nutrients

MDPI

Review

The Core Human Microbiome: Does It Exist and How Can We Find It? A Critical Review of the Concept

Itai Sharon [1,2,*], Narciso Martín Quijada [3,4], Edoardo Pasolli [5,6], Marco Fabbrini [7,8], Francesco Vitali [9], Valeria Agamennone [10], Andreas Dötsch [11], Evelyne Selberherr [3], José Horacio Grau [12,13], Martin Meixner [12], Karsten Liere [12], Danilo Ercolini [5,6], Carlotta de Filippo [9], Giovanna Caderni [14], Patrizia Brigidi [8] and Silvia Turroni [7]

1 Migal-Galilee Research Institute, P.O. Box 831, Kiryat Shmona 11016, Israel
2 Faculty of Sciences and Technology, Tel-Hai Academic College, Upper Galilee 1220800, Israel
3 Unit of Food Microbiology, Institute of Food Safety, Food Technology and Veterinary Public Health, Department for Farm Animals and Veterinary Public Health, University of Veterinary Medicine Vienna, A-1210 Vienna, Austria; narciso.m.quijada@vetmeduni.ac.at (N.M.Q.); evelyne.selberherr@vetmeduni.ac.at (E.S.)
4 Austrian Competence Centre for Feed and Food Quality, Safety and Innovation, FFoQSI GmbH, A-3430 Tulln an der Donau, Austria
5 Department of Agricultural Sciences, Division of Microbiology, University of Naples Federico II, 80055 Portici, Italy; edoardo.pasolli@unina.it (E.P.); ercolini@unina.it (D.E.)
6 Task Force on Microbiome Studies, University of Naples Federico II, 80055 Portici, Italy
7 Unit of Microbiome Science and Biotechnology, Department of Pharmacy and Biotechnology, University of Bologna, 40126 Bologna, Italy; m.fabbrini@unibo.it (M.F.); silvia.turroni@unibo.it (S.T.)
8 Microbiomics Unit, Department of Medical and Surgical Sciences, University of Bologna, 40138 Bologna, Italy; patrizia.brigidi@unibo.it
9 Institute of Agricultural Biology and Biotechnology (IBBA), National Research Council (CNR), Via Moruzzi 1, 56124 Pisa, Italy; francesco.vitali@ibba.cnr.it (F.V.); carlotta.defilippo@ibba.cnr.it (C.d.F.)
10 Microbiology and Systems Biology, Netherlands Organization for Applied Scientific Research (TNO), Utrechtseweg 48, 3704 HE Zeist, The Netherlands; valeria.agamennone@tno.nl
11 Department of Physiology and Biochemistry of Nutrition, Max Rubner-Institut (MRI)-Federal Research Institute of Nutrition and Food, 76131 Karlsruhe, Germany; andreas.doetsch@mri.bund.de
12 Amedes Genetics, Amedes Medizinische Dienstleistungen GmbH, 10117 Berlin, Germany; josehoracio.graujipoulou@amedes-group.com (J.H.G.); martin.meixner@amedes-group.com (M.M.); karsten.liere@amedes-group.com (K.L.)
13 Center for Species Survival, Smithsonian Conservation Biology Institute, Washington, DC 20008, USA
14 NEUROFARBA Department, Pharmacology and Toxicology Section, University of Florence, Viale Pieraccini 6, 50139 Florence, Italy; giovanna.caderni@unifi.it
* Correspondence: itaish@migal.org.il

Abstract: The core microbiome, which refers to a set of consistent microbial features across populations, is of major interest in microbiome research and has been addressed by numerous studies. Understanding the core microbiome can help identify elements that lead to dysbiosis, and lead to treatments for microbiome-related health states. However, defining the core microbiome is a complex task at several levels. In this review, we consider the current state of core human microbiome research. We consider the knowledge that has been gained, the factors limiting our ability to achieve a reliable description of the core human microbiome, and the fields most likely to improve that ability. DNA sequencing technologies and the methods for analyzing metagenomics and amplicon data will most likely facilitate higher accuracy and resolution in describing the microbiome. However, more effort should be invested in characterizing the microbiome's interactions with its human host, including the immune system and nutrition. Other components of this holobiontic system should also be emphasized, such as fungi, protists, lower eukaryotes, viruses, and phages. Most importantly, a collaborative effort of experts in microbiology, nutrition, immunology, medicine, systems biology, bioinformatics, and machine learning is probably required to identify the traits of the core human microbiome.

check for updates

Citation: Sharon, I.; Quijada, N.M.; Pasolli, E.; Fabbrini, M.; Vitali, F.; Agamennone, V.; Dötsch, A.; Selberherr, E.; Grau, J.H.; Meixner, M.; et al. The Core Human Microbiome: Does It Exist and How Can We Find It? A Critical Review of the Concept. *Nutrients* **2022**, *14*, 2872. https://doi.org/10.3390/nu14142872

Academic Editor: Connie Weaver

Received: 1 June 2022
Accepted: 8 July 2022
Published: 13 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Keywords: core microbiome; healthy microbiome; immune system; diet; prokaryotes; eukaryotes; virome; gut; NGS sequencing; omics

1. Introduction

The importance of the human microbiome to our health has been demonstrated in numerous studies in recent years. Our microbiome is involved in various processes in our body including nutrient uptake [1,2], defense against pathogens [3], and the development of the immune system in newborns [4]. Microbes that inhabit our gut metabolize carbohydrates, proteins, vitamins, and other nutrients [5]. The human microbiome has been linked to different health conditions such as inflammatory bowel diseases (IBD), in which changes in the gut microbiome were observed in IBD patients, as well as mutations in genes related to the interactions between the immune system and the microbiome [6]. Links between the gut microbiome and the brain (the gut–brain axis) have been reported for multiple conditions such as depression and autism, and animal-based experiments show that fecal transplant can affect neuropsychiatric conditions [7]. While it is still not clear whether the human microbiome is causally related to these and other health conditions, it is clear that our microbiome is central to our health. The human microbiome has been the subject of significant works by thousands of researchers worldwide and multiple large national and international projects such as the Human Microbiome Project (HMP) [8,9] and MetaHIT [10].

Identifying a common set of stable components—a core microbiome—has been a major goal in human microbiome research for decades [9]. The main contribution of the core microbiome may be in the functions it provides to the human host. Further, dysbiosis may be caused by impaired functionality of the core microbiome in the holobiontic system, which leads to disease states [11]. From the ecological perspective, members of the core microbiome may not interact directly with the human host but contribute to the stability of the (core) microbiome. This can occur through metabolite processing, nutrient synthesis or in other ways. The human microbiome is highly individualized, with similarities observed between close family members [12,13]. At the individual level, the microbiome changes over time but certain species and strains may occupy the body for years [14,15]. Strains are transferred from mothers to their children during labor and later through breastfeeding [16]. Characterization of the core microbiome has been attempted repeatedly [12,13,17–20]. Nonetheless, it remains unclear whether patterns relating to this concept can be observed at the community composition, function, ecology, or some other level. Consequently, the core microbiome remains an elusive concept that is difficult to evaluate and investigate in microbiome research. In fact, it is difficult to assess whether the core microbiome even exists.

In this review, we consider the state of research on the core human microbiome. Other reviews on the topic focused on the concepts and terminology that should be used to define the core human microbiome in terms of common sets of taxa or functions [9,19,21–24]. Actual attempts to define the core and healthy human microbiome are based on genes and community compositions derived from 16S and metagenomics data [12,13,17,20]. In contrast, we consider the different information sources that can help achieve an evidence-based model of the core human microbiome. We assess the technological, methodological, and conceptual gaps that should be addressed to consider these levels of information. Specifically, we consider the present state and limitations of omics technologies and bioinformatics methods, and estimate the microbial and host dimensions that cannot currently be evaluated or that do not receive sufficient attention, and that may contribute to core human microbiome research.

2. Defining the Core Microbiome

The concept of the core human microbiome describes the stable components of the microbiome, both over time and between individuals. This contrasts with the transient aspects of the microbiome, which vary according to fluctuations in environmental and individual conditions such as diet, host genetics, and other, mainly exogenous factors. The stability of a microbiome is considered in its capacity to resist changes caused by a perturbation (resistance) and in the capacity to return to equilibrium after a perturbation (resilience) [23,24]. Stability is maintained at the level of the individual over a relatively short-term and over the long-term evolutionary scale at the population/species level, including changes in the microbiome during the evolution of the host species [18]. Early attempts to define the core microbiome focused on the set of genes present in the vast majority of humans [9]. However, the term has evolved as studies have offered different perspectives on the nature of the microbiome [25], including community composition and ecological parameters. For host–microbiome systems, the core microbiome is considered in terms of important microbes and functions for the host [26]. Any definition of a core microbiome should account for the diverse habitats for microbial life in the human body, which harbor their distinct microbiota [27]. Table 1 summarizes the main approaches for defining the core human microbiome. We focus on the two most commonly used approaches, community composition, and functional profile.

Table 1. Approaches for defining the core human microbiome.

Approach	Pros	Cons	Examples
Community composition: the core microbiome is described in terms of shared taxa	Relatively simple to implement; can be applied to amplicon studies	Common taxa are usually identified at high taxonomic levels only	[17,28–31]
Functional profile: definition relies on a set of common functions	Captures the contribution of the core human microbiome to the host and the community	It is difficult to distinguish between human-specific and broad core functions	[9,12,13,26]
Ecology: include taxon abundance, interactions, co-occurrence, and other community-level patterns	Can capture complex patterns in community structure; may be more realistic than community composition alone	Less clear which patterns should be considered; no standard methods and programs are available	[19]
Stability: consider factors that maintain community stability and resilience	Stability is a critical characteristic of the core microbiome that is not captured through community composition alone	Definition is vague; there are no widely accepted methods for evaluating stability and resilience	[32]

A **community-based** definition of the core microbiome searches for taxa consistently found in the studied host population. Accordingly, a core microbiome consists of members of the community that are common to multiple communities across different hosts. Other than the mere existence or nonexistence of specific taxa, features that can be considered include abundance, the occurrence of related taxa (as a proxy for functional redundancy), periodic re-occurrence (persistence), and correlation and anti-correlation of taxa (connectivity) [19]. A community-based definition of a core human microbiome assumes that its members contribute to the host's health, either directly, through functions beneficial to the host, or indirectly, through contribution to community stability. Of particular interest are keystone species that play crucial roles in the ecological structure of their microbial communities and that are therefore essential for maintaining the organization and functions of these microbiomes [33]. Because of the disproportionate effect of keystone species on microbial ecosystems, the loss of these species can dramatically change the ecological niches of a microbiome. This may lead to dysbiosis and, for host-microbe systems, affect the capacity to perform functions that are important for the host; the clinical implications may be significant [34].

A **function-based** description of the core microbiome focuses on the set of functions that are consistent across populations at the level of genes or pathways. This approach is potentially more realistic than a community-based approach because it considers that multiple species can sometimes fill the same niche and provide the same functions. Functional redundancy ensures that certain functions in a microbiome may be performed by different species [35]. For example, in the human gut, multiple microbial species are responsible for the degradation of complex carbohydrates and the production of specific metabolites, ensuring the stability of the microbiome [36]. In this case, a specific functional capacity rather than the presence of certain taxa could be the crucial element that defines the core.

To illustrate the challenges in defining the core microbiome, we consider the community and functional profiling of a relatively small dataset composed of 370 human and 277 animal samples from nine cohorts of fecal samples. These include healthy Westerners from the US (HMP phases 1, 2, and 3) [9,37] and Denmark (MetaHIT), individuals with IBD from Spain (MetaHIT) [28], and a group of hunter-gatherers and traditional Bantu agriculturalists from the Central Africa Republic [38]. As controls, we consider fecal samples from gorillas [38], mice [39], and chickens [40]. Community composition was evaluated using MetaPhlan3, and functional profiles were generated using HUMAnN3 [41] (see Supplementary Information for more details). Overall, 107 species were detected in at least 70% of at least one cohort (Figure 1a). As expected, abundant species profiles were more similar between the human cohorts than between humans and mice, gorillas, and chicken. An ordination plot highlights remarkable clustering of microbial composition in the function of the host species; Westernized and non-Westernized human populations are distinct. Moreover, human microbiomes tend to be more similar to the microbiomes of gorillas than to those of mice and chickens (Figure 1b). However, only a few species were typically identified in more than 90% of the samples within each cohort, and many of the highly prevalent species in one cohort were not as prevalent in other human cohorts (Figure 1c). Only one species, *Faecalibacterium prausnitzii*, was detected in at least 90% of the samples in all six human cohorts. Even this result may be questioned, because MetaPhlAn3 falsely identifies multiple *Faecalibacterium* species as *F. prausnitzii* and may therefore lead to false positive detection of this species [42]. These results show the inadequacy of a simplistic species-based approach for describing a community-based core microbiome. Considering the function-based approach, we identified 134 pathways shared among >90% of the samples in all the human cohorts, of which 85 were also shared with gorillas, mice, and chickens (Figure 1d and Supplementary Figure S1). Several studies identified a number of functions that are shared among many individuals as the possible core microbiome [9,12,13]. However, as many of these functions are also shared with other organisms, defining human-specific core microbial functions is difficult. Various genes that encode widespread functions may be unrelated to the interactions between the microbiome and its human host or to community stability, for example, housekeeping genes.

The difficulties in identifying a set of universal taxa or human-specific functions that can be regarded as a core human microbiome were also noted in previous studies on the topic [25]. These difficulties may arise because of the algorithms and databases available, or because of our focus on the genomic level instead of other, more relevant levels. It is also possible that our definitions of the core microbiome do not capture its essence. Next, we consider each of these aspects and try to evaluate their importance for the study of the core human microbiome. Table 2 summarizes the issues considered in this review. The rest of the paper is organized as follows. In Section 3, we consider the sources of data that are currently available; in Section 4, we describe the state of 16S and metagenomics bioinformatics and data sharing; in Section 5, we go beyond prokaryotic meta-omics to evaluate other fields that may contribute to our understanding of the core human microbiome; finally, in Sections 6 and 7 we discuss the path forward in core human microbiome research.

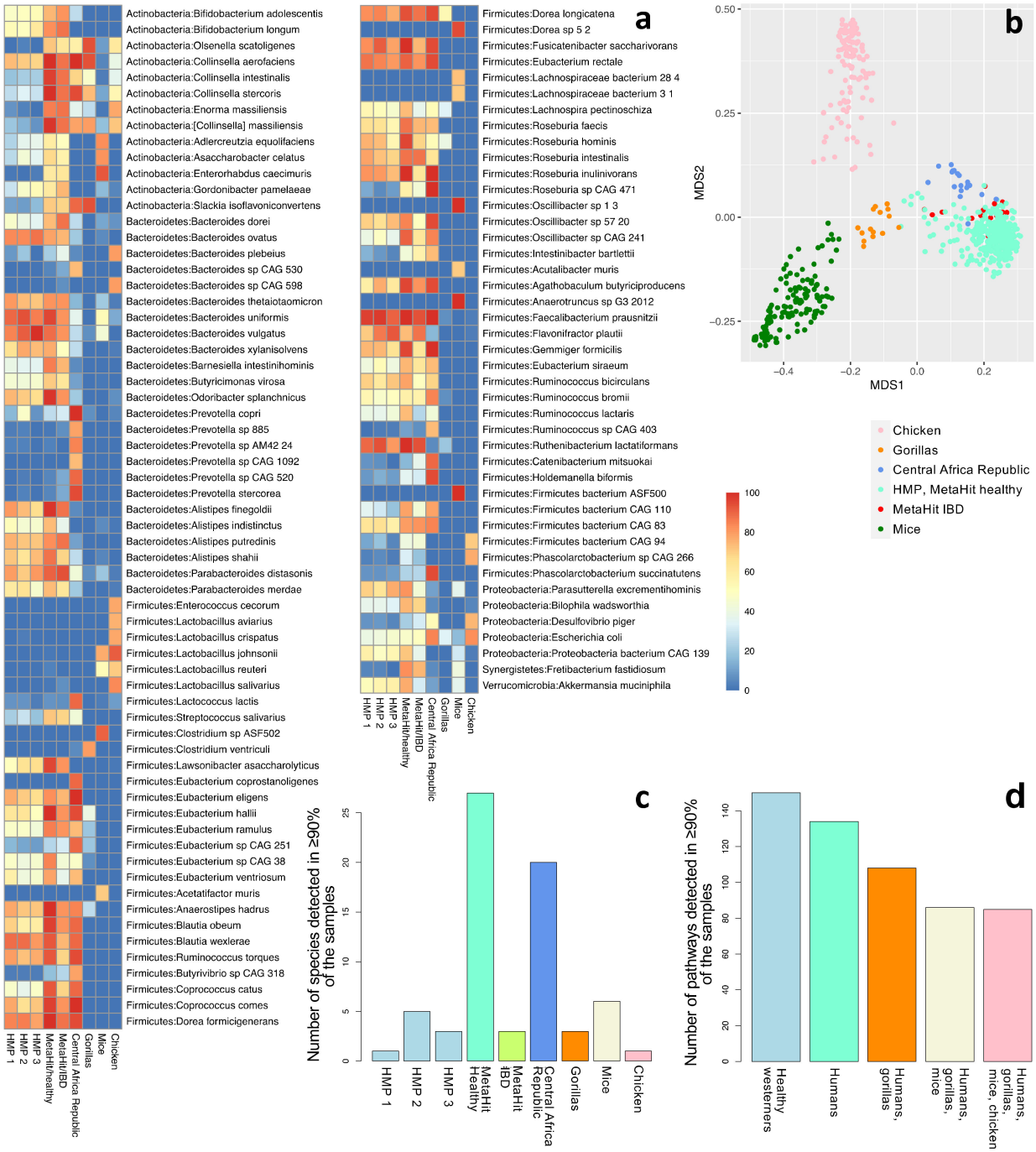

Figure 1. Community profiling of nine cohorts: HMP phases 1 (n = 138), 2 (n = 91), and 3 (n = 42); healthy individuals from Denmark (n = 64); individuals with IBD from Spain (n = 16); hunter-gatherers and traditional agriculturalists (n = 19); gorillas (n = 15); mice (n = 141); and chickens (n = 121). (**a**) The fraction of samples that contain each species, for species detected in at least 70% of the samples in at least one cohort (a total of 107 species). Refer to Supplementary Table S3 for a complete list of all the detected species in all cohorts. (**b**) The first two components of an unweighted UniFrac-based MDS analysis considering all samples. (**c**) The number of species detected in at least 90% of the samples of each cohort. Only one species (*F. prausnitzii*) was detected in 90% or more of the samples across all healthy human Western (n = 4) and all human (n = 6) datasets. (**d**) The number of pathways detected in >90% of the samples in all healthy Western human datasets (n = 4); all human datasets (n = 6); human and gorilla datasets (n = 7); human, gorilla, and mouse datasets (n = 8); and all datasets, including chickens (n = 9). See also Supplementary Tables S4 and S5.

Table 2. A summary of the challenges that affect our ability to characterize the core human microbiome.

Data (Section 3)	
Technology	Hundreds of thousands of human-associated 16S and metagenomics samples. Only thousands of other meta-omics samples
Population	Mostly Westerners. Agricultural and traditional populations are significantly underrepresented
Body part	Mostly fecal samples. The gut environment consists of multiple niches, each may have its own core microbiome
Methods (Section 4)	
16S rRNA surveys	Hundreds of thousands of public 16S samples are available. The method is inexpensive with established lab procedures and bioinformatics pipelines. The data provides only low taxonomic-resolution community composition and no functional information
Metagenomics bioinformatics	Tens of thousands of public human-associated metagenomes available. The data can provide strain-level and functional information. Bioinformatics analysis is complex, reference databases still lack a significant portion of human-associated microbial species.
Difficult to mine public databases	Available metadata is typically partial, performing meta-analysis requires significant manual effort
Beyond prokaryotic omics (Section 5)	
Diet Microbiome–immune system interactions	Both diet and the immune system shape the gut microbiome and are related to the functions it provides. Small amounts of data are available for both, with no high-throughput methods available for data collection
Non-prokaryotic members of the microbiome	Eukaryotes and viruses may be either part of the core human microbiome or related to the functions it provides. Both groups are invisible in 16S studies, current metagenomics bioinformatics mostly ignores them

3. The Data: Mostly Genomics, Western-Focused, and Fecal-Based

3.1. Technology: Mostly Genomics

Since the development of culturing-free techniques for studying microbial communities, microbiome research has shifted chiefly to relying on amplicon-based community surveys and metagenomics. As of March 2022, the European Nucleotide Archive (ENA) contains more than 610K human microbiome samples, of which more than 85% are estimated to be 16S rRNA amplicon samples (see Supplementary Information for for the analysis information). These, together with numerous samples available through other platforms, provide an amount of data that could potentially have been sufficient for evaluating the core microbiome of humans. The vast majority of available data was achieved using **amplicon-based community surveys**, which aim—through PCR and a specific set of primers—at a conserved region in the genomes of the target population. In microbiome surveys, one or more variable regions of the 16S rRNA gene are usually used to target prokaryotes. **Metagenomics** is based on untargeted ("shotgun") sequencing of potentially all microbial genomes present in a sample [43]. Metagenomics data can be used to assess community composition and functional potential, and even for the reconstruction in silico of the sequences from the original genomes that were present in the samples (also known as "metagenome-assembled genomes" or "MAGs"). Aspects of metagenomics and amplicon data analysis that limit our ability to define the core microbiome will be discussed later. Other omics approaches, such as metatranscriptomics, metaproteomics, and metabolomics, could have been used for characterizing the core human microbiome at levels that represent the actual functional activity of the microbiome. Examples of functions identified at the transcription level and not at the gene level are known [44]. Assuming that the core human microbiome provides specific functions to the human host, these functions, or the lack of them in dysbiosis and disease, may be manifested at the levels of transcription and translation, and not at the genomic level. Currently, only a small amount of omics data beyond the genomics level is available. For example, we were able to identify less than

1500 human microbiome-related RNA-seq samples in ENA, the majority of which have a shallow sequencing depth of <100 Mbp (see Supplementary Information for for the analysis information). To the best of our knowledge, no study to date addressed the core human microbiome topic at a level other than the genomic. Data from large-scale multi-omics projects such as ibdmdb [45] are expected to facilitate such studies.

3.2. Population: Western-Focused

The core human microbiome differs substantially between geographical locations and populations [46]. Differences are not restricted to the gut microbiome but are also observed in other body habitats such as saliva and the vagina. Defining the core human microbiome therefore requires sufficient data for different populations. However, According to a recent estimate, more than 70% of human microbiome data deposited in public databases have been collected from Westerners [47]. Large-scale national and international efforts to study the human microbiome have mostly focused on American and European populations. These include the Human Microbiome Project [8,27,48], the Dutch LifeLines [49], the Belgian Flemish Gut Flora Project [50], and the MetaHIT Project [10]. However, only a few studies evaluated the gut microbiome in rural communities in Western countries whose diet consists of more traditional products, including those obtained from household plots, yet still in a modern environment [51,52]. In addition, as mentioned above, several studies shed light on the peculiarities of the gut microbiome of non-Western populations, including hunter-gatherers and rural agriculturalists [53–61]. While their sample sizes were generally small, these studies consistently reported different composition and functionality of the intestinal microbiome between urbanized and rural populations, with diet, lifestyle, or more generally, the exposome (i.e., the totality of exposures that each individual faces in the course of life) [62] exerting a pivotal role [63–65]. Clearly, the human gut microbiome varies across geography and subsistence practices [66,67]. Therefore, more data for rural and non-Western populations are required to obtain a reliable description of the core human microbiome.

3.3. Sample Source: Mostly Fecal

The various segments of the gastrointestinal tract of mammals differ in anatomy and function; the microbiome composition of these segments also differs [68]. In addition to the differences between gut segments, within a single gut segment, microbial function and prevalence vary between luminal and mucosal bacteria [69–72]. Oxygen gradients have been found to correlate with radial partitioning of microbial phylotypes in the colon [72,73]. Still, most reports on the human gut microbiome rely on fecal samples to represent the gut microbiome [74]. Fecal material is easy to access and does not require invasive technology. However, fecal microbiota represents only the last part of the colon's luminal fraction; microbiota from other segments may elucidate microbe–host interactions and disease progression in the gut [71].

A certain core microbiome may fit the conditions of each specific microenvironment (niche) of the gut, with varying significance to human health. For example, microbes associated with mucosal sites were shown to have more significant effects on immunity and health parameters compared to luminal and fecal microbes [75,76]. Compared to the fecal microbiota, mucosal microbiota also better differentiate individuals with IBD from controls [77]. Dysbiosis, which leads to the loss of keystone community members in the mucosa may be more significant to identify. However, the mucosal microbiome cannot be evaluated according to fecal samples, as differences between the mucosal and fecal microbiome are evident in the healthy state, in IBD, and in irritable bowel syndrome [77–79], as well as in cirrhosis and hepatic encephalopathy [80].

If a core human microbiome exists, we may be able to characterize it in body parts that have fewer microenvironments, or in body parts that are more easily accessible. Compared to the gut, other parts of the human body have been sampled less extensively (see Supplementary Table S1). On the upper end of the digestive tract, the oral cavity

represents another environment, whose complex mixture of habitats comprises distinct microbiomes [17,81]. The mouth is much more accessible than the gut and was sampled at nine distinct sites in the HMP [27]. Moreover, specific databases of the human oral cavity exist, such as the extended Human Oral Microbiome Database. This database contains over 700 bacterial species isolated from the human mouth and aerodigestive tract, including the pharynx, nasal passages, sinuses, and esophagus, thus enabling more accurate sample-specific microbiome annotations. However, attempts to define a core human oral microbiome have focused on saliva [82–85], which also provides a restricted view of the overall diversity.

4. The Methods: Low-Resolution Amplicon Studies, Challenging Metagenomics Bioinformatics, and Difficult to Mine Public Databases

The analysis of microbial communities at the genomic level heavily relies on 16S rRNA gene surveys, which provide community composition information ("who is there?"), and metagenomics, which also provides functional information ("what can they do?"). Community analysis is performed in terms of ecological measures such as the alpha and beta diversity, lists of taxa and functions, and statistical methods for identifying features enriched in certain groups of samples. Genome reconstruction from metagenomics data helps expand our knowledge of species living in different environments and their potential metabolic and physiological features. Furthermore, 16S surveys and metagenomics are key techniques in defining the core human microbiome. However, both exhibit limitations that affect our ability to define the core human microbiome.

4.1. 16S rRNA Gene Surveys: Established Bioinformatics, and Low-Resolution and Partial Information

Amplicon-based studies are the most cost-effective option to investigate microbial populations, and several established computational pipelines for evaluating community composition are available [86–88]. Furthermore, the high sequencing depth achieved by the most popular Illumina platforms enables in-depth characterization of bacterial/archaeal communities present in an environment, including identification of low abundance microorganisms [89]. Additionally, 16S rRNA gene surveys are considered a semi-quantitative approach as, despite the intermediate PCR step during library preparation, the number of reads obtained for a given taxon is proportional to its abundance in the sample [90].

Several significant limitations of 16S rRNA gene surveys hinder their ability to evaluate the core human microbiome. First, 16S rRNA surveys only partially characterize the microbial community based on a set of selected primers. Eukaryotes and viruses are not represented, nor are lineages of archaea and bacteria that do not fit the primers used. Also, the taxonomic resolution of 16S rRNA gene surveys is low and does not exceed the genus level at best. This is due to the short length of the sequenced variable regions, typically at the level of genus at most [91], and precludes distinguishing between closely related species [92,93]. Strategies for improving resolution, including the use of multiple variable regions simultaneously [94], and the use of other marker genes [95], have been proposed but are not yet widely used. In addition, between one and 15 copies of the 16S rRNA gene (median = 5, according to the Ribosomal RNA Database, https://rrndb.umms.med.umich.edu/, accessed on 20 March 2022) are known to exist in a number of bacteria [96], and variability in the sequence of the copies within the same organism has been described [96]. This may lead to false evaluation of relative abundance and false identification of amplicon sequence variants (ASVs) or operational taxonomic units (OTUs) due to intra-organism variation in the gene's sequence [97]. Finally, it is unclear how OTUs and ASVs can be linked to taxonomic levels that facilitate a common language such as strains and species.

These issues, in addition to the lack of relevant functional information, limit the contribution of 16S rRNA gene surveys to the study of the core human microbiome. High-level taxa that consistently appear in various populations can be identified using 16S surveys. However, the current methodology cannot evaluate their potential contribution

to the core human microbiome, and the exact species and genes that contribute to the core microbiome.

4.2. Metagenomics: High-Resolution Taxonomic and Functional Data, Challenging Bioinformatics

While it is still not clear which taxonomic level (if any) should be considered for defining the core human microbiome, studying the microbiome at the species level is desired. Gene content of the microbiome is essential for applying a functional-based definition of the core microbiome. In theory, metagenomics facilitates characterizing the entire microbial community, including both community composition at the strain level and functional profiling, for both previously seen and unseen species [41,98]. Unlike 16S rRNA gene surveys, in metagenomics, the total DNA extracted from the sample is directly subjected to shotgun sequencing without a targeted PCR amplification step [99]. On the other hand, metagenomics is far more expensive than amplicon-sequencing surveys and is challenging to implement when the studied environment is rich with host DNA, as in mucosal-associated microbiome studies [100]. In addition, metagenomics data analysis requires significant computational resources, and knowledge of bioinformatics and programming is required for analyzing the data [101]. As a result, while the number of publicly available metagenomes is small compared to 16S rRNA gene surveys, analyzing these samples to evaluate the core microbiome is currently challenging.

Both taxonomic composition and the functional profile of metagenomics samples need to be evaluated to identify the core human microbiome. Assembly-free algorithms that rely on read-mapping of the raw data against reference databases can potentially accurately describe the community, including its rare members, with relatively low computational resources. Multiple such methods and pipelines are available, including MetaPhlAn3 [41], Kraken and Bracken [102,103], Centrifuge [104], and mOTUs2 [105], for species-level community composition, and HUMAnN3 [41] for functional profiling. The accuracy and sensitivity of these methods rely heavily on the availability of comprehensive reference databases [106]. Half or more of the species in the human microbiome have been estimated to lack representatives in public databases [67]. Moreover, up to 40–70% of the predicted proteins do not have a known function even for well-studied environments such as the human body [107]. Therefore, to close these gaps, increased effort is required in microbial culturing, biochemical investigations, and genome recovery for uncharacterized species. Assembly-based approaches that rely on de novo assembly of the reads into contiguous contigs and scaffolds can be used. Assembly, followed by gene prediction and the *binning* of contigs and scaffolds into MAGs that represent community genomes of microbial community members, can enrich our knowledge of microbes that evade culturing [108,109]. Indeed, hundreds of thousands of MAGs were reconstructed from human-associated metagenomes in recent years [67,110,111], contributing to the characterization of previously unknown lineages in the tree of life. Once incorporated into public databases, these MAGs facilitate large-scale and accurate species-level profiling of the human-associated metagenomes needed to evaluate the core microbiome.

4.3. Public Microbiome Datasets Are Often Hardly Interoperable and Reusable

Regardless of the adopted theoretical definition of the core human microbiome, its large-scale practical definition will inevitably be subjected to extensive cross-study meta-analysis. In defining the core human microbiome, data from a large number of individuals from several studies should be integrated and analyzed. This requires metadata such as ethnicity, geography, and disease state, and information regarding sample processing. Effective data reuse has been widely explored in all fields of science, especially from the advent of "omics" techniques, and as inspired and guided by the 2016 publication of the FAIR data (findability, accessibility, interoperability, and reuse of data) principle [112]. Unfortunately, performing meta-analyses is still challenging due to biases related to sample processing and the lack of metadata.

Sample processing involves multiple steps, which can introduce biases, such as in sample collection and storage, DNA extraction protocols, PCR, the choice of variable regions and primers (in 16S rRNA gene surveys), sequencing technology, and the bioinformatics pipeline used. The biases can range from eliminating taxa to reporting skewed relative abundances [113–116]. Therefore, all the information relevant to sample processing should be deposited with the data, to account for these issues and to select datasets that are comparable (i.e., have undergone similar processing steps). Unfortunately, this information is currently primarily available in the publication of the data, but usually not as part of the metadata deposited with the data.

Making data public does not guarantee their effective reuse. Journals and funding agencies usually require the upload of raw sequences to public databases; this constitutes a *de facto* obligation for the authors to make the data publicly accessible. However, the availability of raw data does not in itself ensure effective data reuse, and metadata (i.e., the diversificated set of data and information describing the subjects of a study) are required, but are often missing or incomplete [117]. For successful data reuse, a rich set of metadata that is standardized using ontologies should always accompany raw data. The ENA [118] and NCBI's Sequence Read Archive (SRA) [119] allow uploading a high volume of metadata as sample attributes, using well-defined checklists derived from templates and suggestions developed by international initiatives as the minimum information about any (x) sequence (MIxS) specification [120]. In addition, users are suggested to employ words derived from well-defined and standardized vocabularies/ontologies such as ENVO (Environment Ontology; [121]), CHEBI (Chemical Entities of Biological Interest; [122]), and DO (Disease Ontology; [123]), for the annotation of metadata of crucial samples. Examples of such data are the material feature that was the sample source (i.e., feces from the human gut environment), the reference to any chemicals used, and the pathology state. Still, the use of most of those metadata is not mandatory, and crucial information for effective meta-analysis, such as the host disease state and demographics, is often missing. We believe that adherence to a few guidelines considering metadata deposition and collection can greatly improve microbiome data reusability. However, given that core microbiome research can greatly benefit from meta-analyses of many datasets, the current state of public microbiome data poses obstacles.

5. Beyond Prokaryotic Omics: Interactions of the Core Human Microbiome with Its Host, and with Other Members of the Microbiome

5.1. Specific Effects of Diet on the Microbiome

Microbial communities comprise a multitude of species interrelated by the trade of metabolites, some acting as primary energy substrates whilst others (e.g., secondary metabolites, by-products, and wastes) as substrates for other members [124]. The stability of a colonizer is determined by a combination of biotic and abiotic factors; biotic factors are represented by interactions with other members of the community, and abiotic factors by the availability of nutrients or the presence of harmful compounds in the environment. Diet is considered the main determinant of the compositional and functional structures of the human oral [125,126] and gut [127–129] microbiomes.

Diet influences the oral and gut microbiomes beginning at birth, with human milk oligosaccharides involved in the assembly and maturation of the microbiota in early childhood [130,131]. During weaning, the introduction of solid foods leads to an increase in microbial richness and complexity, typically associated with adulthood [132,133]. Once reached, the typical microbiome configuration of the adult is relatively stable and resilient over time; yet characterized by alternating periods of stationary dynamics and fluctuations due to host and transient environmental factors, and also diet [15,24,134]. The gut microbiota, specifically, remains overall stable during adulthood and begins to change in older age, with gradual losses in diversity and compositional rearrangements. This is attributable to a progressively poorer diet (especially in fiber), as well as to an increased sedentary lifestyle, immunosenescence, and inflammaging [135,136].

Several studies have compared populations with disparate lifestyles, such as Westerners vs. residents from rural Venezuela [66] and Papua New Guinea [57], and ancestral populations still following the Paleolithic-type diet as the Hadza hunter-gatherers [54]. Significant differences in gut microbiome composition were found to be correlated with subsistence practices and geography [137].

Taken together, understanding the effects of specific nutrients on the microbiome is necessary for evaluating the core human microbiome. Different core functions may be required for different diets, making nutrition a potentially important factor in determining the core human microbiome. Several studies identified means by which various microbial groups respond to different diets. Generally speaking, a fiber-rich plant-derived diet is associated with enriched microorganisms endowed with fibrolytic capabilities, such as *Prevotella* from the Bacteroidetes phylum, and Lachnospiraceae and Ruminococcaceae members from Firmicutes [63,66]. These fiber-related specific signatures appear as clearly associated with healthy dietary patterns such as those associated with vegan or vegetarian diets [138] or the Mediterranean diet [139,140]. Conversely, Western high-protein high-fat diets are typically associated with bile-tolerant and fat-loving microbes (e.g., *Bacteroides*, *Bilophila*, *Collinsella*, etc.), as well as bacteria capable of utilizing host mucus, which serves as a backup food source when dietary fiber is limited or unavailable [127,141,142]. Non-digestible dietary carbohydrates, otherwise known as microbiota-accessible carbohydrates [141], are fermented by the colonic microbiota, thus producing short-chain fatty acids (SCFAs) [143]. These bioactive small molecules are strongly associated with a multitude of beneficial effects for the human host [144–146]. On the other hand, the gut microbiota can metabolize protein substrates into branched-chain fatty acids and phenolic and indole metabolites, which are variously associated with poor health outcomes [146]. A paradigmatic example of a maladaptive diet–microbiota–host interaction is represented by trimethylamine (TMA), which is formed by the microbiota starting from choline and carnitine present in cheese, eggs, red meat, and seafood. Once absorbed from the gut, TMA circulates to the liver where it is oxidized by host enzymes into TMA-N-oxide, a major cardiovascular risk factor [147]. Consistently, lower levels of TMA-N have been found in vegans and vegetarians than in omnivore subjects [139]. The most recent studies performed by metagenomics and the use of the up-to-date genome reconstruction pipelines mentioned above indicate that the diet affects the gut microbiome up to the strain-level dimension. Indeed, diet-sensitive and diet-associated microbial taxa that are widely studied in the human gut microbiome, such as *Prevotella* and *Faecalibacterium*, have demonstrated great genome-level variability according to the type of diet or the level of industrialization of the studied population [42,148,149]. Genome-level analysis can also be useful for tracing the possible occurrence in the gut of microorganisms that can be acquired through the consumption of certain foods such as lactic acid bacteria from fermented foods [42,150].

To the best of our knowledge, nutrition by itself is not usually considered an aspect of core human microbiome studies, except for defining populations that are assumed to have separate core microbiomes. Different diets require different functions from the microbiome, many of which must be fulfilled by the core microbiome. Considering nutrition information in core microbiome analysis is currently challenging, because accurate dietary information of individuals is usually not available, and the effects of specific nutrients on specific microbial species are generally not known. Given the high complexity of defining the activity of a certain dietary component (as a substrate) in each individual, such knowledge is unlikely to become available in the coming years for most nutrients and the microbes that process them.

5.2. The Immune System and Its Interactions with the Microbiome

The human body and our microbiome have co-evolved towards mutualism and homeostasis [151]. This coexistence is fundamental for the fine-tuning of the immune system and requires an intimate relation with commensal microbes, especially during early childhood [152,153]. Therefore, it is not surprising that alterations in these close microbiome–

immunity interactions are associated with increased susceptibility to pathogenic infections, overgrowth of commensals, and their potential spread throughout the body via impaired mucosal barriers [154,155]. These contribute to the onset and development of several disorders, including IBD [45], rheumatic arthritis [156], neurodegenerative diseases [157], and even tumors [158].

The crosstalk between the host's immune system and the microbiome is extremely complex and involves microbial structural components, e.g., lipopolysaccharide, lipoteichoic acid, microbiome-derived metabolites such as SCFAs, indole derivatives acting as aryl hydrocarbon receptor ligands [146], host-derived molecules including non-specific antimicrobial peptides [159,160], mucosal immunoglobulins IgA [161], and miRNAs that act over bacterial transcripts regulating their growth [162].

The hosts' local environments are sensed by commensal microorganisms to induce responses that enable their avoidance, escape, or tampering with the innate immune mechanism of defense [163,164]. Beyond the importance of adaptation to the environment, an array of microbial-derived molecules can promote commensal processes that are beneficial for both the host and the microbiome; and conversely, the host immune system recognizes the commensal antigen patterns and adapts to them. Microbial components are sensed by pattern-recognition receptors, including toll-like receptors and NOD-like receptors. These promote several mechanisms of tolerance, while preventing infections by harmful organisms [4,153].

SCFAs induce the production of anti-inflammatory cytokines, driving the local increase of Tregs [165,166]. In particular, butyrate acts as an inhibitor of histone deacetylase, epigenetically modulating immune system function, including Foxp3 gene expression [167]. Moreover, propionate has been shown to regulate allergic inflammation, bone marrow hematopoiesis, and dendritic cell function [168]. Several other microbiome-derived compounds, such as polyamines, can inhibit the macrophagic production of pro-inflammatory cytokines [169]. Microbial metabolic activity is sensed by the host immune system [170] and the information is converted into immune signaling pathways that can, in turn, stimulate anti-microbial activity or suppress the local immune response, to ultimately maintain the intimate cross-talk required for homeostasis.

Given its significant interactions with the human microbiome, the immune system is likely involved in defining the core microbiome. As in the example of nutrition, the interactions between the microbiome and the immune system are not taken into account in core microbiome studies. Obtaining accurate data that is related to the immune system may become a more feasible task, and high-throughput methods for characterizing the immune response to microbial species have already been successfully implemented [171]. Such data would enable a better evaluation of the core microbiome in the context of the immune system.

5.3. The Less Explored Members of the Microbiome

Eukaryotes, eukaryotic viruses, and phages are all part of the human microbiome but currently, relatively little is known about their contribution to the core human microbiome, both as members and in shaping the core microbiome.

5.3.1. The Role of Eukaryotes in the Human Microbiome

Eukaryotes, especially lower-order ones such as protists, fungi, stramenopiles, and helminths, belong to the human gut ecosystem [172,173]. In the past, any protist or helminth species identified in humans was considered a parasite with an obvious harmful effect on its host. However, a number of human-associated lower eukaryotes are now known to play various roles in both health and disease [174]. For example, some stramenopile species (*Blastocystis* sp.), commonly considered parasites, have been found to be associated with healthy and diverse gut microbiomes [175]; and some helminths and protists are able to stimulate a beneficial immune response in individuals with allergies [176]. Despite their active role in shaping the gut microbiome, unicellular and multicellular eukaryotes are

often neglected from microbiome studies. Some major limitations towards eukaryotes in microbiomes include: (i) a limited number of reference eukaryotic species with sequenced genomes, (ii) the common presentation of eukaryotic species at lower abundances than bacterial species and thus the likelihood that they are missed if the sequencing is not deep, (iii) the high diversity and poor taxonomic knowledge of many eukaryotes [177], especially in fungi and protists, for which accurate genus and species level are less determined [178], (iv) the need to elucidate the role of eukaryotes in maintaining eubiosis and homeostasis [179,180], and (v) the lack of standardized methodologies and sequencing designs [181]. Furthermore, eukaryotes are generally not identified with common bacterial 16S rDNA primers, and most universal 18S rDNA primers are limited in their ability to amplify a broad spectrum of eukaryotes. Nonetheless, specific primer combinations and standardization have proven useful sets for stool-based detection of gut eukaryotes [182]. In addition, approaches using RNA shotgun sequencing or "meta-ribosomalomics" of stool samples have successfully detected intestinal eukaryotes that have often been excluded from traditional routine diagnostic protocols [183,184].

Eukaryotes, such as the fungal genus *Candida* or *Rhodotorula*, are components of the core microbiome or contribute to its shaping. In Africa, intestinal helminths are common and are estimated to affect a large proportion of the population [185]. *Blastocystis* sp. were detected in 24% of the sampled population in the Czech Republic [186], suggesting that certain eukaryotes are also prevalent in Western countries. Current development of better bioinformatics algorithms, and databases with eukaryotic genomes, are expected to increase the feasibility of evaluating the contribution of eukaryotes to the core microbiome.

5.3.2. The Host Eukaryotic Virome and the Gut Microbiome

More than 1000 eukaryotic virus species are known that can enter and live within the human body, forming the human eukaryotic virome [187,188]. Human viruses can be present in the gut as residential members, as transient passengers, as commensals, and as an integral part of the microbiome community. Interactions between the eukaryotic microbiome and the other members of the gut microbiome can be direct or indirect, and either cooperative or antagonistic. Multiple examples of these complex interactions are known.

The presence of eukaryotic enteric viruses can lead to changes in the bacterial gut microbiome, as in the example of rotavirus infections, which can induce changes within the composition of the *Bacteroides* genus [189]. More recently, SARS-CoV-2 attachment to the colonic ACE2 receptor in the presence of *Coprobacillus* was reported to lead to a reduction in the 'good' gut bacteria *Faecalibacterium*, *Roseburia*, *Bacteroides,* and *Lachnospiraceae;* and to increased pathogenic genera such as *Streptococcus*, *Rothia*, and *Veillonella* [190]. Furthermore, these shifts during and even after SARS-CoV-2 infections give rise to opportunistic bacteria such as *Corynebacterium* and *Ruthenibacterium* [191,192]. Eukaryotic viruses can also activate the host immune system that drives the overall composition of the gut microbiome, as in the example of influenza viruses that invade the respiratory tract and induce a systemic type 1 IFN-1 response. This leads to a reduction in anaerobes and improvement of *Enterobacteriaceae* in the gut microbiome [193].

In contrast to the above, the gut microbiome can affect the infection processes of various eukaryotic enteric viruses in the gastrointestinal tract [194] and can induce an immune response that controls the infection process of eukaryotic viruses [195]. The enteric microbiota enhances the production of mucus and the synthesis of potential antiviral compounds, such as lactic acid, acetic acid, reactive oxygen species, and defensins, which inhibit local viral replication [196]. In addition, small bacterial proteins and bacteriocins, e.g., surfactin, a cyclic lipopeptide produced by *Bacillus subtilis*, can destroy eukaryotic virus integrity. This effect has been described for several enveloped viruses, including Chikungunya, Crimean–Congo hemorrhagic fever, Dugbe, Ebola, Zika, and coronaviruses [197].

The core microbiome may be involved in defending the human host against eukaryotic viruses. The stability of the core microbiome depends on resisting these viruses and the direct and indirect ways in which they affect the microbiome. The eukaryotic virome cannot

be studied through 16S rRNA gene surveys and is rarely addressed in metagenomics studies. Improving bioinformatics methods and reference databases should be a target for evaluating the roles and interactions of these viruses with the core microbiome.

5.3.3. The Prokaryotic Virome (Phageome) and Its Relation to the Gut Microbiome

The bacterial virome (phageome) of the gut is composed of free phages (virus-like particles) and prophages and is one of the most complex and poorly understood consortia in the context of the gut microbiome. New scientific work indicates that the pan gut bacterial virome is composed of more than 50,000 candidate viral species [198], and up to 140,000 different phages [199]. The number of bacteria in the microbiome and the number of virus-like particles (up to 10^{12} per gram of feces) are in about equal proportion [200]. Some large bacterial phages can add more than 500 kbp of genetic information to their host's genome [201].

Phages can be involved in the core microbiome in various ways. Phages that infect multiple species can be part of gene flow networks across phylogenetically distant bacterial species [202]. This can result in the distribution of core functions between multiple species and increase the stability of the core microbiome. Lysogenic phages can contribute new traits and physical/biochemical properties to their host [203]. As a result, certain core functions may be contributed by phages. Phages/prophages control bacterial cell death by indirectly gaining information of environmental factors such as community composition, nutrient availability, and local bacterial cell density via bacterial quorum sensing [202,204,205]. Phage-programmed bacterial cell death also leads to redistribution of nutrients [201] and, consequently, sometimes to remarkable shifts in the metabolome composition. To this end, phages can serve as both a threat to the core microbiome and as a means of eliminating transient members of the community. The exact roles of the gut phageome with respect to the core microbiome are poorly understood. Work carried out by researchers and data scientists towards developing models of core and healthy phageomes [206] may shed light on this.

6. Discussion

The core human microbiome has been the subject of multiple studies due to its potential significance in health and disease. Still, there is currently no consensus as to whether the core human microbiome should be defined at the level of taxa or functions, or some other level; the main forces that shape it, and the consequences of an impaired core microbiome. One major limitation in core microbiome research is the predominantly one-dimensional approach, which focuses on taxa and gene-based functions at the genomic level. Considering other data layers, both regarding the microbiome and the host, will likely help expose relations between these layers that are relevant to the roles and definitions of the core microbiome. Evidently, the core human microbiome should be discussed in the context of its environment. Such contexts include body parts, diet, geography, age, health state, and possibly other factors. Once the true nature of the core human microbiome is understood, accurate definitions and terminology can be developed that will help to uncover the relevance of the core human microbiome to biological, ecological, and evolutionary processes that occur in the human host and its microbiome.

Given the efforts invested and the amount of available data, the core human microbiome is unlikely to be defined by a fixed set of taxa, as no such consistent set has been found. A set of universal functions shared among humans may be identified; however, many of these functions are also shared by other organisms, including non-mammals. A set of functions likely exists that distinguishes our microbiome as "human", but more research to characterize those functions is required. Missing annotations to a large proportion of known and newly discovered genes is an obstacle that may be overcome by homology-based clustering of all known genes and proteins, regardless of their annotation. Incorporating the hundreds of thousands of MAGs published in recent years into public reference databases is expected to contribute to the establishment of a comprehensive set

of genes associated with the microbiome of the gut and other parts of the body. Finally, analyzing functions in the context of their genomes can reveal sets of redundant taxa that define a core microbiome. Programs such as HUMMAnN3 already provide this type of information as part of their output, and more comprehensive reference databases can provide a proper basis for such attempts.

The study of the core human microbiome is currently restricted mostly to NGS-based approaches that, together with proper bioinformatics methods, generate large amounts of data. Of the two most used approaches, amplicon surveys and metagenomics, the latter can provide the necessary species- and even strain-level taxonomic and functional resolution required. Long-read technologies have the potential to improve and simplify the metagenomics analysis process by enabling a more accurate genome recovery process for a larger portion of the community. We anticipate that the coming years will witness addressing shortfalls such as incomplete reference databases; failure to assess the abundance of eukaryotes, eukaryotic viruses, and phages; and the insufficient number of samples from non-Western populations. Other technical challenges such as direct sampling from different parts of the gut currently do not have easy technical solutions. Therefore, the availability of highly localized data may remain low in the near future, limiting our ability to accurately evaluate the core microbiome of the various niches in the gut. Collecting, processing, and analyzing meta-omics data at levels other than the genomic can be challenging. Metatranscriptomic data, for example, are dominated by ribosomal RNA, and mRNA is very unstable, both issues requiring technical solutions that may be challenging. In addition, analysis pipelines for such data still require improvement. As for metabolomics, proton nuclear magnetic resonance (1H NMR), one of the main analytical tools used in metabolomics studies, suffers from low sensitivity and may not detect low-abundance metabolites [207,208]. Besides NMR, the chromatographic approach (whether gas or liquid at high performance) has been effectively used coupled with quadrupole-TOF (time of flight) mass spectrometry [209] to cover a high number of metabolites from microbial samples. However, TOF-based detection conventionally requires complex sample preparation and chromatographic fractionation routines while providing only a reasonable resolution in discovery-based metabolomics. To overcome these limitations, novel high-resolution accurate-mass (HRAM) analytical platforms have been implemented in the field of complex microbial community metabolomics. For example, ultra-high performance liquid chromatography (UHPLC) and quadrupole-orbitrap mass spectrometer allowed the untargeted identification of novel potential metabolites [210,211]. The orbitrap mass analyzer—despite being an expensive instrument that requires complex data analysis skills—enhances the separation of unknown and known compounds enabling high-throughput workflows and constitutes to date the most advanced platform used in discovery-based proteomics and metabolomics, providing valuable tools in facing the unknowns in host-microbiome metabolomics [212].

Overall, we expect our knowledge of the transcriptomic, proteomic, and metabolomic levels to significantly improve in the coming years, though the rate of data generation is not expected to achieve that of amplicon and even metagenomics data. Method development for integrating multi-omics data is still required to generate multi-layer models of the core human microbiome.

Challenges in characterizing the core human microbiome may be due to our limited understanding of the exact interactions between the human microbiome and its environment, particularly diet and the immune system. As currently, no high-throughput technology provides the necessary information, accounting for these factors in the context of the core microbiome is difficult. The collection of more non-genomic data necessitates computational and statistical approaches to combine data layers and to identify complex patterns in the data. Machine learning is a promising direction that can help reveal such patterns that are difficult to identify by humans. Machine learning has been used to predict disease states, identify biomarkers, and more [213,214]. Interpretable machine learning algorithms such as decision trees and random forest can be used to identify features that

are most relevant to groups of samples. For example, functions that are most relevant to the human microbiomes may be identified by comparing samples from humans against samples from mice, gorillas, and other animals. These and other algorithms may identify patterns that involve multiple features (e.g., functions) and are therefore more difficult to discover using standard statistical methods. Clustering algorithms may be used to define the core human microbiome in terms of groups of replaceable species by identifying co-occurring or mutually exclusive microbes. In addition, cloud computing and other high-throughput computational technologies make it feasible to analyze huge amounts of data, potentially even all the data stored in public databases [215]. This requires resolving issues related to the availability of metadata for the various samples. Still, domain expertise in nutrition, medicine, and immunology must be considered to interpret and evaluate the results. Therefore, collaborative efforts between experts in nutrition, immunology, human genetics, the microbiome, bioinformaticians, and machine learning can help design holistic studies that will uncover the patterns and forces that shape the core human microbiome.

7. Conclusions

The concept of the core human microbiome should be considered in the context of its environment, including body parts, diet, and geography. More data and better bioinformatics are expected to improve our abilities to assess the core human microbiome based on metagenomics, while the contribution of 16S rRNA gene surveys may remain limited. The analysis of previously less explored members of the microbiome will benefit from these trends and enable a more complete view of the core human microbiome. The genomic data should be integrated with other omics data; the technology to collect these data exists, but the bioinformatics and available public data are still lagging. Non-omics data related to diet and immunology are desired, but achieving such data is expected to remain challenging. Implementation of machine learning approaches on large amounts of data may reveal patterns that are otherwise difficult to identify. A collaborative effort of multidisciplinary teams will likely be needed to advance understanding of the core human microbiome.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/nu14142872/s1, Figure S1: Functional profiling of eight cohorts; Table S1: Summary of information of the data that was used for the community and function analyses; Table S2: Statistics of samples related to the human microbiome; Table S3: fraction of samples in each cohort in which the species was identified; Table S4: UniRef90 clusters that are present at $\geq$90% of the samples of at least one cohort; Table S5: all MetaCyc pathways detected in at least one sample. References [9,28,37–41,216] are cited in the supplementary materials.

Funding: This research was supported by the Joint Action "European Joint Programming Initiative "A Healthy Diet for a Healthy Life" (JPI HDHL)"and the respective national/regional funding organisations: Fund for Scientific Research (FRS—FNRS, Belgium); Research Foundation—Flanders (FWO, Belgium); INSERM Institut National de la Santé et de la Recherche Médicale (France); Federal Ministry of Food and Agriculture (BMEL, Germany) represented by Federal Office for Agriculture and Food (BLE, Germany, grant number 2819ERA10F); Federal Ministry of Education and Research Germany (BMBF, FKZ 01EA1906A, 01EA1906B) Ministry of Education, University and Research (MIUR), Ministry of agricultural, food and forestry policies (MiPAAF), Ministry of University and Research (MUR, Grant Number JTC-2017-7, Project MeaTIc) National Institute of Health (ISS) on behalf of Ministry of Health (Italy); National Institute of Health Carlos III (Spain); The Netherlands Organisation for Health Research and Development (ZonMw, The Netherlands, grant number 529051024), Austrian Research Promotion Agency (FFG) on behalf of the Austrian Federal Ministry for Education, Science and Research (BMBWF), Ministry of Science and Technology (Israel, grant number 3-16034), Formas (Sweden). BIOMIS project funded by PON RICERCA E INNOVAZIONE 2014-2020 (Cod. ARS01_01220) (Italy). Federal Ministry of Education and Research Germany (BMBF, FKZ01EA1906B). G.C: This study was funded by the Joint Programming Initiative a Healthy Diet for a Healthy Life-Intestinal Microbiomics (JPI HDHL-INTIMIC) Call for Joint Transnational Research Proposals on "Interrelation of the Intestinal Microbiome, Diet and Health" (Reference Number JTC-

2017-7) and by the University of Florence (Fondo ex-60%), Italy. K.L, M.M and J.G: amedes genetics acknowledges funding by BMBF/DLR, grant number 01EA1906G. E.P: The study has been also supported by the JPI HDHL-INTIMIC - Knowledge Platform of Food, Diet, Intestinal Microbiomics and Human Health, granted by the Italian Ministry of Agricultural, Food and Forestry and Tourism Policies with ID 790.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the Supplementary text.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Krajmalnik-Brown, R.; Ilhan, Z.-E.; Kang, D.-W.; DiBaise, J.K. Effects of gut microbes on nutrient absorption and energy regulation. *Nutr. Clin. Pract.* **2012**, *27*, 201–214. [CrossRef] [PubMed]
2. Oliphant, K.; Allen-Vercoe, E. Macronutrient metabolism by the human gut microbiome: Major fermentation by-products and their impact on host health. *Microbiome* **2019**, *7*, 91. [CrossRef] [PubMed]
3. Iacob, S.; Iacob, D.G.; Luminos, L.M. Intestinal microbiota as a host defense mechanism to infectious threats. *Front. Microbiol.* **2018**, *9*, 3328. [CrossRef] [PubMed]
4. Zheng, D.; Liwinski, T.; Elinav, E. Interaction between microbiota and immunity in health and disease. *Cell Res.* **2020**, *30*, 492–506. [CrossRef] [PubMed]
5. Rowland, I.; Gibson, G.; Heinken, A.; Scott, K.; Swann, J.; Thiele, I.; Tuohy, K. Gut microbiota functions: Metabolism of nutrients and other food components. *Eur. J. Nutr.* **2018**, *57*, 1–24. [CrossRef]
6. Glassner, K.L.; Abraham, B.P.; Quigley, E.M.M. The microbiome and inflammatory bowel disease. *J. Allergy Clin. Immunol.* **2020**, *145*, 16–27. [CrossRef]
7. Cenit, M.C.; Sanz, Y.; Codoñer-Franch, P. Influence of gut microbiota on neuropsychiatric disorders. *World J. Gastroenterol.* **2017**, *23*, 5486–5498. [CrossRef]
8. Integrative HMP (iHMP) Research Network Consortium The integrative human microbiome project. *Nature* **2019**, *569*, 641–648. [CrossRef]
9. Turnbaugh, P.J.; Ley, R.E.; Hamady, M.; Fraser-Liggett, C.M.; Knight, R.; Gordon, J.I. The human microbiome project. *Nature* **2007**, *449*, 804–810. [CrossRef]
10. The MetaHIT Consortium; Ehrlich, S.D. Metahit: The european union project on metagenomics of the human intestinal tract. In *Metagenomics of the Human Body*; Nelson, K.E., Ed.; Springer: New York, NY, USA, 2011; pp. 307–316. ISBN 978-1-4419-7088-6.
11. Pitlik, S.D.; Koren, O. How holobionts get sick-toward a unifying scheme of disease. *Microbiome* **2017**, *5*, 64. [CrossRef]
12. Turnbaugh, P.J.; Gordon, J.I. The core gut microbiome, energy balance and obesity. *J. Physiol.* **2009**, *587*, 4153–4158. [CrossRef] [PubMed]
13. Turnbaugh, P.J.; Hamady, M.; Yatsunenko, T.; Cantarel, B.L.; Duncan, A.; Ley, R.E.; Sogin, M.L.; Jones, W.J.; Roe, B.A.; Affourtit, J.P.; et al. A core gut microbiome in obese and lean twins. *Nature* **2009**, *457*, 480–484. [CrossRef] [PubMed]
14. Hildebrand, F.; Gossmann, T.I.; Frioux, C.; Özkurt, E.; Myers, P.N.; Ferretti, P.; Kuhn, M.; Bahram, M.; Nielsen, H.B.; Bork, P. Dispersal strategies shape persistence and evolution of human gut bacteria. *Cell Host Microbe* **2021**, *29*, 1167–1176.e9. [CrossRef] [PubMed]
15. Faith, J.J.; Guruge, J.L.; Charbonneau, M.; Subramanian, S.; Seedorf, H.; Goodman, A.L.; Clemente, J.C.; Knight, R.; Heath, A.C.; Leibel, R.L.; et al. The long-term stability of the human gut microbiota. *Science* **2013**, *341*, 1237439. [CrossRef]
16. Yassour, M.; Jason, E.; Hogstrom, L.J.; Arthur, T.D.; Tripathi, S.; Siljander, H.; Selvenius, J.; Oikarinen, S.; Hyöty, H.; Virtanen, S.M.; et al. Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. *Cell Host Microbe* **2018**, *24*, 146–154.e4. [CrossRef]
17. Zaura, E.; Keijser, B.J.F.; Huse, S.M.; Crielaard, W. Defining the healthy "core microbiome" of oral microbial communities. *BMC Microbiol.* **2009**, *9*, 259. [CrossRef]
18. Björk, J.R.; O'Hara, R.B.; Ribes, M.; Coma, R.; Montoya, J.M. The dynamic core microbiome: Structure, stability and resistance. *BioRxiv* **2018**, 137855. [CrossRef]
19. Shade, A.; Handelsman, J. Beyond the Venn diagram: The hunt for a core microbiome. *Environ. Microbiol.* **2012**, *14*, 4–12. [CrossRef]
20. Huse, S.M.; Ye, Y.; Zhou, Y.; Fodor, A.A. A core human microbiome as viewed through 16S rRNA sequence clusters. *PLoS ONE* **2012**, *7*, e34242. [CrossRef]
21. Lloyd-Price, J.; Abu-Ali, G.; Huttenhower, C. The healthy human microbiome. *Genome Med.* **2016**, *8*, 51. [CrossRef]
22. Risely, A. Applying the core microbiome to understand host-microbe systems. *J. Anim. Ecol.* **2020**, *89*, 1549–1558. [CrossRef] [PubMed]

23. Bäckhed, F.; Fraser, C.M.; Ringel, Y.; Sanders, M.E.; Sartor, R.B.; Sherman, P.M.; Versalovic, J.; Young, V.; Finlay, B.B. Defining a healthy human gut microbiome: Current concepts, future directions, and clinical applications. *Cell Host Microbe* **2012**, *12*, 611–622. [CrossRef] [PubMed]
24. Fassarella, M.; Blaak, E.E.; Penders, J.; Nauta, A.; Smidt, H.; Zoetendal, E.G. Gut microbiome stability and resilience: Elucidating the response to perturbations in order to modulate gut health. *Gut* **2021**, *70*, 595–605. [CrossRef] [PubMed]
25. Berg, G.; Rybakova, D.; Fischer, D.; Cernava, T.; Vergès, M.-C.C.; Charles, T.; Chen, X.; Cocolin, L.; Eversole, K.; Corral, G.H.; et al. Microbiome definition re-visited: Old concepts and new challenges. *Microbiome* **2020**, *8*, 103. [CrossRef] [PubMed]
26. Lemanceau, P.; Blouin, M.; Muller, D.; Moënne-Loccoz, Y. Let the core microbiota be functional. *Trends Plant Sci.* **2017**, *22*, 583–595. [CrossRef] [PubMed]
27. Human Microbiome Project Consortium Structure, function and diversity of the healthy human microbiome. *Nature* **2012**, *486*, 207–214. [CrossRef] [PubMed]
28. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K.S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010**, *464*, 59–65. [CrossRef]
29. Tap, J.; Mondot, S.; Levenez, F.; Pelletier, E.; Caron, C.; Furet, J.-P.; Ugarte, E.; Muñoz-Tamayo, R.; Paslier, D.L.E.; Nalin, R.; et al. Towards the human intestinal microbiota phylogenetic core. *Environ. Microbiol.* **2009**, *11*, 2574–2584. [CrossRef]
30. Arumugam, M.; Raes, J.; Pelletier, E.; Le Paslier, D.; Yamada, T.; Mende, D.R.; Fernandes, G.R.; Tap, J.; Bruls, T.; Batto, J.-M.; et al. Enterotypes of the human gut microbiome. *Nature* **2011**, *473*, 174–180. [CrossRef]
31. Costello, E.K.; Lauber, C.L.; Hamady, M.; Fierer, N.; Gordon, J.I.; Knight, R. Bacterial community variation in human body habitats across space and time. *Science* **2009**, *326*, 1694–1697. [CrossRef]
32. Li, K.; Bihan, M.; Methé, B.A. Analyses of the stability and core taxonomic memberships of the human microbiome. *PLoS ONE* **2013**, *8*, e63139. [CrossRef] [PubMed]
33. Banerjee, S.; Schlaeppi, K.; van der Heijden, M.G.A. Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* **2018**, *16*, 567–576. [CrossRef]
34. Ze, X.; Duncan, S.H.; Louis, P.; Flint, H.J. Ruminococcus bromii is a keystone species for the degradation of resistant starch in the human colon. *ISME J.* **2012**, *6*, 1535–1543. [CrossRef] [PubMed]
35. Lozupone, C.A.; Stombaugh, J.I.; Gordon, J.I.; Jansson, J.K.; Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **2012**, *489*, 220–230. [CrossRef]
36. Gutiérrez, N.; Garrido, D. Species Deletions from Microbiome Consortia Reveal Key Metabolic Interactions between Gut Microbes. *mSystems* **2019**, *4*, e00185-19. [CrossRef] [PubMed]
37. Lloyd-Price, J.; Mahurkar, A.; Rahnavard, G.; Crabtree, J.; Orvis, J.; Hall, A.B.; Brady, A.; Creasy, H.H.; McCracken, C.; Giglio, M.G.; et al. Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **2017**, *550*, 61–66. [CrossRef]
38. Sharma, A.K.; Petrzelkova, K.; Pafco, B.; Jost Robinson, C.A.; Fuh, T.; Wilson, B.A.; Stumpf, R.M.; Torralba, M.G.; Blekhman, R.; White, B.; et al. Traditional human populations and nonhuman primates show parallel gut microbiome adaptations to analogous ecological conditions. *mSystems* **2020**, *5*, e00815-20. [CrossRef]
39. Xiao, L.; Feng, Q.; Liang, S.; Sonne, S.B.; Xia, Z.; Qiu, X.; Li, X.; Long, H.; Zhang, J.; Zhang, D.; et al. A catalog of the mouse gut metagenome. *Nat. Biotechnol.* **2015**, *33*, 1103–1108. [CrossRef]
40. Feng, Y.; Wang, Y.; Zhu, B.; Gao, G.F.; Guo, Y.; Hu, Y. Metagenome-assembled genomes and gene catalog from the chicken gut microbiome aid in deciphering antibiotic resistomes. *Commun. Biol.* **2021**, *4*, 1305. [CrossRef]
41. Beghini, F.; McIver, L.J.; Blanco-Míguez, A.; Dubois, L.; Asnicar, F.; Maharjan, S.; Mailyan, A.; Manghi, P.; Scholz, M.; Thomas, A.M.; et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* **2021**, *10*, e65088. [CrossRef]
42. De Filippis, F.; Pasolli, E.; Ercolini, D. Newly explored faecalibacterium diversity is connected to age, lifestyle, geography, and disease. *Curr. Biol.* **2020**, *30*, 4932–4943.e4. [CrossRef] [PubMed]
43. Quince, C.; Walker, A.W.; Simpson, J.T.; Loman, N.J.; Segata, N. Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* **2017**, *35*, 833–844. [CrossRef] [PubMed]
44. Bashiardes, S.; Zilberman-Schapira, G.; Elinav, E. Use of metatranscriptomics in microbiome research. *Bioinform. Biol. Insights* **2016**, *10*, 19–25. [CrossRef]
45. Lloyd-Price, J.; Arze, C.; Ananthakrishnan, A.N.; Schirmer, M.; Avila-Pacheco, J.; Poon, T.W.; Andrews, E.; Ajami, N.J.; Bonham, K.S.; Brislawn, C.J.; et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **2019**, *569*, 655–662. [CrossRef]
46. Gupta, V.K.; Paul, S.; Dutta, C. Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity. *Front. Microbiol.* **2017**, *8*, 1162. [CrossRef] [PubMed]
47. Abdill, R.J.; Adamowicz, E.M.; Blekhman, R. Public human microbiome data are dominated by highly developed countries. *PLoS Biol.* **2022**, *20*, e3001536. [CrossRef] [PubMed]
48. Human Microbiome Project Consortium A framework for human microbiome research. *Nature* **2012**, *486*, 215–221. [CrossRef]
49. Zijlema, W.L.; Smidt, N.; Klijs, B.; Morley, D.W.; Gulliver, J.; de Hoogh, K.; Scholtens, S.; Rosmalen, J.G.M.; Stolk, R.P. The LifeLines Cohort Study: A resource providing new opportunities for environmental epidemiology. *Arch. Public Health* **2016**, *74*, 32. [CrossRef]

50. Falony, G.; Joossens, M.; Vieira-Silva, S.; Wang, J.; Darzi, Y.; Faust, K.; Kurilshikov, A.; Bonder, M.J.; Valles-Colomer, M.; Vandeputte, D.; et al. Population-level analysis of gut microbiome variation. *Science* **2016**, *352*, 560–564. [CrossRef]
51. Tyakht, A.V.; Alexeev, D.G.; Popenko, A.S.; Kostryukova, E.S.; Govorun, V.M. Rural and urban microbiota: To be or not to be? *Gut Microbes* **2014**, *5*, 351–356. [CrossRef]
52. Saarenpää, M.; Roslund, M.I.; Puhakka, R.; Grönroos, M.; Parajuli, A.; Hui, N.; Nurminen, N.; Laitinen, O.H.; Hyöty, H.; Cinek, O.; et al. The Adele Research Group Do Rural Second Homes Shape Commensal Microbiota of Urban Dwellers? A Pilot Study among Urban Elderly in Finland. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3742. [CrossRef] [PubMed]
53. Schnorr, S.L.; Candela, M.; Rampelli, S.; Centanni, M.; Consolandi, C.; Basaglia, G.; Turroni, S.; Biagi, E.; Peano, C.; Severgnini, M.; et al. Gut microbiome of the Hadza hunter-gatherers. *Nat. Commun.* **2014**, *5*, 3654. [CrossRef] [PubMed]
54. Rampelli, S.; Schnorr, S.L.; Consolandi, C.; Turroni, S.; Severgnini, M.; Peano, C.; Brigidi, P.; Crittenden, A.N.; Henry, A.G.; Candela, M. Metagenome Sequencing of the Hadza Hunter-Gatherer Gut Microbiota. *Curr. Biol.* **2015**, *25*, 1682–1693. [CrossRef] [PubMed]
55. Iebba, V.; Santangelo, F.; Totino, V.; Pantanella, F.; Monsia, A.; Di Cristanziano, V.; Di Cave, D.; Schippa, S.; Berrilli, F.; D'Alfonso, R. Gut microbiota related to Giardia duodenalis, Entamoeba spp. and Blastocystis hominis infections in humans from Côte d'Ivoire. *J. Infect. Dev. Ctries.* **2016**, *10*, 1035–1041. [CrossRef]
56. Gomez, A.; Petrzelkova, K.J.; Burns, M.B.; Yeoman, C.J.; Amato, K.R.; Vlckova, K.; Modry, D.; Todd, A.; Jost Robinson, C.A.; Remis, M.J.; et al. Gut microbiome of coexisting baaka pygmies and bantu reflects gradients of traditional subsistence patterns. *Cell Rep.* **2016**, *14*, 2142–2153. [CrossRef] [PubMed]
57. Martínez, I.; Stegen, J.C.; Maldonado-Gómez, M.X.; Eren, A.M.; Siba, P.M.; Greenhill, A.R.; Walter, J. The gut microbiota of rural papua new guineans: Composition, diversity patterns, and ecological processes. *Cell Rep.* **2015**, *11*, 527–538. [CrossRef]
58. Clemente, J.C.; Pehrsson, E.C.; Blaser, M.J.; Sandhu, K.; Gao, Z.; Wang, B.; Magris, M.; Hidalgo, G.; Contreras, M.; Noya-Alarcón, Ó.; et al. The microbiome of uncontacted Amerindians. *Sci. Adv.* **2015**, *1*, e1500183. [CrossRef]
59. Das, B.; Ghosh, T.S.; Kedia, S.; Rampal, R.; Saxena, S.; Bag, S.; Mitra, R.; Dayal, M.; Mehta, O.; Surendranath, A.; 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. [CrossRef]
60. Ayeni, F.A.; Biagi, E.; Rampelli, S.; Fiori, J.; Soverini, M.; Audu, H.J.; Cristino, S.; Caporali, L.; Schnorr, S.L.; Carelli, V.; et al. Infant and Adult Gut Microbiome and Metabolome in Rural Bassa and Urban Settlers from Nigeria. *Cell Rep.* **2018**, *23*, 3056–3067. [CrossRef]
61. Obregon-Tito, A.J.; Tito, R.Y.; Metcalf, J.; Sankaranarayanan, K.; Clemente, J.C.; Ursell, L.K.; Zech Xu, Z.; Van Treuren, W.; Knight, R.; Gaffney, P.M.; et al. Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat. Commun.* **2015**, *6*, 6505. [CrossRef]
62. Wild, C.P. Complementing the genome with an "exposome": The outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol. Biomarkers Prev.* **2005**, *14*, 1847–1850. [CrossRef] [PubMed]
63. De Filippo, C.; Cavalieri, D.; Di Paola, M.; Ramazzotti, M.; Poullet, J.B.; Massart, S.; Collini, S.; Pieraccini, G.; Lionetti, P. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14691–14696. [CrossRef] [PubMed]
64. De Filippo, C.; Di Paola, M.; Ramazzotti, M.; Albanese, D.; Pieraccini, G.; Banci, E.; Miglietta, F.; Cavalieri, D.; Lionetti, P. Diet, environments, and gut microbiota. A preliminary investigation in children living in rural and urban burkina faso and italy. *Front. Microbiol.* **2017**, *8*, 1979. [CrossRef]
65. Nakayama, J.; Yamamoto, A.; Palermo-Conde, L.A.; Higashi, K.; Sonomoto, K.; Tan, J.; Lee, Y.-K. Impact of westernized diet on gut microbiota in children on leyte island. *Front. Microbiol.* **2017**, *8*, 197. [CrossRef]
66. Yatsunenko, T.; Rey, F.E.; Manary, M.J.; Trehan, I.; Dominguez-Bello, M.G.; Contreras, M.; Magris, M.; Hidalgo, G.; Baldassano, R.N.; Anokhin, A.P.; et al. Human gut microbiome viewed across age and geography. *Nature* **2012**, *486*, 222–227. [CrossRef]
67. Pasolli, E.; Asnicar, F.; Manara, S.; Zolfo, M.; Karcher, N.; Armanini, F.; Beghini, F.; Manghi, P.; Tett, A.; Ghensi, P.; et al. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell* **2019**, *176*, 649–662.e20. [CrossRef] [PubMed]
68. Zmora, N.; Zilberman-Schapira, G.; Suez, J.; Mor, U.; Dori-Bachash, M.; Bashiardes, S.; Kotler, E.; Zur, M.; Regev-Lehavi, D.; Brik, R.B.-Z.; et al. Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. *Cell* **2018**, *174*, 1388–1405.e21. [CrossRef] [PubMed]
69. Eckburg, P.B.; Bik, E.M.; Bernstein, C.N.; Purdom, E.; Dethlefsen, L.; Sargent, M.; Gill, S.R.; Nelson, K.E.; Relman, D.A. Diversity of the human intestinal microbial flora. *Science* **2005**, *308*, 1635–1638. [CrossRef]
70. Stearns, J.C.; Lynch, M.D.J.; Senadheera, D.B.; Tenenbaum, H.C.; Goldberg, M.B.; Cvitkovitch, D.G.; Croitoru, K.; Moreno-Hagelsieb, G.; Neufeld, J.D. Bacterial biogeography of the human digestive tract. *Sci. Rep.* **2011**, *1*, 170. [CrossRef]
71. Zoetendal, E.G.; Raes, J.; van den Bogert, B.; Arumugam, M.; Booijink, C.C.G.M.; Troost, F.J.; Bork, P.; Wels, M.; de Vos, W.M.; Kleerebezem, M. The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J.* **2012**, *6*, 1415–1426. [CrossRef]
72. Albenberg, L.; Esipova, T.V.; Judge, C.P.; Bittinger, K.; Chen, J.; Laughlin, A.; Grunberg, S.; Baldassano, R.N.; Lewis, J.D.; Li, H.; et al. Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology* **2014**, *147*, 1055–1063.e8. [CrossRef] [PubMed]

73. Espey, M.G. Role of oxygen gradients in shaping redox relationships between the human intestine and its microbiota. *Free Radic. Biol. Med.* **2013**, *55*, 130–140. [CrossRef] [PubMed]
74. Liu, Q.; Georgieva, D.C.; Egli, D.; Wang, K. NanoMod: A computational tool to detect DNA modifications using Nanopore long-read sequencing data. *BMC Genom.* **2019**, *20*, 78. [CrossRef] [PubMed]
75. Lepage, P.; Seksik, P.; Sutren, M.; de la Cochetière, M.-F.; Jian, R.; Marteau, P.; Doré, J. Biodiversity of the mucosa-associated microbiota is stable along the distal digestive tract in healthy individuals and patients with IBD. *Inflamm. Bowel Dis.* **2005**, *11*, 473–480. [CrossRef]
76. Green, G.L.; Brostoff, J.; Hudspith, B.; Michael, M.; Mylonaki, M.; Rayment, N.; Staines, N.; Sanderson, J.; Rampton, D.S.; Bruce, K.D. Molecular characterization of the bacteria adherent to human colorectal mucosa. *J. Appl. Microbiol.* **2006**, *100*, 460–469. [CrossRef]
77. Altomare, A.; Putignani, L.; Del Chierico, F.; Cocca, S.; Angeletti, S.; Ciccozzi, M.; Tripiciano, C.; Dalla Piccola, B.; Cicala, M.; Guarino, M.P.L. Gut mucosal-associated microbiota better discloses inflammatory bowel disease differential patterns than faecal microbiota. *Dig. Liver Dis.* **2019**, *51*, 648–656. [CrossRef]
78. Sokol, H.; Pigneur, B.; Watterlot, L.; Lakhdari, O.; Bermúdez-Humarán, L.G.; Gratadoux, J.-J.; Blugeon, S.; Bridonneau, C.; Furet, J.-P.; Corthier, G.; et al. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16731–16736. [CrossRef]
79. Lo Presti, A.; Zorzi, F.; Del Chierico, F.; Altomare, A.; Cocca, S.; Avola, A.; De Biasio, F.; Russo, A.; Cella, E.; Reddel, S.; et al. Fecal and mucosal microbiota profiling in irritable bowel syndrome and inflammatory bowel disease. *Front. Microbiol.* **2019**, *10*, 1655. [CrossRef]
80. Bajaj, J.S.; Hylemon, P.B.; Ridlon, J.M.; Heuman, D.M.; Daita, K.; White, M.B.; Monteith, P.; Noble, N.A.; Sikaroodi, M.; Gillevet, P.M. Colonic mucosal microbiome differs from stool microbiome in cirrhosis and hepatic encephalopathy and is linked to cognition and inflammation. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2012**, *303*, G675–G685. [CrossRef]
81. Mark Welch, J.L.; Ramírez-Puebla, S.T.; Borisy, G.G. Oral Microbiome Geography: Micron-Scale Habitat and Niche. *Cell Host Microbe* **2020**, *28*, 160–168. [CrossRef]
82. Nearing, J.T.; DeClercq, V.; Van Limbergen, J.; Langille, M.G.I. Assessing the Variation within the Oral Microbiome of Healthy Adults. *mSphere* **2020**, *5*, e00451-20. [CrossRef] [PubMed]
83. De Filippis, F.; Vannini, L.; La Storia, A.; Laghi, L.; Piombino, P.; Stellato, G.; Serrazanetti, D.I.; Gozzi, G.; Turroni, S.; Ferrocino, I.; et al. The same microbiota and a potentially discriminant metabolome in the saliva of omnivore, ovo-lacto-vegetarian and Vegan individuals. *PLoS ONE* **2014**, *9*, e112373. [CrossRef] [PubMed]
84. Takeshita, T.; Kageyama, S.; Furuta, M.; Tsuboi, H.; Takeuchi, K.; Shibata, Y.; Shimazaki, Y.; Akifusa, S.; Ninomiya, T.; Kiyohara, Y.; et al. Bacterial diversity in saliva and oral health-related conditions: The Hisayama Study. *Sci. Rep.* **2016**, *6*, 22164. [CrossRef] [PubMed]
85. Li, J.; Quinque, D.; Horz, H.-P.; Li, M.; Rzhetskaya, M.; Raff, J.A.; Hayes, M.G.; Stoneking, M. Comparative analysis of the human saliva microbiome from different climate zones: Alaska, Germany, and Africa. *BMC Microbiol.* **2014**, *14*, 316. [CrossRef] [PubMed]
86. Schloss, P.D.; Westcott, S.L.; Ryabin, T.; Hall, J.R.; Hartmann, M.; Hollister, E.B.; Lesniewski, R.A.; Oakley, B.B.; Parks, D.H.; Robinson, C.J.; et al. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **2009**, *75*, 7537–7541. [CrossRef]
87. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857. [CrossRef]
88. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [CrossRef]
89. Ercolini, D. High-throughput sequencing and metagenomics: Moving forward in the culture-independent analysis of food microbial ecology. *Appl. Environ. Microbiol.* **2013**, *79*, 3148–3155. [CrossRef]
90. De Filippis, F.; Parente, E.; Ercolini, D. Recent past, present, and future of the food microbiome. *Annu. Rev. Food Sci. Technol.* **2018**, *9*, 589–608. [CrossRef]
91. Zhang, J.; Chiodini, R.; Badr, A.; Zhang, G. The impact of next-generation sequencing on genomics. *J. Genet. Genomics* **2011**, *38*, 95–109. [CrossRef]
92. Plummer, E.; Twin, J. A Comparison of Three Bioinformatics Pipelines for the Analysis of Preterm Gut Microbiota using 16S rRNA Gene Sequencing Data. *J. Proteom. Bioinform.* **2015**, *8*, 283–291. [CrossRef]
93. Jovel, J.; Patterson, J.; Wang, W.; Hotte, N.; O'Keefe, S.; Mitchel, T.; Perry, T.; Kao, D.; Mason, A.L.; Madsen, K.L.; et al. Characterization of the gut microbiome using 16S or shotgun metagenomics. *Front. Microbiol.* **2016**, *7*, 459. [CrossRef]
94. Fuks, G.; Elgart, M.; Amir, A.; Zeisel, A.; Turnbaugh, P.J.; Soen, Y.; Shental, N. Combining 16S rRNA gene variable regions enables high-resolution microbial community profiling. *Microbiome* **2018**, *6*, 17. [CrossRef] [PubMed]
95. McLean, K.; Rosenthal, C.A.; Sengupta, D.; Owens, J.; Cookson, B.T.; Hoffman, N.G.; Salipante, S.J. Improved Species-Level Clinical Identification of Enterobacteriaceae through Broad-Range dnaJ PCR and Sequencing. *J. Clin. Microbiol.* **2019**, *57*, e00986-19. [CrossRef] [PubMed]
96. Větrovský, T.; Baldrian, P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS ONE* **2013**, *8*, e57923. [CrossRef] [PubMed]

97. Johnson, J.S.; Spakowicz, D.J.; Hong, B.-Y.; Petersen, L.M.; Demkowicz, P.; Chen, L.; Leopold, S.R.; Hanson, B.M.; Agresta, H.O.; Gerstein, M.; et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat. Commun.* **2019**, *10*, 5029. [CrossRef]
98. Stewart, R.D.; Auffret, M.D.; Snelling, T.J.; Roehe, R.; Watson, M. MAGpy: A reproducible pipeline for the downstream analysis of metagenome-assembled genomes (MAGs). *Bioinformatics* **2019**, *35*, 2150–2152. [CrossRef]
99. Quijada, N.M.; Hernández, M.; Rodríguez-Lázaro, D. High-throughput sequencing and food microbiology. *Adv. Food Nutr. Res.* **2020**, *91*, 275–300. [CrossRef]
100. Gu, W.; Miller, S.; Chiu, C.Y. Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. *Annu. Rev. Pathol.* **2019**, *14*, 319–338. [CrossRef]
101. Tamames, J.; Cobo-Simón, M.; Puente-Sánchez, F. Assessing the performance of different approaches for functional and taxonomic annotation of metagenomes. *BMC Genom.* **2019**, *20*, 960. [CrossRef]
102. Wood, D.E.; Salzberg, S.L. Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **2014**, *15*, R46. [CrossRef] [PubMed]
103. Lu, J.; Breitwieser, F.P.; Thielen, P.; Salzberg, S.L. Bracken: Estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* **2017**, *3*, e104. [CrossRef]
104. Kim, D.; Song, L.; Breitwieser, F.P.; Salzberg, S.L. Centrifuge: Rapid and sensitive classification of metagenomic sequences. *Genome Res.* **2016**, *26*, 1721–1729. [CrossRef]
105. Milanese, A.; Mende, D.R.; Paoli, L.; Salazar, G.; Ruscheweyh, H.-J.; Cuenca, M.; Hingamp, P.; Alves, R.; Costea, P.I.; Coelho, L.P.; et al. Microbial abundance, activity and population genomic profiling with mOTUs2. *Nat. Commun.* **2019**, *10*, 1014. [CrossRef]
106. Pignatelli, M.; Aparicio, G.; Blanquer, I.; Hernandez, V.; Moya, A.; Tamames, J. Metagenomics reveals our incomplete knowledge of global diversity. *Bioinformatics* **2008**, *24*, 2124–2125. [CrossRef] [PubMed]
107. Heintz-Buschart, A.; Wilmes, P. Human gut microbiome: Function matters. *Trends Microbiol.* **2018**, *26*, 563–574. [CrossRef]
108. Bowers, R.M.; Kyrpides, N.C.; Stepanauskas, R.; Harmon-Smith, M.; Doud, D.; Reddy, T.B.K.; Schulz, F.; Jarett, J.; Rivers, A.R.; Eloe-Fadrosh, E.A.; et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **2017**, *35*, 725–731. [CrossRef]
109. Sharon, I.; Banfield, J.F. Microbiology. Genomes from metagenomics. *Science* **2013**, *342*, 1057–1058. [CrossRef]
110. Almeida, A.; Nayfach, S.; Boland, M.; Strozzi, F.; Beracochea, M.; Shi, Z.J.; Pollard, K.S.; Sakharova, E.; Parks, D.H.; Hugenholtz, P.; et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **2020**, *39*, 105–114. [CrossRef]
111. Almeida, A.; Mitchell, A.L.; Boland, M.; Forster, S.C.; Gloor, G.B.; Tarkowska, A.; Lawley, T.D.; Finn, R.D. A new genomic blueprint of the human gut microbiota. *Nature* **2019**, *568*, 499–504. [CrossRef]
112. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 160018. [CrossRef] [PubMed]
113. Hiergeist, A.; Reischl, U.; Priority Program 1656 Intestinal Microbiota Consortium/quality assessment participants; Gessner, A. Multicenter quality assessment of 16S ribosomal DNA-sequencing for microbiome analyses reveals high inter-center variability. *Int. J. Med. Microbiol.* **2016**, *306*, 334–342. [CrossRef] [PubMed]
114. Bartolomaeus, T.U.P.; Birkner, T.; Bartolomaeus, H.; Löber, U.; Avery, E.G.; Mähler, A.; Weber, D.; Kochlik, B.; Balogh, A.; Wilck, N.; et al. Quantifying technical confounders in microbiome studies. *Cardiovasc. Res.* **2021**, *117*, 863–875. [CrossRef]
115. Fouhy, F.; Clooney, A.G.; Stanton, C.; Claesson, M.J.; Cotter, P.D. 16S rRNA gene sequencing of mock microbial populations-impact of DNA extraction method, primer choice and sequencing platform. *BMC Microbiol.* **2016**, *16*, 123. [CrossRef] [PubMed]
116. Tremblay, J.; Singh, K.; Fern, A.; Kirton, E.S.; He, S.; Woyke, T.; Lee, J.; Chen, F.; Dangl, J.L.; Tringe, S.G. Primer and platform effects on 16S rRNA tag sequencing. *Front. Microbiol.* **2015**, *6*, 771. [CrossRef]
117. Duvallet, C. Meta-analysis generates and prioritizes hypotheses for translational microbiome research. *Microb. Biotechnol.* **2018**, *11*, 273–276. [CrossRef]
118. Harrison, P.W.; Ahamed, A.; Aslam, R.; Alako, B.T.F.; Burgin, J.; Buso, N.; Courtot, M.; Fan, J.; Gupta, D.; Haseeb, M.; et al. The european nucleotide archive in 2020. *Nucleic Acids Res.* **2021**, *49*, D82–D85. [CrossRef]
119. Kodama, Y.; Shumway, M.; Leinonen, R. International Nucleotide Sequence Database Collaboration The Sequence Read Archive: Explosive growth of sequencing data. *Nucleic Acids Res.* **2012**, *40*, D54–D56. [CrossRef]
120. Yilmaz, P.; Kottmann, R.; Field, D.; Knight, R.; Cole, J.R.; Amaral-Zettler, L.; Gilbert, J.A.; Karsch-Mizrachi, I.; Johnston, A.; Cochrane, G.; et al. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat. Biotechnol.* **2011**, *29*, 415–420. [CrossRef]
121. Buttigieg, P.L.; Morrison, N.; Smith, B.; Mungall, C.J.; Lewis, S.E. ENVO Consortium The environment ontology: Contextualising biological and biomedical entities. *J. Biomed. Semant.* **2013**, *4*, 43. [CrossRef]
122. Degtyarenko, K.; de Matos, P.; Ennis, M.; Hastings, J.; Zbinden, M.; McNaught, A.; Alcántara, R.; Darsow, M.; Guedj, M.; Ashburner, M. ChEBI: A database and ontology for chemical entities of biological interest. *Nucleic Acids Res.* **2008**, *36*, D344–D350. [CrossRef] [PubMed]
123. Schriml, L.M.; Arze, C.; Nadendla, S.; Chang, Y.-W.W.; Mazaitis, M.; Felix, V.; Feng, G.; Kibbe, W.A. Disease Ontology: A backbone for disease semantic integration. *Nucleic Acids Res.* **2012**, *40*, D940–D946. [CrossRef]

124. Gralka, M.; Szabo, R.; Stocker, R.; Cordero, O.X. Trophic interactions and the drivers of microbial community assembly. *Curr. Biol.* **2020**, *30*, R1176–R1188. [CrossRef] [PubMed]
125. Ribeiro, A.A.; Azcarate-Peril, M.A.; Cadenas, M.B.; Butz, N.; Paster, B.J.; Chen, T.; Bair, E.; Arnold, R.R. The oral bacterial microbiome of occlusal surfaces in children and its association with diet and caries. *PLoS ONE* **2017**, *12*, e0180621. [CrossRef] [PubMed]
126. Adler, C.J.; Dobney, K.; Weyrich, L.S.; Kaidonis, J.; Walker, A.W.; Haak, W.; Bradshaw, C.J.A.; Townsend, G.; Sołtysiak, A.; Alt, K.W.; et al. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nat. Genet.* **2013**, *45*, 450–455e1. [CrossRef]
127. David, L.A.; Maurice, C.F.; Carmody, R.N.; Gootenberg, D.B.; Button, J.E.; Wolfe, B.E.; Ling, A.V.; Devlin, A.S.; Varma, Y.; Fischbach, M.A.; et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **2014**, *505*, 559–563. [CrossRef]
128. Zmora, N.; Suez, J.; Elinav, E. You are what you eat: Diet, health and the gut microbiota. *Nat. Rev. Gastroenterol. Hepatol.* **2019**, *16*, 35–56. [CrossRef]
129. Kolodziejczyk, A.A.; Zheng, D.; Elinav, E. Diet-microbiota interactions and personalized nutrition. *Nat. Rev. Microbiol.* **2019**, *17*, 742–753. [CrossRef]
130. Charbonneau, M.R.; O'Donnell, D.; Blanton, L.V.; Totten, S.M.; Davis, J.C.C.; Barratt, M.J.; Cheng, J.; Guruge, J.; Talcott, M.; Bain, J.R.; et al. Sialylated Milk Oligosaccharides Promote Microbiota-Dependent Growth in Models of Infant Undernutrition. *Cell* **2016**, *164*, 859–871. [CrossRef]
131. Stewart, C.J.; Ajami, N.J.; O'Brien, J.L.; Hutchinson, D.S.; Smith, D.P.; Wong, M.C.; Ross, M.C.; Lloyd, R.E.; Doddapaneni, H.; Metcalf, G.A.; et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **2018**, *562*, 583–588. [CrossRef]
132. Koenig, J.E.; Spor, A.; Scalfone, N.; Fricker, A.D.; Stombaugh, J.; Knight, R.; Angenent, L.T.; Ley, R.E. Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4578–4585. [CrossRef] [PubMed]
133. Laursen, M.F.; Bahl, M.I.; Michaelsen, K.F.; Licht, T.R. First foods and gut microbes. *Front. Microbiol.* **2017**, *8*, 356. [CrossRef]
134. Palleja, A.; Mikkelsen, K.H.; Forslund, S.K.; Kashani, A.; Allin, K.H.; Nielsen, T.; Hansen, T.H.; Liang, S.; Feng, Q.; Zhang, C.; et al. Recovery of gut microbiota of healthy adults following antibiotic exposure. *Nat. Microbiol.* **2018**, *3*, 1255–1265. [CrossRef] [PubMed]
135. Santoro, A.; Ostan, R.; Candela, M.; Biagi, E.; Brigidi, P.; Capri, M.; Franceschi, C. Gut microbiota changes in the extreme decades of human life: A focus on centenarians. *Cell. Mol. Life Sci.* **2018**, *75*, 129–148. [CrossRef] [PubMed]
136. Claesson, M.J.; Jeffery, I.B.; Conde, S.; Power, S.E.; O'Connor, E.M.; Cusack, S.; Harris, H.M.B.; Coakley, M.; Lakshminarayanan, B.; O'Sullivan, O.; et al. Gut microbiota composition correlates with diet and health in the elderly. *Nature* **2012**, *488*, 178–184. [CrossRef] [PubMed]
137. He, Y.; Wu, W.; Zheng, H.-M.; Li, P.; McDonald, D.; Sheng, H.-F.; Chen, M.-X.; Chen, Z.-H.; Ji, G.-Y.; Zheng, Z.-D.-X.; et al. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. *Nat. Med.* **2018**, *24*, 1532–1535. [CrossRef]
138. Tarallo, S.; Ferrero, G.; De Filippis, F.; Francavilla, A.; Pasolli, E.; Panero, V.; Cordero, F.; Segata, N.; Grioni, S.; Pensa, R.G.; et al. Stool microRNA profiles reflect different dietary and gut microbiome patterns in healthy individuals. *Gut* **2021**, *71*, 1302–1314. [CrossRef]
139. De Filippis, F.; Pellegrini, N.; Vannini, L.; Jeffery, I.B.; La Storia, A.; Laghi, L.; Serrazanetti, D.I.; Di Cagno, R.; Ferrocino, I.; Lazzi, C.; et al. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. *Gut* **2016**, *65*, 1812–1821. [CrossRef]
140. Meslier, V.; Laiola, M.; Roager, H.M.; De Filippis, F.; Roume, H.; Quinquis, B.; Giacco, R.; Mennella, I.; Ferracane, R.; Pons, N.; et al. Mediterranean diet intervention in overweight and obese subjects lowers plasma cholesterol and causes changes in the gut microbiome and metabolome independently of energy intake. *Gut* **2020**, *69*, 1258–1268. [CrossRef]
141. Sonnenburg, E.D.; Sonnenburg, J.L. Starving our microbial self: The deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. *Cell Metab.* **2014**, *20*, 779–786. [CrossRef]
142. Wegner, K.; Just, S.; Gau, L.; Mueller, H.; Gérard, P.; Lepage, P.; Clavel, T.; Rohn, S. Rapid analysis of bile acids in different biological matrices using LC-ESI-MS/MS for the investigation of bile acid transformation by mammalian gut bacteria. *Anal. Bioanal. Chem.* **2017**, *409*, 1231–1245. [CrossRef] [PubMed]
143. Cockburn, D.W.; Koropatkin, N.M. Polysaccharide degradation by the intestinal microbiota and its influence on human health and disease. *J. Mol. Biol.* **2016**, *428*, 3230–3252. [CrossRef] [PubMed]
144. Campos-Perez, W.; Martinez-Lopez, E. Effects of short chain fatty acids on metabolic and inflammatory processes in human health. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2021**, *1866*, 158900. [CrossRef] [PubMed]
145. Koh, A.; De Vadder, F.; Kovatcheva-Datchary, P.; Bäckhed, F. From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* **2016**, *165*, 1332–1345. [CrossRef]
146. Turroni, S.; Brigidi, P.; Cavalli, A.; Candela, M. Microbiota-Host Transgenomic Metabolism, Bioactive Molecules from the Inside. *J. Med. Chem.* **2018**, *61*, 47–61. [CrossRef]
147. Zhu, W.; Gregory, J.C.; Org, E.; Buffa, J.A.; Gupta, N.; Wang, Z.; Li, L.; Fu, X.; Wu, Y.; Mehrabian, M.; et al. Gut microbial metabolite TMAO enhances platelet hyperreactivity and thrombosis risk. *Cell* **2016**, *165*, 111–124. [CrossRef]

148. De Filippis, F.; Pasolli, E.; Tett, A.; Tarallo, S.; Naccarati, A.; De Angelis, M.; Neviani, E.; Cocolin, L.; Gobbetti, M.; Segata, N.; et al. Distinct Genetic and Functional Traits of Human Intestinal Prevotella copri Strains Are Associated with Different Habitual Diets. *Cell Host Microbe* **2019**, *25*, 444–453.e3. [CrossRef]
149. Tett, A.; Huang, K.D.; Asnicar, F.; Fehlner-Peach, H.; Pasolli, E.; Karcher, N.; Armanini, F.; Manghi, P.; Bonham, K.; Zolfo, M.; et al. The Prevotella copri Complex Comprises Four Distinct Clades Underrepresented in Westernized Populations. *Cell Host Microbe* **2019**, *26*, 666–679.e7. [CrossRef]
150. Pasolli, E.; De Filippis, F.; Mauriello, I.E.; Cumbo, F.; Walsh, A.M.; Leech, J.; Cotter, P.D.; Segata, N.; Ercolini, D. Large-scale genome-wide analysis links lactic acid bacteria from food with the gut microbiome. *Nat. Commun.* **2020**, *11*, 2610. [CrossRef]
151. Dethlefsen, L.; McFall-Ngai, M.; Relman, D.A. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* **2007**, *449*, 811–818. [CrossRef]
152. Honda, K.; Littman, D.R. The microbiota in adaptive immune homeostasis and disease. *Nature* **2016**, *535*, 75–84. [CrossRef] [PubMed]
153. Thaiss, C.A.; Zmora, N.; Levy, M.; Elinav, E. The microbiome and innate immunity. *Nature* **2016**, *535*, 65–74. [CrossRef] [PubMed]
154. Schell, S.L.; Schneider, A.M.; Nelson, A.M. Yin and Yang: A disrupted skin microbiome and an aberrant host immune response in hidradenitis suppurativa. *Exp. Dermatol.* **2021**, *30*, 1453–1470. [CrossRef] [PubMed]
155. Wang, H.; Wang, G.; Banerjee, N.; Liang, Y.; Du, X.; Boor, P.J.; Hoffman, K.L.; Khan, M.F. Aberrant gut microbiome contributes to intestinal oxidative stress, barrier dysfunction, inflammation and systemic autoimmune responses in mrl/lpr mice. *Front. Immunol.* **2021**, *12*, 651191. [CrossRef] [PubMed]
156. Maeda, Y.; Takeda, K. Host-microbiota interactions in rheumatoid arthritis. *Exp. Mol. Med.* **2019**, *51*, 1–6. [CrossRef]
157. Main, B.S.; Minter, M.R. Microbial Immuno-Communication in Neurodegenerative Diseases. *Front. Neurosci.* **2017**, *11*, 151. [CrossRef]
158. Gopalakrishnan, V.; Helmink, B.A.; Spencer, C.N.; Reuben, A.; Wargo, J.A. The influence of the gut microbiome on cancer, immunity, and cancer immunotherapy. *Cancer Cell* **2018**, *33*, 570–580. [CrossRef]
159. Meade, K.G.; O'Farrelly, C. β-Defensins: Farming the Microbiome for Homeostasis and Health. *Front. Immunol.* **2018**, *9*, 3072. [CrossRef]
160. Zong, X.; Fu, J.; Xu, B.; Wang, Y.; Jin, M. Interplay between gut microbiota and antimicrobial peptides. *Anim. Nutr.* **2020**, *6*, 389–396. [CrossRef]
161. Pabst, O.; Slack, E. IgA and the intestinal microbiota: The importance of being specific. *Mucosal Immunol.* **2020**, *13*, 12–21. [CrossRef]
162. Liu, S.; da Cunha, A.P.; Rezende, R.M.; Cialic, R.; Wei, Z.; Bry, L.; Comstock, L.E.; Gandhi, R.; Weiner, H.L. The host shapes the gut microbiota via fecal microrna. *Cell Host Microbe* **2016**, *19*, 32–43. [CrossRef] [PubMed]
163. Shi, N.; Li, N.; Duan, X.; Niu, H. Interaction between the gut microbiome and mucosal immune system. *Mil. Med. Res.* **2017**, *4*, 14. [CrossRef] [PubMed]
164. Idris, A.; Hasnain, S.Z.; Huat, L.Z.; Koh, D. Human diseases, immunity and the oral microbiota—Insights gained from metagenomic studies. *Oral Science International* **2017**, *14*, 27–32. [CrossRef]
165. Levy, M.; Thaiss, C.A.; Elinav, E. Metabolites: Messengers between the microbiota and the immune system. *Genes Dev.* **2016**, *30*, 1589–1597. [CrossRef]
166. Blacher, E.; Levy, M.; Tatirovsky, E.; Elinav, E. Microbiome-Modulated Metabolites at the Interface of Host Immunity. *J. Immunol.* **2017**, *198*, 572–580. [CrossRef]
167. Furusawa, Y.; Obata, Y.; Fukuda, S.; Endo, T.A.; Nakato, G.; Takahashi, D.; Nakanishi, Y.; Uetake, C.; Kato, K.; Kato, T.; et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **2013**, *504*, 446–450. [CrossRef]
168. Trompette, A.; Gollwitzer, E.S.; Yadava, K.; Sichelstiel, A.K.; Sprenger, N.; Ngom-Bru, C.; Blanchard, C.; Junt, T.; Nicod, L.P.; Harris, N.L.; et al. Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat. Med.* **2014**, *20*, 159–166. [CrossRef]
169. Postler, T.S.; Ghosh, S. Understanding the holobiont: How microbial metabolites affect human health and shape the immune system. *Cell Metab.* **2017**, *26*, 110–130. [CrossRef]
170. Levy, M.; Blacher, E.; Elinav, E. Microbiome, metabolites and host immunity. *Curr. Opin. Microbiol.* **2017**, *35*, 8–15. [CrossRef]
171. Geva-Zatorsky, N.; Sefik, E.; Kua, L.; Pasman, L.; Tan, T.G.; Ortiz-Lopez, A.; Yanortsang, T.B.; Yang, L.; Jupp, R.; Mathis, D.; et al. Mining the human gut microbiota for immunomodulatory organisms. *Cell* **2017**, *168*, 928–943.e11. [CrossRef]
172. Parfrey, L.W.; Walters, W.A.; Knight, R. Microbial eukaryotes in the human microbiome: Ecology, evolution, and future directions. *Front. Microbiol.* **2011**, *2*, 153. [CrossRef] [PubMed]
173. Bogitsh, B.J.; Carter, C.E.; Oeltmann, T.N. *Human Parasitology*; Academic Press: Cambridge, MA, USA, 2018.
174. Huseyin, C.E.; O'Toole, P.W.; Cotter, P.D.; Scanlan, P.D. Forgotten fungi-the gut mycobiome in human health and disease. *FEMS Microbiol. Rev.* **2017**, *41*, 479–511. [CrossRef]
175. Stensvold, C.R.; van der Giezen, M. Associations between gut microbiota and common luminal intestinal parasites. *Trends Parasitol.* **2018**, *34*, 369–377. [CrossRef] [PubMed]
176. Lukeš, J.; Kuchta, R.; Scholz, T.; Pomajbíková, K. (Self-) infections with parasites: Re-interpretations for the present. *Trends Parasitol.* **2014**, *30*, 377–385. [CrossRef] [PubMed]

177. Durkin, L.; Jansson, T.; Sanchez, M.; Khomich, M.; Ryberg, M.; Kristiansson, E.; Nilsson, R.H. When mycologists describe new species, not all relevant information is provided (clearly enough). *MycoKeys* **2020**, *72*, 109–128. [CrossRef]
178. Hofstetter, V.; Buyck, B.; Eyssartier, G.; Schnee, S.; Gindro, K. The unbearable lightness of sequenced-based identification. *Fungal Divers.* **2019**, *96*, 243–284. [CrossRef]
179. Chin, V.K.; Yong, V.C.; Chong, P.P.; Amin Nordin, S.; Basir, R.; Abdullah, M. Mycobiome in the gut: A multiperspective review. *Mediat. Inflamm.* **2020**, *2020*, 9560684. [CrossRef]
180. Pérez, J.C. Fungi of the human gut microbiota: Roles and significance. *Int. J. Med. Microbiol.* **2021**, *311*, 151490. [CrossRef]
181. Huseyin, C.E.; Rubio, R.C.; O'Sullivan, O.; Cotter, P.D.; Scanlan, P.D. The fungal frontier: A comparative analysis of methods used in the study of the human gut mycobiome. *Front. Microbiol.* **2017**, *8*, 1432. [CrossRef]
182. Kounosu, A.; Murase, K.; Yoshida, A.; Maruyama, H.; Kikuchi, T. Improved 18S and 28S rDNA primer sets for NGS-based parasite detection. *Sci. Rep.* **2019**, *9*, 15789. [CrossRef]
183. Wylezich, C.; Höper, D. Meta-Ribosomalomics: RNA Sequencing Is an Unbiased Method for Parasite Detection of Different Sample Types. *Front. Microbiol.* **2021**, *12*, 614553. [CrossRef] [PubMed]
184. Wylezich, C.; Caccio, S.M.; Walochnik, J.; Beer, M.; Höper, D. Untargeted metagenomics shows a reliable performance for synchronous detection of parasites. *Parasitol. Res.* **2020**, *119*, 2623–2629. [CrossRef] [PubMed]
185. Mirisho, R.; Neizer, M.L.; Sarfo, B. Prevalence of intestinal helminths infestation in children attending princess marie louise children's hospital in accra, ghana. *J. Parasitol. Res.* **2017**, *2017*, 8524985. [CrossRef] [PubMed]
186. Lhotská, Z.; Jirků, M.; Hložková, O.; Brožová, K.; Jirsová, D.; Stensvold, C.R.; Kolísko, M.; Jirků Pomajbíková, K. A Study on the Prevalence and Subtype Diversity of the Intestinal Protist *Blastocystis* sp. in a Gut-Healthy Human Population in the Czech Republic. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 544335. [CrossRef]
187. Ogilvie, L.A.; Jones, B.V. The human gut virome: A multifaceted majority. *Front. Microbiol.* **2015**, *6*, 918. [CrossRef]
188. Lecuit, M.; Eloit, M. The human virome: New tools and concepts. *Trends Microbiol.* **2013**, *21*, 510–515. [CrossRef]
189. Zhang, M.; Zhang, M.; Zhang, C.; Du, H.; Wei, G.; Pang, X.; Zhou, H.; Liu, B.; Zhao, L. Pattern extraction of structural responses of gut microbiota to rotavirus infection via multivariate statistical analysis of clone library data. *FEMS Microbiol. Ecol.* **2009**, *70*, 21–29. [CrossRef]
190. Harper, A.; Vijayakumar, V.; Ouwehand, A.C.; Ter Haar, J.; Obis, D.; Espadaler, J.; Binda, S.; Desiraju, S.; Day, R. Viral infections, the microbiome, and probiotics. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 596166. [CrossRef]
191. Kesika, P.; Sivamaruthi, B.S.; Thangaleela, S.; Chaiyasut, C. The antiviral potential of probiotics—a review on scientific outcomes. *Appl. Sci.* **2021**, *11*, 8687. [CrossRef]
192. Yeoh, Y.K.; Zuo, T.; Lui, G.C.-Y.; Zhang, F.; Liu, Q.; Li, A.Y.; Chung, A.C.; Cheung, C.P.; Tso, E.Y.; Fung, K.S.; et al. Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19. *Gut* **2021**, *70*, 698–706. [CrossRef]
193. Lv, Z.; Xiong, D.; Shi, J.; Long, M.; Chen, Z. The interaction between viruses and intestinal microbiota: A review. *Curr. Microbiol.* **2021**, *78*, 3597–3608. [CrossRef] [PubMed]
194. Robinson, C.M.; Woods Acevedo, M.A.; McCune, B.T.; Pfeiffer, J.K. Related enteric viruses have different requirements for host microbiota in mice. *J. Virol.* **2019**, *93*, e01339-19. [CrossRef] [PubMed]
195. Yaron, J.R.; Ambadapadi, S.; Zhang, L.; Chavan, R.N.; Tibbetts, S.A.; Keinan, S.; Varsani, A.; Maldonado, J.; Kraberger, S.; Tafoya, A.M.; et al. Immune protection is dependent on the gut microbiome in a lethal mouse gammaherpesviral infection. *Sci. Rep.* **2020**, *10*, 2371. [CrossRef] [PubMed]
196. Monedero, V.; Buesa, J.; Rodríguez-Díaz, J. The Interactions between Host Glycobiology, Bacterial Microbiota, and Viruses in the Gut. *Viruses* **2018**, *10*, 96. [CrossRef]
197. Neu, U.; Mainou, B.A. Virus interactions with bacteria: Partners in the infectious dance. *PLoS Pathog.* **2020**, *16*, e1008234. [CrossRef]
198. Nayfach, S.; Camargo, A.P.; Schulz, F.; Eloe-Fadrosh, E.; Roux, S.; Kyrpides, N.C. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat. Biotechnol.* **2021**, *39*, 578–585. [CrossRef]
199. Camarillo-Guerrero, L.F.; Almeida, A.; Rangel-Pineros, G.; Finn, R.D.; Lawley, T.D. Massive expansion of human gut bacteriophage diversity. *Cell* **2021**, *184*, 1098–1109.e9. [CrossRef]
200. Moreno-Gallego, J.L.; Chou, S.-P.; Di Rienzi, S.C.; Goodrich, J.K.; Spector, T.D.; Bell, J.T.; Youngblut, N.D.; Hewson, I.; Reyes, A.; Ley, R.E. Virome Diversity Correlates with Intestinal Microbiome Diversity in Adult Monozygotic Twins. *Cell Host Microbe* **2019**, *25*, 261–272.e5. [CrossRef]
201. Devoto, A.E.; Santini, J.M.; Olm, M.R.; Anantharaman, K.; Munk, P.; Tung, J.; Archie, E.A.; Turnbaugh, P.J.; Seed, K.D.; Blekhman, R.; et al. Megaphages infect Prevotella and variants are widespread in gut microbiomes. *Nat. Microbiol.* **2019**, *4*, 693–700. [CrossRef]
202. Maxwell, K.L. Phages tune in to host cell quorum sensing. *Cell* **2019**, *176*, 7–8. [CrossRef]
203. Leigh, B.A. Cooperation among Conflict: Prophages Protect Bacteria from Phagocytosis. *Cell Host Microbe* **2019**, *26*, 450–452. [CrossRef] [PubMed]
204. Harms, A.; Diard, M. Crowd Controlled-Host Quorum Sensing Drives Phage Decision. *Cell Host Microbe* **2019**, *25*, 179–181. [CrossRef] [PubMed]

205. Jancheva, M.; Böttcher, T. A Metabolite of Pseudomonas Triggers Prophage-Selective Lysogenic to Lytic Conversion in Staphylococcus aureus. *J. Am. Chem. Soc.* **2021**, *143*, 8344–8351. [CrossRef] [PubMed]
206. Adriaenssens, E.M. Phage diversity in the human gut microbiome: A taxonomist's perspective. *mSystems* **2021**, *6*, e0079921. [CrossRef] [PubMed]
207. Wang, P.; Ng, Q.X.; Zhang, H.; Zhang, B.; Ong, C.N.; He, Y. Metabolite changes behind faster growth and less reproduction of Daphnia similis exposed to low-dose silver nanoparticles. *Ecotoxicol. Environ. Saf.* **2018**, *163*, 266–273. [CrossRef]
208. Nagana Gowda, G.A.; Raftery, D. Can NMR solve some significant challenges in metabolomics? *J. Magn. Reson.* **2015**, *260*, 144–160. [CrossRef]
209. Callejón-Leblic, B.; Selma-Royo, M.; Collado, M.C.; Gómez-Ariza, J.L.; Abril, N.; García-Barrera, T. Untargeted Gut Metabolomics to Delve the Interplay between Selenium Supplementation and Gut Microbiota. *J. Proteome Res.* **2022**, *21*, 758–767. [CrossRef]
210. Wang, L.-M.; Wang, P.; Teka, T.; Zhang, Y.-C.; Yang, W.-Z.; Zhang, Y.; Wang, T.; Liu, L.-X.; Han, L.-F.; Liu, C.-X. 1H NMR and UHPLC/Q-Orbitrap-MS-Based Metabolomics Combined with 16S rRNA Gut Microbiota Analysis Revealed the Potential Regulation Mechanism of Nuciferine in Hyperuricemia Rats. *J. Agric. Food Chem.* **2020**, *68*, 14059–14070. [CrossRef]
211. Lai, Y.; Liu, C.-W.; Yang, Y.; Hsiao, Y.-C.; Ru, H.; Lu, K. High-coverage metabolomics uncovers microbiota-driven biochemical landscape of interorgan transport and gut-brain communication in mice. *Nat. Commun.* **2021**, *12*, 6000. [CrossRef]
212. Peisl, B.Y.L.; Schymanski, E.L.; Wilmes, P. Dark matter in host-microbiome metabolomics: Tackling the unknowns-A review. *Anal. Chim. Acta* **2018**, *1037*, 13–27. [CrossRef]
213. Ghannam, R.B.; Techtmann, S.M. Machine learning applications in microbial ecology, human microbiome studies, and environmental monitoring. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 1092–1107. [CrossRef]
214. Marcos-Zambrano, L.J.; Karaduzovic-Hadziabdic, K.; Loncar Turukalo, T.; Przymus, P.; Trajkovik, V.; Aasmets, O.; Berland, M.; Gruca, A.; Hasic, J.; Hron, K.; et al. Applications of machine learning in human microbiome studies: A review on feature selection, biomarker identification, disease prediction and treatment. *Front. Microbiol.* **2021**, *12*, 634511. [CrossRef] [PubMed]
215. Edgar, R.C.; Taylor, J.; Lin, V.; Altman, T.; Barbera, P.; Meleshko, D.; Lohr, D.; Novakovsky, G.; Buchfink, B.; Al-Shayeb, B.; et al. Petabase-scale sequence alignment catalyses viral discovery. *Nature* **2022**, *602*, 142–147. [CrossRef] [PubMed]
216. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef] [PubMed]