JGH Open Open Access An open access journal of gastroenterology and hepatology

JGHF Journal of Gastroenterology and Hepatology Foundation

doi:10.1002/jgh3.12902

REVIEW ARTICLE

Human gut microbiome: A primer for the clinician

Saurabh Kedia and Vineet Ahuja

Department of Gastroenterology, All India Institute of Medical Sciences, New Delhi, India

Key words
biological factors, marker gene, metagenomics, microbiome, technical factors.

Accepted for publication 1 April 2023.

Correspondence
Professor Vineet Ahuja, Department of Gastroenterology, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India.
Email: vineet.aiims@gmail.com

Declaration of conflict of interest: Prof. Vineet Ahuja is an Editorial Board member of *JGH Open* and a co-author of this article. To minimize bias, he was excluded from all editorial decision making related to the acceptance of this article for publication.
Author contribution: Vineet Ahuja contributed to the concept and design, supervision, and critical revision for important intellectual content. Saurabh Kedia contributed to drafting of the manuscript.
Financial support: This work was supported by the Indian Council of Medial Research: Center for Advanced Research and Excellence in Intestinal Diseases (Grant Number: 55/4/11/CARE-ID/2018-NCD-II) awarded to Prof. Vineet Ahuja.

Funding support: Indian Council of Medial Research: Center for Advanced Research and Excellence in intestinal diseases55/4/11/CARE-ID/2018-NCD-II

Abstract

The human host gets tremendously influenced by a genetically and phenotypically distinct and heterogeneous constellation of microbial species—the human microbiome—the gut being one of the most densely populated and characterized site for these organisms. Microbiome science has advanced rapidly, technically with respect to the analytical methods and biologically with respect to its mechanistic influence in health and disease states. A clinician conducting a microbiome study should be aware of the nuances related to microbiome research, especially with respect to the technical and biological factors that can influence the interpretation of research outcomes. Hence, this review is an attempt to detail these aspects of the human gut microbiome, with emphasis on its determinants in a healthy state.

Introduction

Humans are inhabited by a second genetically distinct organ, the gut microbiome, often called the "second metabolic organ".[1] The microbial cells and genes are believed to outnumber the human cells and genes by an order of 10 and 100, respectively, although two studies have challenged this proportion and have equated the number of bacterial and human cells.[2,3] The gut microbiome exerts enormous influence on human health,[4] including competition with pathogens through niche exclusion[5] and production of antimicrobial peptides, participating in metabolism and energy harvest with thousands of enzymes,[6] education/development of the immune system,[7] bile salt metabolism, synthesis of vitamins, neurotransmitters, and other metabolites, as well as xenobiotic degradation.[8] The human gut is inhabited by all three domains of life: Archaea, Bacteria, and Eukarya, although bacteria dominate the gut environment. The change in quality and quantity of the gut flora or dysbiosis has been associated with several inflammatory and metabolic disorders, including inflammatory bowel disease,[9,10] metabolic syndrome and obesity,[11] liver diseases, and neurological disorders.[12] However, one needs to understand the basics of the healthy human gut microbiome and its determinants before understanding and planning studies on the interaction of the gut microbiome and these disorders. This review describes the functional and structural characteristics of the human gut microbiome and discusses the factors related to its variability and determinants.

Vocabulary of human microbiome

The terms used to define various aspects of the human microbiome are used interchangeably and are quite confusing. We therefore start with a description of terminology related to microbiome so that the rest of this review can be better understood.[13,14] While α-diversity is a measure of microbiome diversity within the sample, β-diversity indicates the difference in taxonomic diversity between different samples. α-diversity represents the number and relative abundance of the different species; a high α-diversity is associated with a healthy gut microbiome, whereas a reduced α-diversity is associated with various disease states such as *Clostridioides Difficile* colitis, inflammatory bowel disease (IBD), metabolic syndrome, and liver disease. Samples with similar α-diversity might have different relative abundances of various taxa, and this difference is captured by β-diversity, represented by separate clustering of different samples (Fig. 1 and Table 1).

Analyzing the human microbiome

Interpreting the microbiome depends upon two major factors: technical factors related to sample collection, processing, and analysis; and biological factors related to age, genetics, diet, drugs, geography, and environment. Knowledge of the technical factors is equally important for a clinician for optimum planning and execution of a microbiome-related study.[14]

Technical factors

Types of samples for microbiome study. There is significant variability between mucosal and fecal flora, which has important implications for any gut microbiome study.[15] The stool sample reflects the luminal microbiome composition, is easy to collect and useful for longitudinal studies where multiple samples are required at different time points, and can be easily obtained from healthy individuals. However, fecal samples are not exactly representative of the intestinal flora, do not reflect the physiological variations occurring across the intestinal chemical, nutrient, and oxygen gradient, and cannot capture the variability because of differences in intestinal segments. Endoscopic biopsy samples can be used to investigate the mucosal microbiota across different segments of the gastrointestinal (GI) tract, as they more closely reflect the intestinal microbiota composition. However, it is prone to bias because of bowel preparation, contamination due to GI luminal fluid in the biopsy channel, lack of sufficient material for multi-omics studies, and contamination with the host DNA. Because of the invasive nature of biopsy, these samples are not suitable for longitudinal studies and studies on healthy individuals. Multiple studies in healthy controls and across various disorders including irritable bowel syndrome (IBS), IBD, and colorectal cancer have demonstrated the variability between the two sampling locations,[16] with greater compositional difference in the fecal sample (between healthy controls [HCs] and disease) in IBS (as compared to mucosal) *versus* higher difference in mucosal sample in IBD (as compared to fecal).[17–21] Further, a study in HCs has demonstrated that fecal samples do not reflect the composition and metagenomic function of mucosa-associated microbiota distributed across multiple sites of the intestine.[22] The specific methodology and sample type related to gut microbiome studies in various intestinal diseases depend upon the research question (Table 2).

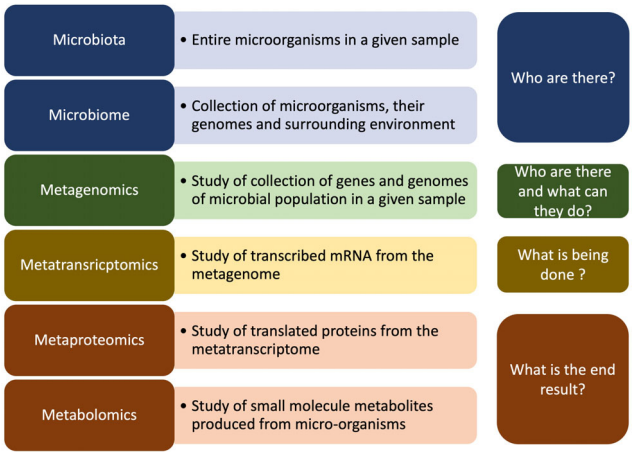

Figure 1 Multi-omics terminologies related to the analysis of gut microbial composition and function.

Collection, transport, processing, and storage of fecal samples

Collection and transport. The sample needs to be collected in a clean container or a clean plastic sheet over the toilet seat. The subsequent transfer to the laboratory depends upon the mode of storage and the study design.[23] The collected specimen needs to be transported to the laboratory as soon as possible, at room temperature within 4 h; if longer, then it should be done at 4°C within 24–48 h. With a DNA stabilizer solution (RNA later, 95% ethanol, Omnigene-Gut R), the sample can be transported over a longer time; however, preservatives can affect the metabolites in the fecal sample and such samples cannot be sub-sampled for culture or used for fecal microbiota transplantation (FMT).

Storage. In the laboratory, samples should be stored at −80°C (such samples remain viable for up to 2 years) (Fig. 2).[24] However, before storage, samples should be aliquoted as per the study design and requirement, as this prevents unnecessary freeze–thaw cycles, which are detrimental to the microbial content in the fecal sample. Sample consistency (as per the Bristol stool chart) also needs to be noted down before storage, but there is no consensus on fecal sample homogenization before aliquoting.

DNA extraction. It is one of the most important factors influencing gut microbiome composition. Available kits differ on their method of bacterial lysis (enzymatic, chemical, mechanical), with most utilizing a combination. Different DNA extraction protocols introduce an inherent bias, thus generating variable results, and hence it is recommended to employ a uniform protocol for the study.[25]

Sequencing and bioinformatics. Purified DNA after extraction is sequenced on various types of platforms and further channelized through various bioinformatic pipelines to get a meaningful output in decoding the gut microbiome.[26] Of the different sequencing platforms available, Illumina-based Miseq is the most commonly used platform, although other platforms,

Table 1 Vocabulary used for describing the microbiome composition and that used during microbial analysis

Measures	Definitions
α-Diversity	Measure of variability of the microbiome diversity within a sample, reflected by the richness (number of species within a sample) and evenness (relative abundance of different species within a sample) of bacterial species. Various indices for α-diversity include • Shannon index: Measures both richness and evenness with more weightage on richness • Simpson index: Measures both richness and evenness with more weightage on evenness • Chao: Nonparametric measure of species richness, giving more weightage to low-abundance species • Abundance-based coverage estimator: Nonparametric measure of species richness • Faith's phylogenic diversity: Measure that also incorporates phylogenetic difference between the species
β-Diversity	Represents the difference in taxonomic diversity between different samples, and can be expressed with (weighted) or without (unweighted) considering the relative abundance of individual taxa • Non-phylogeny-based: Takes into account abundance of various taxa within samples being compared, for example, Bray–Curtis, Euclidean, or Jaccard distance matrices • Phylogeny-based: Considers relative phylogenetic distances between various taxa and also ○ Unweighted Unifrac, which considers only the presence or absence of taxa ○ Weighted Unifrac, which also takes into account the relative abundance information of various taxa
Operational taxonomic unit (OTU)	DNA sequences with a definite level of similarity (>95%, 97%, or 99%)
Amplicon	Target sequence or gene that is amplified
Amplicon sequence variants	Refers to single DNA sequences recovered from a high-throughput marker gene analysis. Provides finer sequence resolution than OTU-based analysis (at the level of single nucleotide change) and is more reproducible, precise, and comprehensive.
Assembly	Alignment and merger of short DNA sequences to form longer DNA fragments
Contig	Contiguous DNA sequences formed by assembly of short DNA fragments
Scaffold	Longer continuous DNA sequences formed by assembly of contigs
Binning	Grouping of DNA sequences or contigs on the basis of their similarity with further assignment into taxa
Annotation	Assignment of functional categories to genes/sequences by mapping to reference genomes
Core microbiome	Microbiome that is present in a definite percentage (50–80%) of the population at a fixed level of abundance (0.01–0.1%)

such as Ion Torrent PGM, Pacific Biosciences, NanoPore, Roche 454 GS, and FLX Titanium, are also available.[27] Importantly, for a given study, the same sequencing platform should be used for all the samples. Though the advent of next-generation high-throughput sequencing has revolutionized the sample analysis workflow, the large amount of data produced poses significant analytical challenges, which, however, have been mitigated through advancements in bioinformatics platforms and algorithms. Different commercial and open-source bioinformatic pipelines are available, which differ in their statistical approaches, computational requirements, data handling, and classification accuracy. The classification algorithm can be composition- or comparison-based depending upon whether they compare sequence features (such as GC content) or homology-based sequences to a reference database. BLAST homology search is one of the commonly used comparison-based method, but recently hybrid methodologies that combine both approaches have been employed, one of the popularly used tools being MetaPhlAn2. Reference databases are equally important and are used for mapping the sample reads to known sequences for taxonomic classification. Several such databases exist, such as SILVA, Greengenes, and the Ribosomal Database Product (RDP)[26] (Fig. 3).

Sample analysis. The gut microbiome composition can be evaluated either by culturing the organisms or by molecular techniques that either detect specific sequences (marker gene analysis) or sequence the entire genome (metagenomics) (Table 3 and Fig. 3).

Culture-based techniques. The advent of molecular techniques in the late 20th and early 21st century phased out anaerobic culture techniques[28] as most species remained uncultured. Although culture-based techniques detect the "real" organism and can inform about the exact abundance of a particular species, they are limited by poor sensitivity and can define only 20–40% of the intestinal bacterial community.

Molecular techniques

- Marker gene analysis (Table 3)

Marker gene analysis is based on the detection of a specific sequence or gene (marker) of an organism, which is present in all organisms of the same type and different from others. 16S rRNA-based sequencing (for bacteria) has revealed more complex gut microbiota than culture-based techniques, with newer sequences and a larger number of species, most of them being assigned to three major phylogenetic linkages (*Bacteroides, Clostridium coccoides*, and *Clostridium leptum* groups).[29] Marker gene sequencing began with Sanger sequencing, which is a labor-intensive, slow, and expensive technique, which gradually was replaced with next-generation sequencing

Table 2 Specific methodology for microbiome studies related to inflammatory bowel disease, irritable bowel syndrome, and colorectal cancer[†]

Disease type	Questions that need to be answered	Study design	Type of sample	Marker gene *versus* metagenomics	Functional studies
Inflammatory bowel disease	Identification of disease-associated microbiota	Cross-sectional	Fecal; Mucosal if region-specific difference and host bacterial interaction need to be studied	Both can be done with fecal sample; Marker gene for mucosal sample	For identification of differences in bacterial gene expression and metabolites produced
	Development of biomarker for disease development	Longitudinal with healthy asymptomatic individuals at risk for disease development	Fecal	Both can be done	To describe specific functional feature associated with disease development
	Development of biomarker for predicting outcomes and response to treatment	Longitudinal in diseased individuals	Fecal	Both can be done	To describe specific functional feature associated with disease variability
	Mechanistic studies to evaluate effect of microbiota on host physiology	Gnotobiotic mice models Patient-derived organoid models to evaluate interaction of microbiota/metabolites			
Irritable bowel syndrome	Identification of disease-associated microbiota	Cross-sectional	Fecal; Mucosal if region-specific difference and host bacterial interaction need to be studied (e.g. small intestinal for patients with IBS)	Both can be done with fecal sample. Marker gene with mucosal sample	For identification of differences in bacterial gene expression and metabolites produced
	Microbiota stability between IBS and controls	Longitudinal	Fecal	Both	
	Mechanistic studies	Germ-free and gnotobiotic mice models, manipulation of gut microbes and microbial metabolites related to IBS-associated physiology			
Colorectal cancer (CRC)	Development of biomarkers for risk stratification and CRC screening	Longitudinal	Fecal	Both depending upon the depth of sequencing queried	Not required
	Detection of adenomas/CRC during surveillance	Longitudinal	Fecal	Both depending upon the depth of sequencing queried	Not required
	Identification of bacteria associated with adenoma formation and CRC	Cross-sectional	Mucosal	Marker gene	Metabolomics may reveal functional differences between normal tissue and CRC

[†]Other important aspects for well-conducted microbiome study: Robust clinical metadata, uniform sample collection and bioinformatics pipeline, rigorous statistical testing, power calculation, and correction for multiple hypothesis, and adjustment for other covariates such as diet, drugs, ethnicity, and environment.

(NGS), a cost-effective technique with massive parallel sequencing throughput brought about by amplification of 16S rRNA genes using primers containing adaptors. Marker gene analysis is convenient for taxonomic classification, easy to perform and interpret, and cheaper than metagenomics. However, it is not appropriate for functional analysis, provides low resolution at the species level, and is difficult to detect low-abundant taxa.

- Metagenomics (Table 3)

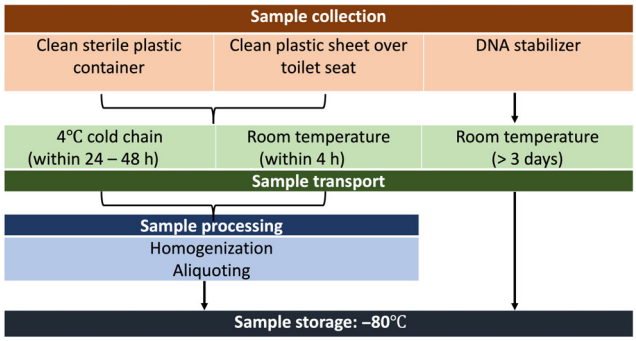

Figure 2 Means of fecal sample collection, transport, processing, and storage for microbial analysis.

Shotgun sequencing was further developed to sequence the entire bacterial DNA (metagenomics) and involves the random sequencing (e.g. whole-genome shotgun sequencing) of the total extracted bacterial DNA and then matching the sequences with previously annotated genes and pathways for taxonomic and functional analysis.[30]

Metagenomics provides higher resolution than marker gene analysis and its functionality can be predicted, but it requires high levels of expertise, computational overheads, and sequencing costs.

Culturomics. Though the gut bacterial diversity was revealed largely through metagenomics, most of these bacteria remained uncultured until culturing techniques resurfaced through "culturomics." Described initially by environmental microbiologists, culturomics incorporates multiple culture conditions, matrix-assisted laser desorption time-of-flight (MALDI-TOF) mass spectrometry, and 16S rRNA sequencing for the identification of bacterial species.[31]

Functional analysis. Functional potential or exact functionality can be assessed indirectly through the metagenomics approach (what can they do?), through the measurement of gene expression (metatranscriptomics, what is being done?), or through the quantification of the proteins (metaproteomics) or metabolites produced (metabolomics) (what is the end result?). Analysis of functional potential is possible either through metagenomic or 16SrRNA sequencing followed by function prediction through specialized pathways (PICRUSt, Table 3) against references databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) and the MetaCyc database, or functionally annotated orthologous groups such as eggNOG. Metatranscriptomics, through analysis of gene expression (RNA sequencing), informs about the active bacterial species present, providing more functionally relevant picture of the gut microbiome than metagenomics. Metaproteomics and metabolomics, through techniques such as liquid chromatography-mass spectroscopy (LC-MS), provide a picture of the actual phenotype by measuring what exactly is produced by the microbiota. Metabolomics is the most sensitive technique with respect to functional resolution of

Table 3 Methods of compositional and functional analysis of the gut microbiome

	Description	Advantages	Disadvantages
Marker gene analysis	Based on targeting an amplicon of one gene (marker: DNA sequence that identifies the genome that contains it without the need to identify the entire genome) that is present in every member of population, and is different between individuals with different genomes.	Convenient for taxonomic classification; Cheaper than metagenomics; Easy to perform and interpret	Low resolution at species level; Difficulty in detecting low-abundant taxa
• 16S rRNA	For bacterial analysis. Located in the 30S subunit of prokaryotic ribosome. Nine variable regions, each flanked by highly conserved DNA sequence, which provides primer sites for amplification		
• Internal transcribed spacer region	For fungal analysis		
• 18S rRNA	For fungal and parasitic analysis		
Meta-genomic shotgun sequencing	Total extracted DNA is fragmented and randomly sequenced. Reveals functions of encoded microbial DNA. Taxonomic classification is achieved through comparison with previously annotated genes	Higher resolution than marker gene analysis Functional analysis possible	Requires high levels of expertise, computational overheads, and high sequencing costs
PICRUSt	Phylogenetic Investigation of Communities Using Reconstruction of Unobserved States	Can assign functional pathways to 16srRNA-based genes, based on their mapping to previous databases	
Quantitative microbial profiling	Combination of flow cytometric bacterial cell counts with qPCR targeting the 16srRNA gene	Gives absolute counts of microbiome in a given sample, rather than relative abundance	

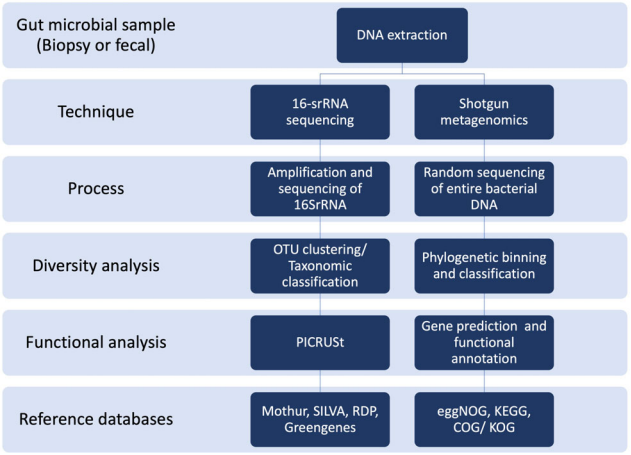

Figure 3 Comparison of analytical techniques for 16SrRNA analysis and shotgun metagenomics sequencing for gut microbial analysis.

the microbiome, although 90% of measured metabolite features may be unknown.[32]

Characteristics of healthy human gut microbiome

Composition of the human gut microbiome. The journey of the human intestinal microflora began with the landmark discovery of Theodor Escherich, who in 1885 described the properties of bacterial population in infant feces, termed "bacterium coli commune," currently known as *Escherichia coli*.[33] The approximate bacterial load in the human intestine (2×10^{11} bacterial cells per gram of feces) and the number of species (~100) were described initially by culture-based techniques, with *Bacteroides, Eubacterium, Clostridium*, and *Ruminococcus*, as the predominant genera.[34] Large-scale 16S rRNA-based analysis of mucosal and fecal flora revealed novel sequences belonging to archeal (*Methanobrevibacter Smithii*) and bacterial phylotypes (*Firmicutes, Bacteriodetes, Proteobacteria, Actinobacteria, Fusobacteria*, and *Verrucomicrobia* in decreasing order of abundance).[35] Metagenomic analysis further expanded this complexity, increasing the number of predicted bacterial genes to 0.5 million in the meta-HIT study of 124 Europeans.[30,36] Similarly, the taxonomic alignment of these genes into species or phylotypes increased to 1000. The composition matched previous reports, with >99% genes belonging to bacteria and archea, with Bacteriodes and Firmicutes having the highest abundance. Less than 0.1% was of eukaryotic or viral origin. The Human Microbiome Project (HMP), which commenced in 2007, characterized the microbiome of 242 screened and phenotyped North Americans, covering five major body areas (oral cavity and oropharynx, nasal cavity, skin, gut, and vagina). Oral and stool microbiota had the highest α- and β-diversities. The microbiota at each body habitat exhibited relationships with various driving physical factors such as oxygen, pH, moisture, host immunological factors, as well as inter-microbial relationships such as mutualism or competition. The recently published expanded dataset of HMP (whole metagenome sequencing of 1631 new samples) characterized new bacterial species, eukaryotes, archaea, and viruses, demonstrating the co-occurrence of nonbacterial organisms with bacterial species.[37,38] An integrated catalog of human gut microbiome was established in 2014 by combining 1267 samples across three continents, and covering strains with diverse occurrences, frequencies, and abundances[39] (Fig. S1).

Functional structure of the human gut microbiome. The gut microbiome is enriched in pathways for metabolism of plant polysaccharides (resulting in the production of short-chain fatty acids [SCFAs; acetate, propionate, butyrate] and gases), methanogenesis, synthesis of essential amino acids and vitamins, detoxification of xenobiotics, and deglucuronidation of bile salt.[6,40]

The gene catalog of the meta-HIT study has been mapped to ~19 000 different functions, which can be classified into the "minimal gut genome" (functions necessary for bacterial survival) and the "minimal gut metagenome" (involved in homeostasis of entire ecosystem: metabolism of plant polysaccharides and synthesis of amino acids and vitamins).[36]

Stability and variability of the human gut microbiome

Inter-individual variability and intra-individual stability. In contrast to more than 99.5% genetic similarity between different individuals, the gut microbiome of every individual is personalized, being significantly different from those of other individuals.[29] Though the percentage of *Bacteridetes* and *Firmicutes* per individual varies from 10% to 90%, their combined percentage remains at about 95%. In the meta-HIT study, even for the ubiquitous species (present in >90% individuals) the variability ranged from 12- to 2187-fold,[36] and in HMP, the within-subject variation over time was much smaller than the between-subject variation; similar findings have been replicated in the expanded HMP.[37,38] Further, several longitudinal studies including 2–37 individuals over a time period as long as 5 years have suggested the temporal stability of the gut microbiome.[41]

Species versus *functional variability.* In the expanded HMP, the species-level dynamics was more personalized than at the pathway level, with a greater between-subject species variability than functional variability.[38] Another study on six adult twin pairs and their mothers suggested a core microbiome at the gene level rather than taxonomy, and >93% functional characteristics were shared between individuals, who otherwise shared less than 0.5% bacterial species.[11]

Development and determinants of healthy microbiota

Aging

Infant microbiome. Though controversial, the fetus is considered sterile in-utero, and the infant acquires the microbiota during and after birth, gradually progressing to dense colonization with aging and reaching an adult-like configuration by 3 years of age.[42] The mode of delivery (vaginal *vs* caesarean) is the first determinant of an adult-like microbiome composition,[43] the other factors being the mode of feeding (breast *vs* formula feed), change in diet (milk-based to complex plant polysaccharides),[44,45] and

other environmental exposures such as antibiotics (Fig. S2). One important age-related change in the functional pathways involves vitamin biosynthesis, with the infant microbiome being enriched in *de novo* folate biosynthesis pathway and adult microbiome characterized by pathways metabolizing dietary folate and synthesis of vitamin B12, B7, and B1.[46]

Adult microbiome. Puberty is associated with a major physiological transition related to sexual maturation, which is reflected in the microbiome composition with divergence into a gender-specific one.[47] The influence of gut microbiome on the susceptibility to autoimmune disease in females was explained in a non-obese diabetic (NOD) mice study, which showed that the higher susceptibility of female than male mice to disease was diminished in germ-free conditions and that fecal transfer from male to female mice prior to disease onset protected against the development of disease.[47,48] Age-specific changes at this stage include a decline in aerobes and facultative anaerobes and an increase in anaerobes and microbial diversity. Adolescent microbiome is less complex and different from that of adults; the adult microbiome is most stable, complex, and resilient to change, with a core microbiome (present in >50% adults) at the functional level.[49] The adult microbiome is influenced by a multitude of factors including genetics, diet, geography, environment, and others, which are described in detail below.

Microbiome in the elderly. Transition to elderly-hood (>65 years) is associated with loss of physical function and decline in functional capacity of organs related to immunity (immunosenescence: loss of naïve CD4+ T cells and increase in the NFkB pathway), growth, metabolism, and energy homeostasis.[50] It is also reflected in the gut microbiome composition, as evidenced by the dominance of the phylum *Bacteriodetes* (57% population) as the core microbiome and reversal of the *Bacteriodetes* to *Firmicutes* ratio (Fig. S3) in comparison with the microbiome in a healthy young adult.[42,51] The microbiome has also been linked with diet (low-fat/high-fiber *vs* high-fat/low-fiber), residential pattern (community dwelling, outpatient, short-term rehabilitation, long-stay rehabilitation), and health status, with community-dwelling population harboring a significantly different and more diverse microbiome than people in long-stay care (Fig. S4).[52] A study examining centenarians (>100 years of age) revealed an interesting association between microbiota and longevity, with the microbial signature of these individuals being paradoxically associated with health-associated taxa such as *Bifidobacterium*, *Christensenellaceae*, and *Akkermansia*.[53]

Genetics

Twin studies. Small low-powered initial studies have demonstrated statistically insignificant greater similarity between monozygotic (MZ) than dizygotic (DZ) twins,[11,54] which was further confirmed by larger studies.[55,56] Taxonomically, the family *Christensenellaceae* (order Clostridiales) was the most inheritable.

Linkage studies. Linkage studies, by associating specific genetic polymorphism with the gut microbiome composition in disease or health states, have further strengthened host genetics–gut microbiome association.[57] IBD-specific loci in Crohn's disease, such as loss-of-function polymorphism in the FUT2 gene and the NOD2 risk allele, have been correlated with the modulation of energy metabolism in the gut microbiome and relative abundance of *Enterobacteriaceae*.[58] Further, the host quantitative trait loci (QTL) associated with genes related to immunity have linked them with the relative abundances of specific microbial taxa.[59,60]

GWAS studies. Three large genome-wide association studies (GWAS) have characterized the gene–microbiota association from three cohorts: Germany, Canada, and the Netherlands.[61–63] These studies linked several loci/single-nucleotide polymorphism (SNP) (9–53) with the relative abundance of microbial taxa and pathways, with the strongest association for C-type lectin molecules: LCT gene locus with *Bifidobacterium* genus, LINGO2 with *Blautia* and *VDR* (vitamin D receptor) gene. The overall contribution of host genetics to β-diversity in these studies ranged from 10.4% to 33%. Further, the analysis of human contamination reads in the data from HMP and expanded HMP[64,65] also correlated the taxonomic and functional composition of the gut microbiome with host genetic variation.

Metagenomics. Metagenomic analysis of 250 adult twins from the Twins UK Registry also demonstrated a greater degree of microbial SNP sharing in MZ than DZ twins.[66] Further, the SNP similarity between MZ twins decreased with decades of living apart, highlighting the impact of environmental influences. However, a recent genotype and microbiome (metagenomics and 16S rRNA gene sequencing) analysis demonstrated minimal contribution of the host genetics (only 1.9%) to the microbiome variability, challenging the concept of genetic influence on the gut microbiome.[67]

Diet.

Diet is one of the most important environmental factors that shapes the gut microbiome, the evidence being derived from both observational and interventional studies.

Observational studies. The diet–microbiome interaction starts as soon as the baby is colonized, with *Bifidobacterium* dominating the gut microbiome of breast-fed infants, and *Atopobium*, *Bacteriodetes*, and *Enterobacteriacae* being relatively abundant in formula-fed infants.[68] The influence of maternal antenatal and postnatal high-fat diet on the infant microbiome composition was also demonstrated in a recent study.[69] The dietary pattern (low-fat/high-fiber plant-based diet *vs* high-fat/low-fiber "Western diet" rich in animal proteins) has a major impact on the gut microbiome composition in adults, as demonstrated by a significant increase in bacterial diversity from carnivore to omnivore to herbivore.[70] Similar findings have been replicated in several studies that compared different dietary patterns across different populations from different or the same geography (Table 4).[46,71–79]

Enterotypes. The effect of long-term dietary pattern in a cohort of 39 individuals assigned the gut microbiome into three clusters or enterotypes (*Bacteroides, Prevotella*, and *Ruminococcus*) based on the dominance of specific bacterial taxa enriched for specific gene functions. The *Bacteroides* enterotype was associated with a Western-type diet high in proteins and fat, and the

Prevotella enterotype was associated with a diet rich in plant polysaccharides.[80] However, the enterotype concept was challenged subsequently[81,82] following a microbial survey of 200 individuals, showing minimal clustering into *Bacteriodes* and *Prevotella* enterotypes, and another survey showing a continuum of *Bacteroides* abundance across samples rather than distinct clustering.[83] The concept has been further revisited with the accumulation of data and re-analyses providing a balanced approach toward this understanding.[84] Additionally, with the advent of quantitative microbiome profiling (combining amplicon-based qPCR with flow cytometric enumeration of microbial cells), a 10-fold variation in the microbial loads of healthy individuals was observed, which was related to enterotype differentiation, with the identification of a low cell count of *Bacteroides* enterotype (Bact 2, characterized by low proportion of *Fecalibacterium* and high proportion of *Bacteroides*), and was correlated with systemic inflammation and disease states.[85,86]

Interventional studies. The effect of dietary interventions on the gut microbiome was demonstrated initially in mice studies, which documented a decline in the *Bacteroidetes*/*Firmicutes* ratio, increase in *Proteobacteria*, and a rapid shift in gut microbial composition and functional pathways on high-fat diet in normal and humanized mice (germ-free mice populated with human gut microbiota).[87,88] Further, a study of 5 genetically different inbred (wild, MyD88_/_, NOD2_/_, ob/ob_/_, Rag_/_) and >200 outbred mouse strains demonstrated a reproducible and reversible alteration in the gut microbiota on high-fat/high-sugar diet across all inbred mice strains independent of their genotype.[89]

Human studies on diet-induced weight loss in obese individuals (with high *Firmicutes* to *Bacteroidetes* ratio) have demonstrated improvement in the microbial gene richness, increase in the *Bacteriodetes*/*Firmicutes* ratio, reduction in butyrate-producing organisms, increased fecal branched-chain fatty acids, and decline in fecal SCFAs.[90,91] In another study, there was rapid (within 1 day of diet change), reproducible, and reversible alteration in the gut microbiome in response to rapid switches from an animal-based to a plant-based diet.[92] However, these effects of short-term dietary perturbations are short-lasting, and long-term dietary pattern is the major determinant of the gut microbiome, as demonstrated by Wu *et al*. who documented that the gut microbiome remained stable on short-term dietary perturbations (high-fat/low-fiber *vs* low-fat/high-fiber diet).[93]

Geography and environment. The effect of geography on the gut microbiome includes the effects of genetics and ethnicity, diet and lifestyle, and environment and culture, with diet and lifestyle being the most important. The influence of these factors on the gut microbiome has been assessed by analyzing three major subgroups: population resembling Paleolithic society represented by the hunter-gatherer population (primitive lifestyle with diet consisting of tubers, nuts, honey, wild game); population resembling Neolithic society represented by rural agricultural societies (thriving on cultivated crop, dairy, and domestic animals); and modern population represented by the urban towns of European and North American populations (on Western high-protein/high-fat/low-fiber refined diet). In general, the gut microbiome proceeds from highest diversity in the foraging population to lowest in urban population, with agricultural society falling in between (Table 4). Gupta *et al*. recently described the concept of geographically conserved core microbiome, which refers to the set of genera commonly found in a specific body site of all populations irrespective of their geographic location.[94] Regarding the gut microbiome, there are 25 genera that are common in all populations across 12 countries, although the relative abundance of these genera might vary. In a recent study, we investigated the gut microbiome of three healthy Indians communities residing at high and low altitude areas (urban and rural). The gut bacterial composition displayed specific signatures and was observed to be influenced by the topographic location and dietary intake of the individuals. The gut microbiome of individuals living at high altitudes was observed to be significantly similar, with a high representation of *Bacteroidetes* and a low abundance of *Proteobacteria*; in contrast, the gut microbiome of individuals living in low altitude areas harbored higher numbers of *Firmicutes* and *Proteobacteria* and was enriched with microbial xenobiotic degradation pathways.[79] The predicted functional diversity of high-altitude and low-altitude rural microbiome was higher than that of the low-altitude urban microbiome.

Other factors

Season and temperature. Seasonal changes in the gut microbiota occur primarily because of the dietary modifications, and have been documented mainly in populations that are dependent upon the environment for their diet.[95]

The effect of temperature was demonstrated in an elegant study, which showed that adaptation to cold temperature changed the microbiota composition to become sufficiently resistant to cold and resembled the microbiota configuration of obese/high-fat-diet-fed mice: high *Firmicutes* to *Bacteroidetes* ratio and low abundance of *Akkermansia muciniphila*, which increased the ability of the microbiota to harvest energy.[96]

Pregnancy. The effect of pregnancy on the gut microbiome was demonstrated in a study of 91 pregnant women, which characterized the gut microbiome in all trimesters. The microbiome of the first trimester resembled that of healthy, non-pregnant women, and the microbiome of third trimester was characterized by increased abundance of *Proteobacteria* and *Actinobacteria*, high inter-individual variability, and lower richness within each woman with respect to the first trimester.[97]

Drugs. In an elegant study, ciprofloxacin was found to decrease the overall diversity, richness, and evenness of bacterial composition in three individuals and impact the abundance of approximately one-third of bacterial taxa, which for some did not reverse even after 6 months of ciprofloxacin withdrawal.[98] The effect was even reproducible after a second course of ciprofloxacin.[99] Further, antibiotic use, through alterations in gut microbiome, can have long-lasting influences on the host physiology, as demonstrated by an increase in the incidence of allergic and inflammatory disorders such as asthma and IBD in children exposed to antibiotics[100] and mice studies documenting the transition of the gut microbiome toward an inflammatory phenotype, both in the parent mice and in fecal transfer experiments in the germ-free mice.[101,102] Drugs other than antibiotics can also impact the gut microbiome as evidenced a study on drug screen

Table 4 Effect of diet and geography on the gut microbiome across several populations

Author/year/children *versus* adults	Regions studied	Diets compared	Results	Overall remarks
De Filippo et al. (2010)/Children	Rural village in Burkina Faso	Vegetarian, rich in starch, fiber, plant polysachharide	Rich in Actinobacteria and Bacterodetes, genus *Prevotella*, *Xylanibacter*, *Treponema* More SCFAs	Overall higher bacterial richness in African than Italian population. Dominance of diet over geography
	Town in Italy	Rich in animal fat, protein, sugar and starch	Rich in Firmicutes, Proteobacteria, and Enterobacteriaceae genus *Allistepes*, *Bacteriodes*	
Yatsuneko et al. (2012)/Children *versus* adults	USA	Western diet rich in animal fat and proteins	Least diversity. Enriched in Bacteriodetes. Rich in glutamine and other amino acid degradation, simple sugar degradation, vitamin and lipoic acid biosynthesis, bile salt metabolism, protein export	Higher temporal instability of children than adult microbiomes Dominance of age, diet, and geography
	Amerindian population of Venezuela	Ancient diet: corn and cassava	Rich in Prevotella, Vitamin B2 biosynthetic pathway, starch degradation, glutamate synthase	
	Rural Malawian	Agricultural diet: maize, fruits, vegetables		
Ou et al. (2013)/ Adults	African-Americans	Rich in animal fat, protein, sugar, and starch	Predominance of Bacteriodes More gene encoding secondary bile acid production Higher fecal secondary bile acids	Dominance of diet
	Rural Africans	Vegetarian, rich in starch, fiber, plant, polysachharide	Higher diversity Predominance of Prevotella More butyrate producers Higher fecal SCFAs	
Schnorr et al. (2014)/ Adults	Hazda hunter-gatherers, Tanzania (Foragers)	Wild foods: meat, honey, baobab, berries and tubers, and game meat	Rich in Prevotella, Succinivibrio, Treponema and unclassified members of Bacteroidetes, Clostridiales and Ruminococcaceae, Proteobacteria. Absent Bifidobacterium	Higher diversity in Hazda than Italians Sex-related divergence due to difference in dietary composition Dominance of diet and gender
	Bolgona, Italy (modern lifestyle)	Mediterranean diet: abundant plant foods, fresh fruit, pasta, bread and olive oil. Low dairy, poultry, fish and red meat	Higher abundance of Bifidobacterium, Firmicutes (*Blautia*, *Ruminococcus*, and *Faecalibacterium*)	
Obregon Tito et al. (2014)/Adults	Matses: Hunter-gatherer population in Peru	Gathered tubers, invasive plantains, fish, low dairy product	Enriched for Proteobacteria Spirochaetes, Cyanobacteria, Tenericutes, and Euryarchaeota	Higher diversity in Mastes and Tunapoco than Oklahoma; Dominance of diet
	Tunapuco, a traditional agricultural community from the Andean highlands	Stem and root tubers, fruits, guinea pig, pork, lamb, and infrequently cow cheese, rice, and bread	Enriched for Proteobacteria, Spirochaetes, Bacteroidetes, Prevotella	

(*Continues*)

Table 4 (Continued)

Author/year/children *versus* adults	Regions studied	Diets compared	Results	Overall remarks
	Oklahoma, a typical US university community	Processed foods including canned fruits and vegetables, bread, and prepackaged meals	Enriched for Actinobacteria, Firmicutes, Bacteriodes, Ruminococcus, Blautia, Dorea	
Gomez et al. (2016)/ Adults	BaAka rainforest hunter-gatherers	Ancient hunter diet	Increased abundance of Prevotellaceae, Treponema, and Clostridiaceae. Increased abundance of predicted virulence, amino acid, and vitamin metabolism functions	Progressive change in microbiome diversity, composition, and function from hunter to agricultural to urban population
	Bantu neighbors	Agricultural diet	Enriched in Firmicutes and Bacteriodes. Intermediate abundance of Prevotella, Clostridiaceae, and Treponema	
	USA Americans	Western diet	Enriched in Bacteriodes Increased abundance of predictive carbohydrate and xenobiotic metabolic pathways	
Morton et al. (2015)/ Adults	Pygmy hunter-gatherers	Hunter ancient diet: Gathered tubers, invasive plantains, fish, low dairy product	Higher Proteobacteria Succinivibrio, Treponema, and Ruminobacter Lower Lachanospiracae	Presence of Entamoeba, location, subsistence, and ancestry as factors determining microbiome compositon. Parasites had the highest dominance
	Bantu farmers	Agricultural diet: grown cereals, vegetables, and meat	High Firmicutes, Ruminococcus, and Treponema	
	Bantu fishing population	Fishes, meat, dairy products	Lowest Bacteriodes and highest Prevotella and Bifidobacteria	Low Shigella and Escherichia in all three
Zimmer et al. (2012)/ Adults	Germany	Vegetarians *versus* vegans *versus* omnivores	Bacteroides, Bifidobacterium, *Escherichia coli*, Enterobacteriaceae low in vegans. Total microbial counts similar; stool pH lowest in vegans	Effect of diet seen
Wu et al. (2016)/ Adults	Urban USA	Vegans *versus* omnivores	Difference in plasma metabolome Similar fecal SCFAs No difference in gut microbiome	No effect of diet seen
Das et al. (2018)/ Adults	Ballabgarh rural	Predominantly vegetarian Cooking oil: ghee	High alpha and low beta diversity High Firmicutes and Proteobacteria, low Bacteroidetes High Parabacteroides, Blautia, Brevundimonas, Pelomonas, Megamonas, Collinsella	Effect of diet, cooking oil and geography Functional pathways High in Ballabgarh: membrane transport, carbohydrate metabolism, lipid metabolism, ion channels, and signal transduction and xenobiotic metabolism pathways
	Ballabgarh urban	Predominantly vegetarian Cooking oil: ghee	High α- and β-diversity High Firmicutes, low Bacteroidetes	

(*Continues*)

Table 4 (Continued)

Author/year/children *versus* adults	Regions studied	Diets compared	Results	Overall remarks
	Leh rural	Predominantly nonvegetarian Cooking oil: sunflower oil	High Lactobacillus, Bacteroides, Vibrio, Eggerthela and Pseudomonas, Collinsella Low α- and β-diversity High Bacteroidetes, low Proteobacteria High Prevotella, Roseburia, Faecalibacterium, and Lachanospiraceae	High in Leh: vitamin biosynthesis, energy metabolism and anti-inflammatory pathways

of >1000 marketed drugs, of which 24% drugs influenced the gut microbiome.[103] Proton pump inhibitors, because of their acid-suppressing properties, can favor the survival of oral bacteria in the distal segments of GI tract and can modify the flora of all GI segments.[104]

Human gut virome and mycobiome

As compared to "bacterial" microbiome, the human gut virome and mycobiome remain relatively unexplored. Though a detailed description on the characteristics of virome and mycobiome is out of the scope of this review, we provide an introductory primer on the characteristics of these microbial populations in the human gut.

The human gut virome consists primarily of bacteriophages and prophages along with a smaller proportion of eukaryotic viruses. The number of virus-like particles (VLPs) matches the number of bacterial cells. Viruses are most difficult to characterize because of the necessity of a eukaryotic or prokaryotic host, absence of conserved genes, and lack of matches in reference databases.[105] The first description of uncultured virome in human feces was published in 2003, and the majority of phage sequences were temperate phages.[106] Like the microbiome, the virome is unique for each individual, being influenced by diet and the environment, temporally stable, and dominated by *Caudavirales* and *Microviridae*.[107] Recent studies have also characterized the virome in Crohn's disease and ulcerative colitis,[108] and virome characteristics have also been correlated with response to microbial manipulation in *Clostridioides Difficile* colitis.[109]

Human gut "mycobiome" has primarily been explored through culture-based techniques and, recently, with the marker gene analysis, which targets the internal transcribed spacer (ITS) sequence in the fungus.[105] Mycobiome research still remains unclear on the standardization of analytical techniques including the fungal DNA extraction, sequencing, metagenomics (no metagenomic study on fungal composition to date), and bioinformatic pipelines including reference databases.[110] Fungal diversity is significantly less than bacterial diversity (10^5–10^6 fungal cells as compared to 10^{11}–10^{12} bacterial cells), although the fungal cells are 100-fold larger in volume than bacterial cells and fungal diversity is considered uniform across the GI tract.[3] Like for bacteria, a core fungal microbiome of 10 genera has been established, the major phyla being represented by Ascomycota and Basidimycota. Similarly, early life and dietary influences on the mycobiome have also been reported, and progress is being made in understanding the fungal–bacterial relationships, be it mutualism, commensalism, parasitism, or competition. The mycobiome also influences the gut immune system, and several IBD susceptibility genes have been involved in fungal recognition and response to fungi.[111]

Conclusion and future perspective

Lederberg in 2001 coined the term "microbiome", which was meant to include the collective genome of the resident microbes associated with any habitat in the human body, and the definition of healthy and dysbiotic microbiome was based on defining the characteristics of these resident microbes.[112] However, the microbiome census does not only include taxomony or "who is present", but also the functional repertoire or the annotated genes ("what can be done"), the expressed genes or RNA analysis (transcriptomics; "what is being done"), and the synthesized metabolites and proteins ("what is the end result"). This complexity surrounding microbiome analysis and significant variation (even in health states) across individuals, populations, and geography make it difficult to establish a uniform definition of a healthy microbiome,[113] and there could be multiple healthy microbiome configurations instead of a perfectly healthy microbiome.[114] In terms of taxonomy, because of large inter-individual differences, the concept of a healthy core structural microbiome is gradually vanishing, together with the realization of the concept that healthy taxa are individual- and context-specific, such as Akkermansia, which is positively correlated with health in metabolic disorders but negatively in multiple sclerosis.[115] The diversity of the microbiome is better correlated with health states, and a highly diverse microbiome is more stable and resilient (capacity to return to homeostatic state in response to external influences) to perturbations, which further characterizes the healthy microbiome, an ecological state that remains temporally constant even after being disturbed by known and unknown factors.[116] Functionally, the microbiome is more similar between individuals, and there have been consistent functional associations with health and disease states. Further, the upcoming concept of microbial ecology, which has expanded the definition to include the host influence on the microbiome, necessitates the

need to incorporate the concepts of community ecology into the field of microbiome science.[117] It considers the host as the foundation species for the microbiome with its vast influence on the microbial habitat, nutrition, metabolism, and immune function.[118] The host shapes the microbial community toward the dominance of species that are beneficial for the host, the concept entertained as the "germ–organ theory",[119] the prime example being the maintenance of epithelial hypoxia thorough oxygen consumption via mitochondrial oxidative phosphorylation, which facilitates the dominance of anaerobes over facultative and obligate aerobes. Therefore, the definition of a heathy microbiome would include the beneficial microbial species and functions and the host component of epithelial hypoxia that maintains these beneficial microbes. Through the phenomenon of colonization resistance, the beneficial microbiome inhibits the harmful species through nutrient competition and production of antibacterial metabolites, and in this way provides a nonspecific immunity towards the pathogens, called "microbiota-nourishing immunity".[120]

Microbiome science has been advancing at a considerable pace, and so has been the advancement in our knowledge on what constitutes a healthy microbiome. Dissecting out these intricacies of microbial contributions to health and disease states would lead to novel strategies to manipulate the microbiome for disease prevention and therapy.

References

1 Nicholson JK, Holmes E, Kinross J *et al*. Host-gut microbiota metabolic interactions. *Science*. 2012; **336**: 1262–7.
2 Sender R, Fuchs S, Milo R. Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans. *Cell*. 2016; **164**: 337–40.
3 Sender R, Fuchs S, Milo R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.* 2016; **14**: e1002533.
4 Marchesi JR, Adams DH, Fava F *et al*. The gut microbiota and host health: a new clinical frontier. *Gut*. 2016; **65**: 330–9.
5 Bäumler AJ, Sperandio V. Interactions between the microbiota and pathogenic bacteria in the gut. *Nature*. 2016; **535**: 85–93.
6 Tremaroli V, Bäckhed F. Functional interactions between the gut microbiota and host metabolism. *Nature*. 2012; **489**: 242–9.
7 Belkaid Y, Hand TW. Role of the Microbiota in Immunity and Inflammation. *Cell*. 2014; **157**: 121–41.
8 Koppel N, Maini Rekdal V, Balskus EP. Chemical transformation of xenobiotics by the human gut microbiota. *Science*. 2017; **356**: eaag2770.
9 Gevers D, Kugathasan S, Denson LA *et al*. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe*. 2014; **15**: 382–92.
10 Ahuja V. Inventory of a reservoir: friends & foes. *Indian J. Med. Res.* 2015; **142**: 4–6.
11 Turnbaugh PJ, Hamady M, Yatsunenko T *et al*. A core gut microbiome in obese and lean twins. *Nature*. 2009; **457**: 480–4.
12 Davis BC, Bajaj JS. The human gut microbiome in liver diseases. *Semin. Liver Dis.* 2017; **37**: 128–40.
13 Goodrich JK, Di Rienzi SC, Poole AC *et al*. Conducting a microbiome study. *Cell*. 2014; **158**: 250–62.
14 Allaband C, McDonald D, Vázquez-Baeza Y *et al*. Microbiome 101: studying, analyzing, and interpreting gut microbiome data for clinicians. *Clin. Gastroenterol. Hepatol.* 2019; **17**: 218–30.
15 Tang Q, Jin G, Wang G *et al*. Current sampling methods for gut microbiota: a call for more precise devices. *Front. Cell. Infect. Microbiol.* 2020; **10**: 151.
16 Ringel Y, Maharshak N, Ringel-Kulka T, Wolber EA, Sartor RB, Carroll IM. High throughput sequencing reveals distinct microbial populations within the mucosal and luminal niches in healthy individuals. *Gut Microbes*. 2015; **6**: 173–81.
17 Lo Presti A, Zorzi F, Del Chierico F *et al*. Fecal and mucosal microbiota profiling in irritable bowel syndrome and inflammatory bowel disease. *Front. Microbiol.* 2019; **10**: 1655.
18 Tap J, Derrien M, Törnblom H *et al*. Identification of an intestinal microbiota signature associated with severity of irritable bowel syndrome. *Gastroenterology*. 2017; **152**: 111–123.e8.
19 Rezasoltani S, Dabiri H, Asadzadeh-Aghdaei H *et al*. The gut microflora assay in patients with colorectal cancer: in feces or tissue samples? *Iran J. Microbiol.* 2019; **11**: 1–6.
20 Kumari R, Ahuja V, Paul J. Fluctuations in butyrate-producing bacteria in ulcerative colitis patients of North India. *World J. Gastroenterol.* 2013; **19**: 3404–14.
21 Verma R, Verma AK, Ahuja V, Paul J. Real-time analysis of mucosal flora in patients with inflammatory bowel disease in India. *J. Clin. Microbiol.* 2010; **48**: 4279–82.
22 Zmora N, Zilberman-Schapira G, Suez J *et al*. Personalized gut mucosal colonization resistance to empiric probiotics is associated with unique host and microbiome features. *Cell*. 2018; **174**: 1388–1405.e21.
23 Wu W-K, Chen C-C, Panyod S *et al*. Optimization of fecal sample processing for microbiome study—The journey from bathroom to bench. *J. Formos. Med. Assoc.* 2019; **118**: 545–55.
24 Song SJ, Amir A, Metcalf JL *et al*. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. *mSystems*. 2016: **1**(3): e00021–16.
25 Videnska P, Smerkova K, Zwinsova B *et al*. Stool sampling and DNA isolation kits affect DNA quality and bacterial composition following 16S rRNA gene sequencing using MiSeq Illumina platform. *Sci. Rep.* 2019; **9**: 13837.
26 Hopson LM, Singleton SS, David JA *et al*. Bioinformatics and machine learning in gastrointestinal microbiome research and clinical application. *Prog. Mol. Biol. Transl. Sci.* 2020; **176**: 141–78.
27 Allali I, Arnold JW, Roach J *et al*. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiol.* 2017; **17**: 194.
28 Hungate RE. Studies on cellulose fermentation: I. The culture and physiology of an anaerobic cellulose-digesting bacterium. *J. Bacteriol.* 1944; **48**: 499–513.
29 Zoetendal EG, Akkermans ADL, De Vos WM. Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria. *Appl. Environ. Microbiol.* 1998; **64**: 3854–9.
30 Gill SR, Pop M, Deboy RT *et al*. Metagenomic analysis of the human distal gut microbiome. *Science*. 2006; **312**: 1355–9.
31 Bilen M, Dufour J-C, Lagier J-C *et al*. The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species. *Microbiome*. 2018; **6**: 94.
32 Heintz-Buschart A, Wilmes P. Human gut microbiome: function matters. *Trends Microbiol.* 2018; **26**: 563–74.
33 Hacker J, Blum-Oehler G. In appreciation of Theodor Escherich. *Nat. Rev. Microbiol.* 2007; **5**: 902–2.
34 Moore WE, Holdeman LV. Human fecal flora: the normal flora of 20 Japanese-Hawaiians. *Appl. Microbiol.* 1974; **27**: 961–79.
35 Eckburg PB, Bik EM, Bernstein CN *et al*. Diversity of the human intestinal microbial flora. *Science*. 2005; **308**: 1635–8.
36 Qin J, Li R, Raes J *et al*. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*. 2010; **464**: 59–65.
37 Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012; **486**: 207–14.

38 Lloyd-Price J, Mahurkar A, Rahnavard G *et al*. Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature*. 2017; **550**: 61–6.
39 Li J, Jia H, Cai X *et al*. An integrated catalog of reference genes in the human gut microbiome. *Nat. Biotechnol.* 2014; **32**: 834–41.
40 Backhed F. Host-bacterial mutualism in the human intestine. *Science*. 2005; **307**: 1915–20.
41 Faith JJ, Guruge JL, Charbonneau M *et al*. The long-term stability of the human gut microbiota. *Science*. 2013; **341**: 1237439.
42 Kundu P, Blacher E, Elinav E, Pettersson S. Our gut microbiome: the evolving inner self. *Cell*. 2017; **171**: 1481–93.
43 Dominguez-Bello MG, Costello EK, Contreras M *et al*. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci. U. S. A.* 2010; **107**: 11971–5.
44 Palmer C, Bik EM, DiGiulio DB *et al*. Development of the human infant intestinal microbiota. *PLoS Biol.* 2007; **5**: e177.
45 Bokulich NA, Chung J, Battaglia T *et al*. Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* 2016; **8**: 343ra82.
46 Yatsunenko T, Rey FE, Manary MJ *et al*. Human gut microbiome viewed across age and geography. *Nature*. 2012; **486**: 222–7.
47 Markle JGM, Frank DN, Mortin-Toth S *et al*. Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity. *Science*. 2013; **339**: 1084–8.
48 Yurkovetskiy L, Burrows M, Khan AA *et al*. Gender bias in autoimmunity is influenced by microbiota. *Immunity*. 2013; **39**: 400–12.
49 Greenhalgh K, Meyer KM, Aagaard KM, Wilmes P. The human gut microbiome in health: establishment and resilience of microbiota over a lifetime. *Environ. Microbiol.* 2016; **18**: 2103–16.
50 Cambier J. Immunosenescence: a problem of lymphopoiesis, homeostasis, microenvironment, and signaling. *Immunol. Rev.* 2005; **205**: 5–6.
51 O'Toole PW, Jeffery IB. Gut microbiota and aging. *Science*. 2015; **350**: 1214–5.
52 Claesson MJ, Jeffery IB, Conde S *et al*. Gut microbiota composition correlates with diet and health in the elderly. *Nature*. 2012; **488**: 178–84.
53 Biagi E, Nylund L, Candela M *et al*. Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. *PLoS One*. 2010; **5**: e10667.
54 Hansen EE, Lozupone CA, Rey FE *et al*. Pan-genome of the dominant human gut-associated archaeon, Methanobrevibacter smithii, studied in twins. *Proc. Natl. Acad. Sci. U. S. A.* 2011; **108**: 4599–606.
55 Goodrich JK, Waters JL, Poole AC *et al*. Human genetics shape the gut microbiome. *Cell*. 2014; **159**: 789–99.
56 Goodrich JK, Davenport ER, Beaumont M *et al*. Genetic Determinants of the Gut Microbiome in UK Twins. *Cell Host Microbe*. 2016; **19**: 731–43.
57 Kurilshikov A, Wijmenga C, Fu J, Zhernakova A. Host genetics and gut microbiome: challenges and perspectives. *Trends Immunol.* 2017; **38**: 633–47.
58 Knights D, Silverberg MS, Weersma RK *et al*. Complex host genetics influence the microbiome in inflammatory bowel disease. *Genome Med.* 2014; **6**: 107.
59 Benson AK, Kelly SA, Legge R *et al*. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proc. Natl. Acad. Sci. U. S. A.* 2010; **107**: 18933–8.
60 McKnite AM, Perez-Munoz ME, Lu L *et al*. Murine gut microbiota is defined by host genetics and modulates variation of metabolic traits. *PLoS One*. 2012; **7**: e39191.
61 Wang J, Thingholm LB, Skiecevičienė J *et al*. Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. *Nat. Genet.* 2016; **48**: 1396–406.
62 Turpin W, Espin-Garcia O, Xu W *et al*. Association of host genome with intestinal microbial composition in a large healthy cohort. *Nat. Genet.* 2016; **48**: 1413–7.
63 Bonder MJ, Kurilshikov A, Tigchelaar EF *et al*. The effect of host genetics on the gut microbiome. *Nat. Genet.* 2016; **48**: 1407–12.
64 Blekhman R, Goodrich JK, Huang K *et al*. Host genetic variation impacts microbiome composition across human body sites. *Genome Biol.* 2015; **16**: 191.
65 Kolde R, Franzosa EA, Rahnavard G *et al*. Host genetic variation and its microbiome interactions within the Human Microbiome Project. *Genome Med.* 2018; **10**: 6.
66 Xie H, Guo R, Zhong H *et al*. Shotgun metagenomics of 250 adult twins reveals genetic and environmental impacts on the gut microbiome. *Cell Syst*. 2016; **3**: 572–584.e3.
67 Rothschild D, Weissbrod O, Barkan E *et al*. Environment dominates over host genetics in shaping human gut microbiota. *Nature*. 2018; **555**: 210–5.
68 Voreades N, Kozil A, Weir TL. Diet and the development of the human intestinal microbiome. *Front. Microbiol.* 2014; **5**: 494.
69 Chu DM, Antony KM, Ma J *et al*. The early infant gut microbiome varies in association with a maternal high-fat diet. *Genome Med.* 2016; **8**: 8.
70 Muegge BD, Kuczynski J, Knights D *et al*. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science*. 2011; **332**: 970–4.
71 De Filippo C, Cavalieri D, Di Paola M *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.* 2010; **107**: 14691–6.
72 Zimmer J, Lange B, Frick J-S *et al*. A vegan or vegetarian diet substantially alters the human colonic faecal microbiota. *Eur. J. Clin. Nutr.* 2012; **66**: 53–60.
73 Ou J, Carbonero F, Zoetendal EG *et al*. Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *Am. J. Clin. Nutr.* 2013; **98**: 111–20.
74 Obregon-Tito AJ, Tito RY, Metcalf J *et al*. Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat. Commun.* 2015; **6**: 6505.
75 Schnorr SL, Candela M, Rampelli S *et al*. Gut microbiome of the Hadza hunter-gatherers. *Nat. Commun.* 2014; **5**: 3654.
76 Morton ER, Lynch J, Froment A *et al*. Variation in rural African gut microbiota is strongly correlated with colonization by entamoeba and subsistence. *PLoS Genet.* 2015; **11**: e1005658.
77 Gomez A, Petrzelkova KJ, Burns MB *et al*. Gut microbiome of coexisting BaAka pygmies and bantu reflects gradients of traditional subsistence patterns. *Cell Rep.* 2016; **14**: 2142–53.
78 Wu GD, Compher C, Chen EZ *et al*. Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. *Gut*. 2016; **65**: 63–72.
79 Das B, Ghosh TS, Kedia S *et al*. Analysis of the gut microbiome of rural and urban healthy Indians living in sea level and high altitude areas. *Sci. Rep.* 2018; **8**: 10104.
80 Arumugam M, Raes J, Pelletier E *et al*. Enterotypes of the human gut microbiome. *Nature*. 2011; **473**: 174–80.
81 Knights D, Ward TL, McKinlay CE *et al*. Rethinking "enterotypes". *Cell Host Microbe*. 2014; **16**: 433–7.
82 Cheng M, Ning K. Stereotypes About Enterotype: the Old and New Ideas. *Genomics Proteomics Bioinformatics*. 2019; **17**: 4–12.
83 Koren O, Knights D, Gonzalez A *et al*. A guide to enterotypes across the human body: meta-analysis of microbial community structures in human microbiome datasets. *PLoS Comput. Biol.* 2013; **9**: e1002863.

84 Costea PI, Hildebrand F, Arumugam M *et al*. Enterotypes in the landscape of gut microbial community composition. *Nat. Microbiol.* 2018; **3**: 8–16.
85 Vandeputte D, Kathagen G, D'hoe K *et al*. Quantitative microbiome profiling links gut community variation to microbial load. *Nature*. 2017; **551**: 507–11.
86 Vieira-Silva S, Falony G, Belda E *et al*. Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. *Nature*. 2020; **581**: 310–5.
87 Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med.* 2009; **1**: 6ra14.
88 Hildebrandt MA, Hoffmann C, Sherrill-Mix SA *et al*. High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology*. 2009; **137**: 1716–1724.e1-2.
89 Carmody RN, Gerber GK, Luevano JM *et al*. Diet dominates host genotype in shaping the murine gut microbiota. *Cell Host Microbe*. 2015; **17**: 72–84.
90 Cotillard A, Kennedy SP, Kong LC *et al*. Dietary intervention impact on gut microbial gene richness. *Nature*. 2013; **500**: 585–8.
91 Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. *Nature*. 2006; **444**: 1022–3.
92 David LA, Maurice CF, Carmody RN *et al*. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*. 2014; **505**: 559–63.
93 Wu GD, Chen J, Hoffmann C *et al*. Linking long-term dietary patterns with gut microbial enterotypes. *Science*. 2011; **334**: 105–8.
94 Gupta VK, Paul S, Dutta C. Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity. *Front. Microbiol.* 2017; **8**: 8.
95 Smits SA, Leach J, Sonnenburg ED *et al*. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science*. 2017; **357**: 802–6.
96 Chevalier C, Stojanović O, Colin DJ *et al*. Gut microbiota orchestrates energy homeostasis during cold. *Cell*. 2015; **163**: 1360–74.
97 Koren O, Goodrich JK, Cullender TC *et al*. Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell*. 2012; **150**: 470–80.
98 Dethlefsen L, Huse S, Sogin ML, Relman DA. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* 2008; **6**: e280.
99 Dethlefsen L, Relman DA. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci. U. S. A.* 2011; **108**: 4554–61.
100 Marra F, Marra CA, Richardson K *et al*. Antibiotic use in children is associated with increased risk of asthma. *Pediatrics*. 2009; **123**: 1003–10.
101 Cox LM, Yamanishi S, Sohn J *et al*. Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. *Cell*. 2014; **158**: 705–21.
102 Schulfer AF, Battaglia T, Alvarez Y *et al*. Intergenerational transfer of antibiotic-perturbed microbiota enhances colitis in susceptible mice. *Nat. Microbiol.* 2018; **3**: 234–42.
103 Maier L, Pruteanu M, Kuhn M *et al*. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature*. 2018; **555**: 623–8.
104 Bruno G, Zaccari P, Rocco G *et al*. Proton pump inhibitors and dysbiosis: current knowledge and aspects to be clarified. *World J. Gastroenterol.* 2019; **25**: 2706–19.
105 Norman JM, Handley SA, Virgin HW. Kingdom-agnostic metagenomics and the importance of complete characterization of enteric microbial communities. *Gastroenterology*. 2014; **146**: 1459–69.
106 Breitbart M, Hewson I, Felts B *et al*. Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol.* 2003; **185**: 6220–3.
107 Shkoporov AN, Clooney AG, Sutton TDS *et al*. The human gut virome is highly diverse, stable, and individual specific. *Cell Host Microbe*. 2019; **26**: 527–541.e5.
108 Norman JM, Handley SA, Baldridge MT *et al*. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell*. 2015; **160**: 447–60.
109 Zuo T, Wong SH, Lam K *et al*. Bacteriophage transfer during faecal microbiota transplantation in *Clostridium difficile* infection is associated with treatment outcome. *Gut*. 2018; **67**: 634–43.
110 Richard ML, Sokol H. The gut mycobiota: insights into analysis, environmental interactions and role in gastrointestinal diseases. *Nat. Rev. Gastroenterol. Hepatol.* 2019; **16**: 331–45.
111 Iliev ID, Leonardi I. Fungal dysbiosis: immunity and interactions at mucosal barriers. *Nat. Rev. Immunol.* 2017; **17**: 635–46.
112 Prescott SL. History of medicine: origin of the term microbiome and why it matters. *Hum. Microbiome J.* 2017; **4**: 24–5.
113 Eisenstein M. The hunt for a healthy microbiome. *Nature*. 2020; **577**: S6–8.
114 Shanahan F, Ghosh TS, O'Toole PW. The healthy microbiome-what is the definition of a healthy gut microbiome? *Gastroenterology*. 2021; **160**: 483–94.
115 Cani PD. Human gut microbiome: hopes, threats and promises. *Gut*. 2018; **67**: 1716–25.
116 Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. *Nature*. 2012; **489**: 220–30.
117 Gilbert JA, Lynch SV. Community ecology as a framework for human microbiome research. *Nat. Med.* 2019; **25**: 884–9.
118 Tiffany CR, Bäumler AJ. Dysbiosis: from fiction to function. *Am. J. Physiol. Gastrointest. Liver Physiol.* 2019; **317**: G602–8.
119 Byndloss MX, Bäumler AJ. The germ-organ theory of non-communicable diseases. *Nat. Rev. Microbiol.* 2018; **16**: 103–10.
120 Byndloss MX, Litvak Y, Bäumler AJ. Microbiota-nourishing immunity and its relevance for ulcerative colitis. *Inflamm. Bowel Dis.* 2019; **25**: 811–5.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's website:

Figure S1. Phylogenetic distribution of bacterial kingdom with the major phyla associated with human gut microbiome (representative example for each phylum and its corresponding class, order, family, genus and species is provided).

Figure S2. Determinants and succession of the infant gut microbiome.

Figure S3. Physiologic transition of the microbial composition and function from adulthood to elderly population.

Figure S4. Determinants of the microbial signature in the elderly population.