

**SPEAKER 1:** We're very lucky to have Dr. Louis Cohen for his Grand Rounds on Understanding the Human Microbiome and Its Therapeutic Potential. Dr. Cohen completed medical school at Tufts University School of Medicine. He received his training in internal medicine at the Massachusetts General Hospital. Dr. Cohen then went on to complete his gastroenterology training here at Mount Sinai before joining as faculty.

In addition to his clinical training, Dr. Cohen completed a Master's in Translational Research from the Rockefeller University. As an NIH-funded researcher, Dr. Cohen studies host microbial interactions. He's currently involved in clinical trials studying the use of autologous stem cell transplant for the treatment of Crohn's disease. Dr. Cohen's clinical interests are in inflammatory bowel disease and Hermansky-Pudlak syndrome.

He currently serves as an assistant professor here at Mount Sinai, where he practices as a gastroenterologist in the Department of Medicine. Please join me in welcoming Dr. Louis Cohen.

**LOUIS COHEN:** All right. So I just want to definitely thank the Department of Medicine for having me out here today to give you a chance to talk a little bit. So I think every young physician, scientist dreams of standing up here at the end of their career and walking you through this amazing research they've done over 20 years. Painting you through everything, and this is great, and there's where we've come.

My talk today is actually focused more on what actually took me into this field in the first place, which is understanding the microbiome. And I think part of it was that, coming at this field as a physician, it seemed like every single journal I opened for a long time told me that the microbiome was magic, that it had something to do in every single disease to the point where it was almost too good to be true.

And so as a physician, I sat back, and I said that, I know that somewhere in here is the reality of it. Somewhere in here is a little bit of hype. And so I set out to try to understand it. And so the goal of my talk today really is to convince you what I've convinced myself. Is the microbiome for real? And I absolutely think that it is.

And what I want to do, after I tell you I have no disclosures-- at least not yet, hopefully soon-- is to provide you first just a general overview. What is the human microbiome? This idea of fact versus fiction. And then what I'm going to do is walk you through-- I call it a primer for physicians-- these studies about microbial ecology, because chances are, when you've looked in these journals and you've opened up these pages, this is really what you're looking at.

How do we do these studies? How as physicians can we look at this? How can we interpret this data critically the same way that we would look at a clinical trial now? And we've been trained to do this. That'll be the first half of the talk.

And then with the second half of the talk, I'm going to transition you a little bit into the research I do which has to do with not looking at the bacteria that are there, not looking at the ecology, but what they actually do, their functions. And part of that came from my understanding of this research and what I view to be a general limitation of it. I also think that this is potentially interesting because as you understand how bacteria affect host cells, that's a natural transition to developing microbiome therapies. And that's really what the focus of my lab is, is in developing new therapies derived from these bacteria.

But that also creates a new field, something that the FDA has struggled with. So to meet this demand for people that say, this is exciting, this is real, we're going to do this, they've had to create a whole new category which are called live biotherapeutic products. So I'm going to then, in the final part of the talk, lay out for you the foundation of that and what that means. And again, this talk is really going to be at a bird's eye view. I want to provide a foundation. I'm not going to get into the minutia of every single mouse study, but just to tell you how we are thinking as a field and how it's evolved from its day one.

So the microbiome, again, we're talking about bacteria. And we're not just talking about bacteria. We're talking about bacteria and the environment they're in. And we're really not even just talking about bacteria but every organism that lives in that environment.

So the Earth is a microbiome of which humans are a part of it. And then we can get into other microbiomes, such as the oceans or the volcanic vents. And like with all scientific fields, we steal from each other. So a lot of what we've done in the human microbiome we've stolen from scientists that have studied these microbiome before us.

And part of the reason they study these areas is that the bacteria that live in these places have been one of our best sources for drug discovery. You go down to a volcanic vent or some sponge, and you isolate a bacteria, and you have the next tetracycline, vancomycin, tacrolimus. These are all natural products from bugs. And this is something we'll talk a little bit about later that's driving a lot of my interest in the field.

But we've known for a long time that the human microbiome was an area where there are lots of different bacteria there, lots of different fungi. This has been known forever. But I think that what's evolved is our understanding of this relationship. I think that the way that we viewed this for a long time was that this is war. This is us versus them, that we have all these systems in our body that exist to keep them out and to fight them.

And then suddenly, the tide started to shift. And we said, you know what, it's not us versus them. We now understand that these bacteria, these organisms are really a part of our physiology. But that led to, I think, a misunderstanding of what our relationship is. These bacteria are terrible. They are not good.

Everyone here knows it. You do not want these bacteria in your blood. It is still a war, but it's an evolution. The same way that we evolved with the sun, and our skin helps to protect us from the sun, but if there's too much sun, we get damage. But we can also use the sun to help make vitamin D. It's the same relationship with these bacteria. They are terrible. They will damage us, but we've evolved to protect ourselves.

But we've also evolved to use the things that they make and do to be a part of our immune system, our metabolic system's function. And we understand that this isn't just a simple one to one. This is a very dynamic system and involves not just the actual organisms themselves but the things that they make, the things that we make that interact with them, and also our diets and things that will manipulate them.

So as we came to, I think, understand this relationship better and that this had something to do with human health and disease, we endeavored to set out and start to answer questions about it. And I stole these questions from Marty Blaser, who's down at NYU, because it was actually a very elegant way of thinking about how to approach the microbiome science, that if we're going to approach this field as scientists, we're going to go through a set of orderly questions that build on each other.

First, we would say to ourselves, well, what bacteria are there? We'd painstakingly catalog these bacteria. And then we'd be ready to move on to the next phase of this science-- what do the bacteria do? Because once we understand what they do, we can then understand how they affect us and how those functions then translate into changes in human physiology.

And then we can move on to the next phase, which is to understand what sets the tone of this system. What are the factors that regulate how the bacteria change and move? And then, only after all this research over decades and decades, we could then finally get to what we want to know, which is what makes someone unique. How do you compare and contrast patient cohorts to get at specific microbiome functions that are either relating to disease or relating to health?

And we did actually start with number one. We started with number one with a \$200 million outlay from our government. This was the Human Microbiome Project. And we wanted to just understand very basically what bacteria are there in healthy individuals at different sites in the body.

And so as we prepared to look at this and we prepared to do this, the government provided the infrastructure and the resources to begin to answer this question. But then this happened. We got really excited.

So I put this up here for a reason. 2012 is when the Human Microbiome Project was published. Everything prior to this, we had very little knowledge, at least from an ecology standpoint, of what constituted normal. But yet, there were all these studies that were coming out, comparing disease cohorts, comparing different phenotypes, and saying, wow, the microbiome is different.

So it's as though we took number one. And before we could even answer it, we jumped straight to five. And so what ended up happening was what drove me into the field into the first place. I saw all this data. And I was like, oh my god, this is amazing. But what did it really mean? And that's what I'm going to walk you through.

So what the government did was they needed to develop an infrastructure to study microbiome ecology. This is again what bacteria are there. What we traditionally used to do, and I think you'll find there's a theme of this talk, that despite being young, I think I'm a little old school, you would actually go in and grow things. You find an area, you take a sample, you grow the bugs. This was the heart and soul of microbiology and ecology for decades.

But as scientists, we love to just say, well, what if? How do you know what's missed? Just because you go to a place and try to grow a bug and it doesn't grow doesn't mean it's not there. So this was the eternal catch-22.

And then we developed this idea of culture independent analysis of the microbiome. This is what fueled the Human Microbiome Project. This is what the government, again, laid out the groundwork so that we could actually do this type of science.

And when we say culture independent, what we're really doing is it's the study of bacterial DNA. So we don't need to worry about growing this bug because they're just going to leave behind their DNA. And then you use that DNA as a surrogate for ecology.

So if you really looked at why this happened, it was because before 2012, we were building up the techniques to look at DNA. Now, initially this was very hard for anybody in a small lab like myself to do. But then this happened, which is that the cost of sequencing DNA between 2002 and 2012 plummeted, such that any Tiddlywink lab could grab two sets of patients, grab a sample, sequence the DNA, and publish in *Nature*. And that's really what it was like for a long time.

But it was the cost of sequencing became good, so it became a very accessible science. But just because it was accessible doesn't mean that we weren't learning. So to back up, if you're going to study the field the microbiome ecology, and a lot of what you're looking at in your journals really is the field of metagenomics. A metagenome is genetic material that's recovered from an environment.

So in this case, we can take a sample from humans, a respiratory sample, a stool, sample a biopsy. And you isolate just the DNA that's there. That DNA, you can call it eDNA, environmental DNA, or metagenomic DNA. And for a lot of the time, the majority of that DNA, such as in stool, is coming from bacteria.

The next question becomes, once you've isolated that DNA, how do you define a microbe by a gene? What is it that you do? And so what this is based on the concept of molecular evolution, something I think that a lot of people here are familiar with, if not through your medical school courses then maybe through *CSI*.

But molecular evolution is this idea that genes undergo a rate of mutation at a certain interval over time. Therefore, species that evolved a less time ago are going to share less-- are going to have-- share more of their genes. Species that evolved a long time ago have accumulated a lot more mutations, and therefore their genes are different. So this is just this basic concept that gene similarity equals species similarity and tells you something about the evolutionary history.

So just to provide perspective here, because I think sometimes we get a little human-centric in things. And we forget that bacteria have been here a little bit longer than us. So if you were to think about the two most common types of bacteria that you see in your bowel, we think about Bacteroidetes and Firmicutes.

And you look at this, and you say, OK, well, the similarity genetics between them should be like between us and monkeys. We share most of our genome. There's a little bit that makes a big difference. And that's it. But that's not true at all. That's like comparing us to corn.

These bacteria have had a long time. Do not underestimate them. They were made to do their function, and they do it way better than we do. And this is going to come back later when you think about things like probiotics, bacteria that have learned to live in yogurt. Yogurt are not people. These are not things that are meant to be in people. They're not things to infect human biology, but that'll be-- I'll get a little bit more into that later.

So coming back to gene similarity equal species similarity, the next question people have to say is, well, what gene? And so this is where the 16S gene came in. This is, again, something I think that people understand.

And I'll explain why we chose this gene. We chose it for the first reason, which is that most everything has ribosome. So this allows us a standard by which you could compare humans, and bacteria, and fungi because everything has this one gene that again standardizes things.

The second part of this is this gene has a special feature. It has one region of it that's highly conserved and one region of it that's very variable. And what does that mean? So a conserved region means that if it accumulates a mutation, it does not make that gene happy. The protein doesn't function right, and it will not function. And therefore, it will die, and it will not carry on, whereas a variable region means that it can take these mutations and keep going.

And so what does this allow us to do? It allows us to save a lot of money because when you have a metagenomic sample, you have all these different types of 16S RNA genes. But what you do have is you have the ability to use one target, one probe, that will attack all of them by binding to this conserved region. We call this a primer. A primer allows us to then serve as a basis on which we can then sequence forward.

So now you can sequence into the variable region. And then what comes off of our sequencing machines is that variable region. Now we can take that region. And again, we can go to a complex sample just using one probe and get all these different variable regions. And then we compare them to a set, a database.

And in that database, say this variable region is 99% identical to some *Clostridium*. Therefore, it is that clostridial species. But if it's only 90% related, we say, well, it's a related clostridial species. Same genus but maybe a different species, or same family but a different phyla. Who knows?

And if it's not really related, we might just say, oh, it's a bacteria. There you go. That is the basis for the majority of the studies you probably are looking at in the literature.

So what are the problems? What has led to a lot of the misconceptions I think we have about what the microbiome is, and does, and its complexity? The first problem is something that we knew from day one. There's nine different variable regions in this gene.

So what does that mean? Let's say you used one primer that targets one variable region. Each color is a different bacteria. This is going to be your bacterial distribution in the mouth, in your head, or on your skin.

Now, take that exact same DNA, and you target a different variable region. Now suddenly, you have a whole different type of bacteria there. So this is something we understood, and we were trying to build on it. But the science escaped us before we could do a lot of this quality control. So a lot of articles you see before 2012 may have just targeted one region or another. So what you can infer from that, again, it's still real data, but you just have to understand how to look at the literature.

Now we know that we can sequence the whole gene. We can sequence multiple regions. There's ways that we've evolved to get around this, such that 16S is still a very good technique, but we've, again, understood its limitations.

The other problem had to do with sequencing error. So here you see platforms that we use to do DNA sequencing. There's all these companies. If any of you had invested in Illumina 10 years ago, you probably aren't sitting here in this audience now.

But let's say you start with 100 base pairs of DNA. What we can do with our sequencing is each base pair will be predictive potentially 100 times. Each one of those predictions we call a read. But if you have a 1% error rate, that means that each base pair is going to be read wrong at least once among all those reads.

So for instance, a lot of the technology we used up until 2012, which was of the 454 technology, had a 1% error rate. Since then, with Illumina we've gone much less. But if you sequence a million base pairs, even a .001% error rate accumulates.

And so what does that mean? So here's our original gene. This is what we want to copy. So what we do is then we run it through a machine. And these are the sequencing reads we get out. The machine spits out all these little small reads that what we want to then do is piece together and get the original gene, which you can see there's two errors here.

So in one case, this error occurs in a region where there was only one read. That gene was only read one time. And so what happens, because there's only one, it gets transmitted to the eventual gene.

But in a second region, you can see that that error, there's actually two other reads above it. Two of them will say a, one will say t. And then when we stitch everything back together, we can correct for it. And we make it an a.

So we now know that more reads gives us more precision, more chances to correct these errors. And we call that sequencing depth. So this is another important thing, as you look at these studies, to understand, well, how often were they read? And sometimes you'll see this.

It'll say 2x coverage, 10x coverage, 100x coverage. They're telling you generally their prediction for how many reads are current for each base pair of DNA.

Now, this created an issue. So this was published in 2012, and they're already telling you at that time, look, we're smart. We know that there's a lot of variable regions. And what we're doing is we're comparing and contrasting primers that are targeting these variable regions. And we're showing you that, in a lot of samples, the measure of diversity, which I'll talk about, is fairly similar depending on the primers you're using.

Now, they're not telling you anything about the distribution. They're just saying how many different types of bacteria that are there. But in some areas, that distribution appears to be very, very different, like in stool.

Now, then they show out this number. And you look at the y-axis, and you're like, what the heck is an OTU? Well, you actually know what an OTU is. So we've talked about it. This is an Operational Taxonomy Unit.

What they've done is they've just arbitrarily now grouped together things that sequence similarity. So they're saying that all the sequences that are 95% identical to each other-- again, gene similarity, species similarity-- that's a species. So anything that's not 95%, that becomes a new species.

And they're telling you that there's 40,000 species that they detect across humans stool samples, which is absolutely extraordinary. But that was a problem. The problem was is it fulfilled our fantasy. Well, not fantasy. It fulfilled our dream or a thought. Our hypothesis heading into it was that with culture dependent techniques, we couldn't grow much.

And then when we used 16S, we said, aha, we were right. There's thousands of bacteria inside of you. We knew we couldn't culture anything.

And you probably saw this. We can only culture 1% of what's inside of us. We can't do this. It's this untapped place. It's like Mars.

But what a lot of this was sequencing error. When this error gets transmitted, let's say that that's 5% of it, you just discovered a new species. All that it was, though, was that you didn't have a lot of reads there. And so these errors had a chance to get through.

And this is something that Jay Faith showed us. So he's another scientist here. He does a lot of this type of microbiome ecology work. And so what you see here is, using these traditional sequencing methods, bacteria that are at a low abundance, you have incredibly inaccurate ability to sequence them.

So that means that bacteria that are inside your gut but aren't 20% of the population there are just a fraction of that population. Chances are, there are a lot of errors. You overestimate the number of these low abundant bacteria. And then he developed this new method that lets you get in there and get at these lower bacteria and do better.

But it just highlights, again, the fact that, when we really started to understand the limitations of our technology, we found that there were probably closer to about 140 or maybe 150 species inside of a person's stool sample. And then when you look back at that 100, you were like, oh, these old school guys, they really actually weren't that far off.

It makes sense that the human microbiome is not that hard to culture from. This is a great environment for bacteria. You eat a bunch of stuff. There's all sorts of nutrients. It's not that bad. Now, we still can't culture everything, and we're working on it, but it's not what we thought it was.

So then finally, in 2012, this is the \$200 million figure. And that literally is a \$200 million figure. It's mind-blowing. And so this is the Human Microbiome Project. And so what you're seeing here is that here's different sites of the body-- your nose, your mouth, the plaque, the tongue, stool.

Each vertical line represents a patient, and each color represents a different phyla of bacteria. And so when you're thinking about a phyla again, this is a bird's eye view. It's not a species. It's not e. coli. It's five or six steps removed from that.

And even at that level, you can look very quickly at this, and you can understand very fast that no two people are the same, that it took all this time for us to realize-- and it is an incredibly important revelation-- that the microbiome diversity, at least measured like this, is incredibly different.

So now we've started to change our attention. Maybe it's the bacteria that are common to everybody that's important. Maybe it's the bacteria, because if you actually look in your gut, there's probably only five or six bacteria that represent the majority of what's there. Maybe it's those ones, the ones that can reach a huge amount in people, that are interesting.

The other thing is, speaking to an audience of physicians, you know this. What about strains? If your patient comes into you and says, I have e. coli, you're going to yell at them, like great. You and everyone else. But if they say to you, I have e. coli 0157, you're going to send them to my office because chances are they're going to be having bloody diarrhea and be a gigantic mess. We know that this is also important.

But if you want to talk about complexity, so let's look at the possibility that everyone in this audience shares strains. So if you look at a strain level, this basically means 100% identity across the whole genome, the exact same bacteria, no two people in this audience share one strain. I could say that with 100% confidence. Even if you're household contacts, chances are you don't share any strains.

The other thing that was really interesting here is that if you look at yourself over time, you actually still have the same strains. Your microbiome can shift a lot, but the players are actually very, very stable. And that was also revelatory. And this is also work that came from Jay Faith from that same paper.

So just to summarize a little bit of what we've talked about, in terms of thinking about question 1, diversity is the norm. We have to beware of 16S analyses. These are things you have to look at very critically, especially things that came out before 2012. We now know that we need to do a lot more sampling, and we're starting to ask new questions because the original question we had wasn't answered right. There's too much diversity.

We also are now getting into the new thing. Why even bother to do the old when you can develop something different? So now we're looking at fungi, we're looking at viruses, we're looking at phages.

Remember, the human microbiome isn't just bacteria. There's a lot of organisms there. And we have to develop all the same techniques to look at these that we had to do for bacteria to begin with. So it's a big problem.

So coming back to this, I am not putting up this slide so that people look at it and say, this is nonsense. It couldn't be further from the truth. This isn't just a good start; it's an incredible start. It's amazing the work that we've done. But what I think that this thing just tells us is there's nothing else other than the microbiome as dynamic. It has the ability to change.

There's also parts of the microbiome that are stable. It took a long time to figure that out. But this dynamic changes also appear to have something to do with disease, which is important to us.

But I think what we all still want to understand, and this is going to be a recurrent theme through my talk, is what about cause versus effect? These dynamic changes, are they responding to something, meaning bacteria are telling us something about human physiology, they're the greatest diagnostic tool we've ever had? Or are they actually changing human physiology, something that we can tap into to treat disease?

The other thing that I always keep saying is it doesn't make sense. We can't be that different. We're all humans. If you actually look at the bacteria that are inside of us, it's a really narrow range compared to what you might find in the soil, which is infinitely variable. So there's something that's still linking us together.

Now, the concept of ecology as being just about DNA doesn't make sense, that probably microbe identity as we've all known, *e. coli* versus *e. coli* 0157, has to do with function, that there's certain shared functions we still just don't understand that are going to bring this all together that make us all human.

So now coming from number one, since we jumped straight from one to five, and this is something that we're going to continue to see in the literature. Just like we talked about with number one, I want to provide a little primer for how we think about number five. So what makes everyone unique? How did they do these cohort studies?



And the way that we did these cohort studies for the microbiome is the same way that we do cohort studies in people for clinical trials. You identify a patient population. You collect your samples. You process these samples, and then you do your analysis. It is the exact same protocol.

Ultimately, though, for human studies, we spend forever sitting in a room obsessing over confounding variables. You use drug A, you use drug B. You want to make sure you're measuring the effect of drug A on those people.

You don't want to measure the effect of the tacos they ate last night. You don't want to measure the effect of the fact that they go to see more doctors more frequently. You want to control for confounders.

So what are the confounders in the human microbiome? Well, if we had proceeded orderly through this work, we would have realized that all the confounders were the stages we decided to skip over before. The confounders are how we respond to these bacteria. The confounders are the external factors that shape the microbiome, things such as diet.

And we have learned a lot since then about these confounders. Things like diet are probably the most important things that are going to shape the microbiome. Your age will shape the microbiome, and I'll show you a little bit about that. We know the medications you're on as well as your disease. That's what all those studies told us, that diseases are shaping us.

So I'm going to get into what this figure actually represents. But just to show you very quickly, if you think about diet as a confounder, if you look in the US with our diet versus you look at other areas of the world that have not been touched by the grace of the Big Mac, you can actually very quickly separate people who have a very stable plant-based diet versus those of us who have a different type of diet.

Now, what's interesting is that you would think that having a more varied diet with all these different things would make our microbiome more complex, and it's actually the opposite. For some reason, that actually leads to less diversity, but that's something we don't understand. Why is that? How are these bacteria responding? And we're trying to work on that.

Genetics is an interesting one. So if genetics was everything, when you looked at monozygotic versus dizygotic twins, you would expect the monozygotic twins would be much similar to how your own microbiome would relate to itself. And in reality, there's no significant difference. There's no difference between monozygotic and dizygotic twins. And they certainly don't look like what you look like relative to yourself in terms of, again, some measure of difference which we'll talk about.

This thing would suggest, though, that there is a difference between these twins and someone who's unrelated, that genetics has some component. More research has come out actually over the last year that suggests that that's probably also not true, that this probably has to do more with relatedness. But genetics is critically important to microbiome function. So it may not shape the bacteria that are there, but it shapes how you respond to these bacteria.

So now for sample collection, this would be another thing. How often should we collect samples? And the Human Microbiome Project, again, its goal was to teach us this stuff. And it provided us this figure.

It said, well, let's look at a technical replicate. Take the same DNA sample. Just run it twice. You'd expect it to be very similar to itself, which is the zero mark. And it is.

But to give you an idea of how much your microbiome can change, if you look at yourself between two different time points, you have a totally different distribution of the bacteria that are there. Now, they're probably the same bacteria, but the distribution can be incredibly variable. It's so variable such that it's almost the same as you comparing yourself to someone else.

So again, this gets back at this idea that we now understand the microbiome is incredibly dynamic. We now know that if you really want to do these types of ecology studies, you're going to have to take multiple samples over time, that just finding some new disease, running off some sequencing at a cross-section, that's a great start. But that isn't where the field really is right now because of a lot of this work that's been done.

So now we get to sample processing. And this was the heart and soul of what a lot of the Human Microbiome Project had done. This was where a lot of the investment came in.

How do we actually extract the DNA? How do we actually process it? How do we actually analyze it? And this was great because DNA is so stable.

It's one type of molecule. So you can develop one extraction method, and you can get all that DNA. You can develop one type of sequencing, and you can understand its limitations.

Be aware of other -omics. Just like bacteria are no longer trendy, and you want to study phages, or viruses, or fungi in the human microbiome, people want to do not metagenomics anymore. They want to do proteomics. They want to do metabolomics.

This is not one molecule. These are incredibly different molecules. These are incredibly different molecules that have all sorts of different extraction protocols and sequencing protocols. These are fields that are beyond in their infancy. So definitely be aware of those types of studies as they start to come out.

And then the final analysis thing, and I'm going to use this article as a way to walk through it. I think when we see some of these figures, we get scared. We say, well, what are they? What do they mean? And it's actually not too hard to understand it.

Basically, the same way we approach the human microbiome in terms of statistics is the same way that we do it in clinical research. You ask somebody 500 questions, and then you represent that data with a mean, a median, a mode, some sort of statistic. You convert all that information to one data point.

Then you can compare. And you can get a P value or a published value, what gets you actually in a journal. But now you can use your t-test, or ANOVA, whatever you want to do. And you can compare and contrast populations.

We do the same thing in the microbiome world. So what you see here is you, again, from that article I showed you, these different countries. We looked at adults, they looked at kids. And then what they did was they created this statistic that they called the UniFrac distance.

The majority of these statistics that you're going to see in microbiome studies, what they are is they're a measure of diversity or taxonomy. They are not telling you what bacteria are there. In fact, I can have 15 types of proteobacteria. Someone else could have 15 types of bacteroides.

Totally different worlds, humans and corn. But yet our diversity is identical. We both have 15 different types of bacteria there. So our statistics will be the same.

Now this is again an important thing to understand. But you have to understand what the data is telling you, just like you understand what a mean is and the median is. So this is telling you about the diversity.

And interestingly, what they found was we have a lot of diversity when we're little. But then as time goes on, we seem to stabilize, that our microbiome starts to find a happy place and stays there. And that diversity changes.

Once you have this statistic, you can use all the same statistical tests that we normally do to compare and contrast populations, either get one asterisks or three asterisks, move on to your publication.

The principal component analysis is another great example of stealing. These things were developed by psychologists. What defines a human being?

Well, you've got to ask a lot of questions to understand who somebody is. What do you like? What's your cat's name? All of this.

And then they had to develop a technique to understand, what's the common thread to that? All these questions we ask, what are the components that are driving the variability of that system? We do the same thing in the microbiome. You have 100,000 million different species. Well, what are the actual species? What are the OTUs? What are the things that are driving the variability there?

This is really a visualization tool to say what are the differences in the populations. And they do tell you that the first principle component, some collection of these bacteria account for 25% of the variation in the system. Another collection account for 6%.

Well, if I chose other principal components, all the sudden, this data may come together again, and may disappear, and may change. So it just depends on what you look at it. It was a great place to begin, but I think now what we're asking is, what does this mean? What are these components? Why are these accounting for the variation in our system? And again, this is getting back at function.

So to summarize this section on cohorts, one thing I think that we can say for sure now is there are clear population differences. but this is important. I think the other thing we have to understand is that these studies are telling us something about human biology. There's meaning here.

We have the statistics. We have this stuff. We understand the limitations of our data. But what we still don't understand is, again, the cause versus effect, that we're still coming back to the idea that what they're telling us about human biology is what we want to understand.

But what gives us hope are our little kind of eerie friends, our mice, because in mice-- and again, I'm not going to walk you through all of these studies-- I can change a mouse's bacterial ecology, and I can make them fat. I can make them thin.

I could give a mouse diabetes. I can give a mouse cancer. I can give a mouse colitis. It tells us that bacteria are sufficient to cause disease. So that's giving us hope. It also tells us that as mice's physiology changes, I can watch that in the bacteria.

But really what our goal is is now to take all this data and to start to move it into humans, that we know enough to say that there's something here. We're working through these phases, but we want to make this translational step. So thinking about five, we now understand that we really have to start to define these other confounding variables. We still know very little about how bacterial functions affect human physiology.

We understand more and more that there's certain things that shake equilibrium of which probably environment and diet are most important, as well as age. And we still have profound questions about what makes people unique. We don't know what normal means. We have no clue.

We need more patients. We're going to need more samples. We're going to need more time points to start to understand this. We have to think differently about our outcomes. It can't just be DNA and ecology as a surrogate. There has to be another outcome that's going to start to provide a little bit more clarity.

So what I would say here is, that for most people, you really need to stop at this point. We have spent a ton of resources sequencing. We have sequenced, and sequenced, and sequenced, and sequenced to the point where everyone's catching up to this.

The hope that you're just going to find two new diseases, and look at stool, and sequence some DNA and get a publication, those days are over. You'll find someone to publish you, but we now are asking different questions. We're returning to the basics. We want to go back through that order. We want to understand more about these cause and effects, that we understand that the real meaning of ecology is going to have something to do with function.

And so this now brings us to where I think we're at in the field, which is really question number two. What are the microbes doing? We have a sense of their identity. We have enough sequencing data that we can live off of for generations, but we want to understand their function because ultimately, the function is going to inform our previous question as well as all of our future questions.

So this is going to move on to the second phase of the talk, which is we're going to talk about microbial function. And because I do believe that this has really been, for me at least, a real pathway to drug discovery, how we can think about taking these microbiome observations and start to translate them into the clinic.

And then the final thing we'll touch on for the last 4 to 5 minutes of the talk is this idea of this new field of live biotherapeutic products. What does it mean? What are they? And how can we think about them as physicians?

So getting at number one and number two, the way that I looked at things when I was beginning is that we have these incredibly interesting correlative studies in humans. And we can do god knows what to mice, that we can do all these really interesting things.

And while that alone is actually something that we can begin to translate, to me there is this missing question. There was something that was really separating us, I think, from really developing these types of therapies in a widespread way. And that, again, was this basic interaction, what I say is this fundamental question. How does a bacteria interact with a cell? That there's two ways that it can interact. It can affect a change on that cell or it can respond to that cell in a way, that this is the most basic fundamental part of microbiome interactions.

Now, we know this isn't simple, that these interactions are dictated by a whole variety of products, that what we make, what they make, what we eat. But if you think about it differently, if we understand how bacteria affect host cells, there's your therapies. That's new treatments. That's how we develop things that we can translate.

If you understand it in the opposite way, there's your diagnostics. That's how we can start to use bacteria to think about monitoring human disease. And so for me, I was much more interested in this pathway, partly because, again, I became interested in developing novel therapies from the human microbiome.

So in addition to this \$200 million figure, they published a second figure right below it. So we now figured out the microbiome is different. And then they put this up. Here's all the metabolic pathways.

So now, this is probably one of the more misinterpreted figures in the Human Microbiome Project, because they did not put this up there so that people would now say that function is the same. That wasn't its purpose, because if you looked at the functions they were looking at, they said, yes, all the different bacteria in these sites, they metabolize sugar. They make DNA.

What they told you was that, yes, of course, they're bacteria. So therefore, that is why they share all these same fundamental pathways. But the take home point from this, from my standpoint, is that the function of these bacteria is simple. There's not this huge portion of their genome that's dedicated to all these weird other functions, which is actually what we see in other environments, like the soil or a volcanic vent. These bacteria have got to fight. And so they've evolved all these crazy systems that aren't just related to these basic things.

But the other thing is that the fact that they can't tell you at a granular resolution what these functions are also still tells you the function is unknown. And I know that they intend this because of the fact that they also provide this type of data. So this is looking at samples from your tongue. And they're telling you very quickly, for this one bacteria, strep mitis, everyone's got it, that everybody's sample is in the gray. Every single person has strep mitis on their tongue.

But then what they show you very quickly is, if you look at two genes in strep mitis that regulate function, two things that they know have a phenotype associated with them, some people have gene one and gene two. Some people only have one gene and not the other one. That function is actually highly unique, and that if you were to use ecology at this level, at a species level, even to predict function, it probably wouldn't be right, that this is something that we have to take it a step further.

The problem is that function is poorly predicted. So remember, we have the sequencing. Now we have our original gene.

Now what you have to do here is you have to go to a database. So you take that gene. You have some unknown gene. You look in a database. And it tells you that this gene is related to this one. And therefore it's bacteria x.

So the hope is we could do the same thing with function, that for an enzyme, well, this enzyme is related to this enzyme. And it makes molecule Y. And we've discovered the next penicillin. Unfortunately, that's not the case, that our database size for function is incredibly small. It's very poorly annotated.

A lot of the times, the only thing we can tell you is incredibly general functions like I showed you. This has something to do with metabolizing a sugar. Well, what does that even mean? Where is that sugar going?

It makes a lipid. You know how many lipids are out there, and every lipid has a different function? That's really what we can understand about bacterial function at this point.

So when I saw this problem, it led us to go back to thinking about how should we approach this. We're going to not just point out the issues but develop the solutions. You have to think about all the different ways that bacteria in human cells can interact. And not everybody can study everything. You can't look at all the surface receptors, and the products, and this and that.

So again, because we like to cheat, we said that I knew hidden in here are going to be functions that I think have a real therapeutic basis. And so again, you go back to the people that came before you. And a lot of these people, for a long time in environmental sciences, studied what are called small molecules.

These are the things that bacteria secrete that don't have to do with their primary metabolism. It's not part of how they live. It's part of actually what they're doing to interact with the environment.

And why did scientists love this for 100 years or whatever? Well, people who were studying this in the environment knew that this is actually how bacteria set up virulence, how they set up symbiotic relationships with other organisms. And importantly, it was a source of therapeutic discovery. This is how we found tetracycline. This is how we found tacrolimus.

So we decided to take that same approach in humans, to look at this particular type of small molecules. One way you can do it is the traditional method, which is called find and grind. You find a bacteria. You culture it in these massive fermentation flasks.

You extract out all the small molecules using magic. You partition them. And then you assay them.

So if you're interested in finding a new cancer drug, you throw these small molecule mixtures on top of a cancer cell. If it kills it, you then go back through, and through iterative processes of fractionation, you can isolate the molecule. And now you have your leading off point for a new potential drug.

For the human microbiome, we actually thought another method was much more interesting. And this is this method we call functional metagenomics, and this is a lot of what I do in my lab. And so in this case, again, we're using metagenomes, we're using DNA, and we're not isolating bacteria.

We isolate the DNA from a sample, in this case humans. And you create what's called a metagenomic clone. So in a metagenomic clone, what you take is you take a bacteria that's really easy to deal with in the lab. Doesn't kill people, it grows really well and makes all sorts of fun stuff. And then you just add to it small pieces of that DNA.

So you don't add the whole metagenome. You just take little fragments that has maybe 10 genes on it. But we have no clue what those genes do. But what we care about is their function. So then you assay them. You identify a new phenotype that those genes give to that bacteria.

So in this case you, could plate other bacteria on top of it. And if you want to discover a new antibiotic, you look at one of your clones to see if it's now killing those bacteria. So then you can now take that clone, and then this is where the advantage is for the human microbiome.

You can now isolate, just like you normally would, the products encoded by that gene, but you also have the genes. You can begin to form that basic science relationship that a gene makes something that does something. And this is that information that was missing. So we can now start to piece that together. And we can do it in actually a very high throughput way.

So coming back to this, rather than worrying about these databases that don't really exist, we just take these genes. We create a whole lot of these bacterial metagenomic clones. And then we pick different human cell functions that we care about, that can bacteria modulate metabolic functions? Can they modulate immune functions, inflammatory functions?

And then we can set up using high throughput assays to say, OK, well look, this gene that I just put in there is now able to turn on inflammatory pathways in human cells, whereas if you just take that same exact bacteria, the exact bacteria minus that gene, it can't do it. So what we're doing is we're bypassing this database. We're defining functions at the ground floor.

Now let's say we had used this database. So for one of the genes I had found that we were interested in, it told us it was a hemolysin. And my P.I. At that time basically almost had a conniption and was like, we've wasted all of our time on a hemolytic thing. No one's going to care about that from Human Microbiome.

But it didn't make sense. We actually knew that its function was it activates NF-kappa B. None of my cells were dying. They weren't lysing. And we have the ability here, because we have a gene and we have a function to isolate the product.

And so we went after it, and we found this product was actually a lipid. When we looked back at the experiments that defined this gene as a hemolysin, they had basically just turned the gene on. They had no idea what it did, but they turned it onto massive amounts.

And if you take a ton of lipid and you put that on a human cell, it doesn't make a difference what that lipid does. It will burst the cell. So they called it a hemolysin. But that's not what it was doing.

And once we had this structure and we had this function, we made an interesting observation, that the structure of this metabolite sure looked an awful lot like human metabolites, that humans make bioactive lipids to signal. This is how our cells work. And it makes sense.

The Human Microbiome Project told us that these bacteria should not be making crazy things. They should be doing simple things and that we have this evolutionary history. And so if humans and bacteria want to interact, they need to share language. And this shared language are the same signaling molecules that we use.

And what we saw was that these lipids, such as those for the endocannabinoid system, which is also very trendy thanks to the liberalization of our laws here in New York, they act through G-protein coupled receptors. So this provided a basis where we could say, look, we now have a gene. We have a function, and we have a product. We think that this may also be active through GPCRs.

And GPCRs were really interesting to us because, again, we were thinking about this therapeutically. What this figure shows you is that these GPCRs from humans, they regulate each other. They regulate a huge portion of our biology. And guess what. Chemists figured this out a long time ago because they've been targeting GPCRs therapeutically also for a long time. And they're one of the most common therapeutic targets we have.

So now, suddenly we have a system whereby we can start to think about how bacteria are modulating a very therapeutically relevant target. And again, it makes sense. What are the endogenous ligands for GPCRs? They're simple neurotransmitters. They're tiny little molecules. These are not complex things. This is how human signaling works.

And so we were able to screen our molecule against a GPCR. We were able to identify a very specific interaction between our lipid and this receptor. And interestingly, this is a receptor linked to autoimmunity and as well as atherosclerosis. These are conditions that those correlative studies, that our mouse studies suggest are linked to the microbiome.

I'm not telling you that my molecule's doing this, but I actually have a mechanistic hypothesis you can test. This is not where we were a long time ago. But by just starting with the basics, you can start to decipher some of the potential language of the human microbiome.

The other exciting thing is that when you have a gene, a product, and a function, we can now use all this data that's been created to our advantage. I now know that the enzyme that we're dealing with is N-acyltransferase. It does a very simple thing.

I have \$200 million worth of sequencing data at my disposal. We can go back into these bacteria. We can identify all the related genes that also are acyltransferases that likely make related products. We can then synthesize the genes. We can put them into our host bacteria, and we can isolate all the related molecules that are linked to these genes.

And what we see is, again, you have a family of these types of N-acyl lipids. You have a simple head group that varies between all of these. And you have a simple tail. And my hypothesis would be is that, just like humans signaling molecules that share these types of variations, each one of these would have its own GPCR that it's acting through.

And that's exactly what we found. Of course the colors don't show, but on the left side, you have your bacterial ligands. On the right side, you have your human ligands. And they sure as heck look an awful lot like each other.

In the middle, you have the different GPCRs that are very specific to these ligands. These are regulating metabolic endpoints, immunity endpoints, tissue repair, all sorts of interesting functions of the human microbiome that are relevant to disease.

So what do we do next? We've been taking these to ask ourselves, what's happening in vivo? So one example we have, and I can show you a bunch of these, but just for the sake of this talk, this is one of the molecules we're interested in. It's called N-acylserena It interacts with this GPCR called 119.

119 is present on enterendocrine cells in the intestine. We liked this because our bacteria and the GPCR are both found in the duodenum next to each other where enterendocrine cells are very dense. They therapeutically targeted this receptor with small molecule agonists that are in phase two and three studies. And our interaction between our molecule and this receptor was highly specific, again, the way that biologic ligands can work sometimes.



So we now have the advantage that we already had our therapy. We already had our therapeutic system. We had a bacteria that we've engineered to have this gene. We have the exact same bacteria without that gene. This is a great way now to study it in mouse models.

So now we can colonize our mice, either with the bacteria with the gene or the bacteria without the gene. And we can put it under regulation. So we can turn the gene off and on and tell you exactly what happens in the microbiome when you make this molecule from this gene.

And if it was doing what we think it was doing, it should increase GLP-1, it should increase insulin, it should lower post prandial blood glucose. That's exactly what we found, that in our mice where we colonized them, even just after a week, our post prandial blood glucose went down. If we removed the stimulus to actually turn on the gene, we saw it go back to normal. We saw an increase in insulin. We saw an increase in GLP-1.

And what was interesting to me, which I think again highlights this evolutionary relationship, and it's something that I think we have to sit back and think about, is that if you looked at the human molecule-- got that wrong-- human molecule, you looked at the bacterial molecule, ours was actually a little bit more active. It increased GLP-1 a little bit more.

What are the native ligands for GPCRs? They started it a long time ago. This is one of the oldest class of receptors. Again, we have this human-centric focus on things that maybe our infectious disease colleagues don't share, but perhaps these GPCRs, actually the endogenous ligands for some of them are bacterial ligands, that this is part of how we interact with the environment.

So coming back to this, some of you may have looked at my curve and said, ah, that doesn't actually matter. But just to show you, this is actually this synthetic small molecule for that same receptor. This is what encouraged drug companies to go into phase 3 studies. And ours has a very similar effect in terms of delivering a bacteria with a biologic ligand.

So now we have a therapy. We have something that we can engineer a microbe to make something and start to consider how we would want to test it in humans. So that leads to now the final part of the talk, which I'll spend a few minutes on, which is, what is the current landscape for these types of products? How are we going to bring this to market? How actually do we think about the early studies for it?

So when we think about what we now call live biotherapeutic products, so this is something you are going to see in the next five years coming toward your clinic. We actually have studies here already going on with not technically a live biotherapeutic product, which I'll explain, but things that are aimed to shape the microbiome.

So in one case, you might think of a single organism that you give to sick people, and you make them healthy. In another case, maybe it's a consortium of organisms that you've got to give. Well, in these cases, a lot of how we found these things to get go were from mouse models.

And a lot of it was actually fortuitous, that we had bacteria, we had an idea. We just put them into mice, and they did something. And we can say for sure that different bacteria don't do that. But what is cool is that we're actually encapsulating a very complex biology. Sometimes it may be hard to figure out all the different things these bugs are doing. And that's actually very attractive from a therapeutic standpoint. It's also very concerning in some ways.

What we would have hoped is that we can intelligently select these. Go in and sequence people, identify the bacteria that are different in disease, and maybe give those or deplete those. This hasn't worked.

It hasn't worked in part because again of our limitations of understanding how to do these studies. What are the relevant bacteria? Everything I showed you in the beginning of it.

But you can also take another strategy, which I've shown you what my group does, which is we actually go back old school. We identify metabolites that have targets. We engineer them. And this is another type of thing that's a live biotherapeutic product. And these are genetically modified organisms, although putting out a GMO is probably going to cause people to be a little bit crazy.

So to meet the demands, to meet the thirst of biotech, the FDA had to scramble. So the Center for Biologics Evaluation and Research created a designation called live biotherapeutic products, and this is what it means. It has to contain a live organism. It's applicable to the prevention, treatment, or cure of a disease. And it's not a vaccine.

Fecal transplant is a live biotherapeutic product. Don't confuse the fact that for C. diff, you don't actually have to go through the normal regulatory pathways, to some degree. But if you want to use a consortia, you absolutely do. Every probiotic that is in market right now has no disease indication. These are not LBPs.

Like I said to you, they weren't found because someone did these mouse models and then put it out in GNC. They were found because someone just realized there was crap in yogurt, and they could sell yogurt better.

Now, what's happened is, after the fact, they've gone to these probiotics and then tried to figure out, do they actually do something? In some cases, they do. In some cases, some of the newer probiotics, there is a little bit of science. But these are not LBPs.

There are other strategies to manipulate the microbiome that don't have this designation, that there's a lot of companies behind, such as giving prebiotics to increase the growth of good ones, or targeting bad ones with antibiotics, or using things like phages which can again target and manipulate the microbiome.

So when we think about how you discover an LBP, I think it highlights what makes this so exciting, especially when you look at how we do small molecule drug discovery. So with small molecules, you start off with a receptor. You identify this receptor or this disease process by comparing humans. And then you say, from this target, we're going to identify a small molecule using screening.

We're going to then take that small molecule lead. We're going to do medicinal chemistry. We're going to work on its encapsulation.

We're going to add what are called excipients. These are things that stabilize the drug. And we're going to formulate it for delivery. Once we formulated it, now we give it to mice, test it for efficacy and toxicity, and then we give it to humans. This is what's been done for decades.

How did we discover live biotherapeutic products? This actually turned this whole process on its head because we combined every single phase together. We coupled bacteria directly to the disease, directly to the phenotype. We were already in mouse models.

In addition, we're already even working on formulation. By giving these mice these bacteria, you're sort of saying what we know about this. This allows us, again, to just say, well look, we're not totally sure about all the mechanisms, but there's a complex biology here we can recapitulate.

Or in the case of a lot of what I do, which is engineering for these ligands, if you try to tell a medicinal chemist to deliver a lipid, they would look at you like you're crazy, that these lipids are just going to get degraded. They're going to be problematic. But now you've got the world's greatest delivery system, a bacteria that's going to go right to its site of action, and make this molecule, and affect a disease outcome.

When we think about formulating an LDP, it's different. A lot of the times, we don't have mechanisms. And even when we do have mechanisms, we're not thinking about dose in the traditional sense. We're thinking about dose in the terms of just the number of live organisms that are there. Again, something that you're not going to see in your probiotics.

When you think about making them-- this is wine, but I think you guys get the picture-- you're dealing with huge fermentation areas. We've had to develop a lot of fermentation strategies to deal with some of these organisms because they don't like to grow with air or things like that. But also, we have a whole new problem. Small molecules are stable. These things change.

You need quality controls as they're fermenting to make sure the bacteria aren't evolving, they're not changing. You have to periodically sample these phenotypes. It's a lot of like what we have with certain types of biologics that have come to market. We accept contamination with small molecules. You can't accept contamination when you're dealing with bacteria.

Encapsulation strategies are a whole other paradigm. We're not just thinking about surviving the stomach and getting into the blood. We're thinking about, where do you want to deliver? There's a whole dozen companies just engineering these types of strategies.

And now colonization. This is the dream, but this is also the thing that, if you stand in front of the FDA, they will absolutely cut your head off. What's the dream? Give a bacteria to somebody one time. Colonize them. They make something, and you never have to do it again.

God knows, if you practice in a resource poor setting, this is something that has to fascinate. I can tell you Gates is incredibly interested in this concept. Is it even possible, though? A lot of what we know is that it's not so easy to colonize people, that some bacteria, like pathogens, can be infiltrative. They'll get in there, and they'll do their job. Most bacteria we're dealing with will either transiently colonize or they require something else.

So now this gets at this idea of excipients, which change, that we have a whole new strategy for stabilizing these things. Maybe you want to use prebiotics so that your bacteria can get in there better. Maybe you want to clear out space using antibiotics. Or maybe you're going to genetically modify that organism so it can colonize.

And so to finish, our efficacy endpoints are probably going to be the same. Now, how we're going to follow these patients we're going to have to think about, because it's going to be a little different because of these colonization questions. And the other thing is that with humans, this may really be the first time that you can think about your formulation. The mice, as we know, have limitations. But in terms of understanding ecology and colonization, they really inform us poorly in terms of what's going to happen in the human microbiome.

Safety endpoints, this has got a lot of people excited because from FMT studies, we see almost no side effects. But that can't be true. If you really believe that bacteria can effect a change, then it's very possible that they'll have side effects, too.

We're going to think about these safety endpoints a little differently. A small molecule can't go person to person. And while we know that it isn't easy for strains to get communicated, we're choosing bacteria with a very specific function here. And we don't know what's going to happen.

And one of the biggest FDA concerns has to do with genetic transmission, not between individuals but among the bacteria. So if you give someone a bacteria that has a genetic element in there, like an antibiotic resistance, that's no bueno, because these bacteria in the gut, they all share that gene pool. So as a result, you can start disseminating that antibiotic resistance. So genome sequencing of all LBPs is mandatory from the FDA.

So to conclude, I think the future directions that we're looking at here, in terms of live biotherapeutic products, nothing is on market. We have a lot of things that are entering into phase 2, a couple that are entering the phase 3 studies. And the data is fascinating.

The biggest challenge that this field is going to have, though, is when that first product comes to market. These are natural products. Natural products, for intellectual property standpoints, are a huge black box. You're not supposed to actually have intellectual property on something that can be from nature.

So companies are trying to find strategies around it. And this intellectual property is the key to innovation. But it's not going to be until we get to market until we find out how this is going to settle out. And it could destroy the field, or it could help it to grow.

I think we're starting to separate the reality from the hype. I think that there's no question, this is an incredibly important field, but we're starting to learn our best ways of studying it. And I think that a lot of what we want to do in the field and what we do in my lab is returning to basics.

Even just this idea that a bacteria can make a molecule that has a receptor, I think we want to marry science. That's really our goal, not to upend it.

There's no such thing as being disruptive in science. You want to learn from the past. And I think that's what we're doing in the microbiome.

So obviously, this couldn't be done without a whole slew of people. And so I want to thank everybody throughout and open up to questions.

**SPEAKER 2:** We have time for one question [INAUDIBLE]. First of all, a lighthearted point, and the second more serious. The lighthearted is, have you found the relative in your past family history they gave you the gene to make you such a good lecturer?

**LOUIS COHEN:** Not yet. Not yet. I'm still looking for him. [INAUDIBLE]

**SPEAKER 2:** Now for the more serious side. Age is a big deal here, and I'd like you to comment about it. Most of the microbiomes are young, if you believe the Hayflick phenomenon that every cell is limited to 50 reproductions.

And many of the hosts are very old, 70, 80, 90, 100 years old. What is your view of the value of turning some of this research into an age question? And maybe you've already done that.

**LOUIS COHEN:** Yeah. No, it's fascinating. What's really interesting in terms of the age question is I showed you just a little bit of work that shows you babies have this incredibly unstable-- so babies have this incredibly unstable microbiome. It changes rapidly. And then somewhere around age three or four, it settles out.

Well, guess what. There's also a back end to that. As you get older, the microbiome becomes unstable again, that it seems as though this diversity undergoes rapid shifts. Now, why is that? Is it the aging of those systems that are in place that help us to reach this regulatory set point? We don't understand that.

But it's also interesting because, even-- I mean, I remember when -- and I'm not that old-- but C. diff used to be an immune question. Babies have C. diff, and they don't have an infection. Then C. diff suddenly became a bacteria question.

But I think the same thing is with the microbiome, is that we need to understand the factors that lead to that stability. And I think that those factors are clearly in play early in life, and they're clearly in play later. There's really interesting research that suggests that there's certain bacteria or certain bacterial functions that have to occur at different stages in life for you to proceed through that.

So this is work from Jeff Gordon's group at Wash U that says like, from age zero to six months, there are certain bacteria that are important. And then once they've done their job, they get swapped out for another set. Then they do their job, swapped out for another set. And then finally, you get to the point where you have your own stable microbiome that have also shown you is unique to you.

So age, I think, is going to be a really fundamental thing to understand. And I think it's going to inform us the most about, what are the important functions of the microbiome, both young and old? And diapers are really easy to collect stool from.

**SPEAKER 2:** Thanks, Dr. Cohen.

[APPLAUSE]