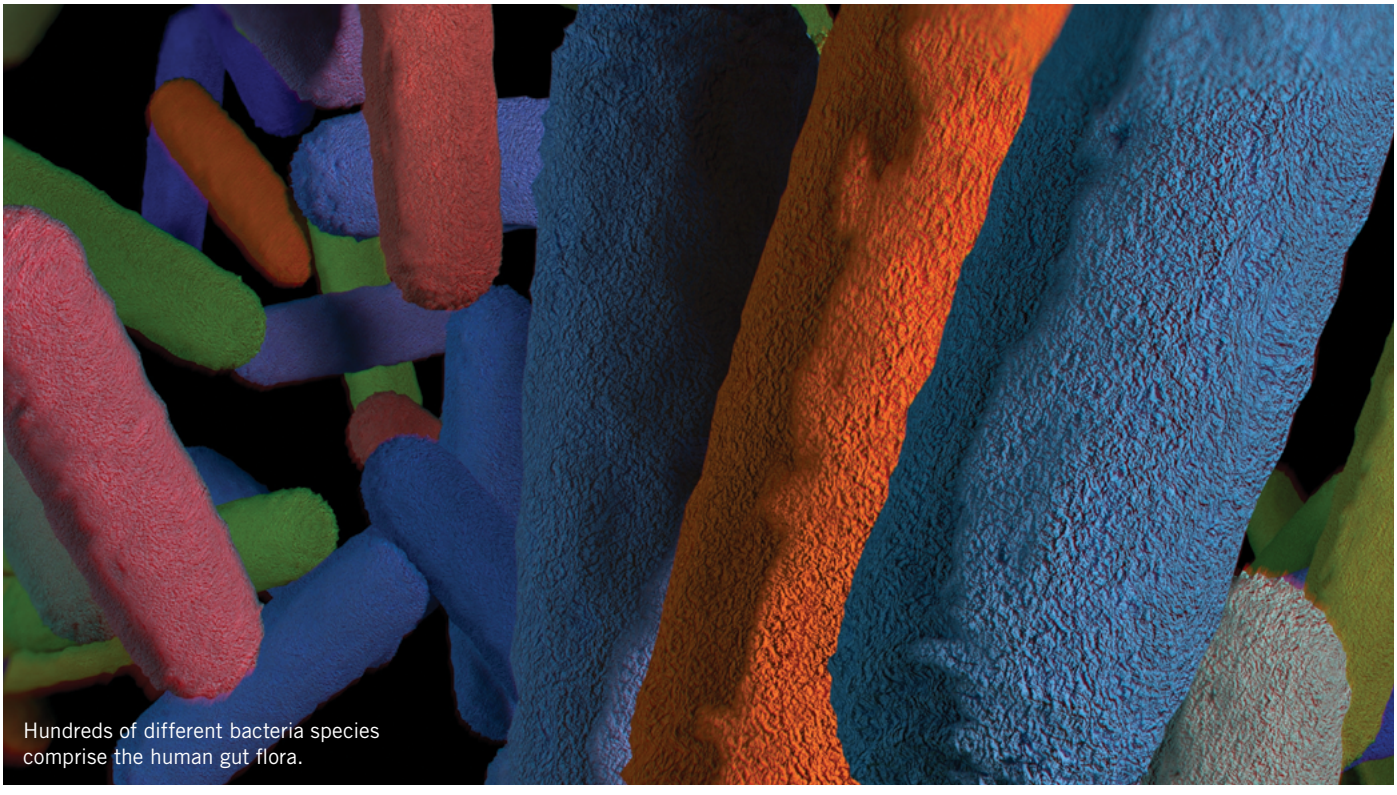


Hundreds of different bacteria species comprise the human gut flora.

Mining microbes: Creating genomic tools to fight disease

With DNA sequencing becoming ever cheaper, genomics has become a popular tool to investigate individual microorganisms or microbiome communities. A variety of sequencing technologies are now available to assist researchers with everything from sample collection to analysis. **By Amber Dance**

In 2015, several patients in the intensive care unit of the Royal Brompton Hospital in London came down with an unexplained illness. Doctors suspected a fungal infection—but those can be difficult to diagnose, and culturing fungi from patient blood samples takes time.

So they called in Jo Rhodes, a genetic epidemiologist at **Imperial College London**. With a miniature DNA-sequencing machine called the MinION, Rhodes was able to identify the pathogen as *Candida auris*, a new species of fungus first seen in Japan in 2009. By comparing sequences between patients, she could tell that the outbreak started with a single source, possibly contaminated equipment. She built a complete genome for the fungus, and by contrasting it to *C. auris* sequences assembled from around the world by the U.S. Centers for Disease Control and Prevention, determined that the hospital strain was related to others from India or Pakistan.

Though the outbreak was contained before its exact origin could be discovered, Rhodes says such rapid sequencing could provide important medical information. For example, sequencing could identify a gene for resistance to the drug fluconazole, which would help doctors avoid a common, but in those cases futile, treatment.

And personalized medicine is not the only application for this kind of microbe DNA-reading. "Cheap sequencing has completely revolutionized microbial genomics," says Jonathan Eisen, an evolutionary microbiologist at the **University of California, Davis,** noting that a full genome for a bacterium or archaeon could cost less than US$100. Researchers are collecting microbial genomes to identify the causes for diverse diseases—from simple viral infections to complex conditions like cancer—that are affected by the microbial community in the gut. They are using those genomes to discover novel treatments—including microbes themselves, or their products. Researchers are also studying the genomes of many other microorganisms, such as those that contaminate water supplies.

Sequencing just the ribosomal RNA (rRNA) genes of microbes can help researchers identify them at the genus level. But many researchers want additional data, for instance, on the different functional genes in a single or group sample, which only full-scale sequencing can provide. Computational tools are then needed to analyze the base pairs. "The heart of what we are doing is bioinformatics," says Pierre Belichard, CEO of **Enterome** in Paris. The company is using fecal samples to develop gut-microbiome–based diagnostics and treatments for inflammatory diseases such as Crohn's disease and cancer. cont.>

Upcoming features

Proteomics: Antibody validation—September 15 ■ Tissue analysis: Liquid tumor biopsies—October 20 ■ DNA/RNA: Single-cell DNA sequencing—November 3

Got poop?

Before researchers like those at Enterome can start crunching data, however, they have a more basic problem—obtaining samples.

At the **Arkansas Children's Research Institute** in Little Rock, clinical researcher John Slattery and colleagues study the causes and physiology of autism spectrum disorder (ASD). Lately, they've been focusing on the gut microbiome, which might be responsible for the gastrointestinal complaints of many with ASD, as well as the mitochondrial dysfunction researchers have observed in them.

Slattery would like to conduct further studies on the gut microbiome of ASD patients, but there's a catch. Stool donations are difficult to acquire—donors must use an unwieldy container called a "collection hat." For both kids and parents, that's a distasteful prospect. Complicating matters, 80% of children with ASD have constipation, diarrhea, or an alternating pattern of both.

That's why Slattery is so keen on the BioCollector, developed by **The BioCollective** in Centennial, Colorado. It hooks conveniently onto the toilet lid and then closes up, "like a terrible present," he says. "You don't ever have to see it."

The BioCollective, founded in 2015, creates an unusual partnership between people willing to send in their stool samples and scientists who want to study them. Members purchase a BioCollector for US$39.95; once they've filled it and mailed it in, their gut microbes will be identified as accurately as possible by rRNA gene sequencing. They are also promised 10% of the profits from every aliquot of their feces that The BioCollective sells to scientists.

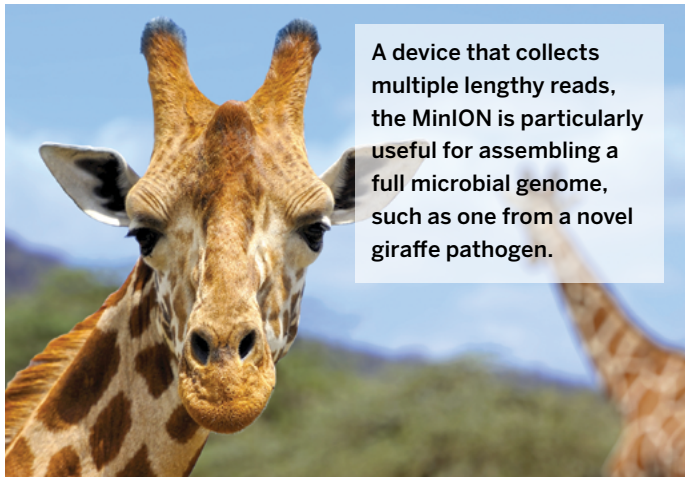
Those researchers, in turn, get easy, efficient, and low-cost access to diverse stool samples. These come complete with donor data, such as antibiotics they've taken or stresses they've endured. The company even hopes to develop a standard "reference" sample—a sort of "poo stew" made of combined healthy samples, says CEO and cofounder Martha Carlin.

Moreover, The BioCollective can provide samples from people with a certain condition or on a specific diet. For example, Noah Zimmerman, a biochemist at **Agro BioSciences** in Wauwatosa, Wisconsin, is interested in how polyphenols—the colorful and beneficial compounds in fruits and vegetables—affect the microbiome. To help Zimmerman, The BioCollective recruited people willing to adopt a high-polyphenol diet for a month before providing their "input."

Once researchers have their samples, they have a couple of options for genomic analysis, and need not perform the work themselves if they don't want to. There are a number of contract research organizations willing to handle sequencing and analysis, such as **CosmosID** of Rockville, Maryland; **Diversigen** of Houston, Texas; and **Second Genome Solutions** in South San Francisco, California.

For purified organisms, one can simply sequence the genome or the transcriptome to identify gene expression patterns. That alone will yield plenty of insights, notes Nur Hasan, vice president and head of R&D at CosmosID. The company can sequence a purified culture from its clients, then perform bioinformatics analyses to provide such information as where it fits in a phylogenetic tree of microorganisms, what antibiotic-resistance genes it harbors, what virulence factors it carries, and more.

For companies developing microorganisms for use in food or medicines, knowing the full sequence of the microbe in question is crucial, notes Jean-Philippe Laine, director of business development at Diversigen.



A device that collects multiple lengthy reads, the MinION is particularly useful for assembling a full microbial genome, such as one from a novel giraffe pathogen.

Giraffe warts

While the next-generation sequencing machines used by Diversigen and others remain hugely popular, another option since 2015 has been the MinION used by Rhodes and produced by **Oxford Nanopore Technologies** in the United Kingdom. It feeds DNA strands through nanopores in a membrane. At the same time, it runs a current through those pores. As adenines, thymines, cytosines, and guanines pass through the pore, each changes the current in a slightly different way, allowing the MinION to read the sequence.

The device, little bigger than a USB stick, can read multiple strands at once through many pores, and collects lengthy reads, unlike shotgun sequencing of short pieces, making the MinION particularly useful for assembling a full microbial genome.

The MinION was the perfect solution for molecular virologists at the **University of Leuven** in Belgium, when veterinarians sent them mystery samples from lesions growing on the faces of giraffes in South African and Danish zoos. Piet Maes and his group prefer to send more than just a few such samples to the traditional next-gen sequencer they typically use—they would rather wait until they can deliver 40–160 samples.

Eager to get the giraffe results, Bert Vanmechelen, a Ph.D. student in Maes' laboratory, fired up the MinION instead. "With the MinION we can do it at our own computers, prepare the sample on day one, and get the results on day two," he explains. In that short amount of time, he was able to sequence the genome of a new kind of papillomavirus, which the researchers named *Giraffa Camelopardalis papillomavirus 1*. There's no cure for this strain of virus, so the giraffes required surgery to remove the lesions—but it was useful to know that the infection was nothing more severe, says Maes.

The MinION also has field capability; it was used to follow the Ebola virus evolution during the recent African epidemic, and has even been rocketed up to the International Space Station.

Get meta

Rhodes and Vanmechelen each had a single microbe to sequence, but for groups of organisms, it's more complicated. Microbiologists seem to be moving away from simply identifying types of microorganisms from rRNA genes, toward more in-depth, whole-genome sequencing, says Sheila Connelly, vice president for research at **Synthetic Biologics,** in Rockville. As she began studying

Featured participants

Agro BioSciences
www.agro-biosciences.com

Arkansas Children's Research Institute
www.archildrens.org/experience-arkansas-childrens-hospital/research/research

Baylor College of Medicine
www.bcm.edu

The BioCollective
www.thebiocollective.com

CosmosID
www.cosmosid.com

Diversigen
diversigen.com

Enterome
www.enterome.com

Imperial College London
www.imperial.ac.uk

Oxford Nanopore Technologies
nanoporetech.com

Second Genome
www.secondgenome.com

Second Genome Solutions
www.secondgenome.com/solutions

Synthetic Biologics
www.syntheticbiologics.com

Texas Children's Microbiome Center
www.bcm.edu/research/centers/texas-childrens-microbiome-center

University of California, Davis
www.ucdavis.edu

University of Leuven
www.kuleuven.be/english

University of Washington
www.washington.edu

how the gut microbiome of pigs is altered by antibiotics—a key medical issue that can allow opportunistic infections such as *Clostridium difficile* to take hold—she hooked up with CosmosID.

The collaborators took a metagenomics approach, shotgun sequencing all the DNA from Connelly's pig fecal samples. The result was a list predicting the bacterial species in the community and their relative abundance, as well as the antibiotic genes likely present in the community as a whole. CosmosID is able to provide this information (e.g., species and strains in a microbial community) to customers due to its carefully curated database of standard microbial genomes, says Hasan.

From the CosmosID data, Connelly and colleagues determined not only how antibiotics change the distribution of bacteria in the pig gut, but also observed the community profile of antibiotic resistance. From that, they could tell how treatment with the antibiotic ceftriaxone allowed organisms that seemed resistant to a variety of drugs to move in, says Connelly.

The pig research helped propel one of Synthetic Biologics' lead products, ribaxamase, an oral medication the company hopes will protect the gut microbiome from the effects of certain intravenous antibiotics that can decimate healthy, beneficial populations and open the door to nastier microorganisms. The drug stays in the gastrointestinal tract, degrading any intravenous antibiotics that reach the gut, but does not penetrate the rest of the body, where those antibiotics are working properly. In human trials, ribaxamase has already been found to reduce *C. difficile* infections.

Like Synthetic Biologics, **Second Genome** aims to turn microbial genomes into ideas for medications, says Mohan Iyer, its chief business officer. For example, Second Genome compared microbes and their functions, analyzing gastrointestinal biopsies of patients with aggravated ulcerative colitis, those with well-controlled colitis, and healthy people.

Through both genome and transcriptome sequencing, the researchers figured out which microbes were present, and which genes were expressed, in healthy versus inflamed guts. From there, they worked out which molecules made by the bacteria may promote or block inflammation. A medication developed by Second Genome to calm inflammation is now in phase 1 human trials.

Fish tacos, anyone?

Companies like Second Genome Solutions, CosmosID, and Diversigen can perform everything from sample prep and sequencing to bioinformatics analysis, enabling scientists to plumb microbiome genomes even if they lack the necessary expertise. But researchers with the right know-how and interest have developed their own novel tools and shared them with the community as well.

The catch is that most analytical tools ask one of two questions: "Who's in the sample?" or, "What genes does the community contain as a group?" Each is only half the story, and it hasn't been easy to integrate these two datasets together, says Elhanan Borenstein, a computational biologist at the **University of Washington** in Seattle. The issue is that different taxonomic groups perform the same function in different peoples' microbiomes. His laboratory recently came up with a solution, a computational method called "Functional Shifts' Taxonomic Contributors" (or FishTaco for short).

For starters, the method surveys all the taxonomic groups in a microbial sample, and based on what's already known about their genomes, infers which genes from the metagenomics sequencing go with which organisms. It can also infer which genes belong with which species when the individual organisms' genomes are unavailable.

Then, it uses that information to help scientists determine which species, or groups of species, are responsible for the crucial differences between the metagenomes of microbiomes they're comparing. For example, Borenstein and colleagues contrasted sequence data from the gut microbiomes of people with type 2 diabetes and healthy people. The metagenomics data showed an overabundance of several sugar transporter genes in the diabetes samples, but which bacteria were responsible for that extra sugar processing? FishTaco determined that members of the genus *Escherichia* were responsible for one type of sugar transport, and members of *Bifidobacterium* contributed a different type.

This kind of information could help scientists envision how to improve health by rebalancing the species that perform desired functions, says Borenstein, for example via antibiotic or probiotic treatments. "It opens the door to a much more tailored and personalized approach for intervention," he says.

It's currently rare for microbial scientists to do this sort of integrated analysis, though they often make educated guesses as to which bacteria perform which functions in a community, says Emily Hollister, a microbial ecologist at the **Baylor College of Medicine** and **Texas Children's Microbiome Center** at Texas Children's Hospital in Houston. She is applying FishTaco to her own studies of microbial imbalance in the gut and respiratory tract.

"The differences we identify may provide insight into potential diagnostic or therapeutic targets," says Hollister.

Amber Dance is a freelance writer living in Los Angeles.

DOI: 10.1126/science.opms.p1700115