

Title	Examining the healthy human microbiome concept
Authors	Joos, Raphaela;Boucher, Katy;Lavelle, Aonghus;Arumugam, Manimozhiyan;Blaser, Martin J.;Claesson, Marcus J.;Clarke, Gerard;Cotter, Paul D.;De Sordi, Luisa;Dominguez-Bello, Maria G.;Dutilh, Bas E.;Ehrlich, Stanislav D.;Ghosh, Tarini Shankar;Hill, Colin;Junot, Christophe;Lahti, Leo;Lawley, Trevor D.;Licht, Tine R.;Maguin, Emmanuelle;Makhalanyane, Thulani P.;Marchesi, Julian R.;Matthijnssens, Jelle;Raes, Jeroen;Ravel, Jacques;Salonen, Anne;Scanlan, Pauline D.;Shkoporov, Andrey;Stanton, Catherine;Thiele, Ines;Tolstoy, Igor;Walter, Jens;Yang, Bo;Yutin, Natalia;Zhernakova, Alexandra;Zwart, Hub;Asnicar, Francesco;Typas, Athanasios;Betsou, Fay;Blottière, Hervé;Bork, Peer;Boutron, Isabelle;Carraturo, Federica;Claesson, Marcus;Cordaillat-Simmons, Magali;Druart, Celine;Fasano, Alessio;Godoy, Yolanda;Haller, Dirk;Hassani, Zahra;Hazenbrink, Diënty H. M. J.;Israelsen, Mads;Iyappan, Anandhi;Jarde, Alexander;Kampshoff, Stephan;Krag, Aleksander;Kriaa, Aicha;Lavelle, Aonghus;Metwaly, Amira;Morozova, Vitalina;Pinto, Federica;Pons, Nicolas;Prost, Pierre-Louis;Ravaud, Philippe;Rhim, Moez;Rodriguez, Julie;Sarati, Arjun;Schierwagen, Robert;Segata, Nicola;Serra, Debora;Trebicka, Jonel;Vecchi, Corrado;Veiga, Patrick;Zitvogel, Laurence;Derosa, Lisa;Doré, Joël;Ross, R. Paul
Publication date	2024-10-23
Original Citation	Joos, R., Boucher, K., Lavelle, A., Arumugam, M., Blaser, M. J., Claesson, M. J., Clarke, G., Cotter, P. D., De Sordi, L., Dominguez-Bello, M. G., Dutilh, B. E., Ehrlich, S. D., Ghosh, T. S., Hill, C., Junot, C., Lahti, L., Lawley, T. D., Licht, T. R., Maguin, E., Makhalanyane, T. P., Marchesi, J. R., Matthijnssens, J., Raes, J., Ravel, J., Salonen, A., Scanlan, P. D., Shkoporov, A., Stanton, C., Thiele, I., Tolstoy, I., Walter, J., Yang, B., Yutin, N., Zhernakova, A., Zwart, H., Doré J. and Ross, R. P. (2024) 'Examining the healthy human microbiome concept', Nature Reviews Microbiology, 23(3), pp.192-205. https://doi.org/10.1038/s41579-024-01107-0
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1038/s41579-024-01107-0

Rights	© 2024, This version of the paper has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1038/s41579-024-01107-0
Download date	2026-10-27 15:15:37
Item downloaded from	https://hdl.handle.net/10468/16696



Examining the healthy human microbiome concept

Raphaëla Joos^{1,2}, Katy Boucher¹, Aonghus Lavelle^{1,3}, Manimozhiyan Arumugam⁴, Martin J. Blaser⁵, Marcus J. Claesson^{1,2}, Gerard Clarke^{1,6}, Paul D. Cotter^{1,7}, Luisa De Sordi⁸, Maria G. Dominguez-Bello⁹, Bas E. Dutilh^{10,11}, Stanislav D. Ehrlich^{12, 13}, Tarini Shankar Ghosh¹⁴, Colin Hill^{1,2}, Christophe Junot¹⁵, Leo Lahti¹⁶, Trevor D. Lawley¹⁷, Tine R. Licht¹⁸, Emmanuelle Maguin¹⁹, Thulani P. Makhalanyane²⁰, Julian R. Marchesi²¹, Jelle Matthijnsens²², Jeroen Raes^{22, 23}, Jacques Ravel^{24, 25}, Anne Salonen²⁶, Pauline D. Scanlan^{1,2}, Andrey Shkorporov^{1,2}, Catherine Stanton^{1,7}, Ines Thiele^{1, 27}, Igor Tolstoy²⁸, Jens Walter^{1,2}, Bo Yang^{29, 30}, Natalia Yutin²⁸, Alexandra Zhernakova³¹, Hub Zwart³², Human Microbiome Action Consortium*, Joël Doré^{12,19}, R. Paul Ross^{1,2,†}

¹APC Microbiome Ireland, University College Cork, Cork, Ireland.

²School of Microbiology, University College Cork, Cork, Ireland.

³Department of Anatomy and Neuroscience, University College Cork, Cork, Ireland.

⁴Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.

⁵Center for Advanced Biotechnology and Medicine, Rutgers University, New Jersey, USA.

⁶Department of Psychiatry and Neurobehavioural Science, University College Cork, Cork, Ireland.

⁷Teagasc Food Research Centre and VistaMilk SFI Research Centre, Moorepark, Fermoy, Cork, Ireland.

⁸Centre de Recherche Saint Antoine, Sorbonne Université, INSERM, Paris, France

⁹Department of Biochemistry and Microbiology, Rutgers University, New Jersey, USA..

¹⁰Institute of Biodiversity, Faculty of Biological Sciences, Cluster of Excellence Balance of the Microverse, Friedrich Schiller University Jena, Jena, Germany.

¹¹Theoretical Biology and Bioinformatics, Department of Biology, Science for Life, Utrecht University, Utrecht, The Netherlands.

¹²Université Paris-Saclay, INRAE, MetaGenoPolis (MGP), Jouy-en-Josas, France.

¹³Department of Clinical and Movement Neurosciences, University College London, London, United Kingdom.

¹⁴Department of Computational Biology, Indraprastha Institute of Information Technology Delhi (IIIT-Delhi), New Delhi, India.

¹⁵Département Médicaments et Technologies pour La Santé (DMTS), Université Paris-Saclay, CEA, INRAE, MetaboHUB, Gif-sur-Yvette, France.

¹⁶Department of Computing, University of Turku, Turku, Finland.

¹⁷Host–Microbiota Interactions Laboratory, Wellcome Sanger Institute, Hinxton, United Kingdom.

¹⁸National Food Institute, Technical University of Denmark, Kgs. Lyngby, Denmark.

¹⁹Université Paris-Saclay, INRAE, AgroParisTech, MICALIS, Jouy-en-Josas, France.

²⁰Department of Microbiology, Faculty of Science, Stellenbosch University, Stellenbosch, South Africa.

²¹Division of Digestive Diseases, Department of Metabolism, Digestion and Reproduction, Imperial College London, London, United Kingdom.

²²KU Leuven, Department of Microbiology, Immunology and Transplantation, Rega Institute, Leuven, Belgium.

²³VIB Center for Microbiology, Leuven, Belgium.

²⁴Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, USA.

²⁵Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, Maryland, USA.

²⁶Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland.

²⁷School of Medicine, University of Ireland, Galway, Ireland.

²⁸National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA.

²⁹State Key Laboratory of Food Science and Resources, Jiangnan University, Jiangsu, Wuxi, China.

³⁰School of Food Science and Technology, Jiangnan University, Jiangsu, Wuxi, China.

³¹Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

³²Erasmus School of Philosophy, Erasmus University Rotterdam, Rotterdam, The Netherlands.

*A list of consortium members appears at the end of the paper.

[†]email: p.ross@ucc.ie

Abstract

Human microbiomes are essential to health throughout the lifespan and are increasingly recognised and studied for their roles in metabolic, immunological and neurological processes. Although the full complexity of these microbial communities is not fully understood, their clinical and industrial exploitation is well advanced and expanding, needing greater oversight guided by a consensus from the research community. One of the most controversial issues in microbiome research is the definition of a 'healthy' human microbiome. This concept is complicated by the microbial variability over different spatial and temporal scales along with the challenge of applying a unified definition to the spectrum of healthy microbiome configurations. In this Perspective, we examine the progress and important gaps that remain to be addressed in order to unlock the full potential of the human microbiome. We propose a roadmap to expand our knowledge of the microbiome–health relationship, incorporating epidemiological approaches informed by the unique ecological characteristics of these communities.

Introduction

The human microbiome, which encompasses communities of microorganisms (microbiota) present at defined niches across the human body and their 'theatre of activity'¹, represents one of the most promising areas of medical research in the coming decades. It is an important driver of host metabolism and immune system activity and is crucial for overall host health and functioning^{2,3}. Recent research has shown that the microbiota plays a role in a variety of diseases including autoimmune and inflammatory diseases, cancer, psychiatric disorders, obesity and type 2 diabetes^{4–12}. It has also been identified as a predictor of all-cause mortality risk over a 15-year follow-up period¹³. These findings have sparked interest in the clinical and industrial sectors, leading to the approval of interventions that use standardised faecal microbiota preparations for treatment of recurrent *Clostridioides difficile* infection (rCDI)^{14,15}. 'Healthy' donor stool remains a subject of investigation in many clinical trials, with health often defined primarily by the absence of overt disease or known pathogens in the donor¹⁶. However, little is known about the features of a 'healthy' microbiota, other than the absence of high-risk species such as extended spectrum β -lactamase (ESBL)-producing *Enterobacteriaceae*^{17,18} and various other commonly accepted overtly pathogenic species¹⁴.

The concept of a 'healthy' microbiome poses a central challenge in microbiome research. It impacts multiple stages of study design, from defining what can be considered a 'difference from control' to translating any findings into microbiome-based diagnostics and therapeutics. Consensus has not been reached on what features best define 'healthy' microbiomes. The lack of consensus is partly due to challenges in characterising the dynamic nature of health and disease states over time, the considerable variability of individual microbiomes and the impact of many environmental factors on microbiome composition¹⁹. These challenges are compounded in the field by poor representation of low and middle-income countries and ethnic minorities^{20–22}, and a focus on gastrointestinal bacterial components over other body sites^{23,24} or viral, protist, fungal and archaeal communities^{25–27}. Additionally, there is comparatively less information on microbiome function^{28–30}, a reliance on single samples, as opposed to high and low resolution time series studies, and a lack of methodological standardisation³¹. Moreover, the cross-sectional nature of many studies precludes identification of microbiome associations with future health states. As such, rather than attempt to define the one ideal 'healthy' microbiome, we need to decide how to best capture as many healthy microbiomes as possible³⁴.

Despite these obstacles defining markers of 'healthy' microbiomes remains essential to advance research, develop reference ranges for microbiome-based diagnostics and therapeutics and for protecting consumer rights. In this Perspective, we summarise specific outputs from the Human Microbiome Action (HMA), a European consortium aiming to standardise and promote cooperation in human microbiome research. We highlight the lack of current consensus on defining healthy microbiomes, emphasise that healthy microbiomes encompass a spectrum of compositions and discuss epidemiology-based approaches that allow healthy microbiome configurations to emerge over time.

Approaches to defining 'healthy' microbiomes

Defining a 'healthy' microbiome is a commonly stated priority and expectation for microbiome science^{29,32} and is shaped by how we define health. In its founding charter, the World Health Organisation (WHO) defined health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity"³³. Despite attempts to revise the definition^{30,32}, it remains a standard mission statement emphasizing a positive and holistic approach to health. Efforts to apply such a definition in microbiome research often distinguish between disease-

related signatures (frequently referred to as the non-scientific term ‘dysbiosis’) and microbial profiles detected in ‘healthy’ or control groups. Such studies require control groups to be defined a priori, selecting individuals based on specified inclusion or exclusion criteria. Studies in which participants are categorised at recruitment constitute most human microbiome research, including the pioneering studies in the field from the METagenomics of the Human Intestinal Tract (MetaHIT) project³⁴ and the Human Microbiome Project²⁴. While this approach has yielded extraordinary insights, it is not without limitations.

A drawback of this binary approach is that ‘healthy’ controls are often obtained from smaller, homogenous convenience groups, such as student volunteers for studies conducted in academic settings³⁵. Overly narrow criteria may exclude healthy individuals (for example, athletes with higher body-mass indices³⁶) and may include controls in pre-clinical stages of various health conditions^{37–39}. Furthermore, excluding chronic medical conditions, medication use or subjects with specific dietary requirements, removes the capacity to match for confounders and potentially contributes to an age bias, as rates of both multi-morbidity and polypharmacy increase with age^{40,41}. The relevance of inclusion and exclusion criteria may also vary depending on the specific microbial niche. For instance, a diagnosis of eczema will be more directly relevant in investigating the skin microbiome, while factors such as current menstrual cycle phase, reproductive and sexual history are crucial co-variables in vaginal microbiome studies. The impact of host health on oral, nasal and skin microbiomes may also be more oblique, making it highly unlikely that a universal set of criteria could be relevant to all microbiomes simultaneously.

Additionally, intrinsic microbiome health is often entirely disregarded, as ecological communities have independent measures of functionality, such as diversity and resilience (FIG. 1a). For instance, antibiotic resistance genes can provide ecological benefit by enhancing the resistance and resilience of microbial communities against perturbation. However, this may be detrimental for the host if such genes are transferred within human microbiomes, leading to the emergence of drug-resistant pathogens^{42,43}. Crucially, microbiomes may not only be markers of current health status but also show potential as predictors of future health and overall longevity^{13,44} (FIG. 1b). Cross-sectional studies may miss latent disease in otherwise healthy individuals, while even in prospective cohort studies, defining health a priori may miss people who respond to management of chronic illness and enter long-term remission. To capture the full spectrum of healthy microbiomes, we must consider disease and its boundaries.

[H2] An epidemiological perspective to a ‘healthy’ microbiome

Epidemiology, the study of the distribution and determinants of disease occurrence in human populations⁴⁵, was originally designed to study and prevent epidemics of infectious diseases. Epidemiological studies consist of large population-level cohorts examined either prospectively or retrospectively, to provide extensive data for the identification of influences, dynamics and associations across aspects of health in a given population. These studies have since been applied to understand risk factors associated with the development of multi-factorial and complex non-communicable diseases (NCDs)⁴⁶, and have identified a range of markers that hold both diagnostic and predictive value in human health (FIG. 1a), such as blood pressure, serum cholesterol, C-reactive protein (CRP) and interleukin (IL)-6⁴⁷.

One of the main goals of microbiome research is to add microbiome-based markers to established clinical biomarkers (FIG 2). Epidemiological approaches have led to the identification of a health-associated microbiome signature^{48,49}, with the convergence of microbiome signatures across diverse disease states, possibly due to the breakdown of host regulation of microbial growth⁵⁰. There are also a number of examples of putative microbiome markers of health and disease, including the association of faecal microbiome gene richness with metabolic health⁵¹, the link of

the *Bacteroides* 2 enterotype with Crohn's disease and systemic inflammation^{52,53}, the vaginal microbiota Community State Type IV (CST IV) associated with bacterial vaginosis using the VALENCIA protocol⁵⁴, the association of *Fusobacterium* spp. and *pks+* *Escherichia coli* with colorectal cancer^{55,56}, the association of microbially derived trimethylamine N-oxide (TMAO) with heart disease in preclinical studies⁵⁷, the associations of the gut microbiome with blood lipid levels⁵⁸ and diverse disease and health markers associated with similar taxonomic patterns in the gut microbiome⁵⁹. A recent study demonstrated that the composition of the gut microbiota can be largely characterized as a combination of five broad species assemblages, or enterosignatures, whose proportions vary across individuals and can be linked to host health variation⁶⁰. Similarly, the so-called succinotypes, which are enterotypical signatures marked by the host- and microbiome-related metabolite succinate, have been established in reliably classifying individuals with, for example, inflammatory bowel disease⁶¹.

Many of these insights into the microbial signatures of health have been derived from population-scale cohort studies (Table 1). In contrast to approaches based on inclusion or exclusion criteria, an unrestricted population-based recruitment refrains from imposing a definition of health, thereby allowing the boundaries of health to 'emerge' from the population and facilitating the refining of such definitions over time.

Nevertheless, the development of large-scale microbiome cohorts and the integration of microbiome epidemiology into study design generates additional challenges, as exemplified by extensive interpersonal variability in taxonomic composition and the plethora of important environmental drivers of community structure⁶². Different distributions of these factors and their interactions within a population (such as urbanization, ethnicity and dietary patterns) may complicate the design of a population study^{63–65}. Some factors, such as diet, not only directly alter the oral and gut microbiota, but also mediate some of their effects on the host⁶⁶, supporting the notion of the microbiota as a proxy for modifiable lifestyle factors mediating long-term disease risk. Importantly, while many of the major environmental determinants of host health also impact the features of the microbiome, some may have a more direct and specific influence on the microbiome, such as pet ownership^{67,68}, medication usage and frequency⁶⁹, gut transit time^{70–72} and cohabitation⁴⁹. Unlike human genomes, the highly plastic nature of microbial communities adds an additional challenge in defining health, with up to 100-fold changes in the abundance of major species within individuals over a short period of time¹⁹. This variability is also reflected in the striking influence of geographic location on microbial composition. Recent research has highlighted major variability in community composition associated with the regional geography within a province of China⁷³. In Western populations, consistencies in composition have been observed⁴⁹, but findings differ greatly when comparing with non-industrialised^{74–76} and transitional populations across urbanisation gradients⁷⁷, or even ethnic minorities within individual countries (for example, Irish Travellers)^{21,74}. Geographical differences are particularly evident from studies showing progressive and persistent changes in composition and function of the gut microbiome following migration^{43,78,79}, which are more pronounced in groups with a younger age at relocation⁸⁰ and with longer duration of residence in the country of resettlement⁸¹. Importantly, while non-industrialised populations show considerable differences in compositional and functional microbial structure, as well as higher microbial diversity within the gut microbiome (which has supported the notion that an overall 'loss of species' is occurring with increasing industrialisation)^{42,82}, these populations often encounter a different set of health conditions⁸³. While non-industrialised populations show lower prevalence and incidence of chronic and non-communicable diseases, their life expectancy and child mortality rates are markedly worse than industrialised populations, often due to infectious diseases mediated by other types of microorganisms⁸⁴. These findings emphasise the importance of understanding variability within

and between countries in establishing baseline health-associated data, as well as geography-relevant confounders and crucial differences in health and disease priorities in non-industrialised and industrialised populations^{85,86}.

Taking these points into consideration, to expand our knowledge of health-associated microbiomes, we emphasise the need for population-level, longitudinal, and ideally, prospective studies that capture a diverse representation of the population and the major host-related environmental determinants of microbiome structure and activity. It is important to consider gaps in current knowledge and the generalisability of findings when identifying populations. While this task may initially seem intractable, it has been applied with great success in several cohorts, with more underway (Table 1). These studies are either standalone microbiome efforts or part of larger epidemiological initiatives, and have already delivered evidence of putative microbiome biomarkers in a variety of health and disease states. We acknowledge that reaching this combined framework will require considerable time, resources and collaboration. However, given the emergence of microbiomes as central aspects of human health and their potential as targets for precision modulation⁸⁷, we believe that these are crucial steps in releasing the potential of the field.

Addressing existing limitations in microbiome research must remain a priority to achieve the maximal impact of large-scale cohorts. For example, the majority of studies in the field to date have been conducted in smaller Western cohorts, typically with a narrow focus on bacterial composition at a genus level within one specific body site and one disease state, employing varying methodologies that limit the comparability of findings. To provide a better foundation for microbiome research, broad improvements to the collection, analysis and reporting of microbiome data and metadata are essential going forward. Combining these efforts with funding calls and initiatives that support the development of global microbiome cohorts and standardisation efforts within the field will also accelerate the identification of the many variants of 'healthy' microbiomes. However, beyond these general aims, reaching consensus on the priorities for future microbiome study design is crucial.

A roadmap for future research

In this section, we discuss how microbiome survey efforts can be directed to maximise the discovery of health- and disease-associated microbial signatures. We emphasise the importance of microbiome-, age-, lifestyle- and population-specific considerations, with the aim of expanding our knowledge of health-associated, or potentially even health-promoting microbiomes relevant to their target populations.

Prioritising population-level microbiome research

Notwithstanding the value of 16S ribosomal RNA gene sequencing, future longitudinal prospective studies should employ shotgun metagenomics at a minimum and ideally would deploy metabolomic and culture-based sampling methods that are fundamental to gain a better understanding of the temporal dynamics of microbiome composition and function. To help mitigate the overall difficulty, cost, and labour involved in establishing microbiome-centred studies, one possible strategy is to integrate microbiome data collection into existing large epidemiological cohorts, as seen with nested microbiome studies within the Dutch LifeLines cohort^{48,49}. Another strategy would be to link microbiome data with clinical metadata available from national health records, as successfully applied in multiple Scandinavian cohorts^{13,88,89}. More recent upcoming and ongoing initiatives such as the French Gut, WeGut⁹⁰, and the Million

Microbiomes from Humans Project aim to apply a similar approach to increasingly larger and more diverse cohorts and improve microbiome research through collaboration and integration with existing projects and national healthcare systems. In addition to enhancing statistical power and representativeness, this approach will also improve the ability to longitudinally track microbiome and related health development. Many population-level microbiome research efforts have highlighted the crucial importance of environmental determinants in shaping the microbiome composition by capturing a wide array of metadata variables (Table 1). Some of these, such as the relevance of transit time to gut communities, remain overlooked and under-captured⁷⁰. A better understanding of the impact of these influences on microbiome function and composition, as well as host–interactions should be a priority for future research. While integration with existing health records represents a path towards an improved quantity of available clinical metadata, improving metadata quality is also essential. For sufficient comparison and interpretation across populations, it is increasingly evident that establishing clear minimal requirements for metadata reporting is essential going forward⁹¹. In particular, improved reporting of complex variables, such as diet and medication usage, is needed to better characterise the full spectrum of factors that can influence a ‘healthy’ microbiome.

Selection of study populations

For the use of this epidemiological approach, the selection of study cohorts should be based on national and international efforts. Given that most insights gained from microbiome cohort studies have been from industrialised and Western populations, a key question is whether microbiomes of individuals in existing cohorts such as those in Japan⁹², Sweden⁸⁹ or Estonia⁶⁹ can be generalised to other countries and continents. Given the association between geographic and individual factors with microbiome composition⁷³, the lack of generalisability is a roadblock for the development and adoption of microbiome-based diagnostics and therapeutics. To address this, strategic national efforts are likely to be required, with regional and international cooperation to best understand these trade-offs and the degree of granularity required. Input from extant microbiome cohorts⁴⁸, epidemiological studies and metagenomic datasets⁹³ will likely be required to help design such studies.

The global gaps in microbiome research highlight the need to include and prioritise sampling of more diverse populations, including geographically remote people, under-sampled regions globally⁹⁴, ethnic groups that have previously been ignored²¹, as well as groups with large developmental changes such as infants, adolescents, and older adults^{95,96}. An improved representation involved enhancing database diversity and forging international partnerships, as it has been proposed in the context of African microbiome research⁸⁵. Intertwined with geographical variation, socioeconomic status represents an understudied influence on human microbiomes both across and within countries and requires more attention in microbiome study design and population selection, given the increasing evidence highlighting the interconnected associations between human microbiomes, socioeconomic status and health^{97–99}.

Another important consideration for selecting populations to define ‘healthy’ microbiomes is age. Over the human life cycle, physiological and metabolic needs vary significantly, as demonstrated by the marked differences between microbial makeup and functionality in different life stages, such as pre- or during infancy, as well as early and later adulthood¹⁰⁰. Starting at birth, the first colonisers of the human body are heavily influenced by maternal health factors, feeding practices and birth mode, with C-section born and formula-fed babies often showing altered microbial composition and functionality. In line with homogeneous feeding practices, the gut microbiomes of infants generally show low diversity and richness compared to adult counterparts^{101–104}.

Furthermore, colonisation is influenced by vertical transmission¹⁰⁵, especially of *Bifidobacterium* and *Bacteroides* strains along with their corresponding bacteriophages, from mother to infant. This process is interrupted by perinatal antibiotics^{106,107}. While this rather sparse *Bifidobacterium* spp.-dominated composition has been considered 'normal' in early life, such a microbiome composition would be detrimental to adult health. A large shift in microbial composition follows the cessation of breast- or bottle-feeding and the introduction of solid foods, which usually occurs around 4 to 6 months of age. This results in a rapid increase in a variety of microbial organisms^{108,109}. By age two to five years, infants gut microbiomes largely resemble adult microbiomes and are thought to remain relatively stable over a life span¹¹⁰. To consider age dependencies of health-associated microbiomes, the development of age-specific reference ranges or criteria for microbial composition and function is necessary. The establishment of baseline profiles of microbial communities at various life stages, as well as longitudinal studies tracking microbiome changes over time are invaluable in providing insight into age-specific dynamics, as well as their impact on health outcomes. The early-life microbiota is particularly important for long-term human health, as reflected in findings from prospective birth cohort studies characterising the microbiome in early life^{90,111,112}. Such work can form the basis for longitudinal studies that are powered to identify microbiome features linked to disease risk, and thus help to define what constitutes a healthy microbiome across the lifespan.

With increasing age, the trajectory of microbiome changes seen in healthy and unhealthy ageing may be different, with frailty, diet and care-setting acting as important co-variables¹¹³. Recent results have suggested that a move from 'core' taxa (such as *Bacteroides* spp.) to rarer microorganisms may be a positive adaptation to physiological changes associated with ageing¹¹⁴. Additionally, medication history, historical physical activity, past dietary habits and specific past laboratory blood parameters also explain a significant fraction of quantitative microbiome variation in older adults, showing the importance of life-long health, lifestyle and exposure on microbiome composition in older age¹¹⁵. In contrast, puberty, which is relatively underexplored, may be more influenced by sex¹¹⁶ or gut-brain axis influences¹¹⁷. Similarly, the onset of menopause and decreases in oestrogen levels with aging represent a window of considerable fluctuation and changes in both vaginal¹¹⁸ and gut¹¹⁹ microbiomes. These different life stages require tailoring of sampling frequency and meta data, considering the respective relevant features.

Refine working definitions for microbiome health

Epidemiological approaches do not require selection criteria as the aim is to sample the population in an unbiased fashion, thereby capturing the true prevalence of health and disease. Although these large studies are crucial in the discovery phase, it will be necessary to causally test putative biomarkers in smaller so-called 'satellite studies' centred around core validation cohorts (FIG. 2), such as case-control studies drawn from these large cohorts evaluating a specific condition or clinical intervention. Within this framework, working definitions of human and microbiome health can be applied and refined over time, as more data are accrued and our understanding deepens. Moreover, as opposed to individual and methodologically different studies, satellite studies benefit from access to a larger population of control subjects. The American Gut Project is a good example of this, as it is able to extensively control for confounding factors¹²⁰. Based on this idea, and in contrary to the current process of biomarker testing, the larger population-level cohort prioritises inclusivity and diversity by reporting, but not excluding, potential participants based on factors such as recent antibiotic usage or co-morbidities, while smaller studies help in establishing causality. A working definition of microbiome health is thereby one which not only emerges from a large cohort study as a set of (quasi-)universal features

(preferably functional) that can be observed in a large set of people in a population, but also one that is set with a specific research question in mind.

Exploration of extra-intestinal niches

There is a need to prioritise microbiome research beyond the gut. While much of the current research on the microbiome has focused on the gut, in particular the large intestine owing to its higher microbial density¹²¹, there is growing recognition that the microbiome plays a vital role in many other areas of the human body. These include the skin, oral cavity and urogenital system, as well as the highly debated presence of resident microbial communities at various low-biomass sites¹²². By expanding the scope of microbiome research to include these other body sites, we can gain a more complete understanding of the role that the microbiome plays in human health. Compared to studies of the gut microbiome, research into the microbiomes of other host niches has predominantly relied on smaller case-control studies. Many of these studies are nested within larger cohorts and have been useful for examining microbiome associations and mechanisms in disease-specific contexts, such as the oral microbiome with different types of cancers^{123,124}, and the vaginal microbiome with HIV acquisition¹²⁵.

Efforts have been made to include a wider range of sampling sites into existing population-level cohort studies. An example is the 4D-SZ cohort that includes faecal¹²⁶, oral¹²⁷ and vaginal¹²⁸ microbiome data. Such research has emphasized the association between host genetic factors and the composition of the oral microbiome^{127,129}, which may account for the limited contribution of environmental factors to oral microbiome composition, as observed in a Canadian cohort¹³⁰. Although analysis of non-faecal microbiomes is gaining focus in population-scale cohorts, within these studies such body sites are not prioritised and often only sampled at lower resolutions in smaller subgroups compared to the gut microbiome. Citizen science initiatives represent an alternative approach to improve sampling of sites beyond the gut, with the Saca La Lengua¹³¹ and Isala¹³² projects focused on studying oral and vaginal microbiomes, respectively (BOX 2).

While population-level cohorts are beginning to provide insights into the composition of the microbiome across multiple host niches, the functional properties of these communities and their role in health remain underexplored¹³³. In the case of the skin microbiome, functional properties differ highly from the gastrointestinal microbiome, with many species enriched for lipid metabolism to utilise host secretion of sebum lipids¹³⁴. Similarly, the exploration of the vaginal microbiome suggests that microbial function may be a better predictor of vaginal inflammation than composition¹³³. In the oral microbiome, shifts in functional capacities have been linked to disease phenotypes such as Parkinson disease¹³⁵. With these opportunities in mind, research must target location-specific differences in functional cores and their impact on varying diseases, enabling a definition of microbiome health that includes body site specificity.

Obtaining new mechanistic insights

For faecal microbiome studies, emerging research indicates that going beyond species to capture strain-level information^{136,137}, as well as genomic detail such as structural variants¹³⁸ and single-nucleotide polymorphisms (SNPs)¹³⁹, is important to understand the association of gut microbiomes and health¹⁴⁷. The complementation of taxonomic analysis with functional approaches is critical for future efforts and may contribute to more generalisable definitions of health, as microbiome functionality has been reported to be more stable and conserved across individuals²⁴ in contrast to the interindividual variability seen in microbiome composition¹⁴⁰. The concept of a 'healthy' functional core microbiome, describing a subset of functions covering essential metabolic and physiological processes governed by the microbiome, not only allows for

more comparability within but also between different populations, thereby opening new opportunities to pinpoint canonical features of microbiome health.

To analyse microbial functionality, shotgun sequencing currently represents a preferable analytic approach, as it allows for direct functional annotation and analysis of metagenomes compared to amplicon sequencing, which can only infer functionality based on taxonomic profiles and associated traits¹⁴¹. Additionally, the sequencing depth achievable with shotgun sequencing allows strain-level resolution (access to the whole genome to identify how conspecific strains differ) for more microbial agents, which is especially important considering strain-specific levels of pathogenicity (for example, in *Escherichia coli*) or antimicrobial resistance, as well as for diagnostics and therapeutics like probiotics in precision medicine¹⁴². To help understand underlying biochemical mechanisms, strain-level analysis combined with functional analysis can link genetic variations with distinct functions that contribute to microbiome health^{143,144} and provide data to support the development of personalised strain-level models of microbial metabolism¹⁴⁵. Finally, including long-read sequencing could lead to new insights into complex genetic variation. As reviewed elsewhere, short-read sequencing is more cost-effective and accurate and, therefore, often more suitable for large studies. However, long-read sequencing can offer improved resolution and genome assembly. Long-read sequencing or hybrid approaches can provide insights into genotypes and phase haplotypes, as well as potentially facilitate de novo assembly of yet uncharacterised genomes, which ultimately could improve the classification of genetic material^{146,147}. While we acknowledge the potential financial burden that such high-resolution sequencing poses, the ever-decreasing cost of such technology, partly tied to not-so-distant expiration dates of industry patents and a general rise in innovation and development in biotechnology, will ensure their feasibility in the future. Nevertheless, it is important to note that the functional annotation of microbial genes is still far from perfect, which sets some limitations on its use as a microbiome descriptor. The presence of any given gene does not necessarily mean that it is functionally expressed, with research showing low correlations between abundances of specific microbial metabolites and the presence of genetic pathways known to produce them¹⁴⁸.

Aside from insight into metabolic outputs of the microbiome with a high enough sequence coverage, shotgun metagenomics provides a broader compositional overview of multiple components of human microbiomes, including viral, prokaryotic and eukaryotic species (BOX 1). However, the different genome sizes or lower abundance of non-bacterial microbiome components available for sequencing, together with limitations in existing fungal and viral reference databases may necessitate integration with other targeted sequencing approaches^{149–151}. Conventional metagenomic sequencing does not provide sufficient fungal genome coverage, hence Internal transcribed spacer (ITS) amplicon sequencing specific to fungi may also need to be integrated to accurately characterise the mycobiome¹⁵². Objective measures of dietary components directly from stool samples have recently become available using either deep shotgun sequencing or chloroplast amplicon sequencing¹⁵³.

Besides metagenomic data, metabolomics technologies are demonstrating great promise in characterising the functional activity of the microbiome by identifying microbiota-derived metabolites¹⁵⁴ and documenting host–microbiota interactions¹⁵⁵. Metabolomics should be considered as adjunctive approaches to sequencing, once scaled to population-level cohorts¹⁵⁶. Similarly, metatranscriptomics holds potential as a tool to explore how microbial functions and expression profiles adapt to various exposures, with analysis from the Men’s Life Validation Study highlighting higher metatranscriptomic variation within individuals over time compared to metagenomic profiles^{157,158}. In contrast, for samples acquired directly from tissue with high levels of host DNA and low microbial biomass, host depletion or targeted amplicon sequencing may be

preferable¹⁵⁹. This was shown from an analysis of maternal breast milk samples using a combined 16S-ITS-23S rRNA gene sequencing approach, which was able to show shared patterns between maternal breast milk and infant oral microbiome composition¹⁶⁰.

Current analysis of microbiome composition typically only covers the relative abundance of taxa, which can mask the full scope of microbial community dynamics and the impact of different influences on overall microbial load^{19,161}. Quantitative microbiome profiling combines sequencing with microbial load estimation using methods such as flow cytometry, quantitative PCR or the addition of 'spike-in' controls during sequencing¹⁶². Measures of absolute abundance may also act as a potential marker of a 'healthy' microbiome itself, as seen in studies that have applied these methods to explore the dynamics of bacterial and fungal community assembly in infants¹⁶¹. The use of absolute abundance has revealed that microbial load represents a key driver of relative abundance alterations observed in patients with Crohn disease⁵² and colorectal cancer¹⁶³. Looking forward, a combination of strain-level taxonomic and functional metagenomics represents an optimal strategy for future research. While there are limitations with current approaches, the integration of sequencing with other technical and analytical tools will play a key role in overcoming these hurdles to better characterise the complexity and diversity of 'healthy' human microbiomes.

Conclusions and outlook

Microbiome science is entering a phase where high-resolution multi-omic analysis is a practical reality. We can now capture unprecedented genetic and functional information and identify deep connections to our physiology and health. To realise the promise of these tools for expanding our knowledge of the interactive and dynamic nature of human health, microbiome surveys must broaden their scope and ambition, prioritising diversity within populations and regions that have been underexplored, as well as addressing knowledge gaps in microbial functions, microbial kingdoms and different niches across the human body. With the rising popularity of various medical and industrial products that offer to support a 'healthy' microbiome and improve microbiome health, comes the responsibility of researchers to engage in responsible promise management¹⁶⁴ and strive towards a shared and practicable definition of a 'healthy' microbiome. To this end, studies rooted in epidemiological principles that account for the unique ecological features of microbiomes have led the way in examining the potential of the microbiome as a reporter of health. Continuing and expanding from these approaches with (inter)national, prospective, longitudinal cohorts will be essential to clarify the additional potential of the microbiome as a predictor of health. Expanding our knowledge of the spectrum of healthy microbiomes will require input across multiple sectors, from governments, funding bodies, industry partners, clinicians and most importantly, the public, patients and populations we aim to help.

References

1. Berg, G. *et al.* Microbiome definition re-visited: old concepts and new challenges. *Microbiome* **8**, 103 (2020).
2. Lynch, S. V. & Pedersen, O. The Human Intestinal Microbiome in Health and Disease. *New England Journal of Medicine* **375**, 2369–2379 (2016).
3. Turnbaugh, P. J. *et al.* The Human Microbiome Project. *Nature* **449**, 804–810 (2007).
4. Larsen, N. *et al.* Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS One* **5**, e9085 (2010).
5. Qin, N. *et al.* Alterations of the human gut microbiome in liver cirrhosis. *Nature* **513**, 59–64 (2014).
6. Wang, T. *et al.* Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME Journal* **6**, 320–329 (2012).
7. Gevers, D. *et al.* The treatment-naïve microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**, 382–392 (2014).
8. Jeffery, I. B. *et al.* Differences in Fecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. *Gastroenterology* **158**, 1016–1028.e8 (2020).
9. McGuinness, A. J. *et al.* A systematic review of gut microbiota composition in observational studies of major depressive disorder, bipolar disorder and schizophrenia. *Mol Psychiatry* **27**, 1920–1935 (2022).
10. Nikolova, V. L. *et al.* Perturbations in Gut Microbiota Composition in Psychiatric Disorders: A Review and Meta-analysis. *JAMA Psychiatry* **78**, 1343–1354 (2021).
11. Zhang, X., Chen, B., Zhao, L. & Li, H. The Gut Microbiota: Emerging Evidence in Autoimmune Diseases. *Trends Mol Med* **26**, 862–873 (2020).
12. Jansen, D. *et al.* Community Types of the Human Gut Virome are Associated with Endoscopic Outcome in Ulcerative Colitis. *J Crohns Colitis* **17**, 1504–1513 (2023).
13. Salosensaari, A. *et al.* Taxonomic signatures of cause-specific mortality risk in human gut microbiome. *Nat Commun* **12**, 2671 (2021).
14. US Food and Drug Administration. Approval Letter - REBYOTA. www.fda.gov/OurSTN:BL125739/0 (2022).
15. Walter, J. & Shanahan, F. Fecal microbiota-based treatment for recurrent *Clostridioides difficile* infection. *Cell* **186**, 1087 (2023).
16. Bibbò, S. *et al.* Fecal microbiota transplantation: Screening and selection to choose the optimal donor. *J Clin Med* **9**, 1–14 (2020).
17. van Duijkeren, E. *et al.* Long-term Carriage of Extended-Spectrum β -Lactamase-Producing *Escherichia coli* and *Klebsiella pneumoniae* in the General Population in The Netherlands. *Clin Infect Dis* **66**, 1368–1376 (2018).
18. Piewngam, P. *et al.* Composition of the intestinal microbiota in extended-spectrum β -lactamase-producing Enterobacteriaceae carriers and non-carriers in Thailand. *Int J Antimicrob Agents* **53**, 435–441 (2019).
19. Vandeputte, D. *et al.* Temporal variability in quantitative human gut microbiome profiles and implications for clinical research. *Nat Commun* **12**, 6740 (2021).
20. Abdill, R. J., Adamowicz, E. M. & Blekhman, R. Public human microbiome data are dominated by highly developed countries. *PLoS Biol* **20**, e3001536 (2022).

21. Keohane, D. M. *et al.* Microbiome and health implications for ethnic minorities after enforced lifestyle changes. *Nat Med* **26**, 1089–1095 (2020).
22. Makhalanyane, T. P. *et al.* African microbiomes matter. *Nature Reviews Microbiology* **21**, 479–481 (2023).
23. Reynoso-García, J. *et al.* A complete guide to human microbiomes: Body niches, transmission, development, dysbiosis, and restoration. *Frontiers in Systems Biology* **2**, 951403(2022).
24. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
25. Vemuri, R., Shankar, E. M., Chieppa, M., Eri, R. & Kavanagh, K. Beyond just bacteria: Functional biomes in the gut ecosystem including virome, mycobiome, archaeome and helminths. *Microorganisms* **8**, 483 (2020).
26. Stockdale, S. R. & Hill, C. Progress and prospects of the healthy human gut virome. *Curr Opin Virol* **51**, 164–171 (2021).
27. Laforest-Lapointe, I. & Arrieta, M.-C. Microbial Eukaryotes: a Missing Link in Gut Microbiome Studies. *mSystems* **3**, e00201-17 (2018).
28. Lloyd-Price, J., Abu-Ali, G. & Huttenhower, C. The healthy human microbiome. *Genome Medicine* **8**, 51 (2016).
29. Shanahan, F., Ghosh, T. S. & O'Toole, P. W. The Healthy Microbiome—What Is the Definition of a Healthy Gut Microbiome? *Gastroenterology* **160**, 483–494 (2021).
30. Najmanová, L., Vídeňská, P. & Cahová, M. Healthy Microbiome – A Mere Idea or a Sound Concept? *Physiol Res* **71**, 719–738 (2022).
31. Vandeputte, D., Tito, R. Y., Vanleeuwen, R., Falony, G. & Raes, J. Practical considerations for large-scale gut microbiome studies. *FEMS Microbiol Rev* **41**, S154–S167 (2017).
32. McBurney, M. I. *et al.* Establishing What Constitutes a Healthy Human Gut Microbiome: State of the Science, Regulatory Considerations, and Future Directions. *Journal of Nutrition* **149**, 1882–1895 (2019).
33. *Constitution of the World Health Organization. American Journal of Public Health and the Nations Health* vol. 36 (1946).
34. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
35. Herd, P., Palloni, A., Rey, F. & Dowd, J. B. Social and population health science approaches to understand the human microbiome. *Nat Hum Behav* **2**, 808–815 (2018).
36. Barton, W. *et al.* The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. *Gut* **67**, 625–633 (2018).
37. Raygoza Garay, J. A. *et al.* Gut Microbiome Composition Is Associated With Future Onset of Crohn's Disease in Healthy First-Degree Relatives. *Gastroenterology* **165**, 670–681 (2023).
38. Ferreiro, A. L. *et al.* Gut microbiome composition may be an indicator of preclinical Alzheimer's disease. *Sci Transl Med* **15**, eabo2984 (2023).
39. Ashton, J. J., Seaby, E. G., Beattie, R. M. & Ennis, S. NOD2 in Crohn's Disease—Unfinished Business. *J Crohns Colitis* **17**, 450–458 (2023).
40. Chowdhury, S. R., Chandra Das, D., Sunna, T. C., Beyene, J. & Hossain, A. Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. *EClinicalMedicine* **57**, 101860(2023).

41. Khezrian, M., McNeil, C. J., Murray, A. D. & Myint, P. K. An overview of prevalence, determinants and health outcomes of polypharmacy. *Ther Adv Drug Saf* **11**, 2042098620933741 (2020).
42. Sonnenburg, E. D. & Sonnenburg, J. L. The ancestral and industrialized gut microbiota and implications for human health. *Nat Rev Microbiol* **17**, 383–390 (2019).
43. Le Bastard, Q., Vangay, P., Batard, E., Knights, D. & Montassier, E. US immigration is associated with rapid and persistent acquisition of antibiotic resistance genes in the gut. *Clinical Infectious Diseases* **71**, 419–421 (2020).
44. Pang, S. *et al.* Longevity of centenarians is reflected by the gut microbiome with youth-associated signatures. *Nat Aging* **3**, 436–449 (2023).
45. Foxman, B. & Riley, L. American Journal of EPIDEMIOLOGY Molecular Epidemiology of Infection Foxman and Riley Molecular Epidemiology: Focus on Infection. (2001).
46. Midha, S., Chawla, S. & Garg, P. K. Modifiable and non-modifiable risk factors for pancreatic cancer: A review. *Cancer Lett* **381**, 269–277 (2016).
47. Pradhan, A. D. *et al.* Inflammatory Biomarkers, Hormone Replacement Therapy, and Incident Coronary Heart Disease Prospective Analysis From the Women's Health Initiative Observational Study. *JAMA* **288**, 980–987 (2002).
48. Zhernakova, A. *et al.* Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* **352**, 565–569 (2016).
49. Gacesa, R. *et al.* Environmental factors shaping the gut microbiome in a Dutch population. *Nature* **604**, 732–739 (2022).
50. Winter, S. E. & Bäumler, A. J. Gut dysbiosis: Ecological causes and causative effects on human disease. *Proc Natl Acad Sci USA* **120**, e2316579120 (2023).
51. Le Chatelier, E. *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541–546 (2013).
52. Vandeputte, D. *et al.* Quantitative microbiome profiling links gut community variation to microbial load. *Nature* **551**, 507–511 (2017).
53. Vieira-Silva, S. *et al.* Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. *Nature* **581**, 310–315 (2020).
54. France, M. T. *et al.* VALENCIA: a nearest centroid classification method for vaginal microbial communities based on composition. *Microbiome* **8**, 166 (2020).
55. Wang, N. & Fang, J. Y. *Fusobacterium nucleatum*, a key pathogenic factor and microbial biomarker for colorectal cancer. *Trends Microbiol* **31**, 159–172 (2023).
56. Pleguezuelos-Manzano, C. *et al.* Mutational signature in colorectal cancer caused by genotoxic pks + *E. coli*. *Nature* **580**, 269–273 (2020).
57. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–65 (2011).
58. Fu, J. *et al.* The gut microbiome contributes to a substantial proportion of the variation in blood lipids. *Circ Res* **117**, 817–824 (2015).
59. Manor, O. *et al.* Health and disease markers correlate with gut microbiome composition across thousands of people. *Nat Commun* **11**, 5206 (2020).
60. Frioux, C. *et al.* Enterosignatures define common bacterial guilds in the human gut microbiome. *Cell Host Microbe* **31**, 1111–1125.e6 (2023).
61. Anthamatten, L. *et al.* Stratification of human gut microbiomes by succinotype is associated with inflammatory bowel disease status. Preprint at <https://www.biorxiv.org/content/10.1101/2023.11.21.568118v1>(2023)

62. Falony, G. *et al.* Population-level analysis of gut microbiome variation. *Science* **352**, 560-564 (2016).
63. Cotillard, A. *et al.* A posteriori dietary patterns better explain variations of the gut microbiome than individual markers in the American Gut Project. *American Journal of Clinical Nutrition* **115**, 432-443 (2022).
64. Walker, R. L. *et al.* Population study of the gut microbiome: associations with diet, lifestyle, and cardiometabolic disease. *Genome Med* **13**, 188 (2021).
65. Lu, J. *et al.* Chinese gut microbiota and its associations with staple food type, ethnicity, and urbanization. *NPJ Biofilms Microbiomes* **7**, 71 (2021).
66. Armet, A. M. *et al.* Rethinking healthy eating in light of the gut microbiome. *Cell Host Microbe* **30**, 764-785 (2022).
67. Kates, A. E. *et al.* Household Pet Ownership and the Microbial Diversity of the Human Gut Microbiota. *Front Cell Infect Microbiol* **10**, 73 (2020).
68. Tun, H. M. *et al.* Exposure to household furry pets influences the gut microbiota of infants at 3-4 months following various birth scenarios. *Microbiome* **5**, 40 (2017).
69. Aasmets, O., Krigul, K. L., Lüll, K., Metspalu, A. & Org, E. Gut metagenome associations with extensive digital health data in a volunteer-based Estonian microbiome cohort. *Nat Commun* **13**, 869 (2022).
70. Procházková, N. *et al.* Advancing human gut microbiota research by considering gut transit time. *Gut* **72**, 180-191 (2023).
71. Roager, H. M. *et al.* Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat Microbiol* **1**, 16093 (2016).
72. Vandeputte, D. *et al.* Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* **65**, 57-62 (2016).
73. He, Y. *et al.* Regional variation limits applications of healthy gut microbiome reference ranges and disease models. *Nat Med* **24**, 1532-1535 (2018).
74. Schnorr, S. L. *et al.* Gut microbiome of the Hadza hunter-gatherers. *Nat Commun* **5**, 3654 (2014).
75. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222-227 (2012).
76. De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A* **107**, 14691-14696 (2010).
77. Tamburini, F. B. *et al.* Short- and long-read metagenomics of urban and rural South African gut microbiomes reveal a transitional composition and undescribed taxa. *Nat Commun* **13**, 926 (2022).
78. Shad, N. S., Shaikh, N. I. & Cunningham, S. A. Migration Spurs Changes in the Human Microbiome: a Review. *J Racial Ethn Health Disparities* (2023)
79. Copeland, J. K. *et al.* The Impact of Migration on the Gut Metagenome of South Asian Canadians. *Gut Microbes* **13**, 1-29 (2021).
80. Kaplan, R. C. *et al.* Gut microbiome composition in the Hispanic Community Health Study/Study of Latinos is shaped by geographic relocation, environmental factors, and obesity. *Genome Biol* **20**, 219 (2019).
81. Vangay, P. *et al.* US Immigration Westernizes the Human Gut Microbiome. *Cell* **175**, 962-972.e10 (2018).
82. Blaser, M. J. The theory of disappearing microbiota and the epidemics of chronic diseases. *Nat Rev Immunol* **17**, 461-463 (2017).

83. Browne, H. P. *et al.* Boosting microbiome science worldwide could save millions of children's lives. *Nature* **625**, 237-240 (2024).
84. Wang, H. *et al.* Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet* **390**, 1084–1150 (2017).
85. Brewster, R. *et al.* Surveying Gut Microbiome Research in Africans: Toward Improved Diversity and Representation. *Trends Microbiol* **27**, 824–835 (2019).
86. Smith, M. I. *et al.* Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* **339**, 548–554 (2013).
87. Deehan, E. C. *et al.* Precision Microbiome Modulation with Discrete Dietary Fiber Structures Directs Short-Chain Fatty Acid Production. *Cell Host Microbe* **27**, 389-404.e6 (2020).
88. Tebani, A. *et al.* Integration of molecular profiles in a longitudinal wellness profiling cohort. *Nat Commun* **11**, 4487 (2020).
89. Dekkers, K. F. *et al.* An online atlas of human plasma metabolite signatures of gut microbiome composition. *Nat Commun* **13**, 5370 (2022).
90. Gou, W. *et al.* Westlake Gut Project: A consortium of microbiome epidemiology for the gut microbiome and health research in China. *Medicine in Microecology* **14**, 100064 (2022).
91. Cernava, T. *et al.* Metadata harmonization—Standards are the key for a better usage of omics data for integrative microbiome analysis. *Environmental Microbiomes* **17**, 33 (2022).
92. Nagata, N. *et al.* Population-level Metagenomics Uncovers Distinct Effects of Multiple Medications on the Human Gut Microbiome. *Gastroenterology* **163**, 1038–1052 (2022).
93. Pasoli, E. *et al.* Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell* **176**, 649-662.e20 (2019).
94. Allali, I. *et al.* Human microbiota research in Africa: a systematic review reveals gaps and priorities for future research. *Microbiome* **9**, 241 (2021).
95. Stewart, C. J. *et al.* Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583–588 (2018).
96. Claesson, M. J. *et al.* Gut microbiota composition correlates with diet and health in the elderly. *Nature* **488**, 178–184 (2012).
97. Amato, K. R. *et al.* The human gut microbiome and health inequities. *Proc Natl Acad Sci U S A* **118**, e2017947118 (2021).
98. Dixon, M., Dunlop, A. L., Corwin, E. J. & Kramer, M. R. Joint effects of individual socioeconomic status and residential neighborhood context on vaginal microbiome composition. *Front Public Health* **11**, 1029741 (2023).
99. Bowyer, R. C. E. *et al.* Socioeconomic status and the gut microbiome: A twinsuk cohort study. *Microorganisms* **7**, 17 (2019).
100. Kaisanlahti, A. *et al.* Maternal microbiota communicates with the fetus through microbiota-derived extracellular vesicles. *Microbiome* **11**, 249 (2023).
101. Mueller, N. T., Differding, M. K., Østbye, T., Hoyo, C. & Benjamin-Neelon, S. E. Association of birth mode of delivery with infant faecal microbiota, potential pathobionts, and short chain fatty acids: a longitudinal study over the first year of life. *BJOG* **128**, 1293–1303 (2021).

102. Coker, M. O. *et al.* Specific class of intrapartum antibiotics relates to maturation of the infant gut microbiota: a prospective cohort study. *BJOG* **127**, 217–227 (2020).
103. Yang, B. *et al.* Bifidobacterium and lactobacillus composition at species level and gut microbiota diversity in infants before 6 weeks. *Int J Mol Sci* **20**, 3306 (2019).
104. Beller, L. *et al.* Successional Stages in Infant Gut Microbiota Maturation. *mBio* **12**, e0185721 (2021).
105. Laursen, M. F. *et al.* Maternal milk microbiota and oligosaccharides contribute to the infant gut microbiota assembly. *ISME Communications* **1**, 21 (2021).
106. Duranti, S. *et al.* Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission. *Microbiome* **5**, 66 (2017).
107. Feehily, C. *et al.* Detailed mapping of Bifidobacterium strain transmission from mother to infant via a dual culture-based and metagenomic approach. *Nat Commun* **14**, 3015 (2023).
108. McKeen, S. *et al.* Adaptation of the infant gut microbiome during the complementary feeding transition. *PLoS One* **17**, e0270213 (2022).
109. Lim, E. S. *et al.* Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nat Med* **21**, 1228–1234 (2015).
110. Laursen, M. F., Bahl, M. I. & Licht, T. R. Settlers of our inner surface-factors shaping the gut microbiota from birth to toddlerhood. *FEMS Microbiol Rev* **45**, fuab001 (2021).
111. Warmink-Perdijk, W. D. B. *et al.* Lifelines NEXT: a prospective birth cohort adding the next generation to the three-generation Lifelines cohort study. *Eur J Epidemiol* **35**, 157–168 (2020).
112. Korpela, K. *et al.* Cohort profile: Finnish Health and Early Life Microbiota (HELMi) longitudinal birth cohort. *BMJ Open* **9**, e028500 (2019).
113. Jeffery, I. B., Lynch, D. B. & O'Toole, P. W. Composition and temporal stability of the gut microbiota in older persons. *ISME Journal* **10**, 170–182 (2016).
114. Wilmanski, T., Gibbons, S. M. & Price, N. D. Healthy aging and the human gut microbiome: why we cannot just turn back the clock. *Nat Aging* **2**, 869–871 (2022).
115. Si, J. *et al.* Long-term life history predicts current gut microbiome in a population-based cohort study. *Nat Aging* **2**, 885–895 (2022).
116. Korpela, K. *et al.* Gut microbiota develop towards an adult profile in a sex-specific manner during puberty. *Sci Rep* **11**, 23297 (2021).
117. McVey Neufeld, K. A., Luczynski, P., Dinan, T. G. & Cryan, J. F. Reframing the teenage wasteland: Adolescent microbiota-gut-brain axis. *Canadian Journal of Psychiatry* **61**, 214–221 (2016).
118. Park, M. G., Cho, S. & Oh, M. M. Menopausal Changes in the Microbiome—A Review Focused on the Genitourinary Microbiome. *Diagnostics* **13**, 1193 (2023).
119. Peters, B. A. *et al.* Menopause Is Associated with an Altered Gut Microbiome and Estrobolome, with Implications for Adverse Cardiometabolic Risk in the Hispanic Community Health Study/Study of Latinos. *mSystems* **7**, e0027322 (2022).
120. McDonald, D. *et al.* American Gut: an Open Platform for Citizen Science Microbiome Research. *mBio* **3**, e00031-18 (2018).
121. McCallum, G. & Tropini, C. The gut microbiota and its biogeography. *Nat Rev Microbiol* **22**, 105–118 (2023).

122. Kennedy, K. M. *et al.* Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies. *Nature* **613**, 639–649 (2023).
123. Hayes, R. B. *et al.* Association of oral microbiome with risk for incident head and neck squamous cell cancer. *JAMA Oncol* **4**, 358–365 (2018).
124. Fan, X. *et al.* Human oral microbiome and prospective risk for pancreatic cancer: A population-based nested case-control study. *Gut* **67**, 120–127 (2018).
125. McClelland, R. S. *et al.* Evaluation of the association between the concentrations of key vaginal bacteria and the increased risk of HIV acquisition in African women from five cohorts: a nested case-control study. *Lancet Infect Dis* **18**, 554–564 (2018).
126. Jie, Z. *et al.* A transomic cohort as a reference point for promoting a healthy human gut microbiome. *Medicine in Microecology* **8**, 100039 (2021).
127. Liu, X. *et al.* Metagenome-genome-wide association studies reveal human genetic impact on the oral microbiome. *Cell Discov* **7**, 117 (2021).
128. Jie, Z. *et al.* Life History Recorded in the Vagino-cervical Microbiome Along with Multi-omes. *Genomics Proteomics Bioinformatics* **20**, 304–321 (2022).
129. Kurushima, Y. *et al.* Host Genotype Links to Salivary and Gut Microbiota by Periodontal Status. *J Dent Res* **102**, 146–156 (2023).
130. Nearing, J. T., DeClercq, V., Van Limbergen, J. & Langille, M. G. I. Assessing the Variation within the Oral Microbiome of Healthy Adults. *mSphere* **5**, e00451-20 (2020).
131. Willis, J. R. *et al.* Citizen-science reveals changes in the oral microbiome in Spain through age and lifestyle factors. *NPJ Biofilms Microbiomes* **8**, 38 (2022).
132. Lebeer, S. *et al.* A citizen-science-enabled catalogue of the vaginal microbiome and associated factors. *Nat Microbiol* **8**, 2183–2195 (2023).
133. Correia, G. D., Marchesi, J. R. & MacIntyre, D. A. Moving beyond DNA: towards functional analysis of the vaginal microbiome by non-sequencing-based methods. *Curr Opin Microbiol* **73**, 102292 (2023).
134. Byrd, A. L., Belkaid, Y. & Segre, J. A. The human skin microbiome. *Nat Rev Microbiol* **16**, 143–155 (2018).
135. Jo, S. *et al.* Oral and gut dysbiosis leads to functional alterations in Parkinson's disease. *NPJ Parkinsons Dis* **8**, 87 (2022).
136. De Filippis, F. *et al.* Distinct Genetic and Functional Traits of Human Intestinal Prevotella copri Strains Are Associated with Different Habitual Diets. *Cell Host Microbe* **25**, 444-453.e3 (2019).
137. Valles-Colomer, M. *et al.* The person-to-person transmission landscape of the gut and oral microbiomes. *Nature* **614**, 125–135 (2023).
138. Zhernakova, D. V. *et al.* Host genetic regulation of human gut microbial structural variation. *Nature* **625**, 813–821 (2024).
139. Zahavi, L. *et al.* Bacterial SNPs in the human gut microbiome associate with host BMI. *Nat Med* **29**, 2785–2792 (2023).
140. Chen, L. *et al.* The long-term genetic stability and individual specificity of the human gut microbiome. *Cell* **184**, 2302-2315.e12 (2021).
141. Jovel, J. *et al.* Characterization of the gut microbiome using 16S or shotgun metagenomics. *Front Microbiol* **7**, 459 (2016).
142. Hu, D., Fuller, N. R., Caterson, I. D., Holmes, A. J. & Reeves, P. R. Single-gene long-read sequencing illuminates Escherichia coli strain dynamics in the human intestinal microbiome. *Cell Rep* **38**, (2022).
143. Lloyd-Price, J. *et al.* Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **550**, 61–66 (2017).

144. Carrow, H. C., Batachari, L. E. & Chu, H. Strain diversity in the microbiome: Lessons from *Bacteroides fragilis*. *PLoS Pathog* **16**, e1009056 (2020).
145. Magnúsdóttir, S. *et al.* Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat Biotechnol* **35**, 81–89 (2017).
146. Gehrig, J. L. *et al.* Finding the right fit: evaluation of short-read and long-read sequencing approaches to maximize the utility of clinical microbiome data. *Microb Genom* **8**, 000794 (2022).
147. Tedersoo, L., Albertsen, M., Anslan, S. & Callahan, B. Perspectives and Benefits of High-Throughput Long-Read Sequencing in Microbial Ecology. *Appl Environ Microbiol* **87**, 1–19 (2021).
148. Pascal Andreu, V. *et al.* gutSMASH predicts specialized primary metabolic pathways from the human gut microbiota. *Nat Biotechnol* **41**, 1416–1423 (2023).
149. Lind, A. L. & Pollard, K. S. Accurate and sensitive detection of microbial eukaryotes from whole metagenome shotgun sequencing. *Microbiome* **9**, 58 (2021).
150. Shkoporov, A. N. *et al.* The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. *Cell Host Microbe* **26**, 527–541.e5 (2019).
151. Begum, N. *et al.* Host-mycobiome metabolic interactions in health and disease. *Gut Microbes* **14**, 2121576(2022).
152. Usyk, M. *et al.* Comprehensive evaluation of shotgun metagenomics, amplicon sequencing, and harmonization of these platforms for epidemiological studies. *Cell Reports Methods* **3**, 100391 (2023).
153. Pollak, S. Plant DNA in feces as a nutritional crystal ball. *Proc Natl Acad Sci USA* **120**, e2309172120 (2023).
154. Quinn, R. A. *et al.* Global chemical effects of the microbiome include new bile-acid conjugations. *Nature* **579**, 123–129 (2020).
155. Guo, C. J. *et al.* Discovery of Reactive Microbiota-Derived Metabolites that Inhibit Host Proteases. *Cell* **168**, 517–526.e18 (2017).
156. Hajjar, G. *et al.* Scaling-up metabolomics: Current state and perspectives. *TrAC - Trends in Analytical Chemistry* **167**, 117225 (2023).
157. Wang, K. *et al.* The gut microbiome modifies the associations of short- and long-term physical activity with body weight changes. *Microbiome* **11**, 121 (2023).
158. Mehta, R. S. *et al.* Stability of the human faecal microbiome in a cohort of adult men. *Nat Microbiol* **3**, 347–355 (2018).
159. Yap, M. *et al.* Evaluation of methods for the reduction of contaminating host reads when performing shotgun metagenomic sequencing of the milk microbiome. *Sci Rep* **10**, 21665 (2020).
160. Spreckels, J. E. *et al.* Analysis of microbial composition and sharing in low-biomass human milk samples: a comparison of DNA isolation and sequencing techniques. *ISME Communications* **3**, 116 (2023).
161. Rao, C. *et al.* Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature* **591**, 633–638 (2021).
162. Galazzo, G. *et al.* How to Count Our Microbes? The Effect of Different Quantitative Microbiome Profiling Approaches. *Front Cell Infect Microbiol* **10**, 403 (2020).
163. Tito, R. Y. *et al.* Microbiome confounders and quantitative profiling challenge predicted microbial targets in colorectal cancer development. *Nat Med* **30**, 1339–1348(2024).

164. Chadwick, R. & Zwart, H. From ELSA to responsible research and Promisomics. *Life Sci Soc Policy* **9**, 3 (2013).
165. Nishijima, S. *et al.* Extensive gut virome variation and its associations with host and environmental factors in a population-level cohort. *Nat Commun* **13**, 5252 (2022).
166. Liu, X. *et al.* Metagenome-genome-wide association studies reveal human genetic impact on the oral microbiome. *Cell Discov* **7**, 117 (2021).
167. Jie, Z. *et al.* A transomic cohort as a reference point for promoting a healthy human gut microbiome. *Medicine in Microecology* **8**, 100039 (2021).
168. Brandwein, M., Katz, I., Katz, A. & Kohen, R. Beyond the gut: Skin microbiome compositional changes are associated with BMI. *Hum Microb J* **13**, 100063 (2019).
169. Hughes, D. A. *et al.* Genome-wide associations of human gut microbiome variation and implications for causal inference analyses. *Nat Microbiol* **5**, 1079–1087 (2020).
170. Koponen, K. K. *et al.* Associations of healthy food choices with gut microbiota profiles. *American Journal of Clinical Nutrition* **114**, 605–616 (2021).
171. Palmu, J. *et al.* Association between the gut microbiota and blood pressure in a population cohort of 6953 individuals. *J Am Heart Assoc* **9**, e016641 (2020).
172. Palmu, J. *et al.* Gut microbiome and atrial fibrillation—results from a large population-based study. *EBioMedicine* **91**, 104583 (2023).
173. Peters, B. A. *et al.* Association of the gut microbiome with kidney function and damage in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). *Gut Microbes* **15**, 2186685 (2023).
174. Nash, A. K. *et al.* The gut mycobiome of the Human Microbiome Project healthy cohort. *Microbiome* **5**, 153 (2017).
175. Zhernakova, A. *et al.* Population-Based Metagenomics Analysis Reveals Markers for Gut Microbiome Composition and Diversity. *Science* **352**, 565–569.
176. Rühlemann, M. C. *et al.* Genome-wide association study in 8,956 German individuals identifies influence of ABO histo-blood groups on gut microbiome. *Nat Genet* **53**, 147–155 (2021).
177. Frost, F. *et al.* Long-term instability of the intestinal microbiome is associated with metabolic liver disease, low microbiota diversity, diabetes mellitus and impaired exocrine pancreatic function. *Gut* **70**, 522–530 (2021).
178. Nogal, A. *et al.* Genetic and gut microbiome determinants of SCFA circulating and fecal levels, postprandial responses and links to chronic and acute inflammation. *Gut Microbes* **15**, 2240050 (2023).
179. Koonin, E. V., Dolja, V. V. & Krupovic, M. The healthy human virome: from virus–host symbiosis to disease. *Current Opinion in Virology* **47**, 86–94 (2021).
180. Kumata, R., Ito, J., Takahashi, K., Suzuki, T. & Sato, K. A tissue level atlas of the healthy human virome. *BMC Biol* **18**, 55 (2020).
181. Dutilh, B. E. *et al.* A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat Commun* **5**, 4498 (2014).
182. Yutin, N. *et al.* Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. *Nat Commun* **12**, 1044 (2021).
183. Edwards, R. A. *et al.* Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. *Nat Microbiol* **4**, 1727–1736 (2019).
184. Clooney, A. G. *et al.* Whole-Virome Analysis Sheds Light on Viral Dark Matter in Inflammatory Bowel Disease. *Cell Host Microbe* **26**, 764–778.e5 (2019).

185. Ma, Y., You, X., Mai, G., Tokuyasu, T. & Liu, C. A human gut phage catalog correlates the gut phageome with type 2 diabetes. *Microbiome* **6**, 24 (2018).
186. Roux, S., Hallam, S. J., Woyke, T. & Sullivan, M. B. Viral dark matter and virus–host interactions resolved from publicly available microbial genomes. *Elife* **4**, e08490 (2015).
187. Beller, L. & Matthijnssens, J. What is (not) known about the dynamics of the human gut virome in health and disease. *Curr Opin Virol* **37**, 52–57 (2019).
188. Chibani, C. M. *et al.* A catalogue of 1,167 genomes from the human gut archaeome. *Nat Microbiol* **7**, 48–61 (2022).
189. Koskinen, K. *et al.* First Insights into the Diverse Human Archaeome: Specific Detection of Archaea in the Gastrointestinal Tract, Lung, and Nose and on Skin. *mBio* **8**, e00824-17 (2017).
190. Beghini, F. *et al.* Large-scale comparative metagenomics of Blastocystis, a common member of the human gut microbiome. *ISME Journal* **11**, 2848–2863 (2017).
191. Scanlan, P. D. *et al.* The microbial eukaryote Blastocystis is a prevalent and diverse member of the healthy human gut microbiota. *FEMS Microbiol Ecol* **90**, 326–330 (2014).
192. Marzano, V. *et al.* “Omic” investigations of protozoa and worms for a deeper understanding of the human gut “parasitome”. *PLoS Negl Trop Dis* **11**, (2017).
193. Armstrong, A. J. S., Parmar, V. & Blaser, M. J. Assessing saliva microbiome collection and processing methods. *NPJ Biofilms Microbiomes* **7**, 81 (2021).
194. Robinson, C. K., Brotman, R. M. & Ravel, J. Intricacies of assessing the human microbiome in epidemiologic studies. *Ann Epidemiol* **26**, 311–321 (2016).
195. Hill, C. J. *et al.* Effect of room temperature transport vials on DNA quality and phylogenetic composition of faecal microbiota of elderly adults and infants. *Microbiome* **4**, 19 (2016).
196. Tang, Q. *et al.* Current Sampling Methods for Gut Microbiota: A Call for More Precise Devices. *Front Cell Infect Microbiol* **10**, 151 (2020).
197. Lourenço, M. *et al.* The Spatial Heterogeneity of the Gut Limits Predation and Fosters Coexistence of Bacteria and Bacteriophages. *Cell Host Microbe* **28**, 390–401.e5 (2020).
198. Badawy, R. *et al.* Metadata Concepts for Advancing the Use of Digital Health Technologies in Clinical Research. *Digit Biomark* **3**, 116–132 (2019).
199. Berry, S. E. *et al.* Human postprandial responses to food and potential for precision nutrition. *Nat Med* **26**, 964–973 (2020).

Acknowledgements

We want to thank all those who helped facilitate the workshop on the 6th of October 2022, including workshop attendees, postdoctoral scribes, audiovisual experts and staff of the River Lee Hotel in Cork. This publication has been supported by the project entitled International Human Microbiome Coordination and Support Action (IHMCSCA), which receives funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 964590.

Author contributions

R.J., K.B. and A.L. wrote the article. All authors contributed substantially to discussion of the content, reviewed and edited the manuscript before submission.

Competing interests statement

M.J.C. and P.D.C. are co-founders of SeqBiome Ltd. The rest of the authors declare no competing interests.

Consortium authorship

Human Microbiome Action Consortium

Manimozhiyan Arumugam, Francesco Asnicar, Athanasios Typas, Fay Betsou, Hervé Blottière, Peer Bork, Katy Boucher, Isabelle Boutron, Federica Carraturo, Marcus Claesson, Magali Cordaillat-Simmons, Joël Doré, Celine Druart, Alessio Fasano, Yolanda Godoy, Dirk Haller, Zahra Hassani, Diënty Hazenbrink, Colin Hill, Mads Israelsen, Anandhi Iyappan, Alexander Jarde, Raphaela Joos, Stephan Kampshoff, Aleksander Krag, Aicha Kriaa, Aonghus Lavelle, Emmanuelle Maguin, Amira Metwaly, Vitalina Morozova, Federica Pinto, Nicolas Pons, Pierre-Louis Prost, Philippe Ravaud, Moez Rhimi, Julie Rodriguez, Paul Ross, Arjun Sarati, Robert Schierwagen, Nicola Segata, Debora Serra, Jonel Trebicka, Corrado Vecchi, Patrick Veiga, Laurence Zitvogel and Hub Zwart, Lisa Derosa

Peer review information

Nature Reviews Microbiology thanks Eugene B. Chang, Orlando DeLeon and the other, anonymous, reviewers for their contribution to the peer review of this work.

Display Items

Table 1. Characteristics of population-level microbiome cohort studies

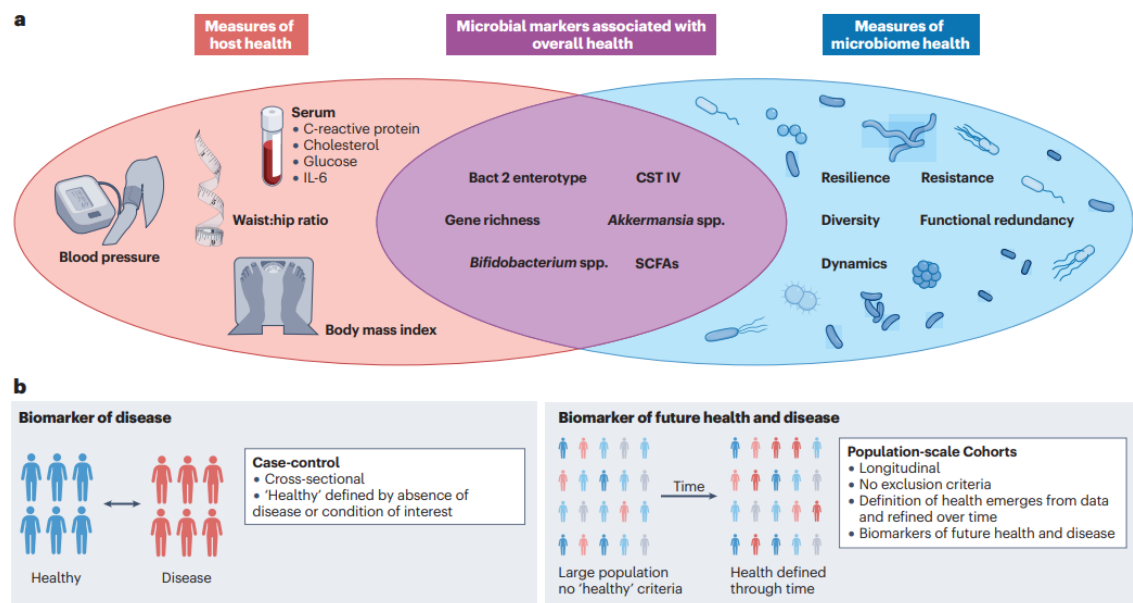
Cohort Name	Participants (n)	Data generated	Study features	References
4D (Disease, Drug, Diet and Daily Life)	4198	WGS Faecal Bacteriome and virome	Japan Longitudinal data for 243 participants	92,165
4D-SZ	2183 (Faecal) 2675 (Oral)	WGS Faecal, oral and vaginal samples	China Multi-omic analysis Exclusion criteria (only for analysis) for pregnancy or the presence of disease	128,166,167
American Gut	11005	16S Faecal, oral, skin and vaginal samples	United States (USA) and United Kingdom (UK) Longitudinal data for 565 participants Exclusion criteria (only for analysis) for BMI outside of 18-30 range, IBD, diabetes and antibiotic usage within 1 year 'Citizen Science' approach	63,120,168
Atlantic Partnership for Tomorrow's Health (PATH)	1049 (Oral)	16S Oral	Canada Subgroup of CanPath cohort Exclusion criteria (only for analysis) for the presence of disease	130
Dutch Microbiome Project	8208	WGS Faecal	Netherlands Multi-omic analysis Subgroup of LifeLines cohort	49
Estonian Microbiome Project (EstMB)	2509	WGS Faecal	Estonia Multi-omic analysis Subgroup of EstBB cohort	69
Flemish Gut Flora Project (FGFP)	5302	16S WGS Metabolomics Faecal	Belgium Adults Multi-omic analysis Longitudinal data from 1000 individuals	62,169
Framingham Heart Study (FHS)	1356	16S Faecal	USA Data from newer cohorts within FHS (Gen3, OMNI2 and New Offspring cohorts)	64
FINRISK	7231	16S and WGS Faecal	Finland Multi-omic analysis Adults (25-74 years)	13,170-172

		Plasma metabolomics		
HCHS/SOL Gut Origins of Latino Diabetes (GOLD)	3297	16S, ITS and WGS Faecal Bacteriome and mycobiome	Specific to Latino population in USA Exclusion criteria (only for analysis) for antibiotic usage within 6 months and relocation to USA after age 26 Subgroup of HCHS/SOL cohort (Hispanic Community Health Study/Study of Latinos)	80,119,152,173
Human Microbiome Project (HMP)	300	16S, ITS and WGS Faecal, oral, skin, nasal and vaginal samples Bacteriome and mycobiome	USA Extensive exclusion criteria Longitudinal data for all participants, with data from 3 timepoints available for 100 participants	3,174
Isala	3196 (Vaginal)	16S Vaginal, saliva and skin samples	Belgium 'Citizen Science' approach	132
LifeLines DEEP	1135	WGS and 16S Faecal	Netherlands Multi-omic analysis Longitudinal data for 338 participants (WGS) Subgroup of LifeLines cohort	140,175
Men's Lifestyle Validation Study (MLVS)	308	WGS Faecal	Specific to male healthcare professionals in USA Multi-omic analysis (Metatranscriptomics) Longitudinal data for all participants Subgroup of Health Professionals Follow-up Study (HPFS)	157,158
Saca La Lengua (SLL2)	1648 (Oral)	16S and fungal culturing Oral Bacteriome and mycobiome	Spain 'Citizen Science' approach	131
SCAPIS	8583	WGS Faecal	Sweden Longitudinal multi-omic data available for 94 participants	88,89
TwinsUK	2507	16S (oral) and WGS (faecal)	Specific to twins in the UK Multi-omic analysis	129,178

Westlake Gut Project – WeGut Consortium	25799	16S ITS and WGS (ongoing) Faecal	China Consortium of existing cohorts with microbiome data	90
---	-------	--	---	----

Studies represent cohorts with limited or no exclusion criteria. Participants numbers are recorded as the largest reported number of adults as of March 2024 (from faecal sampling unless stated otherwise). WGS, whole genome sequencing; 16S, 16S ribosomal RNA sequencing; ITS, internal transcribed spacer; BMI, body mass index; IBD, inflammatory bowel disease.

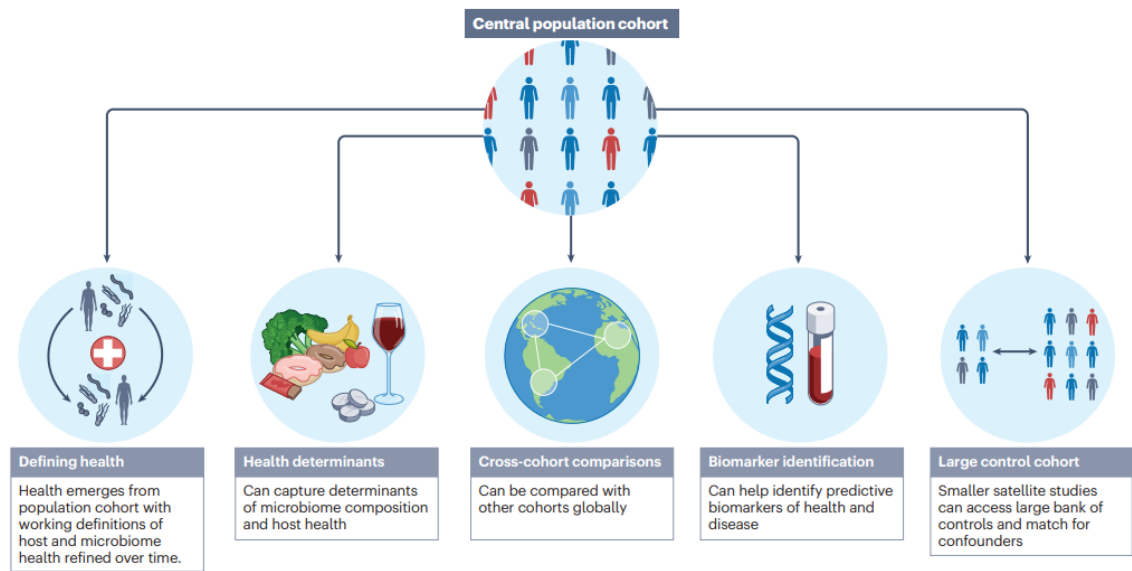
Fig. 1. Using the microbiome as a reporter and predictor of health.



a) A central aim of microbiome research is to identify features of healthy microbial communities that best facilitate the health of the host. This requires the combination of independent measures of both the host and the microbial community health. These measures represent the cumulative impact of genetic and environmental factors determining the overall health of the host and/or the microbiota, such as nutrition, medication or lifestyle choices (for example, smoking and alcohol). Potential microbial biomarkers can range from measures of defined taxa (for example, the presence or abundance of *Bifidobacterium* spp, Bact2 enterotype, *Akkermansia* spp., etc.) or metabolites (for example, short-chain fatty acids (SCFAs)) to overall community composition and functionality. However, unlike existing markers of host health status, these microbial markers are yet to be validated for clinical use, and many potential biomarkers remain unexplored. b) The potential of microbiome data to act as a reporter of current disease status or as a predictor of future health is shaped by choices in study design and how 'healthy' is defined. Cross-sectional case-control studies can determine associations between the microbiome and current disease states but require identifying and stratifying patients into a binary of healthy and disease groups. Following a large population over time without defining health beforehand can reveal how general health adapts over time, including recovery from disease and independent of disease in the case of healthy ageing. Longitudinal population-scale studies can address all potential applications of microbiome data, by capturing a total population without initial exclusion criteria

to define health, which can then be stratified and viewed under different lenses depending on the research question being asked (either through prospective or retrospective analysis).

Fig. 2. Satellite model for microbiome research.



Development and maintenance of longitudinal population-scale cohorts represent the optimal strategy for future microbiome research. These baseline longitudinal 'core' cohorts would consist of representative populations with no exclusion criteria and inclusion restricted only by logistic or geographical factors, created either as independent microbiome-specific cohorts or as subsets of existing non-microbiome epidemiological studies⁹⁵ (as seen with those included in Table 1) and. These cohorts enable the identification of health determinants (including environmental, genetic and lifestyle factors), and thereby a more inclusive and accurate definition of health, over time by using detailed metadata to capture influences of health and disease. This approach facilitates the discovery of novel biomarkers by tracking microbial signatures across entire populations or specific subgroups, and hence supporting longitudinal and comparative analysis across various populations. Additionally, such studies would also support the development of personalised approaches by stratifying and grouping separate sub-populations present within the central cohort. With access to data from a larger centralised cohort to validate findings, insights from independent 'satellite' studies could be quickly and cost-effectively explored at scale to improve comparability and generalisability, as well as provide sufficient statistical power.

Box 1. Non-bacterial components of human microbiomes

Developments in sequencing tools and reference databases have improved the characterisation of viruses, fungi, archaea and other eukaryotes present across various human body sites^{131,149}.

Viruses (virome)

The viral communities inhabiting human body sites are primarily composed of bacteriophages^{28,179} and show considerable inter-individual and geographic diversity²⁶ akin to their bacterial hosts. Eukaryote-infecting viruses have also been observed in humans, often in the absence of clinical disease^{179,180}. Previous population-scale and longitudinal analysis of the gut virome support the idea of a 'persistent personal virome' in healthy individuals, dominated by tailed bacteriophages (class *Caudoviricetes*), including crAss-like bacteriophages (class *Crassvirales*)^{181,182}, whose association to human conditions remains currently ambiguous¹⁸³. While several phage groups have been associated to disease^{184,185}, challenges remain in annotating the unclassified majority of viral 'dark matter' sequences, and specifying the interaction of the virome with both the host and bacteria and their impact on the host and microbiome health^{150,186,187}.

Archaea (archaeome)

Despite archaea relying on identical sequencing tools to bacteria²⁵, they remain significantly underexplored in studies of the human microbiomes. Insights from studies using primer sets optimised to target archaeal species have shown that methanogens such as *Methanobrevibacter smithii* comprise the most prevalent archaeal species at mucosal sites, with considerable variation in abundance attributed to possible geographical and lifestyle differences^{25,188,189}.

Fungi (mycobiome)

Fungal communities have recently been established to be commensal microorganisms across various body sites within the human body²⁵. They are less diverse and abundant compared to bacteria and are dominated by *Saccharomyces*, *Malassezia* and *Candida* genera¹⁷⁴. Despite the relatively low genetic contribution of fungi to the microbiome, they could potentially provide a significant contribution to total microbial metabolic pathways and outputs, which can shape the structure of microbial communities¹⁵¹. However, more research on the topic is needed¹⁵¹.

Other eukaryotes (parasitome)

Characterised using 18S ribosomal RNA amplicon sequencing and more recently metagenomic approaches¹⁹⁰, other previously thought pathogenic eukaryotic species including helminths and protozoa, are now being studied as contributors to healthy microbial ecosystems²⁷. Within the healthy gastrointestinal tract, *Blastocystis* species appear to be prevalent colonisers, while helminths abundance is positively correlated with bacterial diversity, and any pathogenicity is likely subtype-dependent^{191,192}.

Box 2. Novel tools and considerations for population-scale microbiome research

Imitating the successes of human genomes exploration, 'citizen science' emerges as a crucial force in future large-scale data collection, as demonstrated by its current utilisation in large microbiome studies¹²⁰. Participants contribute to the laborious and financial burdens, benefitting from individual advantages of scientific discoveries. New and efficient sampling tools and digital technologies considerably support the application of citizen science, as detailed as follows.

Novel sampling tools

While increasingly used at-home sampling kits for the oral microbiome only minimally impact microbiome composition¹⁹³, faecal sampling proves more challenging and can be greatly impacted by collection method, sample transport and storage¹⁹⁴. The gold standard involves the immediate freezing of fresh samples, demanding additional effort and vigilance³¹. A novel approach allows the preservation of stool samples at room temperature using a preservative buffer¹⁹⁵. OMNIgene GUT is the only Food and Drug Administration-approved kit of this kind. However, buffer-based approaches need further verification as they impede later quantification, isolation and culturing. Besides facilitating sample collection, future research must ensure representative sampling across diverse body sites and niches. While sampling of the oral cavity or skin is considerably straightforward, non-invasive profiling of many regions of the gastrointestinal tract, for instance, the small intestine, remains a challenge¹²¹. As such, the predominantly employed stool samples are often seen as a proxy for the entire gastrointestinal tract yet fail to provide a full view of the spatial dynamics of gut microbial communities^{196,197}. Ingestible sampling devices like CapScan or SIMBA, capable of sampling intestinal niches while passing the digestive tract, present promising initial results.

Digital health technologies

Digital health technologies encompass a wide variety of devices and tools integrated into our everyday lives, including items such as wearable devices (for example, step trackers) or smartphone apps (for example, activity or diet trackers)¹⁹⁸. These technologies represent innovative ways of objective and directly quantifiable metadata collection. They can also substantially reduce recall bias by supplying detailed real-time information. The ZOE-PREDICT1 study¹⁹⁹, employing continuous blood sugar sensors, exemplifies the successful inclusion of digital health technologies. However, ethical concerns around data and consumer protection must be addressed before their application in large population studies.

ToC blurb

In this Perspective, Joos, Ross and colleagues reflect on current approaches and barriers to defining a 'healthy' microbiome and propose an updated framework to help interpret these dynamic ecosystems by adapting ideas from epidemiology to microbiome study design.