An Integrated Multi-Disciplinary Perspective for Addressing Challenges of the Human Gut Microbiome

Rohan M. Shah [1,2] , Elizabeth J. McKenzie [3] , Magda T. Rosin [3], Snehal R. Jadhav [4], Shakuntla V. Gondalia [5], Douglas Rosendale [6] and David J. Beale [2,*]

[1] Department of Chemistry and Biotechnology, Faculty of Science, Engineering and Technology, Swinburne University of Technology, Hawthorn, VIC 3122, Australia; rshah@swin.edu.au
[2] Land and Water, Commonwealth Scientific and Industrial Research Organization (CSIRO), Dutton Park, QLD 4102, Australia
[3] Liggins Institute, The University of Auckland, Grafton, Auckland 1142, New Zealand; liz.mckenzie@auckland.ac.nz (E.J.M.); m.rosin@auckland.ac.nz (M.T.R.)
[4] Centre for Advanced Sensory Science, School of Exercise and Nutrition Sciences, Deakin University, Burwood, VIC 3125, Australia; snehal.jadhav@deakin.edu.au
[5] Centre for Human Psychopharmacology, Swinburne University of Technology, Hawthorn, VIC 3122, Australia; sgondalia@swin.edu.au
[6] Anagenix Ltd., Parnell, Auckland 1052, New Zealand; doug.rosendale@anagenix.com
[*] Correspondence: david.beale@csiro.au; Tel.: +61-73833-5774

Abstract: Our understanding of the human gut microbiome has grown exponentially. Advances in genome sequencing technologies and metagenomics analysis have enabled researchers to study microbial communities and their potential function within the context of a range of human gut related diseases and disorders. However, up until recently, much of this research has focused on characterizing the gut microbiological community structure and understanding its potential through system wide (meta) genomic and transcriptomic-based studies. Thus far, the functional output of these microbiomes, in terms of protein and metabolite expression, and within the broader context of host-gut microbiome interactions, has been limited. Furthermore, these studies highlight our need to address the issues of individual variation, and of samples as proxies. Here we provide a perspective review of the recent literature that focuses on the challenges of exploring the human gut microbiome, with a strong focus on an integrated perspective applied to these themes. In doing so, we contextualize the experimental and technical challenges of undertaking such studies and provide a framework for capitalizing on the breadth of insight such approaches afford. An integrated perspective of the human gut microbiome and the linkages to human health will pave the way forward for delivering against the objectives of precision medicine, which is targeted to specific individuals and addresses the issues and mechanisms in situ.

Keywords: metabolomics; microbiome; omics integration

1. Introduction

Recent advances in culture-independent study techniques of microbial communities, as well as an increasing interest in the role of the gut microbiota in health and disease, have facilitated vast insights into human microbial communities [1]. Identification of prokaryotes mainly with 16S ribosomal RNA (rRNA)-encoding gene sequences [2] and eukaryotes with predominantly Internal Transcribed Spacer (ITS) rDNA sequences [3], coupled with metagenomic analyses, has revealed the ever-increasing diversity list of microorganisms within the human gut [4]. Transcriptomic, proteomic and metabolomic high-throughput tools allow us to begin to grasp the function of the human gut microbiome [5].

Despite these tools, the current state of microbiome research struggles to account for all members of the microbial community, integrate community structure with function, and characteristics of spatial and temporal aspects of the gut microbiota ecosystem [4–6]. Navigating these challenges further complicates our primary task: to design and carry out experiments that generate data which can be integrated, analysed, and interpreted to yield biologically relevant findings for human health.

Most studies tend to focus exclusively on the bacterial component of the human gut, likely omitting the archaea, eukaryotes (including fungi) and viruses within the gastrointestinal tract [7], which has a potential to miss significant microbial intra- and inter-kingdom interaction, and calls into question the relevance of diversity scoring, which is widely reported and used as a consolidated and simplified metric for a vast and complex ecosystem. In the interest of understanding function and mechanisms, microbial metabolites have become a target of investigations, and a sincere international effort has been made to create databases that allow function to be inferred with relatively high confidence from a range of types of sequence data. Inferential tools are invaluable in interpreting high-throughput cost-effective taxonomical datasets, but the gold standard for determining function remains chemometric detection of small molecules in biological samples, which presents additional complications to be accounted for.

Most samples used in the study of the human gut microbiome are home-collected after defecation by the participant rather than endoscopically-collected from the gut since the cost and feasibility of self-collection outweighs in vivo procedure [8]. Faecal samples provide data on the distal microbiome of the gastrointestinal tract (GIT), which is not representative of the proximal microbiome. Although metagenomics of home- and endoscopically-collected samples show little variation between the samples, volatile organic compounds (VOC) in the metabolome differ between the in vivo and ex vivo faecal samples [8], which is important when considering the self-collected faecal samples metabolome as a proxy for the functional reflection of the gut microbiome [9].

Despite these fundamental limitations, either metabolic activity initiated by the host or the gut microbiota can lead to marker metabolites in different biological fluids that allow differentiation between healthy and disease status. Metabolomics may be viewed as a process in which non-targeted metabolomics, or global metabolite profiling, aims to discover a list of candidate metabolites of interest, or biomarkers, which can be further quantified and validated using targeted metabolomics. Accounting for vast individual variation in the microbiome and metabolome has led to the development of alternative experimental designs, such as 'N of 1,' which bypasses the challenge of increasing noise while attempting to increase a signal by increasing the 'N' of a study. Similarly, this approach addresses the challenge of applying traditional statistics based on power calculations and reproducibility to the multivariate and non-parametric datasets that represent microbiome structure and function. At present, most microbiome findings present correlations rather than causation, which can only be overcome with a solid grasp on the functional relationship between host and microbiome.

To further investigate different perspectives of this area, a peer session on 'metabolomics and its application in gut microbiome research' was held during the most recent Australian and New Zealand Metabolomics (ANZMET) conference (Auckland, New Zealand, from 30 August to 1 September 2018). Over 20 metabolomics-based researchers attended the session and participated in the discussion on how functional omics could be used to advance gut microbiome research. In addition to summarizing the key points from the peer session, we provide an overview of the current challenges when researching the gut microbiome, from the perspective of the metabolomics community. In should be noted that challenges relating to specific sampling and storage conditions or techniques and approaches are not captured in this review. Where possible, the interested reader is directed to other published works for this information.

2. Biogeography of the Microbiome and Metabolome: Implications for Faecal Samples as Proxies

Human cells live in coexistence with a vast and diverse collection of symbiont microorganisms referred to as the human microbiota or the microbiome. The GIT is, by far, the most heavily colonized organ in the human body, with a large surface area and a consistent nutrient source of digested food

for microbes to utilise, making it a preferred site for microbial colonisation. The GIT is also one of the most studied microbial ecosystems [10–13]. Although research to date has mainly focused on the bacterial component of the human gut, microorganisms belonging to Archaea and Eukarya domains of life, as well as viruses, also constitute the human microbiome [14–16]. A rich and complex ecosystem of bacteria, fungi, viruses, archaea, protists and (sometimes) helminths flourishes in the human gut.

Among the trillions of microorganisms that make up the gut microbiome, commensal bacteria are predominant and distributed throughout the GIT. Recent endeavours in gut microbiome research using metagenomics has provided a strong understanding of bacterial communities in largely diverse environments [17–23]. Although composed of strict anaerobic bacteria from over 50 different phyla, the *Firmicutes* and the *Bacteroidetes* are the two dominant phyla in the human gut. The members of other phyla such as *Proteobacteria*, *Verrucomicrobia*, *Actinobacteria*, *Fusobacteria*, and *Cyanobacteria* are present in minor proportions. The number of bacteria increases from 10 cells/g of contents in the stomach to 10^{12} cells/g in the colon (Figure 1). The human gut microbiota is not homogenous and as such, a wide variation in microbial composition between these sites is reported (Figure 1).
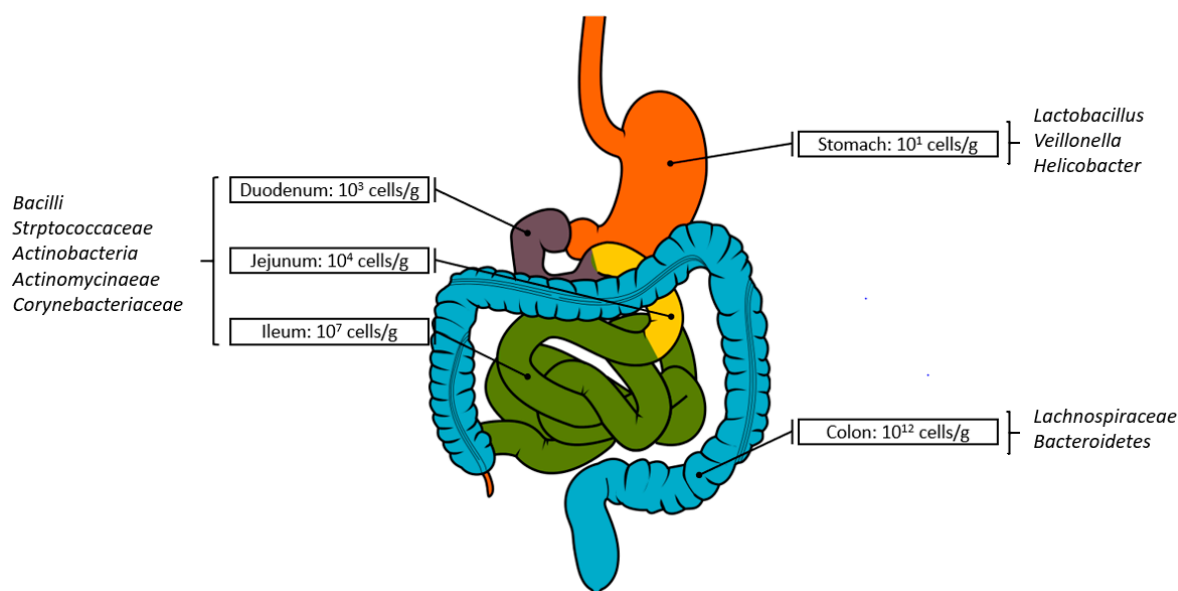

Figure 1. Variations in bacterial number and composition across the length of the gastrointestinal tract (GIT). Image: Olek Remesz (wiki-pl: Orem, commons: Orem) (https://commons.wikimedia.org/wiki/File:GISystem.svg), "GISystem", Text modified/overlaid by Shah et al., https://creativecommons.org/licenses/by-sa/2.5/legalcode.

The degree of richness, complexity and function of the regional gut microbiome is highly dependent on the microenvironmental conditions, nutrients, oxygen and water availability, as well as host site-specific characteristics [24]. The lowest microbial counts are observed in the stomach and upper small intestines, due to localized harsh conditions (i.e., high acidity, high bile acid concentrations and short retention time) [25]. A gradual increase in microbial numbers towards the distal ileum and within the colon reflects a more tolerant microenvironment that permits colonization [26]. Bacteria residing in the human colon are thought to be the most substantial contributors to the total human microbiome population, with an estimate of 3.8×10^{13} cells [27] that influence physiological processes, both in health and disease [2].

Faecal samples are used as proxies for the fermenta within the colon. Highly developed biofilm microbiota closely associated with the intestinal mucosa are generally believed to be more relevant than the planktonic microbes that exist in the lumen gut. As such, faecal samples may not present an accurate snapshot of the mucosal microbiota. This is a necessary compromise in the absence of routinely used endoscopy to sample the contents of the colon, even though we know that endoscopically collected

(in vivo) samples have different microbiome profiles and markedly different metabolome profiles than faecal (ex vivo) samples [8]. Endoscopic biopsies also have their own limitations, such as fasting and colon cleansing prior to endoscopic collection, and the contamination of tissue specimens by luminal microbes during the process. Faecal sampling is non-invasive, and thereby more accessible to researchers. The treatment and handling of samples after collection is a critical aspect of ongoing microbiome studies. Lauber, et al. [28] studied the effects of storage conditions on human faeces using 16s rRNA pyrosequencing. The results indicated that the microbial composition was not significantly affected by short-term storage of up to 14 days at $-80\,°C$, $-20\,°C$, $4\,°C$ and $20\,°C$. In another study, Wu, et al. [29] showed that there was no significant difference between the faecal samples immediately frozen at $-80\,°C$ and those stored on ice for up to 48 h. On the contrary, Roesch, et al. [30] revealed that the stability of faecal samples may be compromised when stored at room temperature for more than 12 h. For more information on the collection and storage of microbiome related samples, the reader is directed to the reports by Panek et al. [31] and Vanderputte et al. [32].

However, despite their availability, we need to consider how and what faecal samples proxy for. Faecal samples are not end-points of a determinative process, they are snapshots of a continuously developing and evolving microbial ecosystem, where the ecological stage encapsulated in that snapshot faecal sample is driven by transit time and stool consistency (water content) [33] Differences in transit represents different rates of passage of material through the colon, and hence different ages of the developing ecology, and these ages in turn are representative of the physical, chemical and microbiological continuum throughout the length of the colon. We summarize these conditions in Figure 2, where faecal snapshots could vertically bisect any given point on the continuum. Thus, faecal samples may proxy for quite different gut conditions.
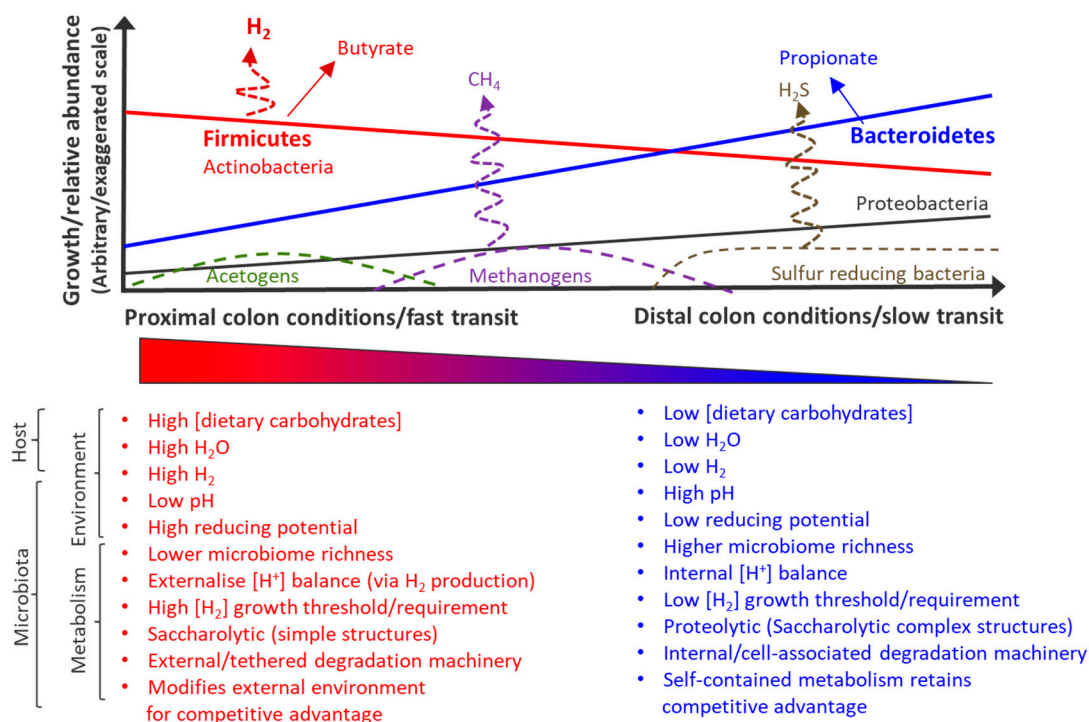

Figure 2. Schematic representation of conditions gradient in the colon determining state of gut ecosystem and how this effects faecal snapshots. Local/regional conditions in the gut graphically represented, left to right, from entry (proximal colon) to exit (distal colon).

The implications of this are that different relative abundances of microbes and their respective metabolites are present. Firstly, this may explain much of the inter-individual differences observed in microbiome research, and furthermore, simply explain much intra-individual differences over time. As an aside, the intersubject variation in microbial metabolites argues in favour of cross over designs in intervention studies and, therefore, individuals can act as their own controls. Nevertheless, if these snapshot dynamics can be modelled and standardized, we would be able to account for much of the observed variation (at least at a metabolic or ecological niche or functional microbial guild level). Potentially, we may unmask common mechanisms thus far obscured by variation. Secondly, this means that with different relative abundances of microbes, metabolite profiles will obviously be different. The most obvious consequence of this at the bacterial phylum level is that a snapshot representative of more proximal colonic conditions rich in Firmicutes would be higher in butyrate and lower in propionate, whilst a snapshot representative of more distal conditions proportionately richer in Bacteroidetes would have higher concentrations of propionate and likely be comparatively depleted in butyrate (Figure 2). Obviously, other metabolites would also be different, as befits the other interconnecting microbial- and host-driven environmental conditions (Figure 2). This is commonly observed with the other fermentation by-products.

To illustrate this, as a part of core metabolism, fermentation of sugars through glycolysis has the cofactor NAD^+ converted to $NADH + H^+$. Regeneration of NAD^+ can become the rate-limiting step in glycolysis, so is balanced by the rapid regeneration and commensurate generation of lactate, acetate, formate, succinate and ethanol [34]. Thus, acid production is a necessity driven by competition for abundant energy resources. Bacteria in a dietary carbohydrate-richer environment will maintain their competitive advantage by rapidly pushing sugar-derived carbons through glycolysis and out as waste products to deny their competitors that same raw resource. Alternative pathways will be employed as variations of this theme, where trade-offs between energy yield and protein cost are made [35]. In vitro experiments show that under the right carbohydrate-rich conditions, a single bacterium can dominate a faecal microbiome, with commensurate impact on the short-chain fatty acid (SCFA) profile [36]. Firmicutes will perform this rapidly and profligately, as they are adapted to surviving in the acidic conditions they generate, consistent with their food-rich environment. In contrast, members of the Bacteroidetes phylum are less acid tolerant, and will metabolise sugars slower, somewhat maintaining this NAD^+-requiring redox balance within their cytoplasm, thereby having less impact on their (food-depleted) environment. Their competitive advantage lies with flexibility [37] and priority [38] in terms of choice of substrate. The take-home message from this recap of fundamental metabolism is that the host-derived conditions (dietary carbohydrate, water) drive the microbiome, which in turn drive the conditions (H^+, pH) in a teleological fashion, in accordance with the scheme outlined in Figure 2. Obviously, secondary metabolites are dependent on this primary metabolism and the microbiome driving it within that faecal snapshot.

How bacteria and their metabolism illustrate the stage and/or location of the gut ecosystem is particularly evident in the case of the H_2-utilizing bacteria and archaea. When these organisms are the dominant H_2 utilizer, they tend to maintain H_2 levels at the lowest threshold they require for growth [39]. Since the methanogen threshold is lower than the acetogen threshold, and the sulphur reducing bacteria threshold is one or more orders of magnitude lower than the methanogen threshold, faecal snapshots really will represent how this cyclic feedback between the factors like transit, water, carbohydrates and the host and microbial influences propagate that local environment at that time and place. Fortunately, this suggests that simply the relative presence or abundance of these H_2-utilizing microbes and their metabolic activities act as biomarkers that, with sufficient experience, calibration or learning data sets, may allow the investigator to back-calculate or project what the conditions in the rest of the colon are/would be like on the basis of this faecal snapshot.

Examples of successful back-calculation using faecal metabolites is the case of the SCFAs acetate, butyrate and propionate. Butyrate, produced by some Firmicutes, is rapidly absorbed and used by the colonic epithelia as an energy source, which, given the high metabolic activity of the gut, translates to as much as 10% of the body's energy coming from microbial butyrate. Butyrate is transported into colonocytes and diffuses into the mitochondria where it undergoes β-oxidation to acetyl-CoA which then enters the TCA cycle resulting in reduction of NAD^+ to NADH, the latter entering the electron transport chain for ATP production [40]. Propionate, predominantly from Bacteroidetes, acts similarly, entering the TCA cycle through succinyl-CoA [40]. The path from ubiquitously produced microbial acetate is directly through acetyl-CoA. This rapid absorption of these acids, particularly butyrate, means that correlations between these and host metabolic markers do not correlate [41]. However, when the uptake or flux of these acids is calculated (parameters obtained through ^{13}C acetate, propionate and butyrate infusion of mice caeca), uptake fluxes correlate linearly with host metabolic markers [41]. Known or estimated SCFA flux rates may have potential to assist with correcting or standardizing other metabolite concentrations from faecal snapshots as proxies for colonic contents.

3. The Role of Microbiome Structure and Function in Human Health and Disease

Intense clinical investigation of the human gut microbiome has revealed a sophisticated interplay between the microbiome and the host immune system and metabolism. The role of gut microbiome in accomplishing protective, structural and metabolic functions in human hosts is well documented. These include, but are not limited to, defence against colonization by harmful or pathogenic organisms [2,42,43], digestion of food [44–47], nutrition [46,48,49] and maintenance of a healthy immune system [49–54]. Any perturbations in the gut microbiome may result in dysbiosis and can further lead to a variety of phenotypes including obesity, inflammatory bowel disease (IBD), type II diabetes, fatty liver disease, cancer and several additional human disease states or disorders (Table 1). The role of the human gut microbiome in disease development and progression has become a growing research field in the recent years, yet the cause and effect of the gut dysbiosis and human health has not been well documented [2].

Table 1. Changes in the gut microbiome associated with disorder/disease.

Disease/Disorder	Positively Implicated Members of Microbiota (↑)	Negatively Implicated Members of Microbiota (↓)	Reference
Colorectal cancer	*Fusobacterium* *Porphyromonas*	*Clostridium* *Bacteroides* *Lachnospiraceae*	[55]
Colitis-associated colorectal cancer	*Bifidobacterium* *E. coli*		[56]
Cirrhosis	*Enterococcaeae* *Staphylococcaceae* *Enterobacteriaceae*	*Clostridiales XIV* *Ruminococcaceae* *Lachnospiraceae* *Veillonellaceae* *Porphyromonadaceae*	[57]
Non-alcoholic fatty liver disease and steatohepatitis	*Ruminococcus* *Dorea*	*Oscillospira*	[58]
Celiac's disease	*Bacteroides vulgatus* *Escherichia coli*	*Clostridium coccoides*	[59]
Gastric cancer	*Helicobacter pylori*		[60]
Autism	*Bacteroidetes* *Proteobacteria*	*Actinobacteria* *Firmicutes*	[61–63]
Parkinson's Disease	*Enterobacteriaceae*	*Prevotellaceae*	[64]
Type 2 diabetes	*Betaproteobacteria*	*Firmicutes* *Clostridia*	[65,66]
IBD – Crohn's Disease	*Bacteroides ovatus* *Bacteroides vulgatus*	*Bacteroides uniformis*	[67]
IBD – Ulcerative colitis	*Gammaproteobacteria* *Deltaproteobacteria* *Actinobacteria* *Proteobacteria*	*Firmicutes*	[68–70]

The prevalence of technologies (sequencing, databases, and analytical paradigms) that characterise the microbiome community, rather than elucidating mechanisms, has led to an abundance of studies that associate patterns, for instance species diversity or species richness, with disease states. However, despite this focus on community rather than mechanism, robust associations have laid the foundation for further exploration. For example, functional gastrointestinal disorders (FGIDs), defined as disorders of the gut-brain interactions, have been linked, at least partly, to altered gut microbiome and immune dysregulation [71]; however, no clear association between different microbial patterns and FGIDs has been drawn [72,73]. The gut and brain appear to communicate via the neural and hormonal signalling, the immune system, and via microbial metabolites such as SCFAs [74,75], and perturbations of this bi-directional relationship between the colonic microbiome and central nervous system, coined the 'gut-brain axis', may manifest in neurological conditions such as anxiety, depression, autism spectrum disorders, and influence mood and social behaviour [74,76–80]. Multifactorial autoimmune diseases have also been explored in the light of the gut microbiome, where the breakdown of intestinal epithelial barrier and failure of the gut mucosal immunity allows for microbial cells or their metabolites to trigger systemic inflammation [81–84]. The gut microbiome may also activate or inhibit natural, systemic anti-tumour immunosurveillance [85,86] or induce the formation of local cancers [87–89], as well as influence the efficacy of some chemotherapy treatments [90,91]. In cardiovascular diseases [92], atherogenic trimethylamine N-oxide (TMAO) produced in the liver from microbially-derived trimethylamine (TMA) and cardioprotective SCFAs [93,94] are of particular interest. An increase in cardiovascular disease, as well as type 2 diabetes risk, has also been linked with metabolic syndrome, which in turn appears to be influenced by the gut microbiome at several levels [95]. The previous notion that a simple change in the ratio of Bacteroidetes to Firmicutes in the human gut contributes to the development of obesity, an aspect of metabolic syndrome, has been recently challenged and appears to be much more complex [96].

4. Analysing Microbiome Structure and Function with Non-Metabolomics Approaches

Most of the microbes residing in the gut are strictly anaerobic, so their isolation and cultivation in laboratory conditions is challenging, with almost 75 percent of the gut microbiome uncultivable [97]. Culture independent methodologies such as small subunit (SSU) rRNA gene amplicon sequencing (16S SSU for bacteria, 18S SSU for eukaryotes and ITS in fungi) and whole-metagenome sequencing have helped overcome this limitation to a great extent [98]. The 16S rRNA sequencing is a more rapid method for assessment of overall phylogeny and diversity of a bacterial community. As such, the 16S rRNA method may provide information on the composition of the gut microbiome; however, it does not always provide a clear link between the microbes identified and their functions in the gut [99]. Recent development of next generation sequencing (NGS) tools has greatly advanced the high-throughput metagenomics approach, with several software platforms for comparative analyses on the gene level developed. These include, Integrated Microbial Genomes with Microbiome Samples (IMG/M) [100], MicrobesOnline [101], Microbial Genomes database (MBGD) [102], Roary [103], EzBioCloud [104], OrtholugeDB [105] and Efficient Database framework for comparative Genome Analyses using BLAST score Ratios (EDGAR) [106]. The IMG/M software is one of the largest platforms containing annotated bacterial, archaeal and metagenomic sequence data [107].

Considering the importance of understanding the functional capacity of the microbiome and the low cost-effectiveness of metagenomics approaches, an alternative could be to use the 16S rRNA gene profiles for predicting the functions of the microbial communities [108]. One such predictive tool is Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUST). This tool is based on over 39,000 reference genomes [109], and relies on the availability of fully characterised bacterial genomes and uses their phylogenetic relationships to predict the functional capacity of other genomes. This tool has been used and validated in the Human Microbiome Project.

Odamaki, et al. [110] used PICRUST analysis to study the age-related changes in gut microbiota composition in healthy individuals of various age groups from newborn to centenarian. It is important to understand that while both metagenomics and PICRUST provide a functional hypothesis of the gut microbiome, it still needs validation via the use of specific primers or metabolomic analysis. Tax4Fun is another such tool that predicts functional capabilities for prokaryotes in the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database [111]. The most recent addition to these 16S data analysis pipelines is Piphillin, that overcomes some of the limitations associated with PICRUST, such as its dependence on outdated functional databases and specific data pre-processing tools [99].

A major weakness in 16S research is that 16S profiling is vulnerable to bias from diverse sources. The universal primers are not truly 'universal'. The universal primer sets tend to underperform when they encounter particular mismatches that undermine hybridization to their target sequence. This may result in the under-amplification of certain organisms. Moreover, they do not capture viruses and archaea [109] or eukaryotes. An overwhelming majority of gut microbiome studies have primarily focused on bacterial flora, to characterize their composition and association with human health and diseases. The gut archeome, mycobiome, virome and eukaryome have received less attention until recently.

Archaea that reside in the human colon are nearly always strictly anaerobic methanogens; most of which belong to the order *Methanobacteriales* and the most common genera being the closely related *Methanobrevibacter* and *Methanosphaera* [112]. To date, three species of methanogenic archaea have been isolated from human faeces [113]: *Methanobrevibacter smithii* [114], *Methanosphaera stadtmanae* [115] and *Methanomassiliicoccus luminyensis* [116]. *M. smithii* has been found to inhabit in almost 95.7% of humans and is the most abundant methanogen in the human gut. When prevalent, it may control H^2 concentrations [39]. Bang, et al. [117] recently reported that *M. smithii* and *M. stadtmanae* induce monocyte-derived dendritic cell maturation; *M. stadtmanae* leads to substantial release of pro-inflammatory cytokines in these cells. Lecours, et al. [118] indicated an increased prevalence of *M. stadtmanae* in IBD. Recent studies support an association of *M. smithii* with leanness [119–121]. An increased prevalence of methanogens may cause chronic constipation [122]. There is strong evidence that there is a lower prevalence of methanogens in patients that tend to have diarrhoea episodes (such as those with IBD) [123].

The human gut mycobiome is a neglected component of microbiota for several reasons, including lack of stability and low abundance and diversity [3,124]. Interactions between fungi and bacteria are common, but are complex and may have dramatic effects on growth and pathogenesis of micro-organisms [125]. Approximately 247 fungal species belonging to 126 genera have been identified in faeces and GI biopsies [126]. Dollive, et al. [127] found *Aspergillus*, *Cryptococcus*, *Penicillium*, *Pneumocystis* and *Saccharomycetaceae* yeasts (*Candida* and *Saccharomyces*) in the GIT of healthy individuals. Fungi have been associated with a number of GIT diseases including IBD [128,129], peptic ulcers [130], irritable bowel syndrome (IBS) [131], antibiotic-associated diarrhoea [132] and chemotherapy-induced enteric disorders [133].

The enteric virome includes viruses that infect host cells, endogenous retroviruses, and viruses that infect the various microbial inhabitants of the GIT, such as bacteria, archaea, and fungi. As such, there is immense complexity in coding potential of gut virome and has received much less attention as compared to bacterial flora [134]. Bacteriophages are the most abundant and diverse members of gut virome and are most likely to have a substantial impact on the host [135]. The gut virome plays an important role in the pathogenesis of dysbiosis [136]. The gut virome has also been associated with intestinal disorders such as IBD [137], Crohn's disease (CD) [138] and colon cancer [139].

Many studies investigating the gut microbiome have used metagenomics. Although this is a powerful technology, alone it suffers from the same limitations as other unintegrated omics technologies: (a) inability to identify microbial sources, (b) expensive and time-consuming, (c) presence of human contaminants in samples and (d) lack of functional annotations of outputs [140], although the latter

is rapidly improving [141]. With rapid development and integration of the other omics techniques, such as metatranscriptomics, metaproteomics and metabolomics, the functional activity of the gut microbiome can be better identified.

A metatranscriptomics approach is used to study gene activity. Gosalbes, et al. [142] investigated the faecal samples from ten healthy individuals and identified the key functions of gut microbiome—carbohydrate metabolism, energy production and synthesis of cellular components. Several housekeeping functions such as amino acid metabolism and lipid metabolism were under-represented in the gut metatranscriptome. Franzosa, et al. [143] collected the stool samples from eight healthy individuals in order to relate the gut metagenome and metatranscriptome. About 59% of microbial transcripts were differentially regulated relative to their genomic abundances. Sporulation and amino acid biosynthesis were consistently downregulated, and ribosome biogenesis and methanogenesis were consistently upregulated.

Functional activity can be studied using a metaproteomics approach. Verberkmoes, et al. [144] investigated faecal samples from a female healthy monozygotic twin pair by shotgun metaproteomics approach. Several proteins required for translation, energy production and carbohydrate metabolism were identified in faecal samples. Erickson, et al. [145] combined shotgun metagenomics and metaproteomics approaches to identify potential functional signatures of CD in stool samples from six twin pairs that were either healthy, or that had CD. Studies have shown higher similarity in gut microbiota between healthy twins than between unrelated individuals. By contrast, twin pairs in which one or both individuals had CD indicated very dissimilar gut microbiome. Integration of omics approaches revealed ileum CD phenotype was associated with alterations in bacterial carbohydrate metabolism, bacterial–host interactions, as well as human host-secreted enzymes. A study by Kolmeder, et al. [146] revealed that the faecal metaproteome in healthy individuals was subject-specific. The functional metaproteome core was stable over a year and was mainly involved in carbohydrate and degradation.

Table 2 outlines many bioinformatic analytical processes applied to metagenomics and metatranscriptomics data. This information may help us undertake a more holistic approach to understanding the functions of the gut in overall human health, especially in case of diseases such as inflammatory bowel disease, irritable bowel syndrome, and obesity.

Table 2. Computational methods for meta-omic analysis (modified from [98]).

Method	Tool	Description	Reference
Assembly	DIME	Combines the DIvide, conquer, and MErge strategies	[147]
	Genovo	Generative probabilistic model of reads	[148]
	Khmer	Probabilistic de Bruijn graphs	[149]
	MAP	OLC (Overlap/Layout/Consensus) strategy for longer reads	[150]
	Meta-IDBA	De Bruijn graph approach	[151]
	metAMOS	A Modular Open-Source Assembler component for metagenomes	[152]
	MetaVelvet	De Bruijn graph approach	[153]
	MOCAT	a metagenomics assembly and gene prediction toolkit	[154]
	SOAPdenovo	Single-genome assembler commonly tuned for metagenomes	[155]
	MetaORFA	Gene-targeted assembly approach	[156]
	MetaPAR	Metagenomic sequence assembly via iterative reclassification	[157]
	XGenovo	An extended Genovo assembler by incorporating paired-end information	[158]

Table 2. *Cont.*

Method	Tool	Description	Reference
Taxonomic profiling	Amphora	Automated pipeline for Phylogenomic Analysis	[159–161]
	CARMA3	Taxonomic classification of metagenomic shotgun sequences	[162,163]
	ClaMS	Classifier for Metagenomic Sequences	[164]
	CLARK	Fast and accurate classification of metagenomic and genomic sequences using discriminative k-mers	[165]
	DiScRIBinATE	Distance Score Ratio for Improved Binning and Taxonomic Estimation	[166]
	FOCUS	An agile composition-based approach using non-negative least squares	[167]
	INDUS	Composition-based approach for rapid and accurate taxonomic classification of metagenomic sequences	[168]
	MARTA	Suite of Java-based tools for assigning taxonomic status to DNA sequences	[169]
	MetaCluster	Binning algorithm for high-throughput sequencing reads	[170]
	MetaPhlAn	Profiles the composition of microbial communities from metagenomic shotgun sequencing data	[171,172]
	MetaPhyler	Taxonomic classifier for metagenomic shotgun reads using phylogenetic marker reference genes	[173]
	MOCAT2	A metagenomic assembly, annotation and profiling framework	[174]
	MTR	Taxonomic annotation of short metagenomic reads using clustering at multiple taxonomic ranks	[175]
	NBC	Naive Bayes Classification tool for taxonomic assignment	[176]
	PaPaRa	Aligning short reads to reference alignments and trees	[177]
	PhyloPythia	Accurate phylogenetic classification of variable-length DNA fragments	[164]
	PhyloSift	Phylogenetic analysis of metagenomic samples	[178]
	Phymm	Classification system designed for metagenomics experiments that assigns taxonomic labels to short DNA Reads	[179]
	RAIphy	Phylogenetic classification of metagenomics samples using iterative refinement of relative abundance index Profiles	[180]
	RITA	Classifying short genomic fragments from novel lineages using composition and homology	[181]
	SOrt-ITEMS	Sequence orthology-based approach for improved taxonomic estimation of metagenomic sequences	[182]
	SPHINX	Algorithm for taxonomic binning of metagenomic sequences	[183]
	TACOA	Taxonomic classification of environmental genomic fragments using a kernelized nearest neighbour approach	[184]
	Treephyler	Fast taxonomic profiling of metagenomes	[185]
Functional profiling	HUMAnN	Determines the presence/absence and abundance of microbial pathways in meta-omic data	[186]
	metaSHARK	web platform for interactive exploration of metabolic networks	[187]
	MOCAT2	A metagenomic assembly, annotation and profiling framework	[174]
	PRMT	Predicted Relative Metabolomic Turnover: determining metabolic turnover from a coastal marine metagenomic dataset	[188]
	RAMMCAP	Rapid analysis of Multiple Metagenomes with Clustering and Annotation Pipeline	[189]

Table 2. *Cont.*

Method	Tool	Description	Reference
Interaction networks	SparCC	Estimates correlation values from compositional data for network inference	[190]
	CCREPE	Predicts microbial relationships within and between microbial habitats for network inference	[191]
Single-cell sequencing	IDBA-UD	De Bruijn graph approach for uneven sequencing depths	[192]
	SmashCell	Software framework for the analysis of single-cell amplified genome sequences	[193]
Simulators	GenSIM	Error-model based simulator of next-generation sequencing data	[194]
	Metasim	A sequencing simulator for genomics and metagenomics	[195]
Statistical tests	Metastats	Statistical analysis software for comparing metagenomic samples	[196]
	LefSe	Nonparametric test for biomarker discovery in proportional microbial community data	[197]
	ShotgunFunctionalizeR	A statistical test based on a Poisson model for metagenomic functional comparisons	[198]
	SourceTracker	A Bayesian approach to identify and quantify contaminants in a given community	[199]
General toolkit	CAMERA	Dashboard for environmental metagenomic and genomic data, metadata, and comparative analysis tools	[200]
	GenBoree	A web-based platform for multi-omic research and data analysis using the latest bioinformatics tools	[201]
	GraPhlAn	Compact graphical representation of phylogenetic data and metadata	[202]
	IMG/M	Integrated metagenome data management and comparative analysis system	[100]
	MEGAN	Software for metagenomic, metatranscriptomic, metaproteomic, and rRNA analysis	[203]
	METAREP	Online storage and analysis environment for meta-omic data	[204]
	MG-RAST	Storage, quality control, annotation and comparison of meta-omic samples	[205]
	Mothur	An open-source software for microbial ecology community analysis	[206]
	QIIME	An open-source bioinformatics pipeline for performing microbiome analysis from raw DNA sequencing data	[207]
	SmashCommunity	Stand-alone annotation and analysis pipeline suitable for meta-omic data	[208]
	STAMP	Comparative meta-omics software package	[209]
	SnoWMan	Pipeline for analysis of microbiome data	[210]
	VAMPS	Visualization and analysis of microbial population structure	[211]

CC BY-NC; Segata et al. Molecular systems biology 2013, 9, 666.

5. Analysing Microbiome Function with Metabolomics

Metabolomics is ideally placed as the foundation for a systems biology approach in the study of the gut microbiome, primarily because metabolites are involved in biological processes at all levels, driving activity [212] and inter-kingdom communication [213] at the level of the proteome, transcriptome, epigenome, and genome.

Metabolomics provides insights into the molecular mechanisms of microbiome-host intersection which have the potential to be exploited as the predictive tool for dysbiosis, microbiome-metabolome disease signatures, and for the discovery of biomarkers to be used to either diagnose the disease or monitor activity of therapeutics [214]. The symbiotic relationship resulting from the coevolution between microbiota and the human host [215] has been particularly illustrated in the production of SCFA from non-digestible dietary fibres that reach the colon by a range of bacterial species [216]. These volatile metabolites are not only important energy sources for microbial communities and the host, but likely play essential roles in maintaining gut epithelial integrity via tight junction regulation [217]; glucose homeostasis, lipid metabolism and short-term appetite suppression via Peptide YY (PYY) and glucagon-like peptide 1 (GLP-1) signalling pathways [218,219]; and immune function regulation [220,221]. Furthermore, gut microbes are involved in the elimination of toxic compounds [222], synthesis of essential vitamins [223] and may metabolize and influence the bioavailability of other nutritive and non-nutritive components of functional foods and prebiotics [224]. This symbiotic relationship between the gut microbiome, their metabolic products and the human host may be disturbed in disease states [225].

Metabolomics and metabolic profiling are increasingly used in the identification of biomarkers of several GIT disorders including IBD [226–228] and colorectal cancer [229]. For instance, increased levels of cadaverine and taurine were found in patients with ulcerative colitis; while higher levels of bile acids and lower concentrations of branched-chain fatty acids were detected in patients with IBS [228]. Finegold, Dowd, Gontcharova, Liu, Henley, Wolcott, Youn, Summanen, Granpeesheh and Dixon [61] identified lower levels of total short-chain fatty acids, including lower levels of acetate, propionate and valerate in children with autism. Some of the main chemical classes that regulate host-gut microbiome interactions are listed in Table 3.

Table 3. Metabolites associated with gut microbiome (modified from [230] and [231]).

Metabolite Class	Metabolites	Related Bacteria	Biological Functions	Disease/Disorder	Reference
Short-chain fatty acids	• acetate, • propionate, • 2-methylpropionate, • butyrate, • isobutyrate, • hexanoate, • valerate, • isovalerate	• All, although capabilities vary amongst phyla, family and species.	• Cholesterol synthesis ↑ (acetate); • Gluconeogenesis ↑ (propionate); • Energy source for colonocytes ↑ (butyrate); • Colonic pH ↓; • Growth of pathogens ↓; • Water and sodium absorption ↑	• Human obesity, insulin resistance and type 2 diabetes, • colorectal cancer; cardiovascular disease, • IBD - ulcerative colitis, • IBD -Crohn's disease, • antibiotic-associated diarrhoea, • metabolic syndrome, • bowel disorders	[40,232,233]
Phenolic, Benzoyl and phenyl derivatives	• benzoate, • hippurate, • phenylacetate, • phenylpropionate, • 3-hydroxycinnamate, • 2-hydroxyhippurate, • 3-hydroxyhippurate, • 2-hydroxybenzoate, • 3-hydroxybenzoate, • 4-hydroxybenzoate, • 4-hydroxyphenylacetate, • 3-hydroxyphenylpropionate, • 4-hydroxyphenylpropionate, • 3,4-dihydroxyphenylpropionate, • 4-methylphenol, • 4-cresol, • 4-cresyl sulfate, • 4-cresyl glucuronide, phenylacetylglycine, • phenylacetylglutamine, phenylacetylglycine • phenylpropionylglycine, cinnamoylglycine	• *Clostridium difficile, Faecalibacterium prausnitzii, Bifidobacterium, Subdoligranulum, Lactobacillus*	• Detoxification of xenobiotics; indicate gut microbial • composition and activity; • utilize polyphenols.	• Hypertension, • Obesity, • Colorectal cancer, • Autism	[230,234–240]

Table 3. *Cont.*

Metabolite Class	Metabolites	Related Bacteria	Biological Functions	Disease/Disorder	Reference
Bile salts	• cholate, • hyocholate, • deoxycholate, • chenodeoxycholate, • α-muricholate, • β-muricholate, • ω-muricholate, • taurocholate, • glycocholate, • taurochenoxycholate, • glycochenodeoxycholate, • taurocholate, • tauro–α–muricholate, • tauro–β–muricholate, • lithocholate, • ursodeoxycholate, • hyodeoxycholate, • glycodeoxylcholate, • taurohyocholate, • taurodeoxylcholate	• *Lactobacillus, Bifidobacterium,* • *Enterobacter, Bacteroides,* • *Clostridium,* • *Enterobacter, Escherichia*	• Absorption of dietary fats and • Lipid-soluble vitamins, • facilitate lipid assimilation, • maintain gut barrier function, • regulate triglycerides, • cholesterol and glucose by endocrine functions, • energy homeostasis.	• Colon cancer	[241–243]
Choline metabolites	• methylamine, • dimethylamine, • trimethylamine, • trimethylamine-N-oxide, • dimethylglycine, • betaine	• *Faecalibacterium prausnitzii,* • *Bifidobacterium*	• Lipid metabolism, • Glucose homeostasis	• Non-alcoholic fatty liver disease, • Dietary-induced obesity, • diabetes, • cardiovascular disease	[244,245]

Table 3. *Cont.*

Metabolite Class	Metabolites	Related Bacteria	Biological Functions	Disease/Disorder	Reference
Indole derivatives	• N-acetyltryptophan • indoleacetate • indoleacetylglycine (IAG) • indole • indoxyl sulphate • indole-3-propionate • melatonin • melatonin 6-sulfate • serotonin • 5-hydroxyindole	• *Clostridium sporogenes*, • *Escherichia coli*	• Protect against stress-induced lesions in the GI tract; • modulate expression of proinflammatory • genes, • increase expression of anti-inflammatory genes, • strengthen epithelial cell barrier • properties	• GI pathologies, brain-gut axis, • Neurological conditions	[246,247]
Vitamins	• Vitamin K • cobalamin (Vitamin B12) • biotin (Vitamin B8) • folate (Vitamin B9) • thiamine (Vitamin B1) • riboflavin (Vitamin B2) • pyridoxine (Vitamin B6) • niacin (Vitamin B3) • pantothenic acid (Vitamin B5)	• *Bifidobacterium* • Commensal *Lactobacill* • *Bacillus* • *Subtilis, Escherichia coli* and Anaerobes • *Bacteroidetes,* • *Fusobacteria* • *Proteobacteria* • *Actinobacteria*	• Cellular metabolism, • Provide complementary endogenous sources of • vitamins, • strengthen immune function, exert epigenetic effects to regulate cell proliferation		[248–251]
Polyamines	• putrescine • cadaverine • spermidine • spermine	• *Campylobacter jejuni* • *Clostridium saccharolyticum*	• Exert genotoxic effects on the host, • anti-inflammatory • In addition, antitumoral effects. • potential tumour markers		[252]

Table 3. *Cont.*

Metabolite Class	Metabolites	Related Bacteria	Biological Functions	Disease/Disorder	Reference
Lipids	• conjugated fatty acids • lipopolysaccharide (LPS) • peptidoglycan • acylglycerols • sphingomyelin • cholesterol • phosphatidylcholines • phosphoethanolamines • triglycerides	• *Bifidobacterium* • *Roseburia,* • *Lactobacillus* • *Klebsiella* • *Enterobacter* • *Citrobacter* • *Clostridium*	• Impact intestinal permeability, activate intestine brain- • liver neural axis to regulate glucose homeostasis; • LPS induces chronic systemic • inflammation; • conjugated fatty acids improve • hyperinsulinemia, • enhance the immune system • and alter lipoprotein profiles. Cholesterol is the basis for sterol and bile acid production.		[253]
Others	• D-lactate • methanol • ethanol • formate • succinate • lysine • glucose • urea • a-ketoisovalerate, • creatine • creatinine • endocannabinoids • 2-arachidonoylglycerol (2-AG) • N-arachidonoylethanolamide	• Bacteroides • Pseudobutyrivibrio • Ruminococcus • Faecalibacterium • Subdoligranulum • Bifidobacterium • Atopobium • Firmicutes • Lactobacillus	• Direct or indirect synthesis or utilization of compounds or modulation of linked • pathways including endocannabinoid system.		[237,254]

Different analytical strategies have been employed for the quantitative analysis of the metabolome, depending on the availability of the technology and research questions [255]. Metabolites of biological samples, such as serum, plasma, urine, faeces and tissues are different in chemical and physiochemical structure and have a large dynamic range of metabolic concentration. Multiple analytical techniques such as gas chromatography (GC), liquid chromatography (LC) and high/ultra-performance liquid chromatography (H/UPLC) coupled to mass spectrometry (MS), and nuclear magnetic resonance spectroscopy (NMR) enable detection, identification and quantification of a wider range of metabolites [256,257]. There are targeted and untargeted metabolomics approaches, and both have their merits and pitfalls.

A targeted metabolomics approach measures a specific list of metabolites, typically focusing on one or more related pathways of interest [258], driven by a specific biochemical questions or hypothesis. This approach can be effective for pharmacokinetics studies of drug discoveries as well as measuring the influence of the intervention on the targeted pathways or metabolic functions [259]. Targeted metabolomics studies offer distinct advantages for metabolite specificity and quantitative reproducibility.

Untargeted metabolomics, also known as global metabolite profiling, attempts to measure as many metabolites as possible. This approach has enabled new discoveries that link cellular pathways to biological mechanisms and are contributing to the understanding of the cellular metabolism, biology, physiology, medicine and host-microbiome interactions [260–262].

In contrast to the targeted metabolomics results, untargeted metabolomics studies generate large amounts of highly complex data. Manual inspection of the thousands of detected picks is impractical and requires instrumental automation with metabolomics software such as MathDAMP, MetAlign, MZMine and XCMS [263–266]. For the identification of individual metabolites in an untargeted approach, a combination of different techniques is applied to ensure good coverage of the metabolome [256]. Table 4 summarizes the online database available to assist with identification of the individual metabolites generated through NMR and MS detection platforms. These databases contain spectral, chemical, molecular and clinical information about the metabolites found in different human biosamples.

Table 4. Web-based databases for identification of metabolites (modified from [267]).

Database	Web Address/URL	Available Since/Reference
Global Metabolome Database (GMD)	http://gmd.mpimp-golm.mpg.de/	2004, Kopka, et al. [268]
METLIN	https://metlin.scripps.edu/	2005, Smith, et al. [269]
Kyoto Encyclopedia of Genes and Genomes (KEGG)	http://www.genome.jp/kegg/	1995, Kanehisa, et al. [270]
Chemicals Entities of Biological Interest (ChEBI)	http://www.ebi.ac.uk/chebi/	2004, Degtyarenko, et al. [271]
Human Metabolome Database (HMDB)	http://www.hmdb.ca/	2007, Wishart et al. [272,273]
Biological Magnetic Resonance Data Bank (BMRB)	http://www.bmrb.wisc.edu/	2007, Ulrich, et al. [274]
Madison Metabolomics Consortium (MMC) Database	http://mmcd.nmrfam.wisc.edu/	2008, Cui, et al. [275]
BiGG (a knowledgebase of Biochemically, Genetically and Genomically structured genome-scale metabolic network reconstructions)	http://bigg.ucsd.edu/	2010, Schellenberger, et al. [276]
MassBank	http://www.massbank.jp/	2010, Horai, et al. [277]
SetupX and BinBase	https://fiehnlab.ucdavis.edu/	2011, Skogerson, et al. [278]

Multidimensional separations based on mass spectrometry are a powerful tool for revealing systems level information [279]. For a multidimensional analysis combining proteomics, metabolomics, lipidomics and glycomics, there are 106 possible proteins, and 200,000 metabolites. If drugs are included, this adds another 1060 compounds, as well as an unknown number of man-made compounds originating from environmental contaminants. Thus, these techniques come with a considerable data burden.

Many microbial metabolites are volatile and their study yields valuable insights into microbial community metabolism, interactions, and inter-kingdom interactions [280,281]. The volatilome is defined as all the volatile compounds that originate from an organism or ecosystem. Studying the blood and faecal volatilome in conjunction with non-volatile metabolites, epigenetic and metagenomic measurements from the same samples can yield valuable insights into metabolism and the interactome. Similarly, the volatile profiles of exhaled breath have been utilized as a technique for phenotyping IBS subjects [282], and breath shows promise as a rapid and non-invasive sample type to rapidly classify phenotype, based on volatile metabolites that may be formed in the gut, transferred to the blood, and then transferred through the lung membrane into the exhaled breath.

6. Integrating Multi-Omics Datasets

Integrated multi-omics approaches are challenging because the data obtained from such research consists of two or more matrices that contain the same sample IDs, but a range of different biological variables such as genes, transcripts, metabolites, proteins etc. Based on whether these studies take into consideration of prior knowledge, they can be classified as statistics-based methods or knowledge driven methods. Statistical approaches use univariate or multivariate analysis to understand the correlations between the different datasets. Meanwhile, knowledge-based approaches decipher the potential mechanistic links by using the significant variables identified in the different omics approaches and associating them with an existing knowledge base. They are often presented as interaction networks e.g., metabolic networks [283]. Ultimately, the goal is to triangulate between different biological samples that indicate absorption, secretion, or excretion to decipher the interactive systems at play internally. Chemometrics are important for the integrated analysis of the gut microbiome. Chemometrics is the application of data-driven statistical and computational methods to extract information from the measurement of chemical systems [284]. However, commonly available statistical methods are inadequate when dealing with the multidimensional omics datasets necessary for analysis of the gut microbiome. Thus, new approaches are required to deal with multi-dimensional data [285].

Considering the multi-omics represented by metabolomic and taxonomic profiles require not only linking of the two datasets but should include incorporation of prior reference information about metabolic capacities of community members [286], environmental factors such as nutritional/dietary information, disease, etc. [287]. Integration of data into systems-wide approaches has long been recognized, and in some cases, attempted [97,221,286–293]. However, various challenges with standardization of methodologies include scale [286], chemical complexity [221], financial and human resources [289], ecological and clinical context [294], diet [288], inter-individual variation and noise [291], interactions with other body tissues [290], in vivo access to relevant sampling sites [292], and time scales [295]. Nevertheless, perhaps the most comprehensive attempt to integrate datasets can be illustrated by the Virtual Metabolic Human database (https://www.vmh.life) [287], encompassing tens of thousands of unique reactions, thousands of unique metabolites and human genes, hundreds of thousands of microbial genes, thousands of food types, hundreds of diseases and hundreds of microbes, cross-referenced to more than thirty external database resources, with a high percentage of coverage. This resource and framework may be a first real step in integrating multi-omics information.

Machine learning using algorithms such as support vector machine, random forest, Adaboost, logitboost, neural networks, decision tree and other hybrid methods [296,297] can be applied to these large datasets to aid with interpretation. Although microbiome data are complex, meaningful biological insights have been drawn when applied successfully [298–302]. Saulnier, et al. [303] used supervised learning to classify different subtypes of IBS with a high success rate of about 96%. Hacılar, Nalbantoğlu and Bakir-Güngör [296] used machine-learning analysis to investigate subset of gut microbiota that is associated with IBD. Machine learning with advanced data visualization techniques can reveal patterns not detectable by traditional statistical techniques. It has been applied in diverse places, from the intestinal interactome [304], to prediction of metabolic pathways [305], to integration of metabolomics, lipidomics and clinical data [306]. However, critical to the success of machine learning approaches is

the size of the training set. This approach cannot be used on small datasets unless a larger dataset with the same characteristics is available.

NJS16 is an extensive database developed by Sung and co-workers [307]. This literature-curated interspecies network of the human gut microbiota consists of ~570 microbial species and 3 human cell types metabolically interacting via ~44,000 small-molecule transport and macromolecule degradation events. Application of a mathematical model approach to the contents of the global metabolic network was used to extract useful information such as biomarker microbial and metabolic features of the gut microbial ecosystem in type 2 diabetes. This study is a step towards integrative investigations of context-specific community-scale analysis.

Pathway Analysis and Imputation to Relate Unknowns in Profiles from MS-based metabolite data (PAIRUP-MS), recently developed by Hsu et al. [308], is a very exciting approach to analysis of new metabolites and new pathways. PAIRUP-MS enables previously infeasible analyses of a significant portion of signals that often goes unidentified as known metabolites and has been excluded from downstream analyses. This tool, that also offers a pathway annotation and enrichment analysis framework, could possibly be used to link metabolite signals to plausible biological functions despite their unknown chemical identities.

The ability to detect latent variables in omics data, and to separate direct from indirect linkages is key to determining the mechanisms and interactions that drive the gut microbiome [309]. Methods also need to allow environmental and clinical information to be incorporated into the model, to allow multi-factor exploration of interactions within the gut microbiome. Kris Sankaran gives a helpful framework and some insights into future directions for latent variable modelling for the microbiome [310]. Promising approaches include sparsity-based methods (SPLS, Graph-Fused Lasso, CCA), and hierarchical clustering [310]. One such approach currently being applied in a number of studies is Hierarchical-All-against-All analysis [311]. HAllA tests for correlation between all pairs of variables in multi-dimensional datasets and can detect multi-level associations, even in non-homogenous datasets. It has been successfully applied to determine the effect of polyphenolic compounds on the gut microbiome [312].

All of these emerging statistical techniques require considerable computing power compared to traditional statistical analysis, and need to be addressed in new ways. For research institutions in developed countries, this typically means making use of shared research computing infrastructure comprising large computing arrays that can be accessed via virtual machines. Another solution could be carrying out computations via collective, cloud-based computing power. The readers are referred to Huang, Chaudhary and Garmire [301] for a more comprehensive account of various data integration tools.

7. Future Directions that Would Accelerate an Integrated Approach

7.1. New Sampling Techniques

Under normal gut conditions, the most prevalent gases include methane, carbon dioxide, hydrogen, hydrogen sulphide and nitric oxide. Production of such gases and their relative concentrations affecting gut function may have pathogenic roles in several GIT disorders. Lack of direct access to these gases in the gut limits our understanding of their physiology and functional capacity. Direct sampling of gases inside the GIT is likely to provide much more accurate gas-related biomarkers for physiological abnormalities of the gut. Indigestible, endoscopy capsules have been developed by integrating pH, pressure and temperature sensors (SmartPill®, Minneapolis, MN, USA). Kalantar-zadeh, et al. [313] developed indigestible, non-invasive, swallowable intestinal gas capsules that can perform in vivo gas measurements and potentially assess putative gas biomarkers in GIT disorders. The so-called 'gas capsules' are obtained by integrating small gas sensors into indigestible capsule platforms. The related clinical procedures are non-invasive and once real-time monitoring of the appropriate locations to

be sampled is perfected, these capsules could help in surveying gas-related biomarkers and their concentrations throughout the length of the GIT.

7.2. New and Emerging Techniques and Disciplines: Culturomics, the Interactome, Foodomics, and the Exposome

Culturomics is a new approach that uses high-throughput, broad spectrum culturing coupled with mass-spectrometry-based identification [314–316]. Culturomics enables exploration of the 'dark' microbiome to levels approaching those of pyrosequencing, and is able to identify or characterize microbes present in low concentrations. The use of culturomics in tandem with metagenomics would greatly advance understanding of the gut microbiome.

One new and emerging concept, the interactome, is critical to the advancement of integrated analysis of the gut microbiome. The interactome is defined as the whole set of molecular interactions in a cell [317]. Studying the interactome requires the use of multi-omics data from metabolomics, proteomics, and genomics. However, analysing multi-omics data necessitates the development of new statistical tools, in order to tease out direct from indirect associations. The development and application of such tools to the analysis of the interactome is key to significantly accelerating our understanding of complex interactions in the gut microbiome [225,291,304,307,318–320].

The other area critical to the advancement of integrated systems analysis of the gut microbiome is in foodomics. Foodomics is the application of -omics technologies to the study of food and drink, and the nutritional effects of consuming them. It is now well accepted that diet is the main regulator of gut health. Yet, currently, ingested food is not routinely subjected to the same rigorous analytical techniques that faeces are. Dietary recall and food diaries are inadequate in that they are unable to precisely link components from diet with gut health, and important dietary components such as the type of liquids ingested are sometimes ignored. In addition, these techniques do not take into account the degree to which the food has been processed and thus stripped of its natural microbiome and micronutrients, or whether it contains man-made contaminants. Thus, current dietary data collection methods do not accurately reflect the micronutrient or metabolite profile of food as it is prepared and ingested. Metabolite and genomic profiling of food allows investigation of compounds in food and its associated microbiome. While there are few studies that have used a systems-based approach [321], such studies hold great therapeutic potential, as demonstrated by the Ni and co-workers [225] study of interactions between the small molecules from diet and the gut bacterial proteome. Foodomics is an emerging discipline that will become increasingly important to an integrated analysis of the gut microbiome.

Equally critical to the study of the gut microbiome is the integration of the impact of the external environment into systems analysis. The exposome is defined as the totality of environmental exposures from conception onwards. While not all of these factors are easily measured, external contaminants or man-made compounds can be detected using trace analytical techniques, such as mass spectrometry. Analysis of trace environmental contaminants can often be undertaken at the same time as analysis of the metabolome, provided untargeted analytical methods are used, and comprehensive mass spectral libraries such as NIST/Wiley are available to search. At present, primarily due to the availability of large mass spectral libraries, GC-MS is still the most reliable method for exposome analysis. However, considerable advances in analytical software and enlarging databases for high-resolution LC-MS/MS have occurred recently and this technique will surpass GC-MS as many more metabolites are amenable to measurement using LC-MS than GC-MS. Alternatively, targeted analysis of large panels of pharmaceuticals, pesticides, plasticizers, and other man-made compounds can be carried out in tandem with metabolite profiling, but will inevitably suffer from bias due to assumptions made around what compounds to target.

Author Contributions: Conceptualization, D.J.B., D.R. and E.J.M.; R.M.S., E.J.M., M.T.R., S.R.J., S.V.G., D.R. and D.J.B. all authors contributed equally in the preparation and writing of this review article. D.J.B. and D.R. coordinated each author's contribution and oversaw the editorial integration and review of the final paper. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. There is no connection between Anagenix Ltd and the subject of this manuscript.

References

1. Robinson, C.J.; Bohannan, B.J.; Young, V.B. From structure to function: The ecology of host-associated microbial communities. *Microbiol. Mol. Biol. Rev.* **2010**, *74*, 453–476. [CrossRef] [PubMed]

2. Shreiner, A.B.; Kao, J.Y.; Young, V.B. The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* **2015**, *31*, 69. [CrossRef] [PubMed]

3. Nash, A.K.; Auchtung, T.A.; Wong, M.C.; Smith, D.P.; Gesell, J.R.; Ross, M.C.; Stewart, C.J.; Metcalf, G.A.; Muzny, D.M.; Gibbs, R.A. The gut mycobiome of the Human Microbiome Project healthy cohort. *Microbiome* **2017**, *5*, 153. [CrossRef] [PubMed]

4. Schmidt, T.S.; Raes, J.; Bork, P. The human gut microbiome: From association to modulation. *Cell* **2018**, *172*, 1198–1215. [CrossRef] [PubMed]

5. Heintz-Buschart, A.; Wilmes, P. Human gut microbiome: Function matters. *Trends Microbiol.* **2018**, *26*, 563–574. [CrossRef] [PubMed]

6. Tropini, C.; Earle, K.A.; Huang, K.C.; Sonnenburg, J.L. The gut microbiome: Connecting spatial organization to function. *Cell Host Microbe* **2017**, *21*, 433–442. [CrossRef]

7. Lynch, S.V.; Pedersen, O. The human intestinal microbiome in health and disease. *N. Engl. J. Med.* **2016**, *375*, 2369–2379. [CrossRef]

8. Couch, R.D.; Navarro, K.; Sikaroodi, M.; Gillevet, P.; Forsyth, C.B.; Mutlu, E.; Engen, P.A.; Keshavarzian, A. The approach to sample acquisition and its impact on the derived human fecal microbiome and VOC metabolome. *PLoS ONE* **2013**, *8*, e81163. [CrossRef]

9. Zierer, J.; Jackson, M.A.; Kastenmüller, G.; Mangino, M.; Long, T.; Telenti, A.; Mohney, R.P.; Small, K.S.; Bell, J.T.; Steves, C.J.; et al. The fecal metabolome as a functional readout of the gut microbiome. *Nat. Genet.* **2018**, *50*, 790–795. [CrossRef]

10. Marchesi, J.R. Prokaryotic and eukaryotic diversity of the human gut. *Adv. Appl. Microbiol.* **2010**, *72*, 43–62.

11. 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] [PubMed]

12. Koh, A.Y. Gastrointestinal colonization of fungi. *Curr. Fungal Infect. Rep.* **2013**, *7*, 144–151. [CrossRef]

13. Underhill, D.M.; Iliev, I.D. The mycobiota: Interactions between commensal fungi and the host immune system. *Nat. Rev. Immunol.* **2014**, *14*, 405–416. [CrossRef] [PubMed]

14. Seed, P.C. The human mycobiome. *Cold Spring Harb. Perspect. Med.* **2015**, *5*, a019810. [CrossRef] [PubMed]

15. Huffnagle, G.B.; Noverr, M.C. The emerging world of the fungal microbiome. *Trends Microbiol.* **2013**, *21*, 334–341. [CrossRef] [PubMed]

16. Wegener Parfrey, L.; Walters, W.A.; Knight, R. Microbial eukaryotes in the human microbiome: Ecology, evolution, and future directions. *Front. Microbiol.* **2011**, *2*, 153. [CrossRef]

17. Ley, R.E.; Peterson, D.A.; Gordon, J.I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **2006**, *124*, 837–848. [CrossRef]

18. Garrett, W.S.; Gordon, J.I.; Glimcher, L.H. Homeostasis and inflammation in the intestine. *Cell* **2010**, *140*, 859–870. [CrossRef]

19. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Lozupone, C.A.; Turnbaugh, P.J.; Fierer, N.; Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4516–4522. [CrossRef]

20. Bolhuis, H.; Cretoiu, M.S.; Stal, L.J. Molecular ecology of microbial mats. *FEMS Microbiol. Ecol.* **2014**, *90*, 335–350.

21. Huttenhower, C.; Knight, R.; Brown, C.T.; Caporaso, J.G.; Clemente, J.C.; Gevers, D.; Franzosa, E.A.; Kelley, S.T.; Knights, D.; Ley, R.E. Advancing the microbiome research community. *Cell* **2014**, *159*, 227–230. [CrossRef] [PubMed]

22. Norman, J.M.; Handley, S.A.; Virgin, H.W. Kingdom-agnostic metagenomics and the importance of complete characterization of enteric microbial communities. *Gastroenterology* **2014**, *146*, 1459–1469. [CrossRef] [PubMed]

23. Yoon, S.S.; Kim, E.-K.; Lee, W.-J. Functional genomic and metagenomic approaches to understanding gut microbiota–animal mutualism. *Curr. Opin. Microbiol.* **2015**, *24*, 38–46. [CrossRef] [PubMed]

24. Hoffmann, C.; Dollive, S.; Grunberg, S.; Chen, J.; Li, H.; Wu, G.D.; Lewis, J.D.; Bushman, F.D. Archaea and fungi of the human gut microbiome: Correlations with diet and bacterial residents. *PLoS ONE* **2013**, *8*, e66019. [CrossRef] [PubMed]

25. Schulze, J.; Sonnenborn, U. Yeasts in the gut: From commensals to infectious agents. *Dtsch. Arztebl. Int.* **2009**, *106*, 837. [CrossRef]

26. Suhr, M.J.; Hallen-Adams, H.E. The human gut mycobiome: Pitfalls and potentials—A mycologist's perspective. *Mycologia* **2015**, *107*, 1057–1073. [CrossRef]

27. Sender, R.; Fuchs, S.; Milo, R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.* **2016**, *14*, e1002533. [CrossRef]

28. Lauber, C.L.; Zhou, N.; Gordon, J.I.; Knight, R.; Fierer, N. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. *FEMS Microbiol. Lett.* **2010**, *307*, 80–86. [CrossRef]

29. Wu, G.D.; Lewis, J.D.; Hoffmann, C.; Chen, Y.-Y.; Knight, R.; Bittinger, K.; Hwang, J.; Chen, J.; Berkowsky, R.; Nessel, L.; et al. Sampling and pyrosequencing methods for characterizing bacterial communities in the human gut using 16S sequence tags. *BMC Microbiol.* **2010**, *10*, 206. [CrossRef]

30. Roesch, L.F.; Casella, G.; Simell, O.; Krischer, J.; Wasserfall, C.H.; Schatz, D.; Atkinson, M.A.; Neu, J.; Triplett, E.W. Influence of fecal sample storage on bacterial community diversity. *Open Microbiol. J.* **2009**, *3*, 40. [CrossRef]

31. Panek, M.; Čipčić Paljetak, H.; Barešić, A.; Perić, M.; Matijašić, M.; Lojkić, I.; Vranešić Bender, D.; Krznarić, Ž.; Verbanac, D. Methodology challenges in studying human gut microbiota—Effects of collection, storage, DNA extraction and next generation sequencing technologies. *Sci. Rep.* **2018**, *8*, 5143. [CrossRef] [PubMed]

32. Vandeputte, D.; Tito, R.Y.; Vanleeuwen, R.; Falony, G.; Raes, J. Practical considerations for large-scale gut microbiome studies. *FEMS Microbiol. Rev.* **2017**, *41*, S154–S167. [CrossRef] [PubMed]

33. Falony, G.; Vieira-Silva, S.; Raes, J. Richness and ecosystem development across faecal snapshots of the gut microbiota. *Nat. Microbiol.* **2018**, *3*, 526–528. [CrossRef] [PubMed]

34. Wolfe, A.J. The acetate switch. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 12–50. [CrossRef] [PubMed]

35. Flamholz, A.; Noor, E.; Bar-Even, A.; Liebermeister, W.; Milo, R. Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 10039–10044. [CrossRef] [PubMed]

36. Herrmann, E.; Young, W.; Rosendale, D.; Reichert-Grimm, V.; Riedel, C.U.; Conrad, R.; Egert, M. RNA-based stable isotope probing suggests *Allobaculum* spp. as particularly active glucose assimilators in a complex murine microbiota cultured in vitro. *BioMed Res. Int.* **2017**. [CrossRef]

37. Mahowald, M.A.; Rey, F.E.; Seedorf, H.; Turnbaugh, P.J.; Fulton, R.S.; Wollam, A.; Shah, N.; Wang, C.; Magrini, V.; Wilson, R.K.; et al. Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5859–5864. [CrossRef]

38. Tuncil, Y.E.; Xiao, Y.; Porter, N.T.; Reuhs, B.L.; Martens, E.C.; Hamaker, B.R. Reciprocal Prioritization to Dietary Glycans by Gut Bacteria in a Competitive Environment Promotes Stable Coexistence. *mBio* **2017**, *8*. [CrossRef]

39. Kim, C. Identification of Rumen Methanolgens, Characterization of Substrate Requirements and Measurement of Hydrogen Thresholds. Master's Thesis, Massey University, Palmerston North, New Zealand, 2012.

40. Donohoe, D.R.; Garge, N.; Zhang, X.; Sun, W.; O'Connell, T.M.; Bunger, M.K.; Bultman, S.J. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab.* **2011**, *13*, 517–526. [CrossRef]

41. den Besten, G.; Havinga, R.; Bleeker, A.; Rao, S.; Gerding, A.; van Eunen, K.; Groen, A.K.; Reijngoud, D.J.; Bakker, B.M. The short-chain fatty acid uptake fluxes by mice on a guar gum supplemented diet associate with amelioration of major biomarkers of the metabolic syndrome. *PLoS ONE* **2014**, *9*, e107392. [CrossRef]

42. Kinross, J.M.; Darzi, A.W.; Nicholson, J.K. Gut microbiome-host interactions in health and disease. *Genome Med.* **2011**, *3*, 14. [CrossRef] [PubMed]

43. Walsh, C.J.; Guinane, C.M.; O'Toole, P.W.; Cotter, P.D. Beneficial modulation of the gut microbiota. *FEBS Lett.* **2014**, *588*, 4120–4130. [CrossRef] [PubMed]

44. Hennessy, A.A.; Ross, R.P.; Fitzgerald, G.F.; Caplice, N.; Stanton, C. Role of the gut in modulating lipoprotein metabolism. *Curr. Cardiol. Rep.* **2014**, *16*, 515. [CrossRef] [PubMed]

45. O'Mahony, S.M.; Clarke, G.; Borre, Y.; Dinan, T.; Cryan, J. Serotonin, tryptophan metabolism and the brain-gut-microbiome axis. *Behav. Brain Res.* **2015**, *277*, 32–48. [CrossRef]

46. Ramakrishna, B.S. Role of the gut microbiota in human nutrition and metabolism. *J. Gastroenterol. Hepatol.* **2013**, *28*, 9–17. [CrossRef]

47. Fuller, M. Determination of protein and amino acid digestibility in foods including implications of gut microbial amino acid synthesis. *Br. J. Nutr.* **2012**, *108*, S238–S246. [CrossRef]

48. Dutton, R.J.; Turnbaugh, P.J. Taking a metagenomic view of human nutrition. *Curr. Opin. Clin. Nutr. Metab. Care* **2012**, *15*, 448–454. [CrossRef]

49. Kau, A.L.; Ahern, P.P.; Griffin, N.W.; Goodman, A.L.; Gordon, J.I. Human nutrition, the gut microbiome and the immune system. *Nature* **2011**, *474*, 327. [CrossRef]

50. Brestoff, J.R.; Artis, D. Commensal bacteria at the interface of host metabolism and the immune system. *Nat. Immunol.* **2013**, *14*, 676. [CrossRef]

51. Hooper, L.V.; Littman, D.R.; Macpherson, A.J. Interactions between the microbiota and the immune system. *Science* **2012**, *336*, 1268–1273. [CrossRef]

52. Lee, Y.K.; Mazmanian, S.K. Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science* **2010**, *330*, 1768–1773. [CrossRef] [PubMed]

53. Maslowski, K.M.; Mackay, C.R. Diet, gut microbiota and immune responses. *Nat. Immunol.* **2010**, *12*, 5. [CrossRef] [PubMed]

54. Thaiss, C.A.; Zmora, N.; Levy, M.; Elinav, E. The microbiome and innate immunity. *Nature* **2016**, *535*, 65. [CrossRef] [PubMed]

55. Zackular, J.P.; Rogers, M.A.; Ruffin, M.T.; Schloss, P.D. The human gut microbiome as a screening tool for colorectal cancer. *Cancer Prev. Res.* **2014**, *7*, 1112–1121. [CrossRef]

56. Feng, Q.; Liang, S.; Jia, H.; Stadlmayr, A.; Tang, L.; Lan, Z.; Zhang, D.; Xia, H.; Xu, X.; Jie, Z. Gut microbiome development along the colorectal adenoma–carcinoma sequence. *Nat. Commun.* **2015**, *6*, 6528. [CrossRef] [PubMed]

57. Bajaj, J.S.; Heuman, D.M.; Hylemon, P.B.; Sanyal, A.J.; White, M.B.; Monteith, P.; Noble, N.A.; Unser, A.B.; Daita, K.; Fisher, A.R. Altered profile of human gut microbiome is associated with cirrhosis and its complications. *J. Hepatol.* **2014**, *60*, 940–947. [CrossRef]

58. Del Chierico, F.; Nobili, V.; Vernocchi, P.; Russo, A.; Stefanis, C.D.; Gnani, D.; Furlanello, C.; Zandonà, A.; Paci, P.; Capuani, G. Gut microbiota profiling of pediatric nonalcoholic fatty liver disease and obese patients unveiled by an integrated meta-omics-based approach. *Hepatology* **2017**, *65*, 451–464. [CrossRef]

59. Schippa, S.; Iebba, V.; Barbato, M.; Di Nardo, G.; Totino, V.; Checchi, M.P.; Longhi, C.; Maiella, G.; Cucchiara, S.; Conte, M.P. A distinctive 'microbial signature' in celiac pediatric patients. *BMC Microbiol.* **2010**, *10*, 175. [CrossRef]

60. Plummer, M.; Franceschi, S.; Vignat, J.; Forman, D.; de Martel, C. Global burden of gastric cancer attributable to Helicobacter pylori. *Int. J. Cancer* **2015**, *136*, 487–490. [CrossRef]

61. Finegold, S.M.; Dowd, S.E.; Gontcharova, V.; Liu, C.; Henley, K.E.; Wolcott, R.D.; Youn, E.; Summanen, P.H.; Granpeesheh, D.; Dixon, D. Pyrosequencing study of fecal microflora of autistic and control children. *Anaerobe* **2010**, *16*, 444–453. [CrossRef]

62. Finegold, S.M.; Downes, J.; Summanen, P.H. Microbiology of regressive autism. *Anaerobe* **2012**, *18*, 260–262. [CrossRef] [PubMed]

63. Mezzelani, A.; Landini, M.; Facchiano, F.; Raggi, M.E.; Villa, L.; Molteni, M.; De Santis, B.; Brera, C.; Caroli, A.M.; Milanesi, L. Environment, dysbiosis, immunity and sex-specific susceptibility: A translational hypothesis for regressive autism pathogenesis. *Nutr. Neurosci.* **2015**, *18*, 145–161. [CrossRef] [PubMed]

64. Scheperjans, F.; Aho, V.; Pereira, P.A.B.; Koskinen, K.; Paulin, L.; Pekkonen, E.; Haapaniemi, E.; Kaakkola, S.; Eerola-Rautio, J.; Pohja, M.; et al. Gut microbiota are related to Parkinson's disease and clinical phenotype. *Mov. Disord.* **2015**, *30*, 350–358. [CrossRef] [PubMed]

65. Larsen, N.; Vogensen, F.K.; Van Den Berg, F.W.J.; Nielsen, D.S.; Andreasen, A.S.; Pedersen, B.K.; Al-Soud, W.A.; Sørensen, S.J.; Hansen, L.H.; Jakobsen, M. Gut Microbiota in Human Adults with Type 2 Diabetes Differs from Non-Diabetic Adults. *PLoS ONE* **2010**, *5*, e9085. [CrossRef]

66. Qin, J.; Li, Y.; Cai, Z.; Li, S.; Zhu, J.; Zhang, F.; Liang, S.; Zhang, W.; Guan, Y.; Shen, D.; et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **2012**, *490*, 55–60. [CrossRef] [PubMed]

67. Dicksved, J.; Halfvarson, J.; Rosenquist, M.; Järnerot, G.; Tysk, C.; Apajalahti, J.; Engstrand, L.; Jansson, J.K. Molecular analysis of the gut microbiota of identical twins with Crohn's disease. *ISME J.* **2008**, *2*, 716. [CrossRef] [PubMed]

68. Sokol, H.; Lepage, P.; Seksik, P.; Doré, J.; Marteau, P. Molecular comparison of dominant microbiota associated with injured versus healthy mucosa in ulcerative colitis. *Gut* **2007**, *56*, 152–154. [CrossRef]

69. Sokol, H.; Seksik, P.; Furet, J.; Firmesse, O.; Nion-Larmurier, I.; Beaugerie, L.; Cosnes, J.; Corthier, G.; Marteau, P.; Doré, J. Low counts of Faecalibacterium prausnitzii in colitis microbiota. *Inflamm. Bowel Dis.* **2009**, *15*, 1183–1189. [CrossRef]

70. Lepage, P.; Häsler, R.; Spehlmann, M.E.; Rehman, A.; Zvirbliene, A.; Begun, A.; Ott, S.; Kupcinskas, L.; Doré, J.; Raedler, A. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology* **2011**, *141*, 227–236. [CrossRef]

71. Drossman, D.A. Functional gastrointestinal disorders: History, pathophysiology, clinical features, and Rome IV. *Gastroenterology* **2016**, *150*, 1262–1279.e1262. [CrossRef]

72. Nafarin, A.R.; Hegar, B.; Sjakti, H.A.; Vandenplas, Y. Gut microbiome pattern in adolescents with functional gastrointestinal disease. *Int. J. Pediatr. Adolesc. Med.* **2019**, *6*, 12–15. [CrossRef] [PubMed]

73. Shin, A.; Preidis, G.A.; Shulman, R.; Kashyap, P. The gut microbiome in adult and pediatric functional gastrointestinal disorders. *Clin. Gastroenterol. Hepatol.* **2019**, *17*, 256–274. [CrossRef] [PubMed]

74. Foster, J.A.; Rinaman, L.; Cryan, J.F. Stress & the gut-brain axis: Regulation by the microbiome. *Neurobiol. Stress* **2017**, *7*, 124–136. [CrossRef] [PubMed]

75. Grenham, S.; Clarke, G.; Cryan, J.F.; Dinan, T.G. Brain–gut–microbe communication in health and disease. *Front. Physiol.* **2011**, *2*, 94. [CrossRef] [PubMed]

76. Sampson, T.R.; Mazmanian, S.K. Control of brain development, function, and behavior by the microbiome. *Cell Host Microbe* **2015**, *17*, 565–576. [CrossRef] [PubMed]

77. Sharon, G.; Sampson, T.R.; Geschwind, D.H.; Mazmanian, S.K. The central nervous system and the gut microbiome. *Cell* **2016**, *167*, 915–932. [CrossRef]

78. Rieder, R.; Wisniewski, P.J.; Alderman, B.L.; Campbell, S.C. Microbes and mental health: A review. *Brain Behav. Immun.* **2017**, *66*, 9–17. [CrossRef]

79. Bruce-Keller, A.J.; Salbaum, J.M.; Berthoud, H.-R. Harnessing gut microbes for mental health: Getting from here to there. *Biol. Psychiatry* **2018**, *83*, 214–223. [CrossRef]

80. Archie, E.A.; Tung, J. Social behavior and the microbiome. *Curr. Opin. Behav. Sci.* **2015**, *6*, 28–34. [CrossRef]

81. Gomez, A.; Luckey, D.; Taneja, V. The gut microbiome in autoimmunity: Sex matters. *Clin. Immunol.* **2015**, *159*, 154–162. [CrossRef]

82. Clemente, J.C.; Manasson, J.; Scher, J.U. The role of the gut microbiome in systemic inflammatory disease. *BMJ* **2018**, *360*, j5145. [CrossRef] [PubMed]

83. Mielcarz, D.W.; Kasper, L.H. The gut microbiome in multiple sclerosis. *Curr. Treat. Options Neurol.* **2015**, *17*, 18. [CrossRef] [PubMed]

84. Weis, M. Impact of the gut microbiome in cardiovascular and autoimmune diseases. *Clin. Sci.* **2018**, *132*, 2387–2389. [CrossRef] [PubMed]

85. Zitvogel, L.; Ayyoub, M.; Routy, B.; Kroemer, G. Microbiome and anticancer immunosurveillance. *Cell* **2016**, *165*, 276–287. [CrossRef] [PubMed]

86. 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] [PubMed]

87. Wong, S.H.; Kwong, T.N.; Wu, C.-Y.; Yu, J. Clinical applications of gut microbiota in cancer biology. *Semin. Cancer Biol.* **2019**, *55*, 28–36. [CrossRef]

88. Bhutia, Y.D.; Ogura, J.; Sivaprakasam, S.; Ganapathy, V. Gut microbiome and colon cancer: Role of bacterial metabolites and their molecular targets in the host. *Curr. Colorectal Cancer Rep.* **2017**, *13*, 111–118. [CrossRef]

89. Gagnière, J.; Raisch, J.; Veziant, J.; Barnich, N.; Bonnet, R.; Buc, E.; Bringer, M.-A.; Pezet, D.; Bonnet, M. Gut microbiota imbalance and colorectal cancer. *World J. Gastroenterol.* **2016**, *22*, 501. [CrossRef]

90. Goubet, A.-G.; Daillère, R.; Routy, B.; Derosa, L.; Roberti, P.M.; Zitvogel, L. The impact of the intestinal microbiota in therapeutic responses against cancer. *Comptes Rendus Boil.* **2018**, *341*, 284–289. [CrossRef]

91. West, N.R.; Powrie, F. Immunotherapy not working? Check your microbiota. *Cancer Cell* **2015**, *28*, 687–689. [CrossRef]

92. Singh, V.; San Yeoh, B.; Vijay-Kumar, M. Gut microbiome as a novel cardiovascular therapeutic target. *Curr. Opin. Pharmacol.* **2016**, *27*, 8–12. [CrossRef] [PubMed]

93. Kitai, T.; Tang, W.W. The Role and Impact of gut microbiota in cardiovascular disease. *Revista Española de Cardiología* **2017**, *70*, 799–800. [CrossRef] [PubMed]

94. Antza, C.; Stabouli, S.; Kotsis, V. Gut microbiota in kidney disease and hypertension. *Pharmacol. Res.* **2018**, *130*, 198–203. [CrossRef] [PubMed]

95. Mazidi, M.; Rezaie, P.; Kengne, A.P.; Mobarhan, M.G.; Ferns, G.A. Gut microbiome and metabolic syndrome. *Diabetes Metab. Syndr. Clin. Res. Rev.* **2016**, *10*, S150–S157. [CrossRef]

96. John, G.K.; Mullin, G.E. The Gut Microbiome and Obesity. *Curr. Oncol. Rep.* **2016**, *18*, 45. [CrossRef]

97. Mondot, S.; Lepage, P. The human gut microbiome and its dysfunctions through the meta-omics prism. *Ann. N. Y. Acad. Sci.* **2016**, *1372*, 9–19. [CrossRef]

98. Segata, N.; Boernigen, D.; Tickle, T.L.; Morgan, X.C.; Garrett, W.S.; Huttenhower, C. Computational meta'omics for microbial community studies. *Mol. Syst. Biol.* **2013**, *9*, 666. [CrossRef]

99. Iwai, S.; Weinmaier, T.; Schmidt, B.L.; Albertson, D.G.; Poloso, N.J.; Dabbagh, K.; DeSantis, T.Z. Piphillin: Improved prediction of metagenomic content by direct inference from human microbiomes. *PLoS ONE* **2016**, *11*, e0166104. [CrossRef]

100. Chen, I.-M.A.; Markowitz, V.M.; Chu, K.; Palaniappan, K.; Szeto, E.; Pillay, M.; Ratner, A.; Huang, J.; Andersen, E.; Huntemann, M.; et al. IMG/M: Integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res.* **2016**, *45*, D507–D516. [CrossRef]

101. Dehal, P.S.; Joachimiak, M.P.; Price, M.N.; Bates, J.T.; Baumohl, J.K.; Chivian, D.; Friedland, G.D.; Huang, K.H.; Keller, K.; Novichkov, P.S.; et al. MicrobesOnline: An integrated portal for comparative and functional genomics. *Nucleic Acids Res.* **2009**, *38*, D396–D400. [CrossRef]

102. Uchiyama, I. MBGD: Microbial genome database for comparative analysis. *Nucleic Acids Res.* **2003**, *31*, 58–62. [CrossRef] [PubMed]

103. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.G.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **2015**, *31*, 3691–3693. [CrossRef] [PubMed]

104. Yoon, S.-H.; Ha, S.-M.; Kwon, S.; Lim, J.; Kim, Y.; Seo, H.; Chun, J. Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 1613. [CrossRef] [PubMed]

105. Whiteside, M.D.; Winsor, G.L.; Laird, M.R.; Brinkman, F.S.L. OrthologeDB: A bacterial and archaeal orthology resource for improved comparative genomic analysis. *Nucleic Acids Res.* **2012**, *41*, D366–D376. [CrossRef] [PubMed]

106. Blom, J.; Kreis, J.; Spänig, S.; Juhre, T.; Bertelli, C.; Ernst, C.; Goesmann, A. EDGAR 2.0: An enhanced software platform for comparative gene content analyses. *Nucleic Acids Res.* **2016**, *44*, W22–W28. [CrossRef] [PubMed]

107. Yu, J.; Blom, J.; Glaeser, S.P.; Jaenicke, S.; Juhre, T.; Rupp, O.; Schwengers, O.; Spänig, S.; Goesmann, A. A review of bioinformatics platforms for comparative genomics. Recent developments of the EDGAR 2.0 platform and its utility for taxonomic and phylogenetic studies. *J. Biotechnol.* **2017**, *261*, 2–9. [CrossRef] [PubMed]

108. Wilkinson, T.J.; Huws, S.A.; Edwards, J.E.; Kingston-Smith, A.H.; Siu-Ting, K.; Hughes, M.; Rubino, F.; Friedersdorff, M.; Creevey, C.J. CowPI: A Rumen Microbiome Focussed Version of the PICRUSt Functional Inference Software. *Front. Microbiol.* **2018**, *9*. [CrossRef]

109. Langille, M.G. Exploring Linkages between Taxonomic and Functional Profiles of the Human Microbiome. *MSystems* **2018**, *3*, e00163-17. [CrossRef]

110. Odamaki, T.; Kato, K.; Sugahara, H.; Hashikura, N.; Takahashi, S.; Xiao, J.-Z.; Abe, F.; Osawa, R. Age-related changes in gut microbiota composition from newborn to centenarian: A cross-sectional study. *BMC Microbiol.* **2016**, *16*, 90. [CrossRef]

111. Aßhauer, K.P.; Wemheuer, B.; Daniel, R.; Meinicke, P. Tax4Fun: Predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* **2015**, *31*, 2882–2884. [CrossRef]

112. Lurie-Weinberger, M.N.; Gophna, U. Archaea in and on the Human Body: Health Implications and Future Directions. *PLOS Pathog.* **2015**, *11*, e1004833. [CrossRef] [PubMed]

113. Koskinen, K.; Pausan, M.R.; Perras, A.K.; Beck, M.; Bang, C.; Mora, M.; Schilhabel, A.; Schmitz, R.; Moissl-Eichinger, C. First insights into the diverse human archaeome: Specific detection of archaea in the gastrointestinal tract, lung, and nose and on skin. *mBio* **2017**, *8*, e00824-17. [CrossRef] [PubMed]

114. Miller, T.L.; Wolin, M.; De Macario, E.C.; Macario, A. Isolation of Methanobrevibacter smithii from human feces. *Appl. Environ. Microbiol.* **1982**, *43*, 227–232. [CrossRef] [PubMed]

115. Miller, T.L.; Wolin, M.J. Methanosphaera stadtmaniae gen. nov., sp. nov.: A species that forms methane by reducing methanol with hydrogen. *Arch. Microbiol.* **1985**, *141*, 116–122. [CrossRef] [PubMed]

116. Dridi, B.; Fardeau, M.-L.; Ollivier, B.; Raoult, D.; Drancourt, M. Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 1902–1907. [CrossRef] [PubMed]

117. Bang, C.; Weidenbach, K.; Gutsmann, T.; Heine, H.; Schmitz, R.A. The intestinal archaea Methanosphaera stadtmanae and Methanobrevibacter smithii activate human dendritic cells. *PLoS ONE* **2014**, *9*, e99411. [CrossRef]

118. Lecours, P.B.; Marsolais, D.; Cormier, Y.; Berberi, M.; Haché, C.; Bourdages, R.; Duchaine, C. Increased prevalence of Methanosphaera stadtmanae in inflammatory bowel diseases. *PLoS ONE* **2014**, *9*, e87734. [CrossRef]

119. Million, M.; Angelakis, E.; Maraninchi, M.; Henry, M.; Giorgi, R.; Valero, R.; Vialettes, B.; Raoult, D. Correlation between body mass index and gut concentrations of Lactobacillus reuteri, Bifidobacterium animalis, Methanobrevibacter smithii and Escherichia coli. *Int. J. Obes.* **2013**, *37*, 1460. [CrossRef]

120. Le Chatelier, E.; Nielsen, T.; Qin, J.; Prifti, E.; Hildebrand, F.; Falony, G.; Almeida, M.; Arumugam, M.; Batto, J.-M.; Kennedy, S. Richness of human gut microbiome correlates with metabolic markers. *Nature* **2013**, *500*, 541. [CrossRef]

121. Schwiertz, A.; Taras, D.; Schäfer, K.; Beijer, S.; Bos, N.A.; Donus, C.; Hardt, P.D. Microbiota and SCFA in Lean and Overweight Healthy Subjects. *Obesity* **2010**, *18*, 190–195. [CrossRef]

122. Pimentel, M.; Gunsalus, R.P.; Rao, S.S.; Zhang, H. Methanogens in human health and disease. *Am. J. Gastroenterol. Suppl.* **2012**, *1*, 28. [CrossRef]

123. McKay, L.F.; Eastwood, M.; Brydon, W. Methane excretion in man–a study of breath, flatus, and faeces. *Gut* **1985**, *26*, 69–74. [CrossRef] [PubMed]

124. Hallen-Adams, H.E.; Suhr, M.J. Fungi in the healthy human gastrointestinal tract. *Virulence* **2017**, *8*, 352–358. [CrossRef] [PubMed]

125. Wang, Z.K.; Yang, Y.S.; Stefka, A.T.; Sun, G.; Peng, L.H. Review article: Fungal microbiota and digestive diseases. *Aliment. Pharmacol. Ther.* **2014**, *39*, 751–766. [CrossRef] [PubMed]

126. Gouba, N.; Drancourt, M. Digestive tract mycobiota: A source of infection. *Médecine et Maladies Infectieuses* **2015**, *45*, 9–16. [CrossRef]

127. Dollive, S.; Peterfreund, G.L.; Sherrill-Mix, S.; Bittinger, K.; Sinha, R.; Hoffmann, C.; Nabel, C.S.; Hill, D.A.; Artis, D.; Bachman, M.A. A tool kit for quantifying eukaryotic rRNA gene sequences from human microbiome samples. *Genome Biol.* **2012**, *13*, R60. [CrossRef]

128. Ott, S.J.; Kühbacher, T.; Musfeldt, M.; Rosenstiel, P.; Hellmig, S.; Rehman, A.; Drews, O.; Weichert, W.; Timmis, K.N.; Schreiber, S. Fungi and inflammatory bowel diseases: Alterations of composition and diversity. *Scand. J. Gastroenterol.* **2008**, *43*, 831–841. [CrossRef]

129. Richard, M.L.; Lamas, B.; Liguori, G.; Hoffmann, T.W.; Sokol, H. Gut Fungal Microbiota: The Yin and Yang of Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* **2014**, *21*, 656–665. [CrossRef]

130. Ramaswamy, K.; Correa, M.; Koshy, A. Non-healing gastric ulcer associated with Candida infection. *Indian J. Med Microbiol.* **2007**, *25*, 57.

131. Santelmann, H.; Howard, J.M. Yeast metabolic products, yeast antigens and yeasts as possible triggers for irritable bowel syndrome. *Eur. J. Gastroenterol. Hepatol.* **2005**, *17*, 21–26. [CrossRef]

132. Krause, R.; Reisinger, E. *Candida* and antibiotic-associated diarrhoea. *Clin. Microbiol. Infect.* **2005**, *11*, 1–2. [CrossRef] [PubMed]

133. Stringer, A.M.; Gibson, R.J.; Logan, R.M.; Bowen, J.M.; Yeoh, A.S.; Hamilton, J.; Keefe, D.M. Gastrointestinal microflora and mucins may play a critical role in the development of 5-fluorouracil-induced gastrointestinal mucositis. *Exp. Biol. Med.* **2009**, *234*, 430–441. [CrossRef] [PubMed]

134. Cadwell, K. Expanding the Role of the Virome: Commensalism in the Gut. *J. Virol.* **2015**, *89*, 1951–1953. [CrossRef] [PubMed]

135. Duerkop, B.A.; Hooper, L.V. Resident viruses and their interactions with the immune system. *Nat. Immunol.* **2013**, *14*, 654. [CrossRef] [PubMed]

136. Foca, A.; Liberto, M.C.; Quirino, A.; Marascio, N.; Zicca, E.; Pavia, G. Gut Inflammation and Immunity: What Is the Role of the Human Gut Virome? *Mediat. Inflamm.* **2015**, *2015*, 7. [CrossRef] [PubMed]

137. Sun, L.; Nava, G.M.; Stappenbeck, T.S. Host genetic susceptibility, dysbiosis and viral triggers in IBD. *Curr. Opin. Gastroenterol.* **2011**, *27*, 321. [CrossRef] [PubMed]

138. Cario, E. Microbiota and innate immunity in intestinal inflammation and neoplasia. *Curr. Opin. Gastroenterol.* **2013**, *29*, 85–91. [CrossRef]

139. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell* **2011**, *144*, 646–674. [CrossRef]

140. Wang, W.-L.; Xu, S.-Y.; Ren, Z.-G.; Tao, L.; Jiang, J.-W.; Zheng, S.-S. Application of metagenomics in the human gut microbiome. *World J. Gastroenterol. WJG* **2015**, *21*, 803. [CrossRef]

141. Zou, Y.; Xue, W.; Luo, G.; Deng, Z.; Qin, P.; Guo, R.; Sun, H.; Xia, Y.; Liang, S.; Dai, Y.; et al. 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat. Biotechnol.* **2019**, *37*, 179–185. [CrossRef]

142. Gosalbes, M.J.; Durbán, A.; Pignatelli, M.; Abellan, J.J.; Jiménez-Hernández, N.; Pérez-Cobas, A.E.; Latorre, A.; Moya, A. Metatranscriptomic Approach to Analyze the Functional Human Gut Microbiota. *PLoS ONE* **2011**, *6*, e17447. [CrossRef] [PubMed]

143. Franzosa, E.A.; Morgan, X.C.; Segata, N.; Waldron, L.; Reyes, J.; Earl, A.M.; Giannoukos, G.; Boylan, M.R.; Ciulla, D.; Gevers, D.; et al. Relating the metatranscriptome and metagenome of the human gut. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E2329–E2338. [CrossRef] [PubMed]

144. Verberkmoes, N.C.; Russell, A.L.; Shah, M.; Godzik, A.; Rosenquist, M.; Halfvarson, J.; Lefsrud, M.G.; Apajalahti, J.; Tysk, C.; Hettich, R.L. Shotgun metaproteomics of the human distal gut microbiota. *ISME J.* **2009**, *3*, 179. [CrossRef] [PubMed]

145. Erickson, A.R.; Cantarel, B.L.; Lamendella, R.; Darzi, Y.; Mongodin, E.F.; Pan, C.; Shah, M.; Halfvarson, J.; Tysk, C.; Henrissat, B. Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PLoS ONE* **2012**, *7*, e49138. [CrossRef] [PubMed]

146. Kolmeder, C.A.; De Been, M.; Nikkilä, J.; Ritamo, I.; Mättö, J.; Valmu, L.; Salojärvi, J.; Palva, A.; Salonen, A.; de Vos, W.M. Comparative metaproteomics and diversity analysis of human intestinal microbiota testifies for its temporal stability and expression of core functions. *PLoS ONE* **2012**, *7*, e29913. [CrossRef] [PubMed]

147. Guo, X.; Yu, N.; Ding, X.; Wang, J.; Pan, Y. Dime: A novel framework for de novo metagenomic sequence assembly. *J. Comput. Biol.* **2015**, *22*, 159–177. [CrossRef]

148. Laserson, J.; Jojic, V.; Koller, D. Genovo: De novo assembly for metagenomes. *J. Comput. Biol.* **2011**, *18*, 429–443. [CrossRef]

149. Pell, J.; Hintze, A.; Canino-Koning, R.; Howe, A.; Tiedje, J.M.; Brown, C.T. Scaling metagenome sequence assembly with probabilistic de Bruijn graphs. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 13272–13277. [CrossRef]

150. Lai, B.; Ding, R.; Li, Y.; Duan, L.; Zhu, H. A de novo metagenomic assembly program for shotgun DNA reads. *Bioinformatics* **2012**, *28*, 1455–1462. [CrossRef]

151. Peng, Y.; Leung, H.C.M.; Yiu, S.M.; Chin, F.Y.L. Meta-IDBA: A de Novo assembler for metagenomic data. *Bioinformatics* **2011**, *27*, i94–i101. [CrossRef]

152. Treangen, T.J.; Koren, S.; Astrovskaya, I.; Sommer, D.; Liu, B.; Pop, M. MetAMOS: A metagenomic assembly and analysis pipeline for AMOS. *Genome Biol.* **2011**, *12*, P25. [CrossRef]

153. Namiki, T.; Hachiya, T.; Tanaka, H.; Sakakibara, Y. MetaVelvet: An extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res.* **2012**, *40*, e155. [CrossRef] [PubMed]

154. Kultima, J.R.; Sunagawa, S.; Li, J.; Chen, W.; Chen, H.; Mende, D.R.; Arumugam, M.; Pan, Q.; Liu, B.; Qin, J. MOCAT: A metagenomics assembly and gene prediction toolkit. *PLoS ONE* **2012**, *7*, e47656. [CrossRef] [PubMed]

155. Li, R.; Zhu, H.; Ruan, J.; Qian, W.; Fang, X.; Shi, Z.; Li, Y.; Li, S.; Shan, G.; Kristiansen, K.; et al. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **2010**, *20*, 265–272. [CrossRef] [PubMed]

156. Ye, Y.; Tang, H. An ORFome assembly approach to metagenomics sequences analysis. *J. Bioinform. Comput. Biol.* **2009**, *7*, 455–471. [CrossRef]

157. Kim, M.; Ligo, J.G.; Emad, A.; Farnoud, F.; Milenkovic, O.; Veeravalli, V.V. MetaPar: Metagenomic sequence assembly via iterative reclassification. In Proceedings of the 2013 IEEE Global Conference on Signal and Information Processing, Austin, TX, USA, 3–5 December 2013; pp. 43–46.

158. Afiahayati; Sato, K.; Sakakibara, Y. An extended genovo metagenomic assembler by incorporating paired-end information. *PeerJ* **2013**, *1*, e196. [CrossRef]

159. Wu, M.; Eisen, J.A. A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* **2008**, *9*, R151. [CrossRef]

160. Wu, M.; Scott, A.J. Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **2012**, *28*, 1033–1034. [CrossRef]

161. Kerepesi, C.; Banky, D.; Grolmusz, V. AmphoraNet: The webserver implementation of the AMPHORA2 metagenomic workflow suite. *Gene* **2014**, *533*, 538–540. [CrossRef]

162. Gerlach, W.; Stoye, J. Taxonomic classification of metagenomic shotgun sequences with CARMA3. *Nucleic Acids Res.* **2011**, *39*, e91. [CrossRef]

163. Gerlach, W.; Stoye, J. Taxonomic classification of metagenomic shotgun sequences with CARMA3. In *Encyclopedia of Metagenomics: Genes, Genomes and Metagenomes: Basics, Methods, Databases and Tools*; Nelson, K.E., Ed.; Springer: New York, NY, USA, 2015; pp. 653–660.

164. Patil, K.R.; Haider, P.; Pope, P.B.; Turnbaugh, P.J.; Morrison, M.; Scheffer, T.; McHardy, A.C. Taxonomic metagenome sequence assignment with structured output models. *Nat. Methods* **2011**, *8*, 191. [CrossRef] [PubMed]

165. Ounit, R.; Wanamaker, S.; Close, T.J.; Lonardi, S. CLARK: Fast and accurate classification of metagenomic and genomic sequences using discriminative k-mers. *BMC Genom.* **2015**, *16*, 236. [CrossRef] [PubMed]

166. Ghosh, T.S.; Haque, M.; Mande, S.S. DiScRIBinATE: A rapid method for accurate taxonomic classification of metagenomic sequences. *BMC Bioinform.* **2010**, *11*, S14. [CrossRef] [PubMed]

167. Silva, G.G.Z.; Cuevas, D.A.; Dutilh, B.E.; Edwards, R.A. FOCUS: An alignment-free model to identify organisms in metagenomes using non-negative least squares. *PeerJ* **2014**, *2*, e425. [CrossRef] [PubMed]

168. Mohammed, M.H.; Ghosh, T.S.; Reddy, R.M.; Reddy, C.V.S.K.; Singh, N.K.; Mande, S.S. INDUS-a composition-based approach for rapid and accurate taxonomic classification of metagenomic sequences. *BMC Genom.* **2011**, *12*, S4. [CrossRef]

169. Horton, M.; Bodenhausen, N.; Bergelson, J. MARTA: A suite of Java-based tools for assigning taxonomic status to DNA sequences. *Bioinformatics* **2009**, *26*, 568–569. [CrossRef]

170. Wang, Y.; Leung, H.C.; Yiu, S.-M.; Chin, F.Y. MetaCluster 5.0: A two-round binning approach for metagenomic data for low-abundance species in a noisy sample. *Bioinformatics* **2012**, *28*, i356–i362. [CrossRef]

171. Segata, N.; Waldron, L.; Ballarini, A.; Narasimhan, V.; Jousson, O.; Huttenhower, C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* **2012**, *9*, 811. [CrossRef]

172. Truong, D.T.; Franzosa, E.A.; Tickle, T.L.; Scholz, M.; Weingart, G.; Pasolli, E.; Tett, A.; Huttenhower, C.; Segata, N. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **2015**, *12*, 902. [CrossRef]

173. Liu, B.; Gibbons, T.; Ghodsi, M.; Pop, M. MetaPhyler: Taxonomic profiling for metagenomic sequences. In Proceedings of the 2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Hong Kong, China, 18–21 December 2010; pp. 95–100.

174. Kultima, J.R.; Coelho, L.P.; Forslund, K.; Huerta-Cepas, J.; Li, S.S.; Driessen, M.; Voigt, A.Y.; Zeller, G.; Sunagawa, S.; Bork, P. MOCAT2: A metagenomic assembly, annotation and profiling framework. *Bioinformatics* **2016**, *32*, 2520–2523. [CrossRef]

175. Gori, F.; Folino, G.; Jetten, M.S.; Marchiori, E. MTR: Taxonomic annotation of short metagenomic reads using clustering at multiple taxonomic ranks. *Bioinformatics* **2010**, *27*, 196–203. [CrossRef] [PubMed]

176. Rosen, G.L.; Reichenberger, E.R.; Rosenfeld, A.M. NBC: The Naive Bayes Classification tool webserver for taxonomic classification of metagenomic reads. *Bioinformatics* **2010**, *27*, 127–129. [CrossRef] [PubMed]

177. Berger, S.A.; Stamatakis, A. PaPaRa 2.0: A vectorized algorithm for probabilistic phylogeny-aware alignment extension. 2012. Available online: https://pdfs.semanticscholar.org/2b04/11608d4b9fe622ea7aa4df57a1913c625530.pdf (accessed on 1 March 2020).

178. Darling, A.E.; Jospin, G.; Lowe, E.; Matsen, F.A.I.V.; Bik, H.M.; Eisen, J.A. PhyloSift: Phylogenetic analysis of genomes and metagenomes. *PeerJ* **2014**, *2*, e243. [CrossRef] [PubMed]

179. Brady, A.; Salzberg, S.L. Phymm and PhymmBL: Metagenomic phylogenetic classification with interpolated Markov models. *Nat. Methods* **2009**, *6*, 673. [CrossRef]

180. Nalbantoglu, O.U.; Way, S.F.; Hinrichs, S.H.; Sayood, K. RAIphy: Phylogenetic classification of metagenomics samples using iterative refinement of relative abundance index profiles. *BMC Bioinform.* **2011**, *12*, 41. [CrossRef] [PubMed]

181. MacDonald, N.J.; Parks, D.H.; Beiko, R.G. RITA: Rapid identification of high-confidence taxonomic assignments for metagenomic data. *Nucleic Acids Res.* **2012**, *40*, e111. [CrossRef] [PubMed]

182. Monzoorul Haque, M.; Ghosh, T.S.; Komanduri, D.; Mande, S.S. SOrt-ITEMS: Sequence orthology based approach for improved taxonomic estimation of metagenomic sequences. *Bioinformatics* **2009**, *25*, 1722–1730. [CrossRef]

183. Mohammed, M.H.; Ghosh, T.S.; Singh, N.K.; Mande, S.S. SPHINX—An algorithm for taxonomic binning of metagenomic sequences. *Bioinformatics* **2010**, *27*, 22–30. [CrossRef]

184. Diaz, N.N.; Krause, L.; Goesmann, A.; Niehaus, K.; Nattkemper, T.W. TACOA–Taxonomic classification of environmental genomic fragments using a kernelized nearest neighbor approach. *BMC Bioinform.* **2009**, *10*, 56. [CrossRef]

185. Schreiber, F.; Gumrich, P.; Daniel, R.; Meinicke, P. Treephyler: Fast taxonomic profiling of metagenomes. *Bioinformatics* **2010**, *26*, 960–961. [CrossRef]

186. Abubucker, S.; Segata, N.; Goll, J.; Schubert, A.M.; Izard, J.; Cantarel, B.L.; Rodriguez-Mueller, B.; Zucker, J.; Thiagarajan, M.; Henrissat, B. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput. Biol.* **2012**, *8*, e1002358. [CrossRef] [PubMed]

187. Hyland, C.; Pinney, J.W.; McConkey, G.A.; Westhead, D.R. metaSHARK: A WWW platform for interactive exploration of metabolic networks. *Nucleic Acids Res.* **2006**, *34*, W725–W728. [CrossRef] [PubMed]

188. Larsen, P.E.; Collart, F.R.; Field, D.; Meyer, F.; Keegan, K.P.; Henry, C.S.; McGrath, J.; Quinn, J.; Gilbert, J.A. Predicted Relative Metabolomic Turnover (PRMT): Determining metabolic turnover from a coastal marine metagenomic dataset. *Microb. Inform. Exp.* **2011**, *1*, 4. [CrossRef] [PubMed]

189. Li, W. Analysis and comparison of very large metagenomes with fast clustering and functional annotation. *BMC Bioinform.* **2009**, *10*, 359. [CrossRef]

190. Friedman, J.; Alm, E.J. Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* **2012**, *8*, e1002687. [CrossRef]

191. Schwager, E.; Weingart, G.; Bielski, C.; Huttenhower, C. *CCREPE: Compositionality Corrected by Permutation and Renormalization*. 2014. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.683.4029&rep=rep1&type=pdf (accessed on 1 March 2020).

192. Peng, Y.; Leung, H.C.M.; Yiu, S.M.; Chin, F.Y.L. IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **2012**, *28*, 1420–1428. [CrossRef]

193. Harrington, E.D.; Arumugam, M.; Raes, J.; Bork, P.; Relman, D.A. SmashCell: A software framework for the analysis of single-cell amplified genome sequences. *Bioinformatics* **2010**, *26*, 2979–2980. [CrossRef]

194. Khosrovian, K.; Pfahl, D.; Garousi, V. GENSIM 2.0: A customizable process simulation model for software process evaluation. In Proceedings of the International Conference on Software Process, Leipzig, Germany, 10–11 May 2008; pp. 294–306.

195. Richter, D.C.; Ott, F.; Auch, A.F.; Schmid, R.; Huson, D.H. MetaSim—A Sequencing Simulator for Genomics and Metagenomics. *PLoS ONE* **2008**, *3*, e3373. [CrossRef]

196. White, J.R.; Nagarajan, N.; Pop, M. Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Comput. Biol.* **2009**, *5*, e1000352. [CrossRef]

197. Segata, N.; Izard, J.; Waldron, L.; Gevers, D.; Miropolsky, L.; Garrett, W.S.; Huttenhower, C. Metagenomic biomarker discovery and explanation. *Genome Biol.* **2011**, *12*, R60. [CrossRef]

198. Kristiansson, E.; Hugenholtz, P.; Dalevi, D. ShotgunFunctionalizeR: An R-package for functional comparison of metagenomes. *Bioinformatics* **2009**, *25*, 2737–2738. [CrossRef] [PubMed]

199. Knights, D.; Kuczynski, J.; Charlson, E.S.; Zaneveld, J.; Mozer, M.C.; Collman, R.G.; Bushman, F.D.; Knight, R.; Kelley, S.T. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **2011**, *8*, 761. [CrossRef] [PubMed]

200. Seshadri, R.; Kravitz, S.A.; Smarr, L.; Gilna, P.; Frazier, M. CAMERA: A community resource for metagenomics. *PLoS Biol.* **2007**, *5*, e75. [CrossRef] [PubMed]

201. Riehle, K.; Coarfa, C.; Jackson, A.; Ma, J.; Tandon, A.; Paithankar, S.; Raghuraman, S.; Mistretta, T.-A.; Saulnier, D.; Raza, S. The Genboree Microbiome Toolset and the analysis of 16S rRNA microbial sequences. *BMC Bioinform.* **2012**, *13*, S11. [CrossRef] [PubMed]

202. Asnicar, F.; Weingart, G.; Tickle, T.L.; Huttenhower, C.; Segata, N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* **2015**, *3*, e1029. [CrossRef] [PubMed]

203. Huson, D.H.; Auch, A.F.; Qi, J.; Schuster, S.C. MEGAN analysis of metagenomic data. *Genome Res.* **2007**, *17*, 377–386. [CrossRef]

204. Goll, J.; Rusch, D.B.; Tanenbaum, D.M.; Thiagarajan, M.; Li, K.; Methé, B.A.; Yooseph, S. METAREP: JCVI metagenomics reports—an open source tool for high-performance comparative metagenomics. *Bioinformatics* **2010**, *26*, 2631–2632. [CrossRef]

205. Meyer, F.; Paarmann, D.; D'Souza, M.; Olson, R.; Glass, E.M.; Kubal, M.; Paczian, T.; Rodriguez, A.; Stevens, R.; Wilke, A. The metagenomics RAST server—A public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinform.* **2008**, *9*, 386. [CrossRef]

206. 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. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **2009**, *75*, 7537–7541. [CrossRef]

207. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Pena, A.G.; Goodrich, J.K.; Gordon, J.I. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335. [CrossRef]

208. Arumugam, M.; Harrington, E.D.; Foerstner, K.U.; Raes, J.; Bork, P. SmashCommunity: A metagenomic annotation and analysis tool. *Bioinformatics* **2010**, *26*, 2977–2978. [CrossRef] [PubMed]

209. Parks, D.H.; Beiko, R.G. STAMP: Statistical analysis of metagenomic profiles. In *Encyclopedia of Metagenomics: Genes, Genomes and Metagenomes. Basics, Methods, Databases and Tools*; Springer: Berlin, Germany, 2015; pp. 641–645.

210. Stocker, S.; Snajder, R.; Rainer, J.; Trajanoski, S.; Gorkiewicz, G.; Trajanoski, Z.; Thallinger, G.G. SnoWMAn: High-throughput phylotyping, analysis and comparison of microbial communities. 2011; Under Revision.

211. Huse, S.M.; Welch, D.B.M.; Voorhis, A.; Shipunova, A.; Morrison, H.G.; Eren, A.M.; Sogin, M.L. VAMPS: A website for visualization and analysis of microbial population structures. *BMC Bioinform.* **2014**, *15*, 41. [CrossRef] [PubMed]

212. Rinschen, M.M.; Ivanisevic, J.; Giera, M.; Siuzdak, G. Identification of bioactive metabolites using activity metabolomics. *Nat. Rev. Mol. Cell Biol.* **2019**. [CrossRef]

213. Piazza, I.; Kochanowski, K.; Cappelletti, V.; Fuhrer, T.; Noor, E.; Sauer, U.; Picotti, P. A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication. *Cell* **2018**, *172*, 358–372.e323. [CrossRef]

214. Larsen, P.E.; Dai, Y. Metabolome of human gut microbiome is predictive of host dysbiosis. *GigaScience* **2015**, *4*, 42. [CrossRef]

215. Holmes, E.; Li, J.V.; Marchesi, J.R.; Nicholson, J.K. Gut microbiota composition and activity in relation to host metabolic phenotype and disease risk. *Cell Metab.* **2012**, *16*, 559–564. [CrossRef]

216. Rooks, M.G.; Garrett, W.S. Gut microbiota, metabolites and host immunity. *Nat. Rev. Immunol.* **2016**, *16*, 341. [CrossRef]

217. Morrison, D.J.; Preston, T. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes* **2016**, *7*, 189–200. [CrossRef]

218. Psichas, A.; Sleeth, M.; Murphy, K.; Brooks, L.; Bewick, G.; Hanyaloglu, A.; Ghatei, M.; Bloom, S.; Frost, G. The short chain fatty acid propionate stimulates GLP-1 and PYY secretion via free fatty acid receptor 2 in rodents. *Int. J. Obes.* **2015**, *39*, 424. [CrossRef]

219. Chambers, E.S.; Morrison, D.J.; Frost, G. Control of appetite and energy intake by SCFA: What are the potential underlying mechanisms? *Proc. Nutr. Soc.* **2015**, *74*, 328–336. [CrossRef] [PubMed]

220. Corrêa-Oliveira, R.; Fachi, J.L.; Vieira, A.; Sato, F.T.; Vinolo, M.A.R. Regulation of immune cell function by short-chain fatty acids. *Clin. Transl. Immunol.* **2016**, *5*, e73. [CrossRef] [PubMed]

221. Levy, M.; Blacher, E.; Elinav, E. Microbiome, metabolites and host immunity. *Curr. Opin. Microbiol.* **2017**, *35*, 8–15. [CrossRef] [PubMed]

222. Claus, S.P.; Guillou, H.; Ellero-Simatos, S. The gut microbiota: A major player in the toxicity of environmental pollutants? *NPJ Biofilms Microbiomes* **2016**, *2*, 16003. [CrossRef] [PubMed]

223. LeBlanc, J.G.; Milani, C.; De Giori, G.S.; Sesma, F.; Van Sinderen, D.; Ventura, M. Bacteria as vitamin suppliers to their host: A gut microbiota perspective. *Curr. Opin. Biotechnol.* **2013**, *24*, 160–168. [CrossRef] [PubMed]

224. Laparra, J.M.; Sanz, Y. Interactions of gut microbiota with functional food components and nutraceuticals. *Pharmacol. Res.* **2010**, *61*, 219–225. [CrossRef] [PubMed]

225. Carding, S.; Verbeke, K.; Vipond, D.T.; Corfe, B.M.; Owen, L.J. Dysbiosis of the gut microbiota in disease. *Microb. Ecol. Health Dis.* **2015**, *26*, 26191. [CrossRef]

226. Yau, Y.; Leong, R.W.; Zeng, M.; Wasinger, V.C. Proteomics and metabolomics in inflammatory bowel disease. *J. Gastroenterol. Hepatol.* **2013**, *28*, 1076–1086. [CrossRef]

227. Lin, H.M.; Helsby, N.A.; Rowan, D.D.; Ferguson, L.R. Using metabolomic analysis to understand inflammatory bowel diseases. *Inflamm. Bowel Dis.* **2011**, *17*, 1021–1029. [CrossRef]

228. Le Gall, G.; Noor, S.O.; Ridgway, K.; Scovell, L.; Jamieson, C.; Johnson, I.T.; Colquhoun, I.J.; Kemsley, E.K.; Narbad, A. Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. *J. Proteome Res.* **2011**, *10*, 4208–4218. [CrossRef]

229. Phua, L.C.; Chue, X.P.; Koh, P.K.; Cheah, P.Y.; Ho, H.K.; Chan, E.C.Y. Non-invasive fecal metabonomic detection of colorectal cancer. *Cancer Biol. Ther.* **2014**, *15*, 389–397. [CrossRef] [PubMed]

230. Nicholson, J.K.; Holmes, E.; Kinross, J.; Burcelin, R.; Gibson, G.; Jia, W.; Pettersson, S. Host-gut microbiota metabolic interactions. *Science* **2012**, *336*, 1262–1267. [CrossRef] [PubMed]

231. Vernocchi, P.; Del Chierico, F.; Putignani, L. Gut Microbiota Profiling: Metabolomics Based Approach to Unravel Compounds Affecting Human Health. *Front. Microbiol.* **2016**, *7*. [CrossRef] [PubMed]

232. Binder, H.J. Role of colonic short-chain fatty acid transport in diarrhea. *Annu. Rev. Physiol.* **2010**, *72*, 297–313. [CrossRef] [PubMed]

233. Chambers, E.S.; Viardot, A.; Psichas, A.; Morrison, D.J.; Murphy, K.G.; Zac-Varghese, S.E.; MacDougall, K.; Preston, T.; Tedford, C.; Finlayson, G.S. Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. *Gut* **2015**, *64*, 1744–1754. [CrossRef]

234. Qiu, Y.; Cai, G.; Su, M.; Chen, T.; Liu, Y.; Xu, Y.; Ni, Y.; Zhao, A.; Cai, S.; Xu, L.X. Urinary metabonomic study on colorectal cancer. *J. Proteome Res.* **2010**, *9*, 1627–1634. [CrossRef]

235. Calvani, R.; Miccheli, A.; Capuani, G.; Miccheli, A.T.; Puccetti, C.; Delfini, M.; Iaconelli, A.; Nanni, G.; Mingrone, G. Gut microbiome-derived metabolites characterize a peculiar obese urinary metabotype. *Int. J. Obes.* **2010**, *34*, 1095. [CrossRef]

236. Zhao, L.; Liu, X.; Xie, L.; Gao, H.; Lin, D. 1H NMR-based metabonomic analysis of metabolic changes in streptozotocin-induced diabetic rats. *Anal. Sci.* **2010**, *26*, 1277–1282. [CrossRef]

237. Zheng, X.; Xie, G.; Zhao, A.; Zhao, L.; Yao, C.; Chiu, N.H.; Zhou, Z.; Bao, Y.; Jia, W.; Nicholson, J.K. The footprints of gut microbial–mammalian co-metabolism. *J. Proteome Res.* **2011**, *10*, 5512–5522. [CrossRef]

238. Amaretti, A.; Raimondi, S.; Leonardi, A.; Quartieri, A.; Rossi, M. Hydrolysis of the rutinose-conjugates flavonoids rutin and hesperidin by the gut microbiota and bifidobacteria. *Nutrients* **2015**, *7*, 2788–2800. [CrossRef]

239. Taverniti, V.; Guglielmetti, S. Health-promoting properties of Lactobacillus helveticus. *Front. Microbiol.* **2012**, *3*, 392. [CrossRef] [PubMed]

240. Marín, L.; Miguélez, E.M.; Villar, C.J.; Lombó, F. Bioavailability of dietary polyphenols and gut microbiota metabolism: Antimicrobial properties. *BioMed Res. Int.* **2015**, *2015*. [CrossRef] [PubMed]

241. Dawson, P.A.; Lan, T.; Rao, A. Bile acid transporters. *J. Lipid Res.* **2009**, *50*, 2340–2357. [CrossRef] [PubMed]

242. Suhre, K.; Meisinger, C.; Döring, A.; Altmaier, E.; Belcredi, P.; Gieger, C.; Chang, D.; Milburn, M.V.; Gall, W.E.; Weinberger, K.M. Metabolic footprint of diabetes: A multiplatform metabolomics study in an epidemiological setting. *PLoS ONE* **2010**, *5*, e13953. [CrossRef]

243. Swann, J.R.; Want, E.J.; Geier, F.M.; Spagou, K.; Wilson, I.D.; Sidaway, J.E.; Nicholson, J.K.; Holmes, E. Systemic gut microbial modulation of bile acid metabolism in host tissue compartments. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4523–4530. [CrossRef]

244. Wang, Z.; Klipfell, E.; Bennett, B.J.; Koeth, R.; Levison, B.S.; DuGar, B.; Feldstein, A.E.; Britt, E.B.; Fu, X.; Chung, Y.-M. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **2011**, *472*, 57. [CrossRef]

245. Martin, F.-P.J.; Sprenger, N.; Montoliu, I.; Rezzi, S.; Kochhar, S.; Nicholson, J.K. Dietary modulation of gut functional ecology studied by fecal metabonomics. *J. Proteome Res.* **2010**, *9*, 5284–5295. [CrossRef]

246. Bercik, P.; Denou, E.; Collins, J.; Jackson, W.; Lu, J.; Jury, J.; Deng, Y.; Blennerhassett, P.; Macri, J.; McCoy, K.D. The intestinal microbiota affect central levels of brain-derived neurotropic factor and behavior in mice. *Gastroenterology* **2011**, *141*, 599–609.e593. [CrossRef]

247. Keszthelyi, D.; Troost, F.; Masclee, A. Understanding the role of tryptophan and serotonin metabolism in gastrointestinal function. *Neurogastroenterol. Motil.* **2009**, *21*, 1239–1249. [CrossRef]

248. 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]

249. Said, H.M. Intestinal absorption of water-soluble vitamins in health and disease. *Biochem. J.* **2011**, *437*, 357–372. [CrossRef] [PubMed]

250. Rossi, M.; Amaretti, A. *Probiotic Properties of Bifidobacteria*; Caister Academic Press: Norfolk, UK, 2010.

251. Magnúsdóttir, S.; Ravcheev, D.; De Crécy-Lagard, V.; Thiele, I. Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Front. Genet.* **2015**, *6*, 148. [CrossRef] [PubMed]

252. Hanfrey, C.C.; Pearson, B.M.; Hazeldine, S.; Lee, J.; Gaskin, D.J.; Woster, P.M.; Phillips, M.A.; Michael, A.J. Alternative spermidine biosynthetic route is critical for growth of Campylobacter jejuni and is the dominant polyamine pathway in human gut microbiota. *J. Biol. Chem.* **2011**, *286*, 43301–43312. [CrossRef] [PubMed]

253. Serino, M.; Luche, E.; Gres, S.; Baylac, A.; Bergé, M.; Cenac, C.; Waget, A.; Klopp, P.; Iacovoni, J.; Klopp, C. Metabolic adaptation to a high-fat diet is associated with a change in the gut microbiota. *Gut* **2012**, *61*, 543–553. [CrossRef] [PubMed]

254. Muccioli, G.G.; Naslain, D.; Bäckhed, F.; Reigstad, C.S.; Lambert, D.M.; Delzenne, N.M.; Cani, P.D. The endocannabinoid system links gut microbiota to adipogenesis. *Mol. Syst. Biol.* **2010**, *6*, 392. [CrossRef] [PubMed]

255. Fiehn, O. Metabolomics—the link between genotypes and phenotypes. *Plant Mol. Biol.* **2002**, *48*, 155–171. [CrossRef]

256. Weckwerth, W.; Morgenthal, K. Metabolomics: From pattern recognition to biological interpretation. *Drug Discov. Today* **2005**, *10*, 1551–1558. [CrossRef]

257. De Preter, V.; Van Staeyen, G.; Esser, D.; Rutgeerts, P.; Verbeke, K. Development of a screening method to determine the pattern of fermentation metabolites in faecal samples using on-line purge-and-trap gas chromatographic–mass spectrometric analysis. *J. Chromatogr. A* **2009**, *1216*, 1476–1483. [CrossRef]

258. Dudley, E.; Yousef, M.; Wang, Y.; Griffiths, W.J. Targeted metabolomics and mass spectrometry. *Adv. Protein Chem. Struct. Biol.* **2010**, *80*, 45–83. [CrossRef]

259. Nicholson, J.K.; Connelly, J.; Lindon, J.C.; Holmes, E. Metabonomics: A platform for studying drug toxicity and gene function. *Nat. Rev. Drug Discov.* **2002**, *1*, 153. [CrossRef]

260. Wikoff, W.R.; Anfora, A.T.; Liu, J.; Schultz, P.G.; Lesley, S.A.; Peters, E.C.; Siuzdak, G. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 3698–3703. [CrossRef] [PubMed]

261. Panopoulos, A.D.; Yanes, O.; Ruiz, S.; Kida, Y.S.; Diep, D.; Tautenhahn, R.; Herrerías, A.; Batchelder, E.M.; Plongthongkum, N.; Lutz, M.; et al. The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. *Cell Res.* **2011**, *22*, 168. [CrossRef] [PubMed]

262. Wu, J.; Gao, Y. Physiological conditions can be reflected in human urine proteome and metabolome. *Expert Rev. Proteom.* **2015**, *12*, 623–636. [CrossRef] [PubMed]

263. Smith, C.A.; Want, E.J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **2006**, *78*, 779–787. [CrossRef]

264. Baran, R.; Kochi, H.; Saito, N.; Suematsu, M.; Soga, T.; Nishioka, T.; Robert, M.; Tomita, M. MathDAMP: A package for differential analysis of metabolite profiles. *BMC Bioinform.* **2006**, *7*, 530. [CrossRef]

265. Lommen, A. MetAlign: Interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. *Anal. Chem.* **2009**, *81*, 3079–3086. [CrossRef]

266. Katajamaa, M.; Miettinen, J.; Orešič, M. MZmine: Toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* **2006**, *22*, 634–636. [CrossRef]

267. De Preter, V.; Verbeke, K. Metabolomics as a diagnostic tool in gastroenterology. *World J. Gastrointest. Pharmacol. Ther.* **2013**, *4*, 97. [CrossRef]

268. Kopka, J.; Schauer, N.; Krueger, S.; Birkemeyer, C.; Usadel, B.; Bergmüller, E.; Dörmann, P.; Weckwerth, W.; Gibon, Y.; Stitt, M.; et al. GMD@CSB.DB: The Golm Metabolome Database. *Bioinformatics* **2004**, *21*, 1635–1638. [CrossRef]

269. Smith, C.A.; O'Maille, G.; Want, E.J.; Qin, C.; Trauger, S.A.; Brandon, T.R.; Custodio, D.E.; Abagyan, R.; Siuzdak, G. METLIN: A metabolite mass spectral database. *Ther. Drug Monit.* **2005**, *27*, 747–751. [CrossRef]

270. Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **2007**, *36*, D480–D484. [CrossRef] [PubMed]

271. 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.* **2007**, *36*, D344–D350. [CrossRef] [PubMed]

272. Wishart, D.S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A.C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; et al. HMDB: The Human Metabolome Database. *Nucleic Acids Res.* **2007**, *35*, D521–D526. [CrossRef] [PubMed]

273. Wishart, D.S.; Feunang, Y.D.; Marcu, A.; Guo, A.C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; et al. HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Res.* **2017**, *46*, D608–D617. [CrossRef] [PubMed]

274. Ulrich, E.L.; Akutsu, H.; Doreleijers, J.F.; Harano, Y.; Ioannidis, Y.E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z. BioMagResBank. *Nucleic Acids Res.* **2007**, *36*, D402–D408. [CrossRef] [PubMed]

275. Cui, Q.; Lewis, I.A.; Hegeman, A.D.; Anderson, M.E.; Li, J.; Schulte, C.F.; Westler, W.M.; Eghbalnia, H.R.; Sussman, M.R.; Markley, J.L. Metabolite identification via the Madison Metabolomics Consortium Database. *Nat. Biotechnol.* **2008**, *26*, 162–164. [CrossRef]

276. Schellenberger, J.; Park, J.O.; Conrad, T.M.; Palsson, B.Ø. BiGG: A Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. *BMC Bioinform.* **2010**, *11*, 213. [CrossRef]

277. Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K. MassBank: A public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **2010**, *45*, 703–714. [CrossRef]

278. Skogerson, K.; Wohlgemuth, G.; Barupal, D.K.; Fiehn, O. The volatile compound BinBase mass spectral database. *BMC Bioinform.* **2011**, *12*, 321. [CrossRef]

279. May, J.C.; McLean, J.A. Advanced Multidimensional Separations in Mass Spectrometry: Navigating the Big Data Deluge. *Annu. Rev. Anal. Chem.* **2016**, *9*, 387–409. [CrossRef]

280. Schmidt, R.; Etalo, D.W.; De Jager, V.; Gerards, S.; Zweers, H.; De Boer, W.; Garbeva, P. Microbial small talk: Volatiles in fungal–bacterial interactions. *Front. Microbiol.* **2016**, *6*, 1495. [CrossRef] [PubMed]

281. Young, G. Fungal pathogenesis: Fungal communication gets volatile. *Nat. Rev. Microbiol.* **2009**, *7*, 6. [CrossRef]

282. Baranska, A.; Mujagic, Z.; Smolinska, A.; Dallinga, J.; Jonkers, D.; Tigchelaar, E.; Dekens, J.; Zhernakova, A.; Ludwig, T.; Masclee, A. Volatile organic compounds in breath as markers for irritable bowel syndrome: A metabolomic approach. *Aliment. Pharmacol. Ther.* **2016**, *44*, 45–56. [CrossRef] [PubMed]

283. Chong, J.; Xia, J. Computational approaches for integrative analysis of the metabolome and microbiome. *Metabolites* **2017**, *7*, 62. [CrossRef] [PubMed]

284. Kowalski, B.R. *Chemometrics: Theory and Application*; ACS Publications: Washington, DC, USA, 1977.

285. Tauler, R.; Parastar, H. Big (bio)chemical data mining using Chemometric methods: A need for chemists. *Angew. Chem. Int. Ed.* **2018**. [CrossRef]

286. Noecker, C.; Eng, A.; Srinivasan, S.; Theriot, C.M.; Young, V.B.; Jansson, J.K.; Fredricks, D.N.; Borenstein, E. Metabolic model-based integration of microbiome taxonomic and metabolomic profiles elucidates mechanistic links between ecological and metabolic variation. *MSystems* **2016**, *1*, e00013–e00015. [CrossRef]

287. Noronha, A.; Modamio, J.; Jarosz, Y.; Guerard, E.; Sompairac, N.; Preciat, G.; Daníelsdóttir, A.D.; Krecke, M.; Merten, D.; Haraldsdóttir, H.S. The Virtual Metabolic Human database: Integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic Acids Res.* **2018**, *47*, D614–D624.

288. Aguiar-Pulido, V.; Huang, W.; Suarez-Ulloa, V.; Cickovski, T.; Mathee, K.; Narasimhan, G. Metagenomics, Metatranscriptomics, and Metabolomics Approaches for Microbiome Analysis:Supplementary Issue: Bioinformatics Methods and Applications for Big Metagenomics Data. *Evol. Bioinform.* **2016**, *12s1*, EBO.S36436. [CrossRef]

289. Kaput, J.; Van Ommen, B.; Kremer, B.; Priami, C.; Monteiro, J.P.; Morine, M.; Pepping, F.; Diaz, Z.; Fenech, M.; He, Y. Consensus statement understanding health and malnutrition through a systems approach: The ENOUGH program for early life. *Genes Nutr.* **2014**, *9*, 378. [CrossRef]

290. Shoaie, S.; Nielsen, J. Elucidating the interactions between the human gut microbiota and its host through metabolic modeling. *Front. Genet.* **2014**, *5*, 86. [CrossRef]

291. Jacobsen, U.P.; Nielsen, H.B.; Hildebrand, F.; Raes, J.; Sicheritz-Ponten, T.; Kouskoumvekaki, I.; Panagiotou, G. The chemical interactome space between the human host and the genetically defined gut metabotypes. *ISME J.* **2012**, *7*, 730. [CrossRef] [PubMed]

292. McHardy, I.H.; Goudarzi, M.; Tong, M.; Ruegger, P.M.; Schwager, E.; Weger, J.R.; Graeber, T.G.; Sonnenburg, J.L.; Horvath, S.; Huttenhower, C.; et al. Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. *Microbiome* **2013**, *1*, 17. [CrossRef] [PubMed]

293. Thaiss, C.A.; Elinav, E. The remedy within: Will the microbiome fulfill its therapeutic promise? *J. Mol. Med.* **2017**, *95*, 1021–1027. [CrossRef] [PubMed]

294. Young, V.B.; Kahn, S.A.; Schmidt, T.M.; Chang, E.B. Studying the Enteric Microbiome in Inflammatory Bowel Diseases: Getting through the Growing Pains and Moving Forward. *Front. Microbiol.* **2011**, *2*, 144. [CrossRef] [PubMed]

295. Uhr, G.T.; Dohnalová, L.; Thaiss, C.A. The dimension of time in host-microbiome interactions. *MSystems* **2019**, *4*, e00216-18. [CrossRef] [PubMed]

296. Hacılar, H.; Nalbantoğlu, O.U.; Bakir-Güngör, B. Machine Learning Analysis of Inflammatory Bowel Disease-Associated Metagenomics Dataset. In Proceedings of the 2018 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo, Bosnia-Herzegovina, 20–23 September 2018; pp. 434–438.

297. Dave, M.; Higgins, P.D.; Middha, S.; Rioux, K.P. The human gut microbiome: Current knowledge, challenges, and future directions. *Transl. Res.* **2012**, *160*, 246–257. [CrossRef] [PubMed]

298. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M.Y.; Geiger, T.; Mann, M.; Cox, J. The Perseus computational platform for comprehensive analysis of (prote) omics data. *Nat. Methods* **2016**, *13*, 731. [CrossRef] [PubMed]

299. Lin, E.; Lane, H.-Y. Machine learning and systems genomics approaches for multi-omics data. *Biomark. Res.* **2017**, *5*, 2. [CrossRef]

300. Swan, A.L.; Stekel, D.J.; Hodgman, C.; Allaway, D.; Alqahtani, M.H.; Mobasheri, A.; Bacardit, J. A machine learning heuristic to identify biologically relevant and minimal biomarker panels from omics data. *BMC Genom.* **2015**, *16*, S2. [CrossRef]

301. Huang, S.; Chaudhary, K.; Garmire, L.X. More is better: Recent progress in multi-omics data integration methods. *Front. Genet.* **2017**, *8*, 84. [CrossRef]

302. Dubourg-Felonneau, G.; Cannings, T.; Cotter, F.; Thompson, H.; Patel, N.; Cassidy, J.W.; Clifford, H.W. A Framework for Implementing Machine Learning on Omics Data. *arXiv* **2018**, arXiv:1811.10455.

303. Saulnier, D.M.; Riehle, K.; Mistretta, T.A.; Diaz, M.A.; Mandal, D.; Raza, S.; Weidler, E.M.; Qin, X.; Coarfa, C.; Milosavljevic, A.; et al. Gastrointestinal Microbiome Signatures of Pediatric Patients With Irritable Bowel Syndrome. *Gastroenterology* **2011**, *141*, 1782–1791. [CrossRef] [PubMed]

304. Christley, S.; Cockrell, C.; An, G. Computational Studies of the Intestinal Host-Microbiota Interactome. *Computation* **2015**, *3*, 2–28. [CrossRef]

305. Costello, Z.; Martin, H.G. A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *NPJ Syst. Biol. Appl.* **2018**, *4*, 19. [CrossRef] [PubMed]

306. Acharjee, A.; Ament, Z.; West, J.A.; Stanley, E.; Griffin, J.L. Integration of metabolomics, lipidomics and clinical data using a machine learning method. *BMC Bioinform.* **2016**, *17*, 440. [CrossRef]

307. Sung, J.; Kim, S.; Cabatbat, J.J.T.; Jang, S.; Jin, Y.-S.; Jung, G.Y.; Chia, N.; Kim, P.-J. Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis. *Nat. Commun.* **2017**, *8*, 15393. [CrossRef]

308. Hsu, Y.-H.H.; Churchhouse, C.; Pers, T.H.; Mercader, J.M.; Metspalu, A.; Fischer, K.; Fortney, K.; Morgen, E.K.; Gonzalez, C.; Gonzalez, M.E.; et al. PAIRUP-MS: Pathway analysis and imputation to relate unknowns in profiles from mass spectrometry-based metabolite data. *PLoS Comput. Biol.* **2019**, *15*, e1006734. [CrossRef]

309. Menon, R.; Ramanan, V.; Korolev, K.S. Interactions between species introduce spurious associations in microbiome studies. *PLoS Comput. Biol.* **2018**, *14*, e1005939. [CrossRef]

310. Sankaran, K.; Holmes, S. Interactive Visualization of Hierarchically Structured Data. *J. Comput. Graph. Stat.* **2018**, *27*, 553–563. [CrossRef]

311. Rahnavard, G.; Franzosa, E.A.; McIver, L.J.; Schwager, E.; Weingart, G.; Moon, Y.S.; Morgan, X.C.; Waldron, L.; Huttenhower, C. High-Sensitivity Pattern Discovery in Large Multi'omic Datasets. 2017. Available online: https://huttenhower.sph.harvard.edu/halla (accessed on 1 March 2020).

312. Wang, J.; Tang, L.; Zhou, H.; Zhou, J.; Glenn, T.C.; Shen, C.-L.; Wang, J.-S. Long-term treatment with green tea polyphenols modifies the gut microbiome of female sprague-dawley rats. *J. Nutr. Biochem.* **2018**, *56*, 55–64. [CrossRef]

313. Kalantar-zadeh, K.; Yao, C.K.; Berean, K.J.; Ha, N.; Ou, J.Z.; Ward, S.A.; Pillai, N.; Hill, J.; Cottrell, J.J.; Dunshea, F.R.; et al. Intestinal Gas Capsules: A Proof-of-Concept Demonstration. *Gastroenterology* **2016**, *150*, 37–39. [CrossRef] [PubMed]

314. Amrane, S.; Raoult, D.; Lagier, J.-C. Metagenomics, culturomics, and the human gut microbiota. *Expert Rev. Anti-Infect. Ther.* **2018**, *16*, 373–375. [CrossRef] [PubMed]

315. Bilen, M.; Dufour, J.-C.; Lagier, J.-C.; Cadoret, F.; Daoud, Z.; Dubourg, G.; Raoult, D. The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species. *Microbiome* **2018**, *6*, 94. [CrossRef] [PubMed]

316. Lagier, J.-C.; Hugon, P.; Khelaifia, S.; Fournier, P.-E.; La Scola, B.; Raoult, D. The Rebirth of Culture in Microbiology through the Example of Culturomics To Study Human Gut Microbiota. *Clin. Microbiol. Rev.* **2015**, *28*, 237–264. [CrossRef] [PubMed]

317. Sanchez, C.; Lachaize, C.; Janody, F.; Bellon, B.; Röder, L.; Euzenat, J.; Rechenmann, F.; Jacq, B. Grasping at molecular interactions and genetic networks in Drosophila melanogaster using FlyNets, an Internet database. *Nucleic Acids Res.* **1999**, *27*, 89–94. [CrossRef]

318. Gundogdu, A.; Nalbantoglu, U. Human genome-microbiome interaction: Metagenomics frontiers for the aetiopathology of autoimmune diseases. *Microb. Genom.* **2017**, *3*, e000112. [CrossRef]

319. Guven-Maiorov, E.; Tsai, C.-J.; Nussinov, R. Structural host-microbiota interaction networks. *PLoS Comput. Biol.* **2017**, *13*, e1005579. [CrossRef]

320. Yuan, C.; Burns, M.; Subramanian, S.; Blekhman, R. Interaction between Host MicroRNAs and the Gut Microbiota in Colorectal Cancer. *bioRxiv* **2017**, 192401. [CrossRef]

321. Tomás-Barberán, F.A.; Espín, J.C. Effect of Food Structure and Processing on (Poly)phenol–Gut Microbiota Interactions and the Effects on Human Health. *Annu. Rev. Food Sci. Technol.* **2019**. [CrossRef]