Finding diversity in the microbiome

The National Institutes of Health Human Microbiome Project is coming to a close, offering an opportunity to reflect on its legacy and the urgent need to understand the microbiome of underrepresented populations.

In 2008, the National Institutes of Health (NIH) launched the Common Fund Human Microbiome Project (HMP) with the mission of generating resources to enable the comprehensive characterization of the human microbiome and analysis of its role in human health and disease. The first phase of the project (HMP1), completed in 2012, represented a critical turning point in this emerging field of research. Capitalizing on the rapid progress of sequencing technologies, it provided an invaluable reference dataset of microbial genomic diversity associated with different sites of the body in healthy humans. It also set the technical standards and a framework for this type of study — from sample collection through computational analyses and data sharing.

HMP1 also showed that understanding the influence of the microbiome on human health and disease would require more than generating high-level catalogues of the genomes of microbes in the gut or on the skin. A more holistic view of host–microbe interaction was needed, and this need served as the foundation of the second phase of HMP, the integrative HMP (iHMP; Proctor, L. M. et al. *Nature* https://doi.org/10.1038/s41586-019-1238-8, 2019), whose main results are now reported in *Nature* and *Nature Medicine*. The three arms of this second phase focused on common human conditions on which the microbiome is known to exert an influence: pregnancy and preterm birth, inflammatory bowel diseases and stressors associated with prediabetes. A wealth of multi-omics data for each of these cohorts is now available to the community to be mined by future work (https://portal.hmpdacc.org/).

As progress is made towards understanding the contribution of the microbiome to human health and disease, there is a growing expectation that this knowledge will eventually be translated into medical advances for diagnosis, prevention or treatment of diseases. Although there have been isolated examples of highly successful microbial interventions — such as fecal microbiome transplants in patients with *Clostridium difficile* infection — at present, translating the microbiome remains an unfulfilled promise. The results of the iHMP Pregnancy Initiative (MOMS-PI) point to another challenge: advancing microbiome-based therapeutics will require taking into account differences in the baseline composition and function of the microbiome across populations with different ancestry and socioeconomic makeup.

The MOMS-PI followed over 1,500 pregnant women longitudinally in the United States through their pregnancies. Samples from multiple body sites and fluids of the mothers and their babies were collected, and a subset of these underwent deep profiling. As a trend, the vaginal microbiome changes during pregnancy converged towards a more homogeneous microbiome dominated by *Lactobacillus* bacteria by the second trimester. The vaginal microbiomes of women who experienced spontaneous preterm delivery were less likely to show this trend. A critical observation is that pregnancy-associated changes in the microbiome are population-specific and were most pronounced in women of African ancestry and with a lower socioeconomic profile in this cohort. In a second cohort of women also in the United States who had uncomplicated pregnancies, again enrichment of the vaginal microbiome for microbes associated with women's vaginal health was only significant in women of African and Latin American descent.

It is too early to say whether these new results will enable the development of biomarkers to predict the risk of preterm birth in any population. However, as the microbiome field enters the clinical space, it is concerning that studies such as the MOMS-PI, which included individuals of non-European descent, are the exception. In fact, availability of metagenomic data across the whole spectrum of human diversity is still fairly limited — perhaps mirroring a long-standing problem in its parental field of genomics. The majority of studies of genetic association with disease have been performed in individuals of European descent, and the lack of diversity in genomic databases has been recognized as a major barrier to translating precision medicine research into practice (Sirugo, G., Williams, S. M. & Tishkoff, T. *Cell* **177**, 26–31, 2019).

Similar studies quantifying the diversity of human microbiome data currently available are lacking and may be difficult to perform, considering that there is not a dedicated repository for deposition of microbiome data. Anecdotally, in the original HMP1 cohort, often used as a 'healthy microbiome' reference, 85% of the subjects enrolled are self-reported white. In the American Gut cohort, another large ongoing initiative for microbiome research, 89% of the participants are self-reported Caucasian. While very recent studies are beginning to document the contribution of ethnicity to variations in the microbiome, and to assess the diversity of the microbiome at a global level (Pasolli, E. et al. *Cell* **176**, 649–662, 2019), they have been outpaced by the speed at which clinical development is progressing. At the time this issue went to press, more than 1,400 trials were already registered in clinicaltrials.gov either documenting a microbiome intervention or listing assessment of the microbiome as an endpoint. How many of those were designed to take into account genetic ancestry and socioeconomic factors as sources of variation in the microbiome and its functional capacity is unclear.

As opposed to 11 years ago when the HMP launched, funding for microbiome sciences now spans multiple government agencies, philanthropic organizations, industry and public–private partnerships. At this new turning point, it is crucial that the critical stakeholders come together to address the problem of lack of human diversity in microbiome studies. Researchers have the responsibility of identifying sensible strategies to engage underrepresented groups and to educate them about the value of participating in research. Funders can help by dedicating resources to studies that increase our knowledge of the microbiome of underrepresented populations, and by specifying inclusion standards or justification of lack of diversity in the award of research grants. Editors can encourage submission of work that addresses gaps in diversity, as well as enforce responsible and transparent data sharing. The real-life impact of microbiome research will require collaboration and the understanding of all of our microbiomes. ❑