Review

Human Microbiome Acquisition and Bioinformatic Challenges in Metagenomic Studies

Valeria D'Argenio [1,2,3]

1   CEINGE-Biotecnologie Avanzate, via G. Salvatore 486, 80145 Naples, Italy; dargenio@ceinge.unina.it;
     Tel.: +39-081-373-7909
2   Department of Molecular Medicine and Medical Biotechnologies, University of Naples Federico II,
     via Pansini 5, 80131 Naples, Italy
3   Task Force on Microbiome Studies, University of Naples Federico II, 80131 Naples, Italy

Abstract: The study of the human microbiome has become a very popular topic. Our microbial counterpart, in fact, appears to play an important role in human physiology and health maintenance. Accordingly, microbiome alterations have been reported in an increasing number of human diseases. Despite the huge amount of data produced to date, less is known on how a microbial dysbiosis effectively contributes to a specific pathology. To fill in this gap, other approaches for microbiome study, more comprehensive than 16S rRNA gene sequencing, i.e., shotgun metagenomics and metatranscriptomics, are becoming more widely used. Methods standardization and the development of specific pipelines for data analysis are required to contribute to and increase our understanding of the human microbiome relationship with health and disease status.

Keywords: human microbiome; 16S rRNA analysis; metagenomics; metrascriptomics; data analysis; bioinformatics

---

1. Introduction

The advent of Next-Generation Sequencing (NGS) technologies has given rise to metagenomic research during the last ten years. Indeed, NGS-based methods have provided the capability of simultaneously analyzing communities of microorganisms without the need of an a priori knowledge of their composition, and overcoming cultivation steps [1,2].

Based on these technological advances, a plethora of studies have been carried out to characterize the microbial communities of specific environments, including human beings [3–9]. In particular, the mutually beneficial relationship between the microbiome and the human host has been largely investigated [2,10–14]. Just in the last ten years, more than 35,000 scientific reports have been published on metagenomic aspects, about 30,000 of which focus on the human microbiome, and it is easy to predict that this number will progressively increase in a short time.

As already defined elsewhere [2], the human microbiota is the community of microorganisms that lives on the surface and inside the human host (mucosal surfaces, skin, tooth surface, respiratory airways, gut, urogenital system, etc.). In this view, the so-called human holobiont, or superorganism, is the ensemble of human and microbial cells present in a specific ecosystem, whose heterogeneity depends on the body site and on the host's genetic background [15]. Even if the initial statement assessing that bacterial cells overcome human cells by a ratio of 10:1 has been recently revised [16], tuning it down to a 1:1 ratio, it means that we are made at least for a half of bacteria. Based on these data, it is clear that the human microbiota is associated with physiological functions required for the human host healthy status acquisition and maintenance. Thus, if these functions are impaired, because of a microbial dysbiosis, this could lead to pathological phenomena. According to this statement,

microbial alterations have been investigated and identified in a number of diseases [17–23]. Quantitative and/or qualitative microbial alterations, in fact, not only could play a role in diseases onset, thus increasing our knowledge on diseases pathogenetic mechanisms, but could represent also novel, sometimes not invasive (depending on the samples in which they have to be investigated), diagnostic biomarkers or targets for novel therapeutic approaches aiming at dysbiosis correction. However, despite the initial enthusiasm and the interesting findings reported to date, most studies remain at an observational level, while less is known about the complex interplays among the human host and his microbiome, both in health and, consequently, in specific diseases.

The aim of this paper is to underscore the relevance of the gut microbiome for human physiology and health maintenance. In addition, since the increasing capability of data production imposes new challenges for their accurate analysis, the main bioinformatic tools for metagenomics will be examined.

2. Human Microbiome Acquisition and Healthy Status Maintenance

Humans' relationships with their microbiome begin already before birth. Up to a few years ago, we thought that humans were born sterile and that microbial colonization began immediately at birth, being influenced by the kind of delivery [2,24]. In fact, a central dogma of reproductive medicine assesses that the fetus stays in a sterile environment and that the newborn acquires its own microbiota only after birth [25].

Increasing evidence suggests that bacterial colonization may be initiated in utero [24,25]. Collado et al. studied the microbiome present in the maternal feces, placenta, amniotic fluid, colostrum, meconium, and infant feces collected from women undergoing elective caesarean delivery and their babies [26]. Interestingly, not only they found bacterial communities in all the analyzed samples, but also assessed that the placenta and the amniotic fluid share common features, being characterized by a lower diversity and high levels of Proteobacteria compared to the other maternal and fetal samples. In addition, after 3–4 days of life, the infant fecal microbiome appears to become more similar to that of colostrum [26]. Thus, the progressive human gut microbial colonization probably starts already during prenatal life by a specific microbiota resident in the placenta and amniotic fluid and is then continued at delivery and after birth by the feeding habits. The origins of the placental and meconium microbiota are not well known. Different potential sites (including maternal vagina, mouth, and gut) have been suggested, even if the most probable hypothesis is that different bacteria may have different origins [27,28]. Consequently, the maternal microbiota during pregnancy might influence fetal development. Therefore, a maternal dysbiosis occurring during pregnancy may be silent in the mother, but may have consequences on the fetus and the placenta. A maternal dysbiosis, in fact, could trigger inflammatory reactions and may induce an anomalous immune system activation potentially affecting fetal development or inducing preterm delivery [27–29]. These findings suggest that the prenatal microbial "invasion" is not only crucial to modulate fetal development but may also influence the adult healthy status [24–29].

After birth, infants are suddenly exposed to the external environment (including surfaces, air, food) and, being a "naive" environment, they are like an attractive pole for bacteria: this process appears to be essential for immune system development [30]. Despite this invasion, it has been noticed that a strong selection occurs according to the body sites to maintain tissue-specific composition [24,25,30–32]. Several environmental factors can modify infants' microbiome composition, mainly depending on maternal habits [2,24,25,30–32]. In particular, the effects of the delivery mode on microbiome composition have been investigated [33,34]. It has been described that children born by cesarean section, especially by elective cesarean section, have a reduced microbial richness and a reduced quantity of health-associated bacterial species (i.e., *Lactobacilli*) [33–35]. Dominguez-Bello et al. analyzed the microbial composition in the feces of babies born by vaginal and cesarean delivery and compared these samples with the microbial communities present in the maternal mouth, skin, and vagina [34]. First of all, they confirmed the presence of a different microbiota in infant fecal samples depending on the delivery mode. Next, a diversity analysis revealed a tight clusterization

between maternal skin and the gut microbiomes of cesarean-delivered babies, as well as between maternal vagina and the gut microbiomes of vaginal-delivered babies [34]. In disagreement with previous studies, a recent work from Chu et al. reported that even if the infant microbial composition was influenced by the delivery mode at birth, these differences were not present at 6 weeks of age. Interestingly, they found that, unlike the other neonatal body sites, the meconium microbiota at birth is not influenced by the mode of delivery and speculated that the microbial transmission to the fetus may begin at a very early gestational age [35].

Infants' feeding mode is another main determinant of the microbiome composition [30]. While it is expected that differences occur between breast- and formula-fed babies, we have also to consider that the human milk is not sterile, *Staphylococcus*, *Streptococcus*, and *Pseudomonas* being its most represented bacteria [30]. In addition, the microbial composition of the maternal milk itself is influenced by several factors, such as diet, delivery mode, time post-partum, maternal age, metabolic status, hormone status, and previous medical interventions [30]. Despite these differences, Hunt et al. revealed that a core human milk microbiome is present, and that inter-women variability is higher than intra-woman variability [36].

Considering these factors all together, the so-called first 1000 days of life, from conception until two years of age, are critical for the development of the fetus and infant and can modify the risk profile for diseases onset later in life. After the age of two–three years, the gut microbiome becomes more stable and similar to that of adulthood [2,24,25,30–32].

The adult gut microbiome is featured by increased microbial richness and complexity as a consequence of the gradual development of the gut [37]. It is commonly stated that the gut microbial composition is almost similar across adult healthy individuals, because of the development of a core community of microbes that ensure a common, microbial-dependent, functional capability and protect from exogenous insults [37]. Interestingly, this shared microbial composition is present at phylum level, but not at lower phylogenetic levels: a great variability (both qualitative and quantitative) has been described at species and strain levels, suggesting that the gut microbial composition is unique in each individual, even if the microbial functions as a whole result the same [15]. Despite its stability, a number of factors are known to be able to affect the adult gut microbiome, including host genetics, diet, climate, travel, hormonal status, physical activity, and lifestyle [2,15,37]. To solve this discrepancy, it has been recently supposed that the human adult gut microbiome may have two components: one, the core microbiome, almost stable and hardly influenced by external stimuli, and the other, more dynamic and able to change quickly, responsible for microbial plasticity in different conditions and, as a consequence, susceptible to environmental or lifestyle changes [37].

Understanding the role of the gut microbiome is crucial to elucidate human host physiology, since there is growing evidence that it is involved in healthy status acquisition and maintenance [2,15,37,38]. Indeed, even if microbes have been described everywhere in humans, the majority of our microbiota lives in the gut where it is involved in multiple functions, including (but not limited to) the extraction of nutrients from indigestible components of food, the synthesis of essential vitamins and amino acids, the detoxification of potentially harmful molecules, and the defense against pathogens [2,15]. In particular, the gut microbiome has been linked to human physiology through the metabolism of bile acids, choline, and short-chain fatty acids (thus influencing the immune system development) [39]. In addition, it plays a role in gut mucosal permeability, immune system development and self-recognition, energy expenditure, hormonal status, and gut–brain signaling [40]. As a consequence, it is easy to speculate that microbial alterations, or dysbiosis, able to impair these essential functions, may alter human physiology and affect human health status. Accordingly, alterations of the gut microbiome have been related to a number of diseases, both gastrointestinal and non-gastrointestinal, so much so that a provocative report has recently proposed that gut microbiota may be at the intersection of everything [41]. In the attempt to try to generalize knowledge regarding the associations between human gut microbiome and diseases, Duvallet et al. recently carried out a meta-analysis of 28 published gut microbiome studies covering 10 different diseases [42]. This study highlights

that, while some diseases are featured by the presence of potentially pathogenic microbes, others are related to the reduction of helpful microbes. Interestingly, these associations seem to be a common pathogenetic mechanism, rather than disease-specific [42].

The challenge of future research in this field will be to assess whether these diseases-related alterations are bystanders, consequences of pathological conditions, or may play a role as triggers or risk factors for specific diseases. In this context, the "common ground hypothesis" has been also conceived to try to establish if the observed gut microbiota dysbiosis is a cause or a consequence of complex, multifactorial diseases. This hypothesis is based on the observation that several factors (endogenous, exogenous and, more often, a combination of them) may affect the gut mucosa inducing permeability alterations and inflammation, also by promoting gut microbiota alterations. In individuals genetically predisposed to a specific disease, such alterations may drive the colonization of the gut by opportunistic and pathobiont microbes that are responsible for local and systemic pathogenic functional modifications contributing to the disease [15].

Finally, it has to be noticed that, even if the term microbiota refers to all the microbes present in an environment, our knowledge on the human microbiota is largely limited to its bacterial fraction, while less is known about other microbial kingdoms (i.e., fungi, archaea, and viruses), their contribution to human health and disease being, actually, largely underestimated. Recent studies, focused on the characterization of fungal communities, have shown that also fungi influence human host physiology and can be modified by different stimuli, including illness [43]. In particular, the human mycobiota seems to play a role in human immunity development and host responses at mucosal level [43]. Similarly, *Archaea* have been indicated as key elements in gut metabolic functions, and modifications in their composition have been related to some diseases [44]. Finally, even viruses could be our symbiotic inhabitants and contribute to our health and disease status [45–47]. Therefore, our comprehension of human physiology (and, consequently, of human pathology) must take into account also these aspects and how these microbial groups interact with one another and with the human host. More efforts and future studies are required to address these issues.

3. Data Generation, Bioinformatic Issues, and Challenges in Metagenomic Studies

The rapid development of NGS technologies has prompted the study of human microbiome at DNA (16S rRNA and Internal Transcribed Spacers—ITS—sequencing, and shotgun sequencing) and RNA level (metatranscriptomics). Despite the great progresses in this field, technical issues still exist mainly related to the need of continuously updated databases, specific bioinformatic tools, and functional correlations.

In fact, since sequencing technologies have rapidly increased their throughput and reduced the cost of sequencing/base, a great amount of data is piling up imposing a careful evaluation regarding their use, analysis, and storage. Below, the main bioinformatic tools today available for conducting a microbiome study are briefly reported (Table 1) [48].

For a description of features and issues related to molecular, NGS-based analyses, refer to specific papers on these topics [2,49].

3.1. 16S Rrnas and ITS Sequencing

Bacterial 16S ribosomal RNA (rRNA) gene analysis allows the fast characterization of the microbial composition of specific environments. The principle of this approach is that the 16S rRNA gene is a bacterial-specific gene with hyper-variable regions in its sequence that are useful for phylogenetic purposes. Specific bioinformatic tools have been developed for the analysis of 16S rRNA sequences including different functions, like taxonomic assignment, diversity analysis, and functional prediction (Table 1).

Table 1. Bioinformatic tools for metagenomic datasets analyses.

Function	Pipeline	Reference
16S Databases	Ribosomal Database Project Greengenes Silva	(Cole et al., 2014) [50] (DeSantis et al., 2006) [51] (Quast et al., 2013) [52]
OTUs clustering	Cd-hit UCLUST	(Li et al., 2006) [53] (Edgar et al., 2010) [54]
Sub-OTU methods	DADA2 UNOISE2 Deblur	(Callahan et al., 2016) [55] (Edgar, 2016) [56] (Amir et al., 2017) [57]
Communities diversity comparison	UniFrac	(Lozupone et al., 2006) [58]
Functional profiles prediction	PICRUSt Piphillin Tax4Fun	(Langille et al., 2013) [59] (Iwai et al., 2016) [60] (Aßhauer et al., 2015) [61]
Metagenome assembly	SOAPdenovo Velvet MetaVelvet Meta-IBDA Genovo Bambus2 Ray-Meta ViromeScan PanPhlAn	(Li et al., 2010) [62] (Zerbino et al., 2008) [63] (Namiki et al., 2012) [64] (Peng et al., 2011) [65] (Laserson et al., 2011) [66] (Koren et al., 2011) [67] (Boisvert et al., 2012) [68] (Rampelli et al., 2016) [69] (Scholz et al., 2016) [70]
Gene identification	FragGeneScan MetaGeneMark Glimmer-MG	(Rho et al., 2010) [71] (Zhu et al., 2010) [72] (Kelley et al., 2012) [73]
Functional assignment	IMG database MetaRef dbCAN HUMAnN	(Markowitz et al., 2014) [74] (Huang et al., 2014) [75] (Yin et al., 2012) [76] (Abubucker et al., 2012) [77]

Taxonomic assignment is based on the comparison of clustered reads with specific databases of known 16S sequences, i.e., Ribosomal Database Project [50], Greengenes [51], and Silva [52]. To date, different approaches have been optimized for reads clustering (Table 1). Some of these pipelines are based on sequence homology for bacterial identification. In brief, sequence reads obtained in experimental settings are aligned to 16S rRNAs specific databases to assign them to the most similar species. Consequently, the limitation of these approaches is that only the bacterial species annotated in the above-mentioned databases can be identified. Thus, unknown bacteria, corresponding to not aligned reads, are not identified. Greengenes is a specifically designed database of full-length 16S rRNA genes that allows taxonomic identification based on de novo tree inference; however, many reads are not classified [51].

Another possible approach, able to overcome this limitation, is the Operational Taxonomic Units (OTUs) clustering based on the comparison of sequences using different algorithms based on sequence length or pairwise alignment [53,54]. This is the reason why some popular tools for metagenomics, as QIIME (Quantitative Insights Into Microbial Ecology), integrate a combinatorial approach for OTUs clustering [78]. First, sequence reads are aligned against the reference databases for taxonomic assignment. Then, not assigned reads are clustered, based on their sequence similarity to each other. In this way, the taxonomic assignment is more efficient and should avoid the lack of information due to bacterial databases limits. To date, the OTUs clustering has been largely used for 16S analysis; however, its biological relevance has been recently questioned. In fact, even if the clustering of reads allows to minimize the effects of sequencing errors, this operation may affect real phylogenetic diversity, because very similar taxa may be not correctly discriminated. Consequently, alternative strategies,

namely sub-OTUs methods, are becoming available. Divisive Amplicon Denoising Algorithm 2 (DADA2) is an open-source algorithm implemented in R, which uses a statistical inference to correct amplicon errors [55]. This package includes all the analytical steps required for amplicon analysis and, by portioning the reads through a denoising algorithm based on the Illumina error model, allows a more accurate bacterial communities resolution with the potential to achieve strain-level information [55]. Similarly, UNOISE2, uses another algorithm optimized for denoising Illumina amplicon reads, which allows reads clustering in zero-radius OTUs (ZOTUs), i.e., OTUs with an identity higher than the conventional 97% [56]. Finally, Deblur promises the resolution of taxa with just one different bp, by using error profiles to infer putative error-free reads [57].

After taxonomic assignment, bacterial communities are compared to highlight significant differences among groups (i.e., patients versus healthy subjects). Actually, one of the most used approaches for this purpose is UniFrac [58]. Unifrac allows communities comparison by measuring the phylogenetic distance between datasets of taxa and implements different tests, such as *p* test and Principal Coordinates Analysis (PCoA) (Table 1). QIIME integrates also these functions [78].

16S rRNA gene sequence analysis provides a rapid and effective way to analyze the entire content of a specific microbial community. Even if it provides taxonomic content and communities comparison and similarity analyses, it does not allow functional evaluations. Recently, specific pipelines are becoming available to infer specific metabolic functions that could be related to a microbial community, suggesting pathways that may be over- or underexpressed in different analyzed samples (Table 1) [59–61]. PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) tool uses 16S rRNA information to infer the functional composition of a metagenome by predicting which gene families are present and their relative abundances [59]. Even if these imputed data need to be experimentally validated, they could provide a primary hypothesis on possible functional alterations associated to a specific disease-related dysbiosis.

Similar issues also concern fungal analysis. In this case, Internal Transcribed Spacers (ITSs), located within the ribosomal gene operon and specific databases, are used as tags for the taxonomic assignment [49].

It has to be noticed that, currently, most of our knowledge related to archaeome is derived from bacterial 16S rRNA studies as a kind of incidental finding [44]. Since this bacterial-targeted approach is not able to detect all the spectrum of *Archaea*, methodological issues have largely limited archaeome qualitative and quantitative characterization [44]. To overcome these limitations, Koskinen et al. proposed a Polymerase Chain Reaction (PCR)-based protocol specifically designed for an in-depth study of the archaeome and assessed the DADA2 pipeline as the most efficient for this kind of studies [44]. These results may be the basis for future investigations specifically designed to study the human archaeome with the aim to characterize it in healt and disease status, unveil its physiological role, and verify a possible contribution of *Archaea* to pathological mechanisms.

No similar approaches are available for other microbial classes (such as viruses), and, also, specific databases are still poor, so that they are usually not taken into account, their contribution to health and disease status being, consequently, actually underestimated.

3.2. Shotgun Sequencing

To overcome 16S rRNA genes sequencing limitations, a shotgun strategy, able to analyze the entire genomic content of a community, can be used. The molecular approach is similar to that used for the analysis of a single bacterial genome [79,80], but in this case sequences from both the host and all the microbes (bacteria, archaea, fungi, and viruses) present in the studied environment are obtained. Consequently, the taxonomic assignment requires highly performing assembly tools and updated databases. In fact, in addition to the assembly issues related to NGS (Next Generation Sequencing) reads length, in metagenomic studies we have to consider that we want to assemble a multitude of different genomes that are present in different proportions within the studied community.

Specific assemblers have been developed for shotgun metagenomics (Table 1), such as SOAPdenovo [62], Velvet [63], MetaVelvet [64], Meta-IBDA [65], Genovo [66], Bambus2 [67], and Ray-Meta [68]. Currently, there is no unique consensus regarding the best assembly strategy, since it has been noticed that the assembler performances depend on both biological and technical factors, and different assemblers work better on different datasets [81,82].

Indeed, while, when assembling a single genome, it is possible to use sequencing coverage to identify repeat regions and sequencing errors, in the assembly of a metagenome the sequencing coverage is a consequence of the abundance of each genome within the analyzed community. Thus, the assembler should be able to obtain long contigs for highly represented genomes, without losing the low represented ones [82]. It has to be noticed that, since human and bacterial DNAs are expected to be the most represented in metagenomic datasets, the assembly of poorly represented microbes may result very hard and could generate a loss of information. Consequently, a high sequencing coverage is required to be sure to obtain informative reads representing the entire community, and computational time and space are required to carry out the assembly [82].

In addition, many of the bioinformatic tools currently available for shotgun metagenomics do not allow efficient viruses estimation. In this context, ViromeScan is an open-source tool, specifically developed to analyze virome taxonomy and relative abundance directly from shotgun metagenomic data [69]. ViromeScan is able to enrich viral reads from a metagenomic dataset and map them on a hierarchical and customizable viral database; however, a further implementation with an assembly step is required to allow the detection of new viruses, currently not considered in this analysis [69].

Another difficulty of the assembly step is represented by the possibility to have in the same metagenome two strains of the same bacterial species, whose genomes may differ for few nucleotides. To address this issue, specific pipelines for strain-level analysis are becoming available [82]. In this context, Scholz et al. designed PanPhlAn (pangenome-based phylogenomic analysis), an open-source tool able to obtain strain-level resolution from metagenomic data, also including novel strains identification [70]. Considering all the above, a prudential strategy may be the use of different assembly pipelines, even if costs and time may increase.

Shotgun approaches also allow a functional analysis; to avoid biases introduced during the assembly, some pipelines have been optimized to use unassembled reads for functional prediction. Briefly, a functional analysis is based on genes prediction to infer their probable functions. Different methods, such as FragGeneScan [71], MetaGeneMark [72], and Glimmer-MG [73], have been developed and optimized for this aim. Once the genes have been identified, specific databases can be used for functional predictions, such as IMG database [74], MetaRef [75], dbCAN [76], and HUMAnN [77] (Table 1).

Since the data obtained with this shotgun approach are more complete than those from simple 16S rRNAs or ITS (Internal Transcribed Spacer) analyses, it is conceivable that, as the cost of sequencing decreases, the use of shotgun metagenomics will increase. Even if some biases can be present related to sample collection and storage, DNA extraction procedures, and PCR amplification errors, the shotgun metagenomics offers a comprehensive view of genomes of all the microbes that compose the microbiota of interest. The shotgun metagenomic approach, however, requires high sequence coverage so that taxa in small proportion are represented in the data. Consequently, a high coverage can increase sequencing costs, and the large amount of data that is generated requires additional data storage and specialized analysis pipelines to efficiently catch medical-sensitive information.

3.3. Metatranscriptomics

Metatranscriptomics is the analysis of the whole transcriptome of a microbial community. Thus, it provides insight into gene expression and possible functions of human microbiomes in a specific organ or tissue and in a specific condition [83]. In addition to the technical issues related to

RNA samples extraction, quality, and purification, comprehensive bioinformatic tools are not still available, and research groups use different in-house developed pipelines. Recently, Westrich et al. described a novel, publicly available pipeline designed ad hoc for metatrascriptomic datasets analysis and gave also useful indications for obtaining good quality data for metatranscriptomics from stool samples [84]. Similarly, Jiang et al. tested their own pipeline on data from different microbiomes showing how, by integrating taxonomic assignment and functional data, it is possible to infer the contribution of specific taxa [85].

Once data analysis will be further optimized and standardized, it is conceivable to suppose that metatranscriptomics may become a very popular approach offering a previously unthinkable way to correlate the presence of specific microbes to specific biochemical and metabolic functions, alterations, or both.

4. Conclusions

Metagenomic studies are highlighting that quantitative and/or qualitative microbial alterations occur in an increasing number of diseases. Since the microbiome appears as an actionable target for the development of both novel diagnostic tests and possible targeted therapies, a growing interest has been observed in this topic. This enthusiasm was tempered by the fact that most of the already published studies on the human gut microbiome just report an association with no data regarding functional effects and the possible role on diseases onset or development. The further development and optimization of novel tools for data analysis, integration, correlation, and visualization will allow in a near future to fill in this gap. In this way, it will be possible to understand the functional, mutually beneficial relationship between humans and their gut microbiome during aging and, consequently, its role in the acquisition and maintenance of a healthy status.

Author Contributions: Valeria D'Argenio conceived, wrote, and approved the manuscript.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

NGS	Next-Generation Sequencing
rRNA	ribosomal RNA
OTUs	Operational Taxonomic Units
QIIME	Quantitative Insights Into Microbial Ecology
DADA2	Divisive Amplicon Denoising Algorithm 2
ZOTUs	zero-radius OTUs
Bp	base pair
PCoA	Principal Coordinates Analysis
PICRUSt	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
ITSs	Internal Transcribed Spacers
PCR	Polymerase Chain Reaction
PanPhlAn	pangenome-based phylogenomic analysis

References

1. Precone, V.; del Monaco, V.; Esposito, M.V.; de Palma, F.D.; Ruocco, A.; Salvatore, F.; D'Argenio, V. Cracking the code of human diseases using next-generation sequencing: Applications, challenges, and perspectives. *BioMed Res. Int.* **2015**, *2015*, 161648. [CrossRef] [PubMed]
2. D'Argenio, V.; Salvatore, F. The role of the gut microbiome in the healthy adult status. *Clin. Chim. Acta* **2015**, *451*, 97–102. [CrossRef] [PubMed]
3. Di Pietro, M.; Filardo, S.; Porpora, M.G.; Recine, N.; Latino, M.A.; Sessa, R. HPV/*Chlamydia trachomatis* co-infection: Metagenomic analysis of cervical microbiota in asymptomatic women. *New Microbiol.* **2018**, *41*, pmid:29313867.

4.   Wang, W.M.; Jin, H.Z. Skin Microbiome: An Actor in the Pathogenesis of Psoriasis. *Chin. Med. J.* **2018**, *131*, 95–98. [CrossRef] [PubMed]

5.   Nycz, B.T.; Dominguez, S.R.; Friedman, D.; Hilden, J.M.; Diana, I.; Robertson, C.E.; Frank, D.N. Evaluation of bloodstream infections, *Clostridium difficile* infections, and gut microbiota in pediatric oncology patients. *PLoS ONE* **2018**, *13*, e0191232. [CrossRef] [PubMed]

6.   Luna, P.N.; Hasegawa, K.; Ajami, N.J.; Espinola, J.A.; Henke, D.M.; Petrosino, J.F.; Piedra, P.A.; Sullivan, A.F.; Camargo, C.A., Jr.; Shaw, C.A.; et al. The association between anterior nares and nasopharyngeal microbiota in infants hospitalized for bronchiolitis. *Microbiome* **2018**, *6*, 2. [CrossRef] [PubMed]

7.   Si, J.; Lee, C.; Ko, G. Oral Microbiota: Microbial Biomarkers of Metabolic Syndrome Independent of Host Genetic Factors. *Front. Cell. Infect. Microbiol* **2017**, *7*, 516. [CrossRef] [PubMed]

8.   Dong, L.N.; Wang, J.P.; Liu, P.; Yang, Y.F.; Feng, J.; Han, Y. Faecal and mucosal microbiota in patients with functional gastrointestinal disorders: Correlation with toll-like receptor 2/toll-like receptor 4 expression. *World J. Gastroenterol.* **2017**, *23*, 6665–6673. [CrossRef] [PubMed]

9.   D'Argenio, V.; Torino, M.; Precone, V.; Casaburi, G.; Esposito, M.V.; Iaffaldano, L.; Malapelle, U.; Troncone, G.; Coto, I.; Cavalcanti, P.; et al. The cause of death of a child in the 18th century solved by bone microbiome typing using laser microdissection and next generation sequencing. *Int. J. Mol. Sci.* **2017**, *18*, 109. [CrossRef] [PubMed]

10.   Sassone-Corsi, M.; Raffatellu, M. No vacancy: How beneficial microbes cooperate with immunity to provide colonization resistance to pathogens. *J. Immunol.* **2015**, *194*, 4081–4087. [CrossRef] [PubMed]

11.   Selber-Hnatiw, S.; Rukundo, B.; Ahmadi, M.; Akoubi, H.; Al-Bizri, H.; Aliu, A.F.; Ambeaghen, T.U.; Avetisyan, L.; Bahar, I.; Baird, A.; et al. Human gut microbiota: Toward an ecology of disease. *Front. Microbiol.* **2017**, *8*, 1265. [CrossRef] [PubMed]

12.   Kantarci, A.; Hasturk, H. Microbes and host response: A relationship in health and disease. *Oral Dis.* **2017**. [CrossRef] [PubMed]

13.   Ni, J.; Wu, G.D.; Albenberg, L.; Tomov, V.T. Gut microbiota and IBD: Causation or correlation? *Nat. Rev. Gastroenterol. Hepatol.* **2017**, *14*, 573–584. [CrossRef] [PubMed]

14.   Duranti, S.; Ferrario, C.; van Sinderen, D.; Ventura, M.; Turroni, F. Obesity and microbiota: An example of an intricate relationship. *Genes Nutr.* **2017**, *12*, 18. [CrossRef] [PubMed]

15.   Lynch, S.V.; Pedersen, O. The Human Intestinal Microbiome in Health and Disease. *N. Engl. J.Med.* **2016**, *375*, 2369–2379. [CrossRef] [PubMed]

16.   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] [PubMed]

17.   D'Argenio, V.; Precone, V.; Casaburi, G.; Miele, E.; Martinelli, M.; Staiano, A.; Salvatore, F.; Sacchetti, L. An altered gut microbiome profile in a child affected by Crohn's disease normalized after nutritional therapy. *Am. J. Gastroenterol.* **2013**, *108*, 851–852. [CrossRef] [PubMed]

18.   D'Argenio, V.; Casaburi, G.; Precone, V.; Pagliuca, C.; Colicchio, R.; Sarnataro, D.; Discepolo, V.; Kim, S.M.; Russo, I.; del Vecchio Blanco, G.; et al. Metagenomics reveals dysbiosis and a potentially pathogenic *N. flavescens* strain in duodenum of adult celiac patients. *Am. J. Gastroenterol.* **2016**, *111*, 879–890. [CrossRef] [PubMed]

19.   D'Argenio, V.; Casaburi, G.; Precone, V.; Pagliuca, C.; Colicchio, R.; Sarnataro, D.; Discepolo, V.; Kim, S.M.; Russo, I.; del Vecchio Blanco, G.; et al. No change in the mucosal gut microbiome is associated with celiac disease-specific microbiome alteration in adult patients. *Am. J. Gastroenterol.* **2016**, *111*, 1659–1661. [CrossRef] [PubMed]

20.   Pham, V.T.; Lacroix, C.; Braegger, C.P.; Chassard, C. Lactate-utilizing community is associated with gut microbiota dysbiosis in colicky infants. *Sci. Rep.* **2017**, *7*, 11176. [CrossRef] [PubMed]

21.   Yan, Q.; Gu, Y.; Li, X.; Yang, W.; Jia, L.; Chen, C.; Han, X.; Huang, Y.; Zhao, L.; Li, P.; et al. Alterations of the Gut Microbiome in Hypertension. *Front. Cell. Infect. Microbiol.* **2017**, *7*, 381. [CrossRef] [PubMed]

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

23.   Eklöf, V.; Löfgren-Burström, A.; Zingmark, C.; Edin, S.; Larsson, P.; Karling, P.; Alexeyev, O.; Rutegård, J.; Wikberg, M.L.; Palmqvist, R. Cancer-associated fecal microbial markers in colorectal cancer detection. *Int. J. Cancer* **2017**, *41*, 2528–2536. [CrossRef] [PubMed]

24. Perez-Muñoz, M.E.; Arrieta, M.C.; Ramer-Tait, A.E.; Walter, J. A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: Implications for research on the pioneer infant microbiome. *Microbiome* **2017**, *5*, 48. [CrossRef] [PubMed]

25. Rodríguez, J.M.; Murphy, K.; Stanton, C.; Ross, R.P.; Kober, O.I.; Juge, N.; Avershina, E.; Rudi, K.; Narbad, A.; Jenmalm, M.C.; et al. The composition of the gut microbiota throughout life, with an emphasis on early life. *Microb. Ecol. Health Dis.* **2015**, *26*, 26050. [CrossRef] [PubMed]

26. Collado, M.C.; Rautava, S.; Aakko, J.; Isolauri, E.; Salminen, S. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. *Sci. Rep.* **2016**, *6*, 23129. [CrossRef] [PubMed]

27. Neu, J. The microbiome during pregnancy and early postnatal life. *Semin. Fetal Neonatal Med.* **2016**, *21*, 373–379. [CrossRef] [PubMed]

28. Stinson, L.F.; Payne, M.S.; Keelan, J.A. Planting the seed: Origins, composition, and postnatal health significance of the fetal gastrointestinal microbiota. *Crit. Rev. Microbiol.* **2017**, *43*, 352–369. [CrossRef] [PubMed]

29. Mor, G.; Aldo, P.; Alvero, A.B. The unique immunological and microbial aspects of pregnancy. *Nat. Rev. Immunol.* **2017**, *17*, 469–482. [CrossRef] [PubMed]

30. Nuriel-Ohayon, M.; Neuman, H.; Koren, O. Microbial changes during pregnancy, birth, and infancy. *Front. Microbiol.* **2016**, *7*, 1031. [CrossRef] [PubMed]

31. Saavedra, J.M.; Dattilo, A.M. Early development of intestinal microbiota: Implications for future health. *Gastroenterol. Clin. N. Am.* **2012**, *41*, 717–731. [CrossRef] [PubMed]

32. Dogra, S.; Sakwinska, O.; Soh, S.E.; Ngom-Bru, C.; Brück, W.M.; Berger, B.; Brüssow, H.; Lee, Y.S.; Yap, F.; Chong, Y.S.; et al. Dynamics of infant gut microbiota are influenced by delivery mode and gestational duration and are associated with subsequent adiposity. *MBio* **2015**, *6*. [CrossRef] [PubMed]

33. Nagpal, R.; Tsuji, H.; Takahashi, T.; Nomoto, K.; Kawashima, K.; Nagata, S.; Yamashiro, Y. Gut dysbiosis following C-section instigates higher colonisation of toxigenic *Clostridium perfringens* in infants. *Benef. Microbes* **2017**, *8*, 353–365. [CrossRef] [PubMed]

34. Dominguez-Bello, M.G.; Costello, E.K.; Contreras, M.; Magris, M.; Hidalgo, G.; Fierer, N.; Knight, R. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11971–11975. [CrossRef] [PubMed]

35. Chu, D.M.; Ma, J.; Prince, A.L.; Antony, K.M.; Seferovic, M.D.; Aagaard, K.M. Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat. Med.* **2017**, *23*, 314–326. [CrossRef] [PubMed]

36. Hunt, K.M.; Foster, J.A.; Forney, L.J.; Schütte, U.M.; Beck, D.L.; Abdo, Z.; Fox, L.K.; Williams, J.E.; McGuire, M.K.; McGuire, M.A. Characterization of the diversity and temporal stability of bacterial communities in human milk. *PLoS ONE* **2011**, *6*, e21313. [CrossRef] [PubMed]

37. Kundu, P.; Blacher, E.; Elinav, E.; Pettersson, S. Our gut microbiome: The evolving inner self. *Cell* **2017**, *171*, 1481–1493. [CrossRef] [PubMed]

38. Erny, D.; Hrabě de Angelis, A.L.; Prinz, M. Communicating systems in the body: How microbiota and microglia cooperate. *Immunology* **2017**, *150*, 7–15. [CrossRef] [PubMed]

39. Martí, J.M.; Martínez-Martínez, D.; Rubio, T.; Gracia, C.; Peña, M.; Latorre, A.; Moya, A.; P Garay, C. Health and disease imprinted in the time variability of the human microbiome. *mSystems* **2017**, *2*. [CrossRef] [PubMed]

40. Wang, J.Z.; Du, W.T.; Xu, Y.L.; Cheng, S.Z.; Liu, Z.J. Gut microbiome-based medical methodologies for early-stage disease prevention. *Microb. Pathog.* **2017**, *105*, 122–130. [CrossRef] [PubMed]

41. Cani, P.D. Gut microbiota—At the intersection of everything? *Nat. Rev. Gastroenterol. Hepatol.* **2017**, *14*, 321–322. [CrossRef] [PubMed]

42. Duvallet, C.; Gibbons, S.M.; Gurry, T.; Irizarry, R.A.; Alm, E.J. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat. Commun.* **2017**, *8*, 1784. [CrossRef] [PubMed]

43. Iliev, I.D.; Leonardi, I. Fungal dysbiosis: Immunity and interactions at mucosal barriers. *Nat. Rev. Immunol.* **2017**, *17*, 635–646. [CrossRef] [PubMed]

44. 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*. [CrossRef] [PubMed]

45.	Virgin, H.W. The virome in mammalian physiology and disease. *Cell* **2014**, *157*, 142–150. [CrossRef] [PubMed]

46.	Cadwell, K. Expanding the role of the virome: Commensalism in the gut. *J. Virol.* **2015**, *89*, 1951–1953. [CrossRef] [PubMed]

47.	Cadwell, K. The virome in host health and disease. *Immunity* **2015**, *42*, 805–813. [CrossRef] [PubMed]

48.	Goodrich, J.K.; Waters, J.L.; Poole, A.C.; Sutter, J.L.; Koren, O.; Blekhman, R.; Beaumont, M.; Van Treuren, W.; Knight, R.; Bell, J.T.; et al. Human genetics shape the gut microbiome. *Cell* **2014**, *159*, 789–799. [CrossRef] [PubMed]

49.	Shankar, J. Insights into study design and statistical analyses in translational microbiome studies. *Ann. Transl. Med.* **2017**, *5*, 249. [CrossRef] [PubMed]

50.	Cole, J.R.; Wang, Q.; Fish, J.A.; Chai, B.; McGarrell, D.M.; Sun, Y.; Brown, C.T.; Porras-Alfaro, A.; Kuske, C.R.; Tiedje, J.M. Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **2014**, *42*, D633–D642. [CrossRef] [PubMed]

51.	DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072. [CrossRef] [PubMed]

52.	Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [CrossRef] [PubMed]

53.	Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659. [CrossRef] [PubMed]

54.	Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461. [CrossRef] [PubMed]

55.	Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.; Holmes, S.P. DADA2: High-resolution sample inference from *Illumina amplicon* data. *Nat. Methods.* **2016**, *13*, 581–583. [CrossRef] [PubMed]

56.	Edgar, R.C. UNOISE2: Improved error-correction for *Illumina* 16S and ITS amplicon sequencing. *BioRxiv* **2017**. [CrossRef]

57.	Amir, A.; McDonald, D.; Navas-Molina, J.A.; Kopylova, E.; Morton, J.T.; Zech Xu, Z.; Kightley, E.P.; Thompson, L.R.; Hyde, E.R.; Gonzalez, A.; et al. Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* **2017**, *7*, 2. [CrossRef] [PubMed]

58.	Lozupone, C.; Hamady, M.; Knight, R. UniFrac—An online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinform.* **2006**, *7*, 371. [CrossRef] [PubMed]

59.	Langille, M.G.I.; Zaneveld, J.; Caporaso, J.G.; McDonald, D.; Knights, D.; Reyes, J.; Clemente, J.C.; Burkepile, D.E.; Vega Thurber, R.L.; Knight, R.; et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **2013**, *8*, 1–10. [CrossRef] [PubMed]

60.	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] [PubMed]

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

62.	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]

63.	Zerbino, D.R.; Birney, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **2008**, *18*, 821–829. [CrossRef] [PubMed]

64.	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]

65.	Peng, Y.; Leung, H.C.; Yiu, S.M.; Chin, F.Y. Meta-IDBA: A de Novo assembler for metagenomic data. *Bioinformatics* **2011**, *27*, 94–101. [CrossRef] [PubMed]

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

67.	Koren, S.; Treangen, T.J.; Pop, M. Bambus 2: Scaffolding metagenomes. *Bioinformatics* **2011**, *27*, 2964–2971. [CrossRef] [PubMed]

68. Boisvert, S.; Raymond, F.; Godzaridis, E.; Laviolette, F.; Corbeil, J. Ray Meta: Scalable de novo metagenome assembly and profiling. *Genome Biol.* **2012**, *13*, R122. [CrossRef] [PubMed]

69. Rampelli, S.; Soverini, M.; Turroni, S.; Quercia, S.; Biagi, E.; Brigidi, P.; Candela, M. ViromeScan: A new tool for metagenomic viral community profiling. *BMC Genom.* **2016**, *17*, 165. [CrossRef] [PubMed]

70. Scholz, M.; Ward, D.V.; Pasolli, E.; Tolio, T.; Zolfo, M.; Asnicar, F.; Truong, D.T.; Tett, A.; Morrow, A.L.; Segata, N. Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nat. Methods* **2016**, *13*, 435–438. [CrossRef] [PubMed]

71. Rho, M.; Tang, H.; Ye, Y. FragGeneScan: Predicting genes in short and error-prone reads. *Nucleic Acids Res.* **2010**, *38*, e191. [CrossRef] [PubMed]

72. Zhu, W.; Lomsadze, A.; Borodovsky, M. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res.* **2010**, *38*, e132. [CrossRef] [PubMed]

73. Kelley, D.R.; Liu, B.; Delcher, A.L.; Pop, M.; Salzberg, S.L. Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. *Nucleic Acids Res.* **2012**, *40*, e9. [CrossRef] [PubMed]

74. Markowitz, V.M.; Chen, I.M.; Chu, K.; Szeto, E.; Palaniappan, K.; Pillay, M.; Ratner, A.; Huang, J.; Pagani, I.; Tringe, S.; et al. IMG/M 4 version of the integrated metagenome comparative analysis system. *Nucleic Acids Res.* **2014**, *42*, D568–D573. [CrossRef] [PubMed]

75. Huang, K.; Brady, A.; Mahurkar, A.; White, O.; Gevers, D.; Huttenhower, C.; Segata, N. MetaRef: A pan-genomic database for comparative and community microbial genomics. *Nucleic Acids Res.* **2014**, *42*, D617–624. [CrossRef] [PubMed]

76. Yin, Y.; Mao, X.; Yang, J.; Chen, X.; Mao, F.; Xu, Y. dbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **2012**, *40*, W445–W451. [CrossRef] [PubMed]

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

78. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335–336. [CrossRef] [PubMed]

79. D'Argenio, V.; Notomista, E.; Petrillo, M.; Cantiello, P.; Cafaro, V.; Izzo, V.; Naso, B.; Cozzuto, L.; Durante, L.; Troncone, L.; et al. Complete sequencing of *Novosphingobium sp. PP1Y* reveals a biotechnologically meaningful metabolic pattern. *BMC Genom.* **2014**, *15*, 384. [CrossRef] [PubMed]

80. D'Argenio, V.; Petrillo, M.; Pasanisi, D.; Pagliarulo, C.; Colicchio, R.; Talà, A.; de Biase, M.S.; Zanfardino, M.; Scolamiero, E.; Pagliuca, C.; et al. The complete 12 Mb genome and transcriptome of *Nonomuraea gerenzanensis* with new insights into its duplicated "magic" RNA polymerase. *Sci. Rep.* **2016**, *6*, 18. [CrossRef] [PubMed]

81. Bradnam, K.R.; Fass, J.N.; Alexandrov, A.; Baranay, P.; Bechner, M.; Birol, I.; Boisvert, S.; Chapman, J.A.; Chapuis, G.; Chikhi, R. Assemblathon 2: Evaluating de novo methods of genome assembly in three vertebrate species. *Gigascience* **2013**, *2*, 10. [CrossRef] [PubMed]

82. Quince, C.; Walker, A.W.; Simpson, J.T.; Loman, N.J.; Segata, N. Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* **2017**, *35*, 833–844. [CrossRef] [PubMed]

83. Heintz-Buschart, A.; Wilmes, P. Human gut microbiome: Function matters. *Trends Microbiol.* **2017**. [CrossRef] [PubMed]

84. Westreich, S.T.; Korf, I.; Mills, D.A.; Lemay, D.G. SAMSA: A comprehensive metatranscriptome analysis pipeline. *BMC Bioinform.* **2016**, *17*, 399. [CrossRef] [PubMed]

85. Jiang, Y.; Xiong, X.; Danska, J.; Parkinson, J. Metatranscriptomic analysis of diverse microbial communities reveals core metabolic pathways and microbiome-specific functionality. *Microbiome* **2016**, *4*, 2. [CrossRef] [PubMed]