NEWS & VIEWS

MICROBIOLOGY

# Tracking humans and microbes

The Human Microbiome Project put the health–associated microbes found in humans on centre stage. The project's second phase shows how microbial disturbance in disease is linked to host processes. SEE PERSPECTIVE P.641 & ARTICLES P.655 & P.663

VERÓNICA LLORÉNS-RICO & JEROEN RAES

Precision medicine can be defined as "an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person" (see go.nature.com/2wv1ch9). Now, studies from the US National Institutes of Health Integrative Human Microbiome Project (iHMP) provide a resource of microbial and human data tracking the progression of two diseases and pregnancy that will help efforts to understand host-associated microorganisms and their interactions with their human host. The resource will be of great value for aiding the precision-medicine approaches of the future. This work includes three flagship studies: two research papers in *Nature*, one by Zhou *et al.*[1] (page 663) and another by Lloyd-Price *et al.*[2] (page 655) and a paper in *Nature Medicine* by Fettweis *et al.*[3]. There is also a Perspective in *Nature* by the iHMP Research Network Consortium[4] (page 641), and another research paper in *Nature Medicine* by Serrano *et al.*[5].

Research on chronic diseases focused exclusively for many years on human cells and processes. Then, almost a decade ago, when analyses of the microbes that live on and in the body (the microbiota) and their collective genomes (the microbiome) were presented in high-profile reports by groups such as the European Union's MetaHIT (Metagenomics of the Human Intestinal Tract) project and the US Human Microbiome Project (HMP), attention shifted to how the microbiota might also be linked to human disease. This work reflected renewed momentum in microbiology research that has been driven by advances in tools for molecular and genomic analysis, and which has continued unabated ever since.

Following the publication of the first wave of major microbiome papers[6–10], knowledge of the microbial inhabitants of our bodies has increased substantially. Over time, the scale of the data presented has been increasing in terms of sample size, depth of genomic coverage and complexity of experimental design. Analysis methods have been developed that are tailored to the specific challenges of handling microbiome samples, including storage, transport and DNA extraction[11], and they have also been designed to deal with the statistical
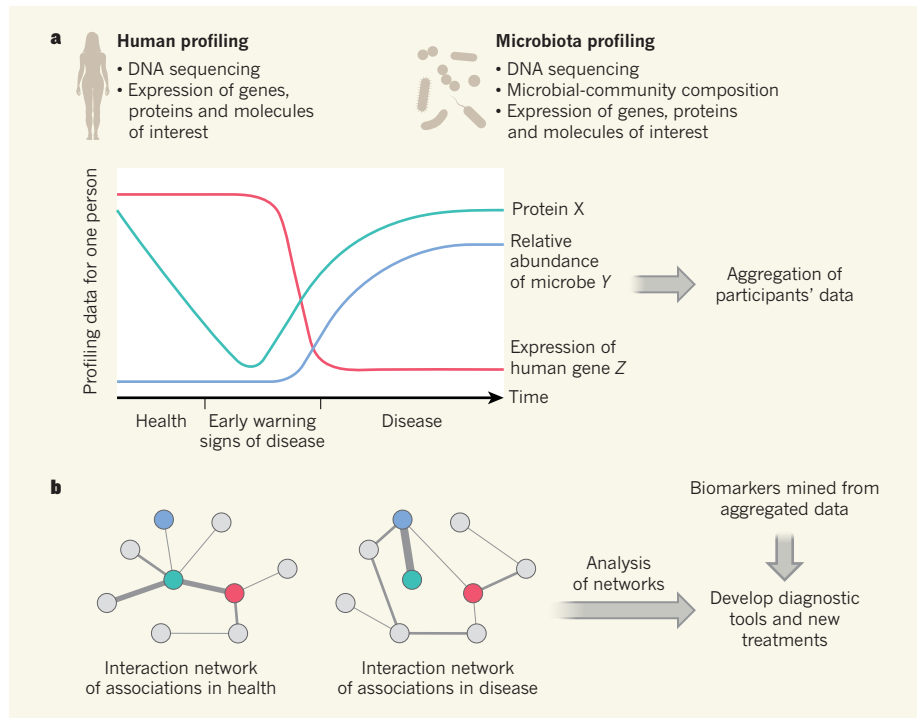
Figure 1 | The integrative Human Microbiome Project (iHMP). Three flagship papers by Zhou *et al.*[1], Lloyd-Price *et al.*[2] and Fettweis *et al.*[3] report flagship studies of the iHMP that tracked people and their resident microorganisms over time. These studies focused on people who were at risk of developing type 2 diabetes, those with inflammatory bowel disease, and women who gave birth prematurely. **a**, The authors of these papers profiled many types of data, including DNA sequences and the expression of genes, proteins and molecules, in individuals and their resident microbes (termed the microbiota). Not all of these types of data were collected for each sample. A hypothetical graph illustrates a small subset of the types of data that might be collected for a person participating in the study. Such data profiles provide a detailed picture of the molecular and microbial changes in a person and their associated microbes over time, in health or disease. The data for individuals were then aggregated for subsequent analysis. **b**, Aggregated participants' data can be used to generate interaction networks for various health or disease statuses that reveal strong associations (thick lines) and weak associations (thin or absent lines) between given measured factors. In the invented networks shown, coloured points match the corresponding factors from the graph for one individual, and white points represent other unique factors. Analysis of such networks, or the identification of biomarkers from other analyses of the data, might lead to the development of innovative diagnostic tools.

complexities of microbiome data[12]. These advances have provided initial glimpses of the complex interactions that occur between our microbes and our bodies, as well as how these change in health and disease, with associations between microbial changes and disease being reported for a wide range of common chronic disorders.

Nevertheless, a big gap remains between our current understanding of microbial patterns and our ability to apply such knowledge to

develop treatments. A transition is in progress that should enable us to move from just having information about associations with disease to understanding the key mechanisms of action that are responsible.

Insufficient knowledge is hindering this transition in three ways. First, most studies have focused on identifying the species composition of the human microbiota and, to a lesser extent, the 'functional potential' that is encoded by the microbiome — for

example, the enzymes that are encoded and the molecules that they might make. However, the actual activity of the microbiota, such as the specific molecules that are being produced or degraded at a given time, is rarely determined in present studies.

Second, most such studies are cross-sectional — that is, they analyse just one sample per individual. Only a few consider the variation of the microbiota over time, and even then, such studies tend to involve just a small number of individuals. Studies that track changes over time are needed for large numbers of people and across a multitude of diseases.

Third, the focus of these studies is often solely the human microbiota. Joint profiling of the human host and the microbiota is rare, yet neglecting this aspect is to omit an important piece of the puzzle that concerns how we interact with our microbial inhabitants. Although some studies overcome one or two of these limitations, studies that address all three are rare. This is where the iHMP data make a difference.

Two years after the original HMP work[7] was published in 2012, the National Institutes of Health established the iHMP to pursue the second phase of this research. Moving from charting the microbiota in health to focusing on disease (Fig. 1), while addressing the aforementioned knowledge gaps, researchers participating in the iHMP tackled three areas of investigation: the onset of type 2 diabetes[1], inflammatory bowel disease[2] and the search for markers of premature birth[3,5].

Together, these projects followed 463 people over periods that ranged from a few months to 4 years. The authors profiled faecal, nasal or vaginal samples on multiple levels — capturing the microbial species that were present and also assessing (depending on the study) the functional potential of the profiled microbiome and the expression patterns of its genes, proteins and metabolites.

In parallel, profiling of the human hosts was also performed, including sequencing protein-coding regions of their genomes, gene-expression analysis, characterizing the presence of immune-system signalling molecules called cytokines, and measuring the levels of clinically relevant factors such as cholesterol. Although the full spectrum of measurements was not captured for all participants at all time points, undeniably, these studies provide the most comprehensive dynamic tracking of the health status of the human host and the characteristics of their microbiota that has been reported so far. The generation of this rich resource is a landmark achievement.

Many interesting patterns were observed in terms of temporal variations in the microbiota that are present at various body sites. Analysis revealed that changes in the composition and function of the vaginal microbiota occur mainly during the first trimester of pregnancy and lead to a predominance of health-associated bacteria from the genus

Lactobacillus, particularly in women who have African or Hispanic ancestry[5]. Interestingly, women who had lower than typical levels of some species of *Lactobacillus* were more likely to deliver babies prematurely[3]. People with inflammatory bowel disease had a higher level of variability in their gut microbiota over time than did people without the condition[2], although microbial-derived molecules such as proteins and metabolites had highly variable levels in all individuals studied.

Although the causes of these examples of microbiota variation over time are not yet known, some evidence suggests that viral infections are a possible trigger of this phenomenon. In inflammatory bowel disease, a rise in the level of viruses was sometimes observed just before the occurrence of abnormalities in the gut microbiota[2], which in some cases were linked to episodes of disease relapse. In the study of the onset of diabetes, intense sampling was undertaken when participants had a viral infection of the respiratory system[1]. This revealed a pattern of coordinated changes in the host immune system and in the gut and nasal microbiota of participants without diabetes.

> "The generation of this rich resource is a landmark achievement."

However, participants who were at risk of developing diabetes had impaired immune responses and unique patterns of change in their microbiota. Future studies will be needed to specifically assess the effects of viral infection on microbiota variation over time, and to determine whether viral infection increases the risk of the emergence or relapse of certain diseases.

The iHMP papers present a breadth of previously unknown correlations between components of the microbiota and between the microbiota and the host. In addition, this vast network of associations is mapped over time. Some of the most interesting connections that were noted in these studies are those between particular microbes and the presence of pro- or anti-inflammatory cytokines. These provide pointers for teasing out the intricate relationships between our immune system and our microbiota and its activity. Such associations will need to be followed up to discern whether direct cause-and-effect relationships can be established.

Even bar-raising studies such as those conducted by the iHMP have weaknesses. In the past few years, key factors that affect microbiota composition have been identified[13], and although some, including age, were considered in the analyses, several other potentially crucial factors, such as diet or taking medication, were not. Work is needed to test whether the associations identified in the iHMP studies are affected by such factors. Another weakness is that, because of the sampling timescale and owing to practical limitations, microbiota variability was typically studied by periodic

sampling, with long intervals between the sampling events. Yet, for the purpose of using such information in the clinic to develop treatments or diagnostic tools, knowledge of the day-to-day variability will be needed.

A further issue is that the data generated are provided in relative units; bacterial abundances are expressed as a proportion of the total microbial population, rather than as the number of cells. To accurately link microbiota activity to host processes, a quantitative analysis framework will be needed to experimentally determine the absolute abundances of the types of microbe that are present in microbiota samples[14]. Finally, only minimal work was performed to replicate the iHMP findings in independently tested populations. Evidence is growing that broad sampling across geographic regions can help to ensure that such findings are robust and more widely applicable[15].

Despite these limitations, one of the most valuable assets of the iHMP work is, undoubtedly, the fantastic trove of data that it provides, most of which is freely available, and which offers a useful resource for future research. These papers also offer prime examples of the comprehensive personalized profiling of samples from humans, for which few precedents exist[16,17]. As such, these works supply some of the much-needed data that could be used to train algorithms or to design future experiments. To develop clinical tools and treatments in this area, it will be essential to compile health-change indications from a wide range of cues that are captured using high-throughput approaches that provide data from both humans and their resident microbes. After all — we're not only bacteria. Right? ∎

Verónica Lloréns-Rico and **Jeroen Raes** *are at the Rega Institute for Medical Research, KU Leuven, 3000 Leuven, Belgium, and at the VIB Center for Microbiology, Leuven. e-mail: jeroen.raes@kuleuven.vib.be*

1. Zhou, W. *et al. Nature* **569**, 663–671 (2019).
2. Lloyd-Price, J. *et al. Nature* **569**, 655–662 (2019).
3. Fettweis, J. M. *et al. Nature Med.* http://10.1038/s41591-019-0450-2 (2019).
4. The Integrative HMP (iHMP) Research Network Consortium. *Nature* **569**, 641–648 (2019).
5. Serrano, M. G. *et al. Nature Med.* http://10.1038/s41591-019-0465-8 (2019).
6. Qin, J. *et al. Nature* **464**, 59–65 (2010).
7. The Human Microbiome Project Consortium. *Nature* **486**, 207–214 (2012).
8. Turnbaugh, P. J. *et al. Nature* **457**, 480–484 (2009).
9. Eckburg, P. B. *et al. Science* **308**, 1635–1638 (2005).
10. Costello, E. K. *et al. Science* **326**, 1694–1697 (2009).
11. Vandeputte, D., Tito, R. Y., Vanleeuwen, R., Falony, G. & Raes, J. *FEMS Microbiol. Rev.* **41**, S154–S167 (2017).
12. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. *Front. Microbiol.* **8**, 2224 (2017).
13. Falony, G. *et al. Science* **352**, 560–564 (2016).
14. Vandeputte, D. *et al. Nature* **551**, 507–511 (2017).
15. He, Y. *et al. Nature Med.* **24**, 1532–1535 (2018).
16. Chen, R. *et al. Cell* **148**, 1293–1307 (2012).
17. Price, N. D. *et al. Nature Biotechnol.* **35**, 747–756 (2017).