Trends in Microbiology

CellPress REVIEWS

Opinion

Human Gut Microbiome: Function Matters

Anna Heintz-Buschart[1,2,3] and Paul Wilmes[1,*,@]

The human gut microbiome represents a complex ecosystem contributing essential functions to its host. Recent large-scale metagenomic studies have provided insights into its structure and functional potential. However, the functional repertoire which is actually contributed to human physiology remains largely unexplored. Here, by leveraging recent omics datasets, we challenge current assumptions regarding key attributes of the functional gut microbiome, in particular with respect to its variability. We further argue that the closing of existing gaps in functional knowledge should be addressed by a most-wanted gene list, the development and application of molecular and cellular high-throughput measurements, the development and sensible use of experimental models, as well as the direct study of observable molecular effects in the human host.

Trends

Functional omics are becoming more accessible, and increasing numbers of studies have employed them, demonstrating their potential in identifying functional traits of the microbiome related to health and disease.

Functional omes display greater variability and sensitivity to perturbation, also in cases of where changes in taxonomic composition are minimal, and they can resolve gut-compartment-specific information.

Methods for resolving functional differences in meta-omic datasets to the taxa contributing them have been developed and are necessary to understand the impact of microbial functions on human physiology.

The Functional Microbiome

The complex assemblages of microorganisms which populate the human gastrointestinal tract are emerging as key players in governing human health and disease. Several essential functions conferred by the gut microbiome on the human host testify to its importance. These include the fermentation of indigestible food components into absorbable metabolites, the synthesis of essential vitamins, the removal of toxic compounds, the outcompetition of pathogens, the strengthening of the intestinal barrier, and the stimulation and regulation of the immune system (see recent reviews [1–7]). Most of these functions are interconnected and tightly intertwined with human physiology. For example, products of microbial fermentation, such as short-chain fatty acids, represent essential substrates for intestinal cells and play important roles in immunomodulatory processes, such as T cell differentiation, which, in turn, may affect the gut microbiome. Although much has been learnt about these tight interrelationships through carefully conducted mechanistic studies, the extensive diversity of microorganisms and molecules in the gut implies that our understanding of this expanse requires a comprehensive toolset to enable new discoveries. More specifically, the emergent functional complement, which is actually contributed to human physiology by the gut microbiome, requires detailed assessment and systematic study.

A widely applied strategy to deconvolute the complex of interactions and to provide avenues to improve human health is constituted by a triad comprising (i) high-resolution, high-fidelity, and high-throughput **omics** (see Glossary) of microbial biomass and comparative analyses, (ii) hypothesis testing in relevant model **experimental systems**, and (iii) **intervention studies** in humans. Ideally, the first type of study should yield testable hypotheses relating to the nature of functions conferred by specific microbiota on human physiology, how and why these functions differ between individuals (most notably between diseased and healthy individuals), and their impact on human health. In this context, much has been described regarding the structural characteristics of the gut microbiota through the application of 16S rRNA gene amplicon sequencing and **metagenomics** [8]. However, to formulate concrete hypotheses for mechanistic studies aimed at understanding dependencies between host and microbes,

[1]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 7 avenue des Hauts-Fourneaux, L-4362 Esch-sur-Alzette, Luxembourg
[2]Current addresses: Helmholtz-Centre for Environmental Research GmbH – UFZ, Department of Soil Ecology, Theodor-Lieser-Str. 4, 06120 Halle (Saale), Germany
[3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany
[@]Twitter: @wilmeslab

*Correspondence: paul.wilmes@uni.lu (P. Wilmes).

observational studies should pinpoint specific functions of the microbiome to specific microbial populations conferring these. In addition, these studies should identify biologically relevant and informative read-outs of health status. Here, functional omics are indispensable.

Variability and Information Content of the Functional Microbiome

The generation and integration of functional omics read-outs derived from **metatranscriptomic**, **metaproteomic**, and **metabolomic** analyses allow a detailed functional assessment of the human gut microbiome [9–18]. It has been observed that the functional omes display greater variability and sensitivity to perturbation when compared to the information content of the metagenome [9–12,14–16,18,19]. Therefore, functional omics are expected to more accurately portray health and disease states [12,13,18,20]. For example, changes in gene expression have been detected in response to dietary interventions, such as fermented milk products [19], and the oral intake of medication [14], despite only minimal changes in observed community structure in both cases. These observations are seemingly in contradiction to the widely accepted [21,22] interpretation of metagenomic data whereby metagenomic **functional profiles** are less variable compared to **taxonomic profiles** [23] (Figure 1A). However, the latter notion may not faithfully reflect reality and may be due to several confounding factors. From a methodological point of view, commonly applied normalization techniques which do not take into account the taxonomic profiles have been shown to underestimate functional variability [24]. In addition, aggregation of genes into broad **functional categories**, such as whole metabolic modules, based primarily on homology irrespective of the direction of metabolic flux, contributes to the impression of stability. Finally, large proportions of the functional genes in a metagenome are not known, and their potential variability is not taken into account at all (see also discussion below). Besides this, we [18] and others [16] have observed that the functional profiles in metatranscriptomes are more variable compared to metagenomic profiles, even when based on very broad functional categories [25] (Figure 1B,C).

The central questions that determine whether functional omics can reveal important functional elements of the microbiome are: (i) is the observed variability biologically meaningful?, and (ii) is a measured microbial functional state informative beyond a single snapshot? For the metagenome, individual-specific taxonomic profiles have been demonstrated [26,27]. It is noteworthy that, also for functional profiles, greater inter-individual than intra-individual variation is observable, at the metagenomic, metatranscriptomic, and metaproteomic levels (Figure 1D and Box 1) [18]. Differences in functional profiles provide direct pointers to the functions involved in microbiome–host interactions. Given the differences between metagenomic and metatranscriptomic profiles, the discriminatory power of metatranscriptomics needs to be assessed. Based on our own datasets [18] and estimation methods for sample sizes necessary to reach a targeted power [28,29], metatranscriptomic functional profiles are at least as (if not more) powerful in resolving differences as metagenomic profiles (Figure 1E,F). Therefore, the functional omes can provide insights into microbial activity and highlight significant microbiome-conferred traits. Our own observations also indicate that the functional omes reflect persistent individual-specific physiology, and that this signal is not dominated by nonspecific momentary fluctuations.

Another fundamental question with respect to microbiome research is: which gut compartment is reflected by meta-omics data obtained from faecal samples? Metagenomic data reflect a mixture of different locations along the gastrointestinal tract as well as spores [30]. By contrast, the relative abundance of housekeeping transcripts has been found to be related to the site of

Glossary

Experimental systems: systems that can be employed as a model for the human microbiome and which are amenable to manipulation; they range from mixed-species and automated culturing, cell-culture-based coculture systems, and, as all animals carry a microbiome, animal models, including germ-free or gnotobiotic animals, including those with a 'humanized' microbiome.
Functional categories: are commonly applied to describe the functions encoded or carried out by the microbiome (the functional complement), usually by grouping genes into metabolism-centered frameworks, such as the orthology put forth by the Kyoto Encyclopedia of Genes and Genomes (KEGG) and the MetaCyc database, or functionally annotated orthologous groups based on sequence similarity, such as eggNOG.
Functional diversity: is a measure of the number (richness) and distribution of different functions within the community. It is related to gene richness but also to phylogenetic diversity, as microbial communities with phylogenetically diverse members often have a wider functional potential. Phylogenetic diversity and functional diversity have been observed as traits of human gut microbiota which are relatively stable over time.
Functional plasticity: the ability of the microbial community or its members to adapt to perturbations by changing gene expression; it can stabilize the taxonomic community structure as well as ecosystem functions.
Functional redundancy: is a measure of the number of different populations within a community that are able to perform the same functions. Functional redundancy can increase functional resilience, in case perturbations affect the taxonomic community structure; this allows for a return to community function, and therefore can increase stability.
Intervention studies: seek to manipulate the human microbiome *in situ*, by means of nutrition, probiotics, antibiotics, or faecal transplants.
Metagenomics: refers to the analysis of genomic DNA for

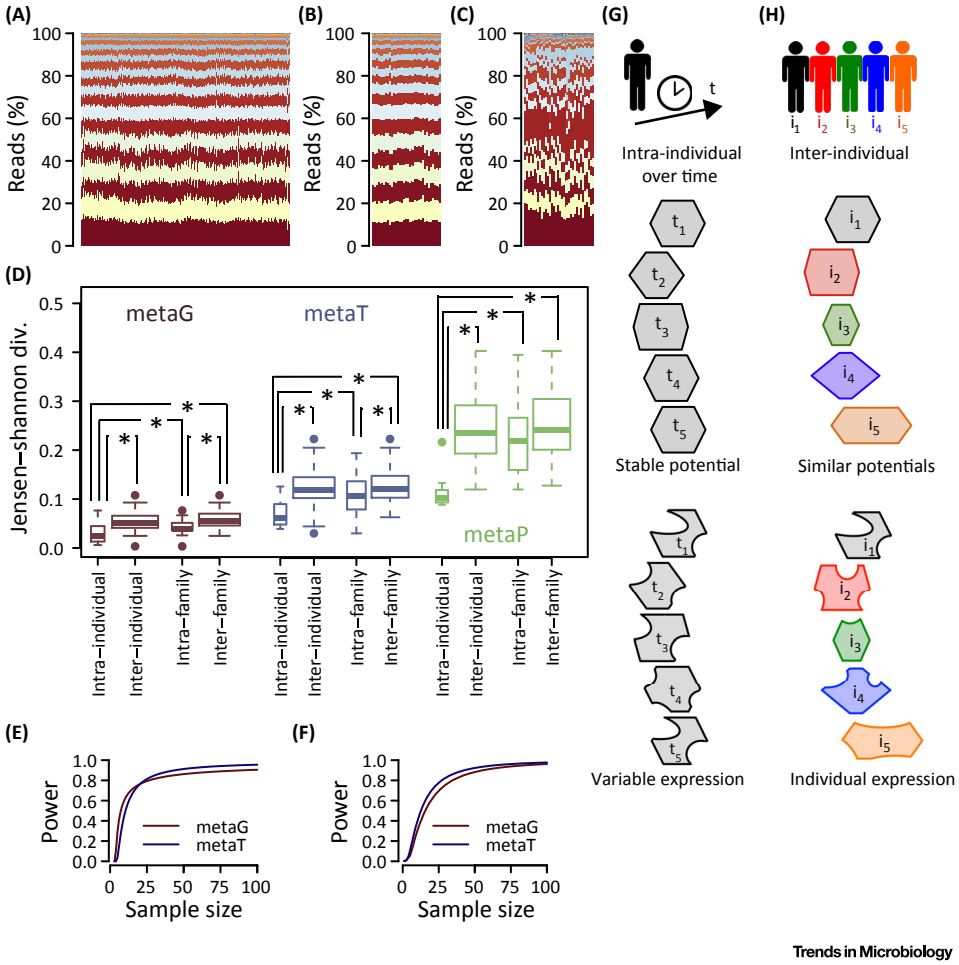

Figure 1. Community-Wide View of the Variability of Encoded and Expressed Functions in the Human Gut Microbiome. (A) High-level functional profiles of 1267 human gut microbiome metagenomes retrieved from the integrated gene catalogue (IGC) of the human gut microbiome [25]. (B) High-level functional profiles from metagenomes of our own smaller integrated multi-omics study [18] annotated using the IGC [25] in comparison to the (C) profiles from metatranscriptomes of the same samples. (D) Comparison of intra-individual to inter-individual and intra-family to inter-family distances (Jensen–Shannon divergence) based on functional metagenomic (MG), metatranscriptomic (MT), and metaproteomic (MP) profiles [18]; *$P < 0.05$, Wilcoxon rank sum test. (E,F) Estimation of power to distinguish functional profiles from members of different families based on metagenome and metatranscriptome measurements [18] applying limma/voom assumptions and the statistical model of Bi *et al.* [29] (E) and van Iterson *et al.* [28] (F). (G) Summarizing scheme, illustrating functional potentials with limited variability (middle) and functional expression profiles with greater plasticity (bottom) within an individual over time (t), compared to (H) the variability between different individuals' (i) microbiomes' functional potentials (middle), and functional expression profiles (bottom).

activity of specific microbial taxa, such that oral species have very low transcript levels in stool samples while colonic organisms are highly active [16,18]. Resolving gene expression to the taxon of origin, and relating this to the overall activity of that taxon, should further help in distinguishing *in situ* activity from noise in functional profiles. Discovery of compartment-specific functional features, which are important in the context of health and disease [31], may therefore be facilitated by metatranscriptomics (but see also Box 1 for a discussion of other omics technologies).

mixtures of (often unknown) species. Its purpose can be to assess the taxonomic composition of a mixed microbial community or to elucidate the functional potential of its members.

(Meta)metabolomics: (also referred to as *metabonomics*, mainly in the context of research on single organisms) technologies that measure intra- and/or extracellular metabolites in and around microbial communities.

Metaproteomics: aims to characterize microbial activity by applying the analysis of proteomes to mixed-species assemblages.

Metatranscriptomics: is the term applied to the analysis of RNA of communities, usually with the aim of inferring activity.

Omics: a group of methodologies that aim at the characterization of the total pool of a class of biomolecules, including metagenomics, (meta) metabolomics, metaproteomics, and metatranscriptomics.

Taxonomic and functional profiling: to quantify the taxa and functions detected in a sample form part of most meta-omic studies of the human microbiome. Increasingly, taxonomic resolution of functions of interest within the microbiome is also achieved.

Box 1. Which Functional Omes to Look at?

Metatranscriptomics, by highlighting changes in expression, reveals a more dynamic picture of the microbiome than metagenomics. The technology allows high sampling depths and high taxonomic resolution of functional processes. Although metatranscriptomic profiles confer essential information on functional gene expression within microbiomes, metaproteomic profiles may be a better indicator of the actual phenotype. However, metaproteomic analyses are not yet able to achieve the information content or sampling depth of sequencing-based technologies. Due to this fact, highly abundant, stably expressed conserved proteins make up the majority of data points, leading to a higher apparent stability of metaproteomic profiles [13]. However, the current limitations may be overcome through improvements in protein preparation [17], identification [120,121], and the adaptation of quantitative methods [122]. Metabolomics, by directly measuring metabolic outcomes, should be the most sensitive with respect to resolving the functional microbiome, and quantitative methods – such as proton nuclear magnetic resonance (NMR) analyses – are able to capture some of the most important, high-abundance microbial metabolites, such as short-chain fatty acids. Although advanced metabolomic methods allow the resolution of thousands of metabolite features from human microbiome samples, current limitations include a very large fraction of unknown metabolites (well in excess of 90% of measured features may be unknowns, even when searching metabolomic data against comprehensive databases [123]) as well as the difficulty in linking specific metabolite features to their microbial provenance. Advances in computational mass spectrometry, as well as in *de novo* metabolic network reconstruction and modelling, will allow some of these limitations to be addressed in the future. Given the different limitations of the single omic levels, as well as their complementary information content, the integration of multi-omic data can also help to close gaps when assessing gut microbial activity *in situ* by bridging genomic content to final phenotype.

The observation that metatranscriptomic functional profiles are more variable than might be inferred based solely on metagenomic information suggests that nonhousekeeping genes, even those with high genomic copy numbers, are not stably expressed *in situ* [10,11,16,18]. We have recently developed an approach which allows taxon-specific resolution of expressed genes [18]. When applying this method to link functional genes to the genomes which encode them, we observed that functions of interest may be contributed to the community-wide phenotype by single or multiple microbial populations in the absence of observable differences in the respective populations' abundances [18]. The identity of these populations may differ in different individuals, as the microbiota may have widely divergent taxonomic compositions [18]. The variability observed at the level of gene expression may very well be a reflection of **functional plasticity** and a prerequisite for stable community function. Consequently, resolving functional differences at multiple omic levels to the taxa contributing them is necessary in order to understand when and how these functions may impact human physiology.

The Unknowns

One challenge for microbiome research in relation to elucidating phenotypic impacts on the host is posed by unknown taxa and functions. While the overall proportion of protein-coding genes for which a molecular function cannot be predicted in the human microbiome (40–70%, depending on the prediction method [18,32,33]), is still generally high, this proportion is higher the rarer a microbial gene is in the human population (Figure 2A). Furthermore, this is especially the case when encoded in taxa which are not well described or even uncharacterized (Figure 2B). In many recent studies, genes without known functions, or those from uncultured taxa, have been completely ignored, because metagenomic data were analysed by mapping to annotated reference genomes. These approaches often make inefficient use of the data [34], are likely to introduce biases in the interpretation [35], and do not have a handle on the large proportion of horizontally transferred functions in the microbiome [36] as well as on strain-specific functional gene complements [37,38] which make up taxa-specific pangenomes. Horizontally transferred and strain-specific genes may be essential [39], in particular when they code for medically relevant functions such as antibiotic resistance [40] or toxins [41]. In this light, the prediction of functional potential [42,43] or even metabolic outcome [44] based on rough (i.e., genus-level) taxonomic profiles must be regarded as questionable.

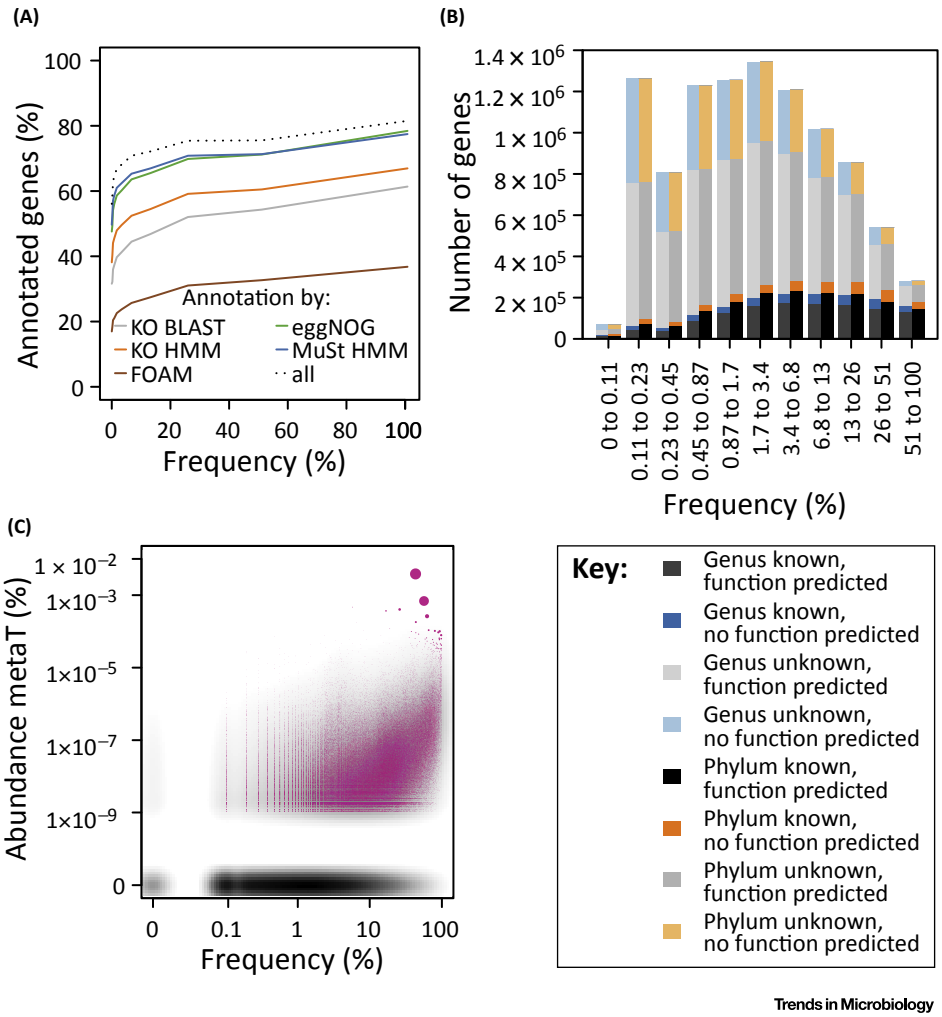

Figure 2. Genes of Unknown Function. (A) Relationship between the fraction of functionally annotated genes and the frequency of their occurrence according to the integrated gene catalogue (IGC) [25]. Annotations: 'KO BLAST': KEGG orthologous group (KO) annotations included in the IGC [25]; 'KO HMM': HMM-based annotations using KOs [18]; 'FOAM': HMM-based annotations using FOAM [32]; 'eggNOG': eggNOG-based [33] annotations included in the IGC [25]; 'MuSt HMM': HMM-based annotations using KOs, Pfam-A-families, TIGR-families, Swiss-Prot- or MetaCyc enzymes [18]; 'all': all annotations by either of the named methods. (B) Relationship between the number of annotated genes (by any of the methods displayed in (A), their relative frequency of occurrence, and the level of taxonomic assignment in the IGC [25]. (C) Frequency of occurrence [25] and maximum observed expression [18] of genes in the IGC. Pink dots highlight genes annotated with orthologous groups or protein domains of unknown function.

Ignoring functional unknowns also limits the potential that metagenomic and metatranscriptomic approaches possess in creating new knowledge. For example, approaches to compare abundances and genomes of uncultured taxa, which contribute approximately 40% of the metagenomic data, are well established [45,46]. Similarly, collections of orthologous groups and protein families without known functions have been established [32,47,48], allowing for cross-sample comparisons. These approaches facilitate the identification of biologically significant entities, for example, because they are found to be enriched or depleted in individuals with a disease or consistently highly abundant and/or expressed. For instance, in our recent multi-omics study, 9% of the differentially abundant transcripts (between families or between

individuals with type 1 diabetes and their healthy family members) were from genes encoding proteins with domains of unknown function. Likewise, in the integrated gene catalogue (IGC) of the human gut microbiome [25], 14% of the genes without predicted known functions can be associated with orthologous groups without known function and, importantly, we found 28% of these genes to be expressed in our own data [18] (Figure 2C).

Several experimental approaches to gain knowledge on 'the dark matter' of the human microbiome have been proposed, in addition to the proven combination of classical microbiological techniques with functional genomics. 'Functional metagenomics' involving the large-scale *in vitro* screening of metagenomic sequences has been developed [49–51], including use of microfluidics to assay millions of metagenomic variants of apparently similar genes [52]. 'Culturomics', the combination of miniaturized cultivation and advanced sequencing approaches, for example, to generate metagenomes from enrichment cultures, allows for the detailed characterization of organisms that are not culturable at a traditional laboratory scale [53,54]. The elucidation of unknowns that differ in health and disease, as well as the specific role they play in microbiome–host interactions, is an important challenge for the coming years.

Beyond Single Functions

Another crucial question regarding the contributions of microbiome-conferred functions to human physiology is related to the dynamics that govern community function and to the functioning of the microbial ecosystem as a whole [55]: does ecosystem functioning, in addition to or independent of specific microbial functions, play a role in human health, and are generalizable patterns discernable from multi-omics? At a fine scale, the gene content of the gastrointestinal microbiome is remarkably different between individuals. For example, the majority of the unique genes in the IGC are not found in more than a few percent of the samples (Figure 3A) [25]. These unique genes, however, carry common functions. In fact, functional annotations in the IGC are usually carried by many unique genes, with respect to both the whole catalogue and the subset present in single samples (Figure 3B). In our recent study [18], we also observed that most functions in microbial metabolism are encoded (Figure 3C) and expressed (Figure 3D) by a number of different microbial populations in any given sample. In addition, we have observed that the expression of genes by the same population can change over time, even when the population's relative abundance does not change. Finally, the relative transcript abundance of a gene function with respect to the whole community is independent of the number of different microbial populations that carry it (Figure 3D). These results imply that microbiome-conferred services need to be explored with respect to their structural and spatial dimensions in relation to their effect on host physiology.

The above observations are likely a reflection of **functional redundancy** within the healthy human microbiome. Functional redundancy can confer resilience [56] and therefore can stabilize ecosystem functionality during perturbations [57], which, in the context of the human microbiome, is generally assumed to lead to both stability and health [58]. However, the actual relationship between functional redundancy and stability has not been studied in the human gut microbiome, in contrast to other microbial ecosystems [59,60]. It is not even known whether there is true redundancy, as different genomic contexts may determine the impact of genes [61] and, within the gut microbiota, the interaction with the host [62,63]. The assumption that functional redundancy of the microbiome is related to human health is primarily based on an apparent relationship between taxonomic stability and the maintenance of taxonomic and functional diversity over time [18,64,65]. However, for the human gut microbiome, it is currently unclear whether diversity is a prerequisite for stability [66], which has been shown in other contexts [67–69]. Functional richness has also been suggested to positively impact human

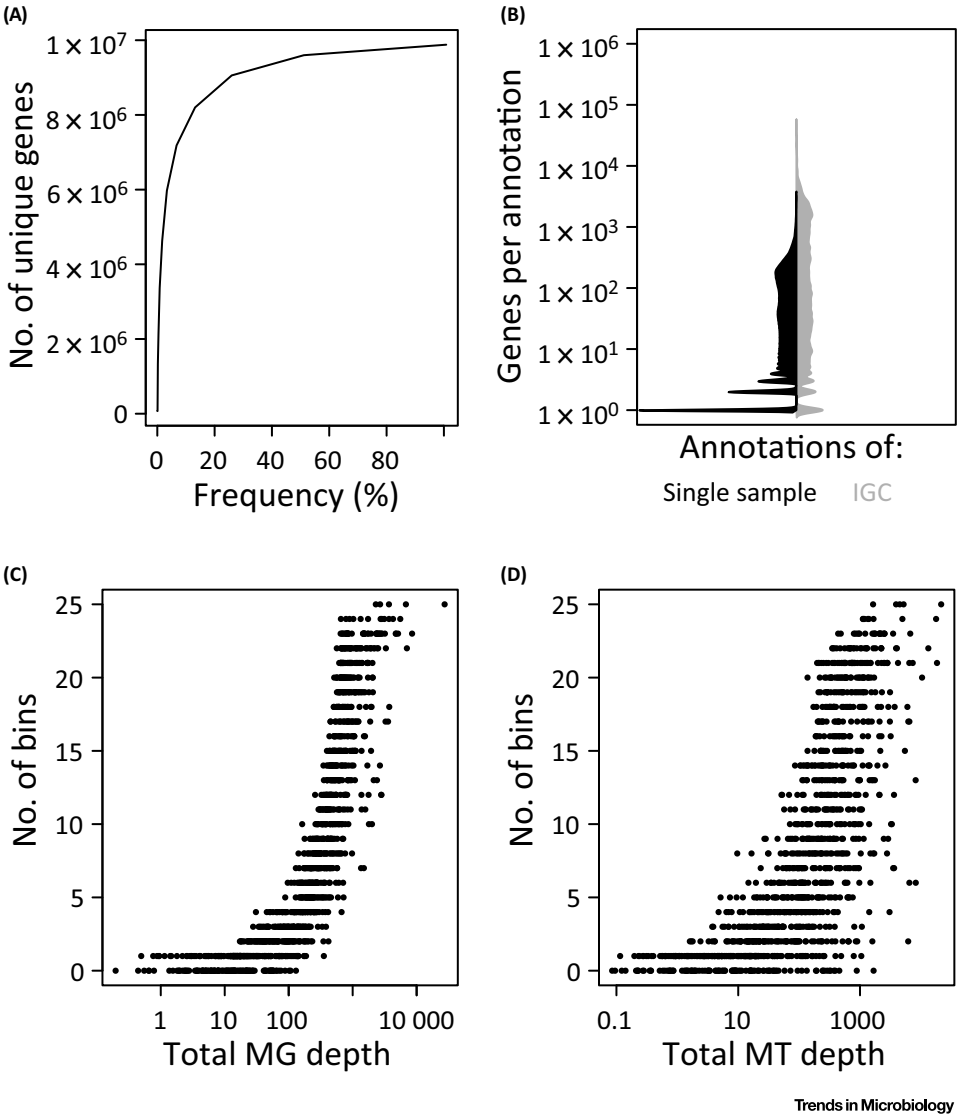

Figure 3. Functional Redundancy in the Human Microbiome (A) Relationship between the cumulative number of unique genes and the frequency of their occurrence. The graph is based on the 1267 human gut microbiome metagenomes retrieved from the integrated gene catalogue (IGC) [25]. (B) Numbers of genes with the same functional annotation, based on KEGG orthologous groups (KO) and eggNOG [33] orthologous groups, as published with the IGC [25]. (C) Relationship between the number of population-level genomes ('bins') containing genes annotated with a function in microbial metabolism and the corresponding cumulative metagenomic (MG) depth of coverage of the genes. (D) Relationship between the number of population-level genomes ('bins') expressing annotated genes and the corresponding cumulative metatranscriptomic (MT) depth of coverage. (C,D) Graphs are based on one representative sample from a healthy individual [18].

health [70], and decreased **functional diversity** has been observed in several diseases [22], although the observed functional richness may also be influenced by colonic transit time [71]. A higher metabolic diversity ensures digestibility of a wider range of nutrients [72] and potentially increases overall energy harvest. Metabolic diversity may also offer a protective potential against environmental toxic substances [3]. Despite the likely importance of functional diversity, the exact mechanism by which the human host benefits from redundant, diverse, and/or stable

communities has not been systematically studied [73–75]. In order to resolve any clear relationships, future analyses of the microbiome in the context of health and disease will have to assess functional redundancy. While most existing studies have examined single time points, the significance of stability of the microbiome will have to be addressed by time series studies during health and disease as well as experimental perturbations. These studies are also necessary to infer causal links and determine whether and how ecosystem functions of the gut microbiome can be shaped by interventions.

Concluding Remarks: A Map to Bring It All Together

Given the potentials and challenges highlighted above, future functional studies will have to integrate and compare reference-based alignments and *de novo* genome reconstructions, the wealth of existing omics datasets, functional knowledge, and orthology-based annotations to home in on the functions that really matter (Figure 4, Key Figure). The functional knowledge

Key Figure

Roadmap for Using Functional Omics to Create New Knowledge

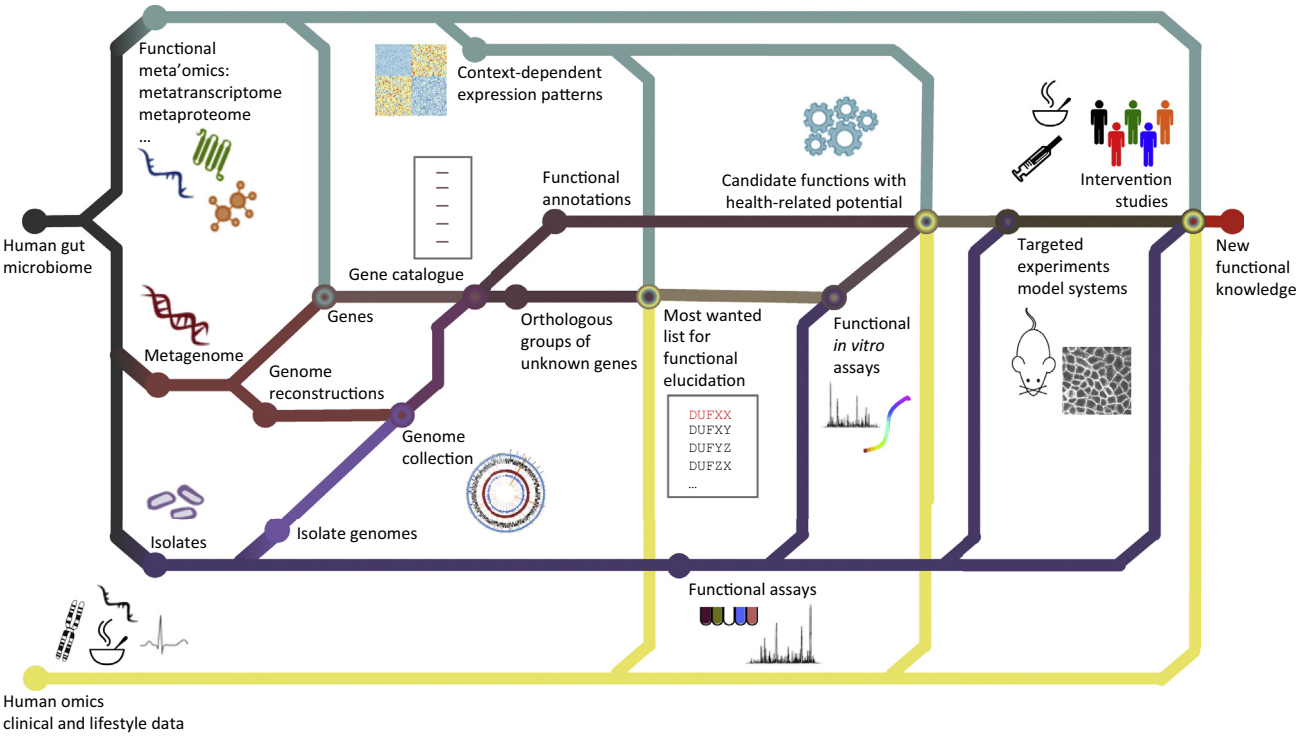

Figure 4. Crucial steps are the integration of reference genomes, metagenomic data collections, and *de novo* gene and genome reconstructions in genome and gene collections or catalogues. Genes should be linked to functions, taxonomic occurrence, and expression in different hosts. Genes without predicted functions can be grouped by orthology to enable comparative analyses and derive a list of 'most wanted' yet to be determined functions. Genes with functions that are likely to affect human health and/or display suggestive patterns of expression in different human hosts should be validated in targeted experiments in model systems and human intervention studies.

should also be systematically linked to the taxonomic structure [76–78] of the analysed samples at strain-level resolution [26,79–84] to explain microbiome-conferred phenotypic traits from a mechanistic point of view [85]. In this context, it will be essential to contrast and identify phenotypic traits which are widely distributed across constituent taxa, that is, those that are redundant, to those which are only encoded and expressed by specific taxa. In either case, identification of functional genes of interest should be performed first, followed by their linking to constituent taxa along the premise of 'form follows function' or 'function first, taxa second'. Analogous to the most-wanted taxa list [86], which was hunted down to a large extent within half a decade [87], a functional most-wanted list should therefore also be established by the community. Such a functional most-wanted list should explicitly take unknowns into account, based on information from omics experiments, such as when or where the genes and products are observed. This information, as well as potential interaction partners [88], should result in hypotheses for assays to elucidate molecular functions [89]. Another needed resource to understand our 'second genome' is an Online Mendelian Inheritance in Man-like framework that would list observed links between (functional) microbial genes and human phentoypes. Such a resource should draw on existing metagenomic or functional meta-omic data [90], in addition to functional reference genome databases [48,91]. Finally, this resource should already anticipate the omes and readouts which are poised to make an impact in the future, such as growth of different populations within the microbiome [92,93], regulatory elements [94,95], and sRNAs [11]. Additionally, several recent studies have demonstrated the power of integrating functional [18,96–102] and genetic data of the human host [103–105], which should likewise be linked to microbiome data in large-scale databases. This knowledge will be essential to understand the interaction between the microbiome and the human host.

A detailed representation and understanding of the functional microbiome is an essential prerequisite for future rational interventions leveraging the gut microbiome to alter host phenotype (see Outstanding Questions). To assess the impact of specific microbial functions on human physiology, and explain their mechanism of action, experiments in representative models will be critical. To model cellular interactions and reach high throughputs, miniaturized *in vitro* models of the human gut interface [106,107] should therefore be employed. Animal models, in spite of notable limitations [108–110], can be used to observe systemic impacts. These studies have yielded remarkable insights through detailed analysis [111–113]. Finally, once useful and safe candidates for improving human health have been established, intervention trials in human cohorts, through diet [114,115] and/or faecal transplants [116,117], faecal components [118], such as small molecules, or faeces-derived selected microbiota [119], could be performed. End-points of these studies should involve the monitoring of health-related physiological markers as well as follow, in detail, the induced changes in the microbiome over time using omics measurements to understand the role of the microbiome-borne functional complement in governing human health and disease.

Outstanding Questions

Which microbial functions impact human physiology, and in which context? Are there specific drivers of, or protective functions against, diseases?

How fast, how strong, and how resilient does the functional microbiome react to perturbations?

What do we not know – what do unknown genes contribute to microbiome functioning and host physiology?

Functional diversity, redundancy, and stability – what is the link to human health? Why is a diverse or a stable microbiome supposed to be beneficial for human health?

(How) can community function and ecology be manipulated?

Acknowledgments

All analyses were performed on the High Performance Computing platform of the University of Luxembourg. This work was supported by Luxembourg National Research Fund (FNR) CORE programme grants (CORE/15/BM/10404093 and CORE/16/BM/11276306) to P.W.

References

1. Karasov, W.H. *et al.* (2011) Ecological physiology of diet and digestive systems. *Annu. Rev. Physiol.* 73, 69–93
2. LeBlanc, J.G. *et al.* (2013) Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Curr. Opin. Biotechnol.* 24, 160–168
3. Claus, S.P. *et al.* (2016) The gut microbiota: a major player in the toxicity of environmental pollutants? *NPJ Biofilms Microbiomes* 2, 16003
4. Kamada, N. *et al.* (2013) Role of the gut microbiota in immunity and inflammatory disease. *Nat. Rev. Immunol.* 13, 321–335

5. Abreu, M.T. (2010) Toll-like receptor signalling in the intestinal epithelium: how bacterial recognition shapes intestinal function. *Nat. Rev. Immunol.* 10, 131–144
6. Sommer, F. and Bäckhed, F. (2013) The gut microbiota – masters of host development and physiology. *Nat. Rev. Microbiol.* 11, 227–238
7. Hooper, L.V. *et al.* (2012) Interactions between the microbiota and the immune system. *Science* 336, 1268–1273
8. Gilbert, J.A. *et al.* (2016) Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature* 535, 94–103
9. Verberkmoes, N.C. *et al.* (2008) Shotgun metaproteomics of the human distal gut microbiota. *ISME J.* 3, 179–189
10. Turnbaugh, P.J. *et al.* (2010) Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. *Proc. Natl. Acad. Sci. U. S. A.* 107, 7503–7508
11. Gosalbes, M.J. *et al.* (2011) Metatranscriptomic approach to analyze the functional human gut microbiota. *PLoS One* 6, e17447
12. Erickson, A.R. *et al.* (2012) Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PLoS One* 7, e49138
13. Ferrer, M. *et al.* (2012) Microbiota from the distal guts of lean and obese adolescents exhibit partial functional redundancy besides clear differences in community structure. *Environ. Microbiol.* 15, 211–226
14. Maurice, C.F. *et al.* (2013) Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell* 152, 39–50
15. Pérez-Cobas, A.E. *et al.* (2013) Gut microbiota disturbance during antibiotic therapy: a multi-omic approach. *Gut* 62, 1591–1601
16. Franzosa, E.A. *et al.* (2014) Relating the metatranscriptome and metagenome of the human gut. *Proc. Natl. Acad. Sci. U. S. A.* 111, E2329–E2338
17. Xiong, W. *et al.* (2015) Development of an enhanced metaproteomic approach for deepening the microbiome characterization of the human infant gut. *J. Proteome Res.* 14, 133–141
18. Heintz-Buschart, A. *et al.* (2016) Integrated multi-omics of the human gut microbiome in a case study of familial type 1 diabetes. *Nat. Microbiol.* 2, 16180
19. McNulty, N.P. *et al.* (2011) The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. *Sci. Transl. Med.* 3, 106ra106
20. Proctor, L.M. (2016) The National Institutes of Health Human Microbiome Project, Seminars in Fetal & Neonatal Medicine. *Semin. Fetal Neonatal. Med.* 21, 368–372
21. Börnigen, D. *et al.* (2013) Functional profiling of the gut microbiome in disease-associated inflammation. *Genome Med.* 5, 65
22. Cho, I. and Blaser, M.J. (2012) The human microbiome: at the interface of health and disease. *Nat. Rev. Genet.* 13, 260–270
23. Huttenhower, C. *et al.* (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214
24. Manor, O. and Borenstein, E. (2017) Revised computational metagenomic processing uncovers hidden and biologically meaningful functional variation in the human microbiome. *Microbiome* 5, 19
25. Li, J. *et al.* (2014) An integrated catalog of reference genes in the human gut microbiome. *Nat. Biotechnol.* 32, 834–841
26. Schloissnig, S. *et al.* (2013) Genomic variation landscape of the human gut microbiome. *Nature* 493, 45–50
27. Franzosa, E.A. *et al.* (2015) Identifying personal microbiomes using metagenomic codes. *Proc. Natl. Acad. Sci. U. S. A.* 112, E2930–E2938
28. van Iterson, M. *et al.* (2013) General power and sample size calculations for high-dimensional genomic data. *Stat. Appl. Genet. Mol. Biol.* 12, 449–467
29. Bi, R. and Liu, P. (2016) Sample size calculation while controlling false discovery rate for differential expression analysis with RNA-sequencing experiments. *BMC Bioinformatics* 17, 146
30. Browne, H.P. *et al.* (2016) Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* 533, 543–546
31. Zhang, S. *et al.* (2017) Sampling strategies for three-dimensional spatial community structures in IBD microbiota research. *Front. Cell. Infect. Microbiol.* 7, 51
32. Prestat, E. *et al.* (2014) FOAM (Functional Ontology Assignments for Metagenomes): a Hidden Markov Model (HMM) database with environmental focus. *Nucleic Acids Res.* 42, e145–e145
33. Powell, S. *et al.* (2012) eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Res.* 40, D284–D289
34. Narayanasamy, S. *et al.* (2016) IMP: a pipeline for reproducible reference-independent integrated metagenomic and metatranscriptomic analyses. *Genome Biol.* 17, 260
35. Timmons, J.A. *et al.* (2015) Multiple sources of bias confound functional enrichment analysis of global -omics data. *Genome Biol.* 16, 186
36. Brito, I.L. *et al.* (2016) Mobile genes in the human microbiome are structured from global to individual scales. *Nature* 535, 435–439
37. Zhang, C. and Zhao, L. (2016) Strain-level dissection of the contribution of the gut microbiome to human metabolic disease. *Genome Med.* 8, 41
38. Antony-Babu, S. *et al.* (2017) Multiple *Streptomyces* species with distinct secondary metabolomes have identical 16S rRNA gene sequences. *Sci. Rep.* 7, 11089
39. McClean, D. *et al.* (2015) Single gene locus changes perturb complex microbial communities as much as apex predator loss. *Nat. Commun.* 6, 8235
40. Sommer, M.O.A. *et al.* (2014) The human microbiome harbors a diverse reservoir of antibiotic resistance genes. *Virulence* 1, 299–303
41. Dutilh, B.E. *et al.* (2013) Screening metatranscriptomes for toxin genes as functional drivers of human colorectal cancer. *Best Pract. Res. Clin. Gastroenterol.* 27, 85–99
42. Langille, M.G.I. *et al.* (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* 31, 814–821
43. Jun, S.-R. *et al.* (2015) PanFP: pangenome-based functional profiles for microbial communities. *BMC Res. Notes* 8, 479
44. Magnúsdóttir, S. *et al.* (2017) Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* 35, 81–89
45. Sunagawa, S. *et al.* (2013) Metagenomic species profiling using universal phylogenetic marker genes. *Nat. Methods* 10, 1196–1199
46. Nielsen, H.B. *et al.* (2014) Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* 32, 822–828
47. Huerta-Cepas, J. *et al.* (2016) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293
48. Mende, D.R. *et al.* (2017) proGenomes: a resource for consistent functional and taxonomic annotations of prokaryotic genomes. *Nucleic Acids Res.* 45, D529–D534
49. Lakhdari, O. *et al.* (2010) Functional metagenomics: a high throughput screening method to decipher microbiota-driven NF-κB modulation in the human gut. *PLoS One* 5, e13092
50. Dobrijevic, D. *et al.* (2013) High-throughput system for the presentation of secreted and surface-exposed proteins from Gram-positive bacteria in functional metagenomics studies. *PLoS One* 8, e65956
51. Lam, K.N. *et al.* (2015) Current and future resources for functional metagenomics. *Front. Microbiol.* 6, 149

52. Colin, P.-Y. *et al.* (2015) Ultrahigh-throughput discovery of promiscuous enzymes by picodroplet functional metagenomics. *Nat. Commun.* 6, 10008
53. Lagier, J.-C. *et al.* (2016) Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol.* 1, 16203
54. Berdy, B. *et al.* (2017) *In situ* cultivation of previously uncultivable microorganisms using the ichip. *Nat. Protoc.* 12, 2232–2242
55. Cadotte, M.W. *et al.* (2011) Beyond species: functional diversity and the maintenance of ecological processes and services. *J. Appl. Ecol.* 48, 1079–1087
56. Walker, B. (1995) Conserving biological diversity through ecosystem resilience. *Conserv. Biol.* 9, 747–752
57. Naeem, S. (1998) Species redundancy and ecosystem reliability. *Conserv. Biol.* 12, 39–45
58. Lozupone, C.A. *et al.* (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* 489, 220–230
59. Allison, S.D. and Martiny, J.B.H. (2008) Colloquium paper: resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* 105, 11512–11519
60. Delgado-Baquerizo, M. *et al.* (2016) Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat. Commun.* 7, 10541
61. Blount, Z.D. *et al.* (2012) Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* 488, 513–518
62. Louwen, R. *et al.* (2012) A novel link between *Campylobacter jejuni* bacteriophage defence, virulence and Guillain–Barré syndrome. *Eur. J. Clin. Microbiol. Infect. Dis.* 32, 207–226
63. Sahl, J.W. *et al.* (2015) Genomic characterization of *Burkholderia pseudomallei* isolates selected for medical countermeasures testing: comparative genomics associated with differential virulence. *PLoS One* 10, e0121052
64. Flores, G.E. *et al.* (2014) Temporal variability is a personalized feature of the human microbiome. *Genome Biol.* 15, 531
65. de Meij, T.G.J. *et al.* (2016) Composition and stability of intestinal microbiota of healthy children within a Dutch population. *FASEB J.* 30, 1512–1522
66. Coyte, K.Z. *et al.* (2015) The ecology of the microbiome: Networks, competition, and stability. *Science* 350, 663–666
67. McCann, K.S. (2000) The diversity–stability debate. *Nature* 405, 228–233
68. Erkus, O. *et al.* (2013) Multifactorial diversity sustains microbial community stability. 7, 2126–2136
69. Carrara, F. *et al.* (2015) Experimental evidence for strong stabilizing forces at high functional diversity of aquatic microbial communities. *Ecology* 96, 1340–1350
70. Le Chatelier, E. *et al.* (2013) Richness of human gut microbiome correlates with metabolic markers. *Nature* 500, 541–546
71. Roager, H.M. *et al.* (2016) Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat. Microbiol.* 1, 16093
72. Tap, J. *et al.* (2015) Gut microbiota richness promotes its stability upon increased dietary fibre intake in healthy adults. *Environ. Microbiol.* 17, 4954–4964
73. Matsen, F.A. (2014) Phylogenetics and the Human Microbiome. *Syst. Biol.* 64, e26–e41
74. Moya, A. and Ferrer, M. (2016) Functional redundancy-induced stability of gut microbiota subjected to disturbance. *Trends Microbiol.* 24, 402–413
75. Sommer, F. *et al.* (2017) The resilience of the intestinal microbiota influences health and disease. *Nat. Rev. Microbiol.* 15, 630–638
76. Jiang, Y. *et al.* (2016) Metatranscriptomic analysis of diverse microbial communities reveals core metabolic pathways and microbiome-specific functionality. *Microbiome* 4, 2
77. Plichta, D.R. *et al.* (2016) Transcriptional interactions suggest niche segregation among microorganisms in the human gut. *Nat. Microbiol.* 1, 16152
78. Manor, O. and Borenstein, E. (2017) Systematic characterization and analysis of the taxonomic drivers of functional shifts in the human microbiome. *Cell Host Microbe* 21, 254–267
79. Luo, C. *et al.* (2015) Constrains identifies microbial strains in metagenomic datasets. *Nat. Biotechnol.* 33, 1045–1052
80. Nayfach, S. *et al.* (2016) An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Res.* 26, 1612–1625
81. Scholz, M. *et al.* (2016) Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nat. Methods* 13, 435–438
82. Truong, D.T. *et al.* (2017) Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res.* 27, 626–638
83. Costea, P.I. *et al.* (2017) metaSNV: A tool for metagenomic strain level analysis. *PLoS One* 12, e0182392
84. Quince, C. *et al.* (2017) DESMAN: a new tool for de novo extraction of strains from metagenomes. *Genome Biol.* 18, 181
85. Olesen, S.W. and Alm, E.J. (2016) Dysbiosis is not an answer. *Nat. Microbiol.* 1, 16228
86. Fodor, A.A. *et al.* (2012) The 'most wanted' taxa from the human microbiome for whole genome sequencing. *PLoS One* 7, e41294
87. Almeida, M. *et al.* (2016) Capturing the most wanted taxa through cross-sample correlations. *ISME J.* 10, 2459–2467
88. Mayers, M.D. *et al.* (2017) Quantitative metaproteomics and activity-based probe enrichment reveals significant alterations in protein expression from a mouse model of inflammatory bowel disease. *J. Proteome Res.* 16, 1014–1026
89. Prosser, G.A. *et al.* (2014) Metabolomic strategies for the identification of new enzyme functions and metabolic pathways. *EMBO Rep.* 15, 657–669
90. Nayfach, S. *et al.* (2015) MetaQuery: a web server for rapid annotation and quantitative analysis of specific genes in the human gut microbiome. *Bioinformatics* 31, 3368–3370
91. Caspi, R. *et al.* (2013) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* 42, D459–D471
92. Korem, T. *et al.* (2015) Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. *Science* 349, 1101–1106
93. Brown, C.T. *et al.* (2016) Measurement of bacterial replication rates in microbial communities. *Nat. Biotechnol.* 34, 1256–1263
94. Nawrocki, E.P. and Eddy, S.R. (2014) Computational identification of functional RNA homologs in metagenomic data. *RNA Biol.* 10, 1170–1179
95. Bao, G. *et al.* (2015) Strand-specific community RNA-seq reveals prevalent and dynamic antisense transcription in human gut microbiota. *Front. Microbiol.* 6, 94
96. Kato, T. *et al.* (2014) Multiple omics uncovers host-gut microbial mutualism during prebiotic fructooligosaccharide supplementation. *DNA Res.* 21, 469–480
97. Palm, N.W. *et al.* (2014) Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* 158, 1000–1010
98. Kolmeder, C.A. *et al.* (2015) Colonic metaproteomic signatures of active bacteria and the host in obesity. *Proteomics* 15, 3544–3552
99. Debyser, G. *et al.* (2016) Faecal proteomics: A tool to investigate dysbiosis and inflammation in patients with cystic fibrosis. *J. Cyst. Fibros.* 15, 242–250
100. Liu, S. *et al.* (2016) The host shapes the gut microbiota via fecal microRNA. *CHOM* 19, 32–43
101. Pedersen, H.K. *et al.* (2016) Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* 535, 376–381
102. Okai, S. *et al.* (2016) High-affinity monoclonal IgA regulates gut microbiota and prevents colitis in mice. *Nat. Microbiol.* 1, 16103

103. Turpin, W. *et al.* (2016) Association of host genome with intestinal microbial composition in a large healthy cohort. *Nat. Genet.* 48, 1413–1417

104. Bonder, M.J. *et al.* (2016) The effect of host genetics on the gut microbiome. *Nat. Genet.* 48, 1407–1412

105. Wang, J. *et al.* (2016) Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. *Nat. Genet.* 48, 1396–1406

106. Kim, H.-J. *et al.* (2016) Contributions of microbiome and mechanical deformation to intestinal bacterial overgrowth and inflammation in a human gut-on-a-chip. *Proc. Natl. Acad. Sci. U. S. A.* 113, E7–E15

107. Shah, P. *et al.* (2016) A microfluidics-based in vitro model of the gastrointestinal human-microbe interface. *Nat. Commun.* 7, 11535

108. Fritz, J.V. *et al.* (2013) From meta-omics to causality: experimental models for human microbiome research. *Microbiome* 1, 14

109. Nguyen, T.L.A. *et al.* (2015) How informative is the mouse for human gut microbiota research? *Dis. Models Mechan.* 8, 1–16

110. Arrieta, M.-C. *et al.* (2016) Human microbiota-associated mice: a model with challenges. *Cell Host Microbe* 19, 575–578

111. Furusawa, Y. *et al.* (2013) Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* 504, 446–450

112. Atarashi, K. *et al.* (2013) Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* 500, 232–236

113. Turroni, F. *et al.* (2016) Deciphering bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics approach. *ISME J.* 10, 1656–1668

114. David, L.A. *et al.* (2013) Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563

115. O'Keefe, S.J.D. *et al.* (2015) Fat, fibre and cancer risk in African Americans and rural Africans. *Nat. Commun.* 6, 6342

116. van Nood, E. *et al.* (2013) Duodenal infusion of donor feces for recurrent *Clostridium difficile*. *N. Engl. J. Med.* 368, 407–415

117. Li, S.S. *et al.* (2016) Durable coexistence of donor and recipient strains after fecal microbiota transplantation. *Science* 352, 586–589

118. Ott, S.J. *et al.* (2017) Efficacy of sterile fecal filtrate transfer for treating patients with *Clostridium difficile* infection. *Gastroenterology* 152, 799–811.e7

119. Petrof, E.O. *et al.* (2013) Stool substitute transplant therapy for the eradication of *Clostridium difficile* infection: 'RePOOPulating' the gut. *Microbiome* 1, 3

120. Tanca, A. *et al.* (2016) The impact of sequence database choice on metaproteomic results in gut microbiota studies. *Microbiome* 4, 51

121. Tang, H. *et al.* (2016) A graph-centric approach for metagenome-guided peptide and protein identification in metaproteomics. *PLoS Comput. Biol.* 12, e1005224

122. Gillet, L.C. *et al.* (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* 11, O111.016717

123. Roume, H. *et al.* (2013) A biomolecular isolation framework for eco-systems biology. *ISME J.* 7, 110–121