Lauerman: What was the concern about releasing that into public databases?

Larry Thompson: So Amy, you want to answer that?

Dr. Amy McGuire: Yes, I can take it. So this is a topic that's actually currently getting a lot of attention in human genetics because there's been lots of debate about sort of how to make human genetic information available to the research community in a way that balances sort of protection of privacy of individuals from whom the DNA came and also sort of the scientific utility of making the information available and the actual privacy risks I think are somewhat uncertain. We do know that if you have human – actually relatively small amounts of human DNA available, there had been some scientific studies that showed that if you have a reference sample

from somebody for example, you can match it to a database and identify that that person is in the database. There's been a lot of policy discussion about how serious of a risk is that and what should be the appropriate response in terms of how broadly to make human DNA data available to the public. So in the area of human genetics who's kind of come to a place now where in order to protect privacy and because of sort of the rapid technological advances that are ongoing and the uncertainty about future privacy risks associated with having information available more broadly, most human DNAs put into a controlled access database and for NIH funded-studies that goes into a database, the database for genotypes and phenotypes or dbGaP where there is a data access committee that you have to go through in order to access the data. So from the beginning of the microbiome project, it was always the policy that all the human DNA that was being collected would go into that controlled access database so this is really a technical matter of being able to effectively filter out the human DNA and the adjustments that needed to be made throughout the project were – and making sure that those filtering approaches were keeping up with sort of the technology.

Larry Thompson: Okay.

John Lauerman: Okay. I have one more question. So did you find any variation? Were you able to in any way determine – you said you found different types of microbiomes in different individuals and was there anything that correlated with that along with the types of individuals and I'm just going to say along racial or ethnic lines or where they came from in the country, geographic lines, age, I don't know?

Larry Thompson: Dr. Birren, would you like to answer that please?

Dr. Bruce Birren: Sure. Well, we did collect a lot of different information about these patients – sorry, subjects – all of whom had been pre-screened to eliminate any extreme health issues. Certainly, we didn't see any dramatic or surprising correlations between different bugs. I think the sort of – one that was not so surprising but strong correlation was the vaginal microbiome and the pH of the environment. Then we did see correlation between ethnicity in the subjects and some of the bacteria seen. That's a very complicated formula. It could be reflect all of the environmental, cultural, dietary, genetic elements so it's really early. We don't know what that really means.

Dr. Eric Green: This is Eric Green. I thought I'd make one comment. I'm sure Bruce would agree with this is that we should also view these publications as sort of the initial analyses of this very large data set. The whole idea of a large project like this much like the Human Genome Project has produced a tremendous amount of data and make it publicly available to the scientific community so that analyses and more creative ideas might bubble up of different way of analyzing this data, so even though the initial set of publications might describe a set of analyses that were done, one can imagine a year from now, two years from now, three years from now down the road, other ideas might bubble up. You go back, reanalyze that data and other correlations might come to the fore.

Larry Thompson: Dr. Proctor, is there something you'd like to add?

Dr. Lita Proctor: Oh, yes, I want to add to Bruce's comments about the correlative data. The way that the healthy cohort study was designed, there were 300 individuals total. I believe about 10% of the subjects self declared a racial

minority and another 20% self declared an ethnic designation but that meant really that the populations of non-white subject was very, very small. So any real effort to look for [forum] correlations or strong relationships between microbiome properties and ethnic or racial groups is going to require much larger cohort studies than what we were able to tackle with the healthy cohort study.

John Lauerman: Cool.

Larry Thompson: Alright.

Dr. Bruce Birren: This is Bruce again. I think any of us who work in human genetics understand that sometimes you get a glimmer and you follow up and you need to look at many more people before you have confidence that that holds true.

Larry Thompson: Makes sense. Alright. Thanks, John. Let's move on to the next question please.

Operator: Next we'll go to Elizabeth Lopatto with Bloomberg News. Please go ahead, your line is open.

Elizabeth Lopatto: Hi, folks. I guess my question is how can this data be used alongside the data from the Human Genome to understand the way that environment influences disease and are there any obvious candidates to this kind of analysis going forward?

Larry Thompson: Phil, do you want to take a shot at that?

Dr. Phillip Tarr: Sure. So whenever you have any microbial human interaction, the outcome of that interaction is a function of the microbes and the human and the human response to it and we learned that over – we learned that the monomicrobials infection legacy to date, one person will get a very mild case of pneumonia, another person will go on to a fatal outcome. Now that we have the potential drivers or a census of the potential drivers on the microbial side including suspects that we never had really – never thought of as being bad actors but now might have to reconsider and now that we have a compendium of the human genes and allelic variations, we can use informatics and develop new statistical paradigms and approaches to correlate what sets of combinations are good for the host and what are bad, so the same technology, two different populations of genes but the same outcome.

Elizabeth Lopatto: If you don't mind, you mentioned you had subjects you'd never thought of as bad actors, could you give me an example?

Dr. Phillip Tarr: Okay. So this is not quite so much a specific pathogen discovery. It would be imbalance of classes of organisms and until we have a little bit better – until we now extend this baseline censusing to our patients, we may see for example a blossom, a bloom or a diminishment in various taxis such as [Firmicutes]. They had ordained outcome so we would be looking at host responses to classes of organisms. Now, the next chapter in the story is going to be what is it in the disease populations compared to this reference group and also compared to continually [enrolled] controls. So we now have a platform which we can build but the answer in a word is yes, human genome data and microbial genome data are partners.

Larry Thompson: Dr. Proctor would like to add something?

Dr. Lita Proctor: Yes. I wanted to comment that I'm sure it's clear to the listeners that the Human Microbiome Project was completely focused on understanding the microbes of the human microbiome. Of course, that's only half the story, right? We didn't really focus on the human host side. So we hope that this will really catalyze future studies where the volunteers can be broadly consented so that we can not only sequence the microbes but also sequence the volunteers so we'd have genetic information about the host as well as genetic information about the microbiome and I think definitely you'll really start to see real emerging ideas about what the relationships are between the microbes and their human hosts.

Larry Thompson: Okay, Elizabeth?

Elizabeth Lopatto: Yes, thank you.

Larry Thompson: Great. Let's move on to the next question from Mark, please?

Operator: Mark Johnson with Milwaukee Journal Sentinel, your line is open.

Mark Johnson: Yes. Thank you. A couple of questions, one is I'm assuming from the way you've been talking about this that this is not the end of the Human Microbiome Project but just an early preliminary stage. I was hoping you could talk about what the next stage of work will be. I'm also interested in whether you could maybe spell out a few scenarios in which the knowledge you've gained would lead to different strategies to defeat diseases.

Larry Thompson: Can I ask Dr. Anderson to start with that please?

Dr. James Anderson: Yes. Let me address where we are with approaching research in the Human Microbiome. What we've talked about today was a coherent trans-NIH approach that was supported and organized by the Common Fund. This was to enable everyone in the community to start to do this work. What we found is that many of the institutes at NIH are now supporting projects in this area. We also find that there's a robust support interest around the world, in Europe and in China and so we need to think about the investments both from the Common Fund and just the enabling ability for everyone else to start work here. So we are at NIH thinking – well, let me back up and say, "Why was this project so successful?" It was really impact by design. We knew how we could get there and what the impact would be if we sequenced – created this reference database, the algorithms and tools to query and recover information from the database and what we're doing in NIH now is review of what is that next opportunity that we could provide and push for the field.

Larry Thompson: Dr. Green, you want to add…

Dr. Eric Green: I wonder if the second question we might want to ask Dr. Tarr?

Dr. Phillip Tarr: Sure. I'm going to revert to the example of C. difficile or Clostridium difficile. So if you take 1,000 adults walking into a hospital for an elective or even non-elective admission, a couple dozen will develop variably severe C. difficile infection and one or two may develop incredibly severe C. difficile infection. We are terrified of this germ. We know how to reduce some of the risk, good hand washing, very good hygiene within the hospital but we have not gotten that risk down to zero. If we can identify when those people walk into that hospital who is at major risk by

examining their microbiome, who is carrying in their gut the community of germs that will either enable this pathogen to flourish or to stay under control, we can anticipate this outcome and prevent it.

Larry Thompson: Dr. Proctor, do you want to [crosstalk]?

Dr. Lita Proctor: Yes, let me just add something to what Phil Tarr just said. I hope that the audience here is realizing sort of one thing when we're talking about variability and diversity and dynamics of the microbiome. Those properties of the microbiome are actually to our benefit. The microbiome is a mutable, changeable property and that actually can be exploited for supporting health and curing illness so I think that's probably a major take home from our discussion this morning.

Larry Thompson: Okay and so I want to observe that we have about 10 minutes left and I know there are at least four questions in the queue so I'd ask you guys to keep your questions short and my colleagues to keep their answers brief. Alright, the questions brief and the answers brief. So let's go to the next question please.

Operator: Matthew Herper with Forbes, please go ahead, your line is open.

Matthew Herper: Okay. Best as I can, how do we know given that we can't culture these, how do we know what we're missing? How much of the microbes are we actually getting to the sequencing method? I mean it sounds like particularly the method only gets bacteria. Can people be identified by their microbiomes?

Larry Thompson: Bruce, you want to give a shot at that and then Amy, you want to answer the identification issue?

Dr. Bruce Birren: Sure. So we have a pretty good sense of what is new by comparing our data to the reference databases of what we've already seen. We don't need to culture these but very often the sequences we see are similar enough to things we know well that we recognize them. There is an effort to make sure we've not only identified new organisms but plug in the technologists who have the ability to isolate those organisms and sequence them even perhaps without needing to culture them to fill in those holes that are information.

Matthew Herper: I'm basically asking if there could be holes in the DNA as the DNA you're not – is there DNA or I don't [crosstalk]?

Dr. Bruce Birren: Sorry, holes in the DNA, yes, but of course, the 16S focuses on the bacteria but the sequencing of all of the DNA doesn't have that particular bias in it and we know for instance – although most of the DNA we saw was bacterial, there were Eukaryotes and Archaea in there in small numbers as well.

Larry Thompson: Amy, you want to discuss identifiability briefly?

Dr. Amy McGuire: Yes, and Matthew, you're asking how identifiable is the microbiome? Is that your question?

Matthew Herper: Yes.

Dr. Amy McGuire: Yes. So there had been some very early studies that have looked at this. Rob Knight's group out of Colorado has looked at sort of if you look at somebody's keyboard for example, can you match sort of the microbial communities that are shed from their fingertips and match it to their own fingertips and they had some very small sample sizes but some early suggestions that we might actually have more unique microbial communities than we might have first thought. So I think this is very much an area that is we're still exploring and it's not entirely clear sort of how unique the microbiome is and then once – if it is in fact sort of unique to individuals then how identifiable it might be and how you would go about identifying people based on their microbiome. So this is something that I think is gaining a lot of attention because of course it does have a lot of implications from an ethical and policy perspective but I don't think we have the evidence yet to really say with certainty how unique or identifiable the microbiome is.

Larry Thompson: Great. Thank you very much, Dr. McGuire. Can we move on to the next question please?

Operator: Next we have Tina Saey with Science News, your line is open.

Tina Saey: Hi. I have a question about some of the numbers of the species inhabiting the various body sites. This is from the data papers in Nature and also [unintelligible] paper in PLoS ONE. There seems to be quite a range of estimates of the number of species and the number of genes that are in each body site. Is there yet a consensus about what's there or should be there?

Larry Thompson: Bruce, do you think you want to try that?

Dr. Bruce Birren: Well, I don't think we're going to force this to a single number because we saw such large variation between people. Some people have many more organisms or many different organisms than other people do so what we've tried to describe is among those group of Western healthy adults, what are all of the organisms we see there.

Larry Thompson: Okay. Alright, can we move to the next question please?

Operator: Next we have Carolyn Johnson with Boston Globe, your line is open.

Carolyn Johnson: Hi, I just wanted to clarify, the metagenomic data, could you just explain, is that – what kinds of organisms would that include? Would it be bacteria and like virus and fungi or is it mainly an only-bacteria – the 16S [unintelligible] would be only?

Larry Thompson: Dr. Birren?

Dr. Bruce Birren: Sure. Carolyn, it's everything we could crack open so it was primarily bacteria but there were also microeukaryotes present, Archaea present and viral sequences present.

Carolyn Johnson: Okay. Thanks.

Larry Thompson: Okay. Let's go to the next question please.

Operator: Next we have Lauran Neergaard with the Associated Press, your line is open.

Lauran Neergaard: Hi, thanks. Is there a core of normal microorganisms that we all carry? I mean just reading these papers, the diversity is so striking that it made me wonder, alright, how can you tell which of all of these microbes are the normal ones?

Larry Thompson: Dr. Birren, I hate to make you work so hard but can you help us on that one?

Dr. Bruce Birren: I'm delighted people are interested. I think the conclusion from looking at all these different body sites that there are core functions clearly but as I said earlier, we don't all have the same bacteria although they all seem to have been organized to do the similar things.

Larry Thompson: Lauran, anything else?

Lauran Neergaard: Well then. how can you tell which ones there are normal and which ones are just waiting to do something abnormal and you just happened upon them in the course of the study?

Dr. Bruce Birren: Well, these communities within individual people seem to be stable, that is we don't see dramatic changes in bloom so the specific thing in the course of healthy life. So it may be that for all the examples we said about our own genetics, our diet, our lifestyle, et cetera, that induces each one of us to have arrived at a solution that works for us about the collection of microbes we coexist with, understanding that other people's microbes may end up doing the same things for them but the difference just as different editors and reporters work together to produce great stories, you don't always need the same person one end or the other.

Larry Thompson: Any closing remarks? Dr. Proctor?

Dr. Lita Proctor: Are we going to close now?

Larry Thompson: I think we're about to wrap it up. Yes, ma'am.

Dr. Lita Proctor: Okay. So I want to emphasize as we close up this tele-briefing that the HMP is always designed to be a community resource of data, tools, scientific projects and so on and our Data Analysis Coordination Center which we refer to as the DACC, tomorrow, in conjunction with the actual publication of the two Nature papers and the virtual HMP collection of papers and PLoS will have posted on their website – so that's dacc.org – all the QC value added data sets that the entire consortium produced together in order to conduct the analysis that's published in the Nature paper. So I invite you to go to the DACC website tomorrow and you'll see many more data sets and you'll see again a full idea of the extent of the data sets. Those are in addition to what's already posted in our public repository which is at NCBI. Amy McGuire mentioned the databases so I invite the public to go to the DACC website to look at the data sets.

Larry Thompson: Okay and we're just about at the top of the hour and everybody's busy I know so I'm going to bring this to a conclusion since we have no more questions but if any of the reporters out there still need anything, please call the Genome Institute's press office and we will help you out in any way that we can. I want to thank everybody for participating in it. All of our experts, I appreciate you taking your time and thank you very much. This concludes the call.

END