Our Second Genome

by Rob Knight, PhD, and Daniel McDonald

In graduate school, Rob Knight was interested in microbes that live in extreme environments, like hydrothermal vents. But he realized that the human body is full of extreme environments from a microbial perspective, and that the communities of microbes that make their home in us are as fascinating as anything in the world outside our bodies. Now a professor of chemistry and biochemistry at the BioFrontiers Institute at the University of Colorado, Knight is a pioneer in the study of the microbiome and how it affects health and disease.

No man is an island. From a microbial perspective, each of us is more like an archipelago. The palm of the hand is an entirely different environment from the surface of a tooth or a tonsil or an intestine, but each of those parts of our body is densely populated with its own microbial inhabitants. We are each like our own Galapagos Islands—a collection of adjacent habitats with very different environments, each with its own community of organisms adapted to life in that locale.

The 100 trillion or so microbial cells that thrive in and on us outnumber our own cells by as many as 10 to 1, although collectively they weigh only a few pounds. The fact that microscopic creatures live on us has been known for over 300 years, since the invention of the microscope. Until recently, our ability to study our co-inhabitants was limited by the fact that only a few percent of them can be grown in a Petri dish. Much has been learned about those microbes and their importance in human health and disease. But with the vast majority invisible to analysis, even basic questions—such as how many different species live on us, how diverse they are genetically, and which of them are common to all humans and which are unique to an individual—were unanswerable.

In recent years new methods of identifying microbes based on their DNA have emerged, and many important questions can now be answered without growing them in culture. Innovations in molecular techniques, sequencing technology, and computational tools have made it possible to analyze vast amounts of DNA sequence data from diverse microbial communities, and to compare the communities within a single individual, across individuals, across time, and across species. In particular, the cost of DNA sequencing has decreased by a factor of about a million in the past decade. It's an incredible time to be interested in microbes.

While our genomes may be 99.9 percent the same, our gut microbiomes can be 100 percent different.

***Escherichia coli* bacteria in culture.**

False-colored scanning electron micrograph of *E. coli* bacteria grown in tissue culture (magnification: x10,000). *E. coli* is one of the best-known bacteria that live in the human intestinal tract, so we may think of it as prevalent, but it is actually a rare member of the gut of most people. We just think of it as common because it is very good at growing in a Petri dish.

NIBSC/SCIENCE SOURCE

What's Normal?

A basic step in understanding the relationship between our microbes and ourselves is to survey the microbial landscape of the healthy human body. Large-scale projects such as the NIH-funded Human Microbiome Project—which examined the microbiota in up to 18 body sites across 300 healthy individuals—have made substantial progress toward this goal. (Note: Researchers use the word *microbiota* to refer to all the microbes in a community and *microbiome* to refer to their collective genes. We can talk about the microbiota and microbiomes of the entire body or of selected regions, like the gut.)

The most complex and densely populated of our personal microbial jungles is the one in our intestinal tract. It plays a key role in many essential processes, including vitamin and amino acid biosynthesis, carbohydrate metabolism, and immune function.

Before we began analyzing the human gut microbiome using genetic tools, we thought all humans would share a similar "core" population of microbes. That expectation turned out to be wrong: there are no shared microbial species across the whole human population! Our microbial communities are in fact highly personalized and specialized to our bodies. Each of us harbors more than 1,000 microbial species. Most of those belong to just a few major taxonomic groups, or phyla; thus, at a high taxonomic level, there is some consistency across individuals. But the relative proportions of those phyla and the species present vary markedly across individuals. While our genomes may be 99.9 percent the same, our gut microbiomes can be 100 percent different.

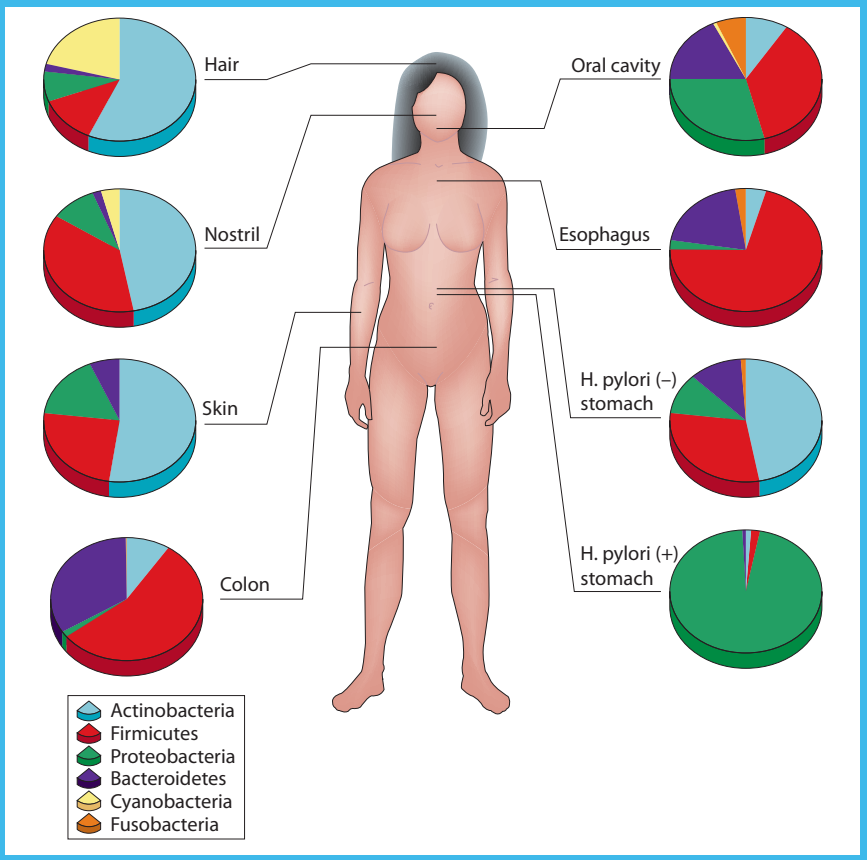
Why do people differ so much in their gut microbial profiles? This is a very active area of research. Parents and their adult children have more similar microbiota profiles than unrelated adults. However, identical and fraternal adult twins have equally similar microbiomes, suggesting that shared environment may drive familial similarities more than genetics.

We do know that diet has a significant impact on gut microbial profiles, with phylum-level differences found between people who consume protein-rich or protein-poor diets. There are also systematic differences in the microbiota of obese versus lean individuals. We can predict with 90 percent accuracy—based on gut microbe profile alone—whether an individual is obese or lean, while predictions based on all known human DNA mutations associated with obesity perform little better than chance. We will likely find many more conditions where the phenotypic differences we see between people arise from differences in their microbiota, rather than differences in their genomes.

Who's There ≠ What's Happening

Knowing the microbial community profile—the species that compose the community—does not translate directly into knowledge of the metabolic activities that community is performing. The human gut microbiome consists of a set of genes about 150 times larger than the human genome. To assess the functional metabolic activities of a given community, we have to use other kinds of assays that look at gene expression and enzymatic activity. And when we look at this level—at the *functional* profile of a microbial community as a whole—we do find an extensive "common core" of metabolic activities that a healthy gut microbial community as a whole performs, even though the microbial species performing those activities might be quite different from one gut to the next.

This finding, although not what we expected, coincides with what we know of large-scale ecosystems. For instance, rainforests around the world are composed of wildly different species—there is no common core of rainforest plants and animals—yet each rainforest community as a whole successfully performs a whole suite of ecological functions, carried out by its unique population. The common core is at the level of functions, not members.



Compositional differences in the microbiome by anatomical site.

The relative proportions of the major microbial phyla in the microbiome of a healthy individual at seven anatomical sites. Certain conditions, such as the presence (+) or absence (–) of *Helicobacter pylori* in the stomach, can lead to marked perturbations in community composition.

ADAPTED BY PERMISSION FROM MACMILLAN PUBLISHERS LTD: *NATURE REVIEWS GENETICS*. "THE HUMAN MICROBIOME: AT THE INTERFACE OF HEALTH AND DISEASE" BY ILSEUNG CHO AND MARTIN J. BLASER, 2012.

AMANDA BIRMINGHAM

Rob Knight collects a bacterial sample from his dog.

Personalized Metabolism

The implications of this diversity in the microbial makeup of healthy individuals are vast. Many of the metabolic activities carried out by particular microbes are outside the "common core" and not essential to the survival of the community, so they can afford to be different from gut to gut. Differences in gut microbe profiles are already helping explain some puzzling variations between individuals in metabolizing foods and drugs.

For example, the health benefits of diets rich in soy have been found to be due to a metabolite produced from soy by bacterial rather than human digestive enzymes. But only 25–30 percent of adults from Western countries produce this metabolite, compared with 50–60 percent of adults from Japan, Korea, or China. The cancer-protective effect of a soy-rich diet that has been described in Asian populations may therefore not be equally true for Western populations. Similarly, gut microbes differ in the way they metabolize the drug acetaminophen (the active ingredient in Tylenol), altering its liver toxicity in different individuals.

This knowledge opens up incredible prospects for personalized medicine. In the future, before prescribing a drug or a supplement, a doctor may be able to prescreen a patient, analyzing not only their human genome but also their gut microbiome, to choose a drug or therapeutic intervention that maximizes the effectiveness and minimizes toxicity and risk to that person.

An Evolving Landscape

How do we acquire our microbiome? And how does it change as we develop? This is another very active area of research with far-reaching implications.

We all begin life germ-free. Contact with skin during birth provides the microbes that dominate our communities as newborns. Over the following months, the composition of our microbial communities undergoes a series of changes—caused in part by the dietary change from milk to solid foods—until a relatively stable, adult-like community is established by age three. The maturation of the gut microbiota is an example of ecological succession.

In the first months and years of life, the GI tract's human cells and microbial community develop together in a masterpiece of cross-species communication. The infant immune system is making lifelong decisions about what is "self" and "other"—the former being immune from attack and the latter being a target. That's a complicated enough process when considering just the human molecules of the body. But in the gut, the immune system must also learn to distinguish "other but helpful" and "other but harmless" from "other and dangerous." Gut microbes communicate directly with gut epithelial and immune cells by secreting signaling molecules. In fact, the maturation of some immune cells depends on receiving microbial signals. As a child's diet changes, the nutrients available to the microbial community change, driving changes in the microbial profile, resulting in changes in the signals the microbes and human cells exchange and in the efficiency of processing different foods. It's an incredible process of interlocking co-development.

That's when all goes well. When it does not, the consequences can be catastrophic. Malnutrition underlies more than one-third of all deaths worldwide of children under age 5. Heartbreakingly, once a child is malnourished, putting them on a diet with enough calories is typically not sufficient to restore them to a healthy state. If the GI tract fails to develop properly, it may not be able to use the nutrients moving through it, or may only be able to use specific types of nutrients.

We have studied this as part of a global initiative by the Bill & Melinda Gates Foundation to understand the underlying biological causes of malnutrition. Together with colleagues from Washington University, we examined children in Malawi, where an acute form of malnutrition

Conducting a Census with DNA

We use DNA sequencing and analysis to determine the relative abundance of different taxonomic groups in a microbial community. Just what does that mean? How do we define microbial species on the basis of DNA sequence?

To conduct a microbial census, we don't sequence the full genomes of everything in a sample. That would be a monumental sequencing and computational task. Instead, we analyze just one representative gene that is found in all bacteria and archaea, the 16S rRNA gene. It was chosen because it is thought to be a good "molecular clock," in which observed changes in DNA sequence can be interpreted as evolutionary distance. It is also distinct from its eukaryotic equivalent, allowing us to cleanly separate bacterial DNA from all the animal, fungal, and plant DNA in our samples. By targeting this gene specifically, human DNA is not sequenced, reducing privacy concerns. We batch sequence all the DNA in a sample at once, giving us an output file with hundreds of millions of DNA sequences.

Once we have all that sequence data, we have to organize it. Most sequences found in humans can be attributed to known bacterial and archaea phyla, but the species may be unknown or not well characterized. However, we can still group the sequences into clusters that approximately represent species. One of the contributions by my lab to the field was to develop a bioinformatics algorithm called UniFrac that relates microbial communities to one another based on the evolutionary history they share. This algorithm lets us compare many communities at once—to see, for example, whether individuals with the same disease tend to resemble each other microbially, or to observe how microbial communities change during different biological processes including development and disease.

called kwashiorkor is common, and found that the gut microbiota of kwashiorkor-affected children had failed to develop in an age-appropriate way. These children could not benefit from normal food because their microbiomes did not have the functional ability to metabolize it. By understanding the role of gut microbiota in malnutrition, we hope to devise targeted, microbially appropriate treatments to restore the health of their gut microbes and to correct the effects of early malnutrition.

So Many Questions

Our new tools for analyzing microbial communities open up so many fascinating and medically important questions. How does our genome interact with our microbiome? What are the connections between the collection of microbes we harbor and our disease susceptibility and prognosis? How do our microbiomes change over our lifetime? How resilient are they to assault—by pathogens, or by things we throw at them, like antibiotics, pollutants, and other chemicals we have added to our environment? How can we make our microbiomes healthier, and more resilient?

As we learn more about our relationship with our microbial environment, we are on the cusp of a paradigm shift in our understanding of health and disease. It appears increasingly likely that our "second genome," as our microbiome is sometimes called, exerts an influence on our health as great as our human genome. And with knowledge comes the power to improve the health of the lives in our care. i

A native of New Zealand, **Rob Knight** grew up amidst some of the planet's most beautiful ecosystems. After college at the University of Otago, he studied ecology and evolutionary biology at Princeton. When he's not in the lab analyzing microbiomes, he's thinking about microbiomes he'd like to analyze.

A graduate student in the Interdisciplinary Quantitative Biology Program (IQ Biology) in the Knight lab, **Daniel McDonald** is a leader of the American Gut Project. Daniel's primary discipline is computer science, and his focus is on finding novel ways to use computation as a tool within the life sciences.

What to Read Next

Some of My Best Friends Are Germs
A fascinating, must-read article by Michael Pollan about his visit to Rob Knight's lab, what he learned about his own gut microbes, and our changing understanding of health.
http://michaelpollan.com/articles

The Sequencing Machine, a 2012 *Nature* profile of Rob Knight
"Faeces, lizards, keyboards, faces—Rob Knight likes to sequence the microbes on anything and everything. Next, he plans to sequence Earth."
www.nature.com/news/microbes-en-masse-the-sequencing-machine-1.10985

Find more by and about Rob Knight at:
https://knightlab.colorado.edu/wordpress

www.hhmi.org/scientists/rob-knight

Online Bioinformatics Tools

Interested in exploring microbiome data yourself? A variety of free, open bioinformatics resources have been created to help people get started.

Our lab helped develop **PyCogent** (pycogent.org), a Python-based software toolkit for genomic biology. It provides easy-to-use methods for obtaining sequences from public databases like **Genbank** (www.ncbi.nlm.nih.gov/genbank), exploring the molecular evolution of sequences, and many other common functions used in investigating molecular sequences.

Rosalind.info is a fun site for learning bioinformatics through problem solving. Another Python-based site, it will take you through many of the classic algorithms in the field.

We also helped develop **Quantitative Insights Into Microbial Ecology (QIIME)**, a software toolkit that facilitates the processing of the raw sequence data and provides useful statistical and visualization tools to help make sense of the data. QIIME allows you to compare different samples. Feel free to download raw data (from www.microbio.me/qiime) and play with it directly. A virtual machine that contains QIIME, PyCogent, and many other wonderful tools is available free for download at www.qiime.org.

Wonder what's in your gut? The American Gut, a crowd-funded project, allows members of the public to have their own microbiome analyzed.

http://humanfoodproject.com/americangut

HUMAN FOOD PROJECT