

GENOMICS

Microbiome privacy risk

The DNA of microorganisms living on a person's body could identify that individual.

BY EWEN CALLAWAY

Call it a 'gut print'. The collective DNA of the microbes that colonize a human body can uniquely identify someone, researchers have found, raising privacy issues.

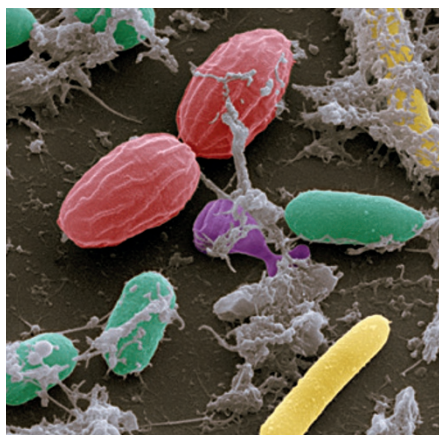
The finding¹, published in *Proceedings of the National Academy of Sciences* on 11 May, suggests that it might be possible to identify a participant in an anonymous study of the body's microbial denizens — its microbiome — and to reveal details about that person's health, diet or ethnicity. A publicly available trove of microbiome DNA maintained by the US National Institutes of Health (NIH), meanwhile, already contains potentially identifiable human DNA, according to a study² published in *Genome Research* on 29 April.

The papers do not name individuals on the basis of their microbiomes — and predict that it would be difficult to do so currently — but they do suggest that those conducting microbiome research should take note.

"Right now, it's a little bit of a Wild West as far as microbiome data management goes," says Curtis Huttenhower, a computational biologist at the Harvard T. H. Chan School of Public Health in Boston, Massachusetts, who led the latest study¹. "As the field develops, we need to make sure there's a realization that our microbiomes are highly unique."

Human-genomics researchers have grappled with privacy concerns for years. In 2013, scientists showed³ that they could name five people who had taken part anonymously in the international 1,000 Genomes project, by cross-referencing their DNA with a genealogy database that also contained ages, locations and surnames.

In recent years, the microbiome's influence on our health and behaviour has become a hot research topic. Data from human-microbiome studies tend to end up in public repositories, but it was not clear whether microbiomes were permanent enough in individuals to identify them over time.



DNA from bacteria in human faeces could be used as a 'gut print' to identify individuals.

Working with publicly available data from the NIH Human Microbiome Project (HMP), Huttenhower's team searched samples taken from body sites, including the gut, mouth, skin and vagina, for combinations of microbial genetic markers that were both unique to a person and stable over time. (Although the HMP does not identify individuals by name, it is possible to compare a participant's first sample with a second one donated weeks or months later.)

Stool samples offered the best microbiome signatures; a person's first sample could be linked to their second sample 86% of the time. By contrast, skin samples could be accurately matched only about one in four times. The researchers note that DNA signatures based on individual strains of microbes did the best job of distinguishing people — much better than those based only on microbial species.

Still, Huttenhower concludes that it would be "exceptionally challenging to do anything with the microbiome data in a single study". The likeliest risk to privacy, he thinks, would come from a scenario in which someone had participated in two different microbiome studies that each contained different pieces of accompanying

information, such as age and health status.

But microbiomes could also pose a privacy risk because they inevitably get jumbled up with human DNA. Although the NIH went to considerable lengths to weed human DNA out of its HMP database, a team led by computational biologist Jonathan Allen of Lawrence Livermore National Laboratory in California has found² that contamination is still rife. For example, the team found sequences known as short tandem repeats that tend to vary between individuals and are used for making DNA matches in forensics. It is not clear whether their presence in microbiome samples could constitute a precise DNA signature, Allen says, but the rise of publicly available DNA databases increases the likelihood. *Genome Research* agreed to publish the paper by Allen and his team only if the NIH removed the known human sequences from its database.

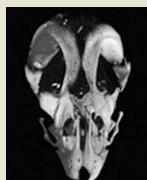
The odds of identifying someone on the basis of their microbiome is low, but researchers should take reasonable steps to protect privacy, says Yaniv Erlich, a computational geneticist at the New York Genome Center who led the team that identified³ participants in the 1,000 Genomes study. Those who took part in the HMP were advised of the risk, says Amy McGuire, a bioethicist at Baylor College of Medicine in Houston, Texas. "I don't think there should be premature panic over this."

An overreaction could slow understanding of the microbiome. Laura Rodriguez, director of policy at the NIH's National Human Genome Research Institute in Bethesda, Maryland, says that as long as protections are in place, such as removing as much human DNA from the HMP as possible, "we would want to keep it in open access because of the value it adds to science". ■

1. Franzosa, E. A. et al. *Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1423854112> (2015).
2. Ames, S. K. et al. *Genome Res.* <http://doi.org/4jt> (2015).
3. Gymrek, M., McGuire, A. L., Golan, D., Halperin, E. & Erlich, Y. *Science* **339**, 321–324 (2013).


**MORE
ONLINE**

TOP STORY



Making chicken embryos with dinosaur-like snouts helps to explain evolution of beak
go.nature.com/zj4wyt

MORE NEWS

- Why the polls got the UK elections wrong go.nature.com/djmijg
- Mobile-phone microscope detects eye parasite go.nature.com/fjykaw
- Five factors that will decide if Philae wakes go.nature.com/secttl

NATURE PODCAST



Latest from the Large Hadron Collider, and neurologist Oliver Sacks's new memoir nature.com/nature/podcast