



Published in final edited form as:

*Nutr Rev.* 2012 August ; 70(Suppl 1): S38–S44. doi:10.1111/j.1753-4887.2012.00493.x.

## Defining the Human Microbiome

Luke K Ursell<sup>1</sup>, Jessica L Metcalf<sup>1</sup>, Laura Wegener Parfrey<sup>1</sup>, and Rob Knight<sup>1,2,\*</sup>

<sup>1</sup>Department of Chemistry & Biochemistry, University of Colorado at Boulder, Boulder, CO, 80309, USA

<sup>2</sup>Howard Hughes Medical Institute, Boulder, CO, 80309, USA

### Abstract

Rapidly developing sequencing methods and analytical techniques are enhancing our ability to understand the human microbiome, and, indeed, how we define the microbiome and its constituents. In this review we highlight recent research that expands our ability to understand the human microbiome on different spatial and temporal scales, including daily timeseries datasets spanning months. Furthermore, we discuss emerging concepts related to defining operational taxonomic units, diversity indices, core versus transient microbiomes and the possibility of enterotypes. Additional advances in sequencing technology and in our understanding of the microbiome will provide exciting prospects for exploiting the microbiota for personalized medicine.

### Introduction

The human microbiota consists of the 10-100 trillion symbiotic microbial cells harbored by each person, primarily bacteria in the gut; the human microbiome consists of the genes these cells harbor [1]. Microbiome projects worldwide have been launched with the goal of understanding the roles that these symbionts play and their impacts on human health [2, 3]. Just as the question, “what *is* it to be human?”, has troubled humans from the beginning of recorded history, the question, “what *is* the human microbiome?” has troubled researchers since the term was coined by Joshua Lederberg in 2001 [4]. Specifying the definition of the human microbiome has been complicated by confusion about terminology: for example, “microbiota” (the microbial taxa associated with humans) and “microbiome” (the catalog of these microbes and their genes) are often used interchangeably. In addition, the term “metagenomics” originally referred to shotgun characterization of total DNA, although now it is increasingly being applied to studies of marker genes such as the 16S rRNA gene. More fundamentally, however, new findings are leading us to question the concepts that are central to establishing the definition of the human microbiome, such as the stability of an individual’s microbiome, the definition of the OTUs (Operational Taxonomic Units) that make up the microbiota, and whether a person has one microbiome or many. In this review, we cover progress towards defining the human microbiome in these different respects.

Studies of the diversity of the human microbiome started with Antonie van Leewenhoek, who, as early as the 1680s, had compared his oral and fecal microbiota. He noted the striking differences in microbes between these two habitats and also between samples from individuals in states of health and disease in both of these sites [5, 6]. Thus, studies of the profound differences in microbes at different body sites, and between health and disease, are as old as microbiology itself. What is new today is not the ability to observe these obvious

\*To whom correspondence should be addressed: Rob Knight, Department of Chemistry & Biochemistry, UCB 215, University of Colorado, Boulder, CO 80309, Tel: 303-492-1984, Fax: 303-492-7744, rob.knight@colorado.edu.

differences, but rather the ability to use powerful molecular techniques to gain insight into why these differences exist, and to understand how we can affect transformations from one state to another.

Culture-independent methods for characterizing the microbiota, together with a molecular phylogenetic approach to organizing life's diversity, provided a fundamental breakthrough in allowing researchers to compare microbial communities across environments within a unified phylogenetic context (reviewed in [7]). Although host-associated microbes are presumably acquired from the environment, the composition of the mammalian microbiota, especially in the gut, is surprisingly different from free-living microbial communities [8]. In fact, an analysis of bacterial diversity from free-living communities in terrestrial, marine, and freshwater environments as well as communities associated with animals suggests that the vertebrate gut is an extreme [8]. In contrast, bacterial communities from environments typically considered extreme, such as acidic hot springs and hydrothermal vents, are similar to communities in many other environments[9]. This suggests that coevolution between vertebrates and their microbial consortia over hundreds of millions of years has selected for a specialized community of microbes that thrive in the gut's warm, eutrophic, and stable environment[8]. In the human gut and across human-associated habitats, bacteria comprise the bulk of the biomass and diversity, though archaea, eukaryotes, and viruses are also present in smaller numbers and should not be neglected[10, 11].

Interestingly, estimates of the human gene catalog and the diversity of the human genome pale in comparison to estimates of the diversity of the microbiome. For example, the Meta-HIT consortium reported a gene catalog of 3.3 million non-redundant genes in the human gut microbiome alone[3], as compared to the ~22,000 genes present in the entire human genome[12]. Similarly, the diversity among the microbiome of individuals is immense compared to genomic variation: individual humans are about 99.9% identical to one another in terms of their host genome[13], but can be 80-90% different from one another in terms of the microbiome of their hand[14] or gut[15]. These findings suggest that employing the variation contained within the microbiome will be much more fruitful in personalized medicine, the use of an individual patient's genetic data to inform healthcare decisions, than approaches that target the relatively constant host genome.

Many fundamental questions about the human microbiome were difficult or impossible to address until recently. Some questions, such as the perennially popular “how many species live in a given body site?”, are still hard to answer, due to problems with definitions of bacterial species and with the rate of sequencing error. Other questions, such as “how does the diversity within a person over time compare to the diversity between people?”, or “how does the diversity between sites on the same person's body compare to the diversity between different people at the same site?”, or “is there a core set of microbial species that we all share?”, can now be answered conclusively. In the next section, we discuss some of the tools that have allowed these long-standing questions to be answered.

## Tools for Microbial Analysis

The drastic reduction in sequencing costs experienced over the past few years has made it possible to identify specific microbial taxa found within the human gut that are difficult or impossible to culture. Researchers are now able to generate millions of sequences per sample in order to assess differences in microbial communities between body sites and individuals. Our increased sequencing power has required the development of equally powerful computational tools to handle the burgeoning amount of sequence data produced by modern technologies[16]. There are several pipelines for analysis of microbial community data such as mothur[17], w.A.T.E.R.S[18], the RDP pyrosequencing tools[19],

and QIIME (pronounced “chime”)[20]. QIIME is a free, open-source platform for the analysis of high-throughput sequencing data that enables users to import raw sequence data and readily produce measures of inter- and intra-sample diversity. Consistency in the identification of operational taxonomic units (OTUs) and establishing agreed-upon measures of diversity within and between samples are crucial for the comparison of results across studies, although the concept of OTU is increasingly problematic as sequence data accumulate and explicitly phylogenetic approaches gain in popularity.

Beta diversity refers to the measurement of the degree of difference in community membership or structure between two samples. A recent review of taxon-based measurements of beta diversity found that some metrics, including Canberra and Gower distances, have increased power for discriminating clusters, while other metrics, such as chi-squared and Pearson correlation distances, are more appropriate for elucidating the effects of environmental gradients on communities[21]. A robust method for comparing the differences between microbial communities is UniFrac, which measures the proportion of shared branch lengths on a phylogenetic tree between samples[22]. Highly similar microbial communities result in UniFrac scores near 0, while two completely independent communities that do not share any branch length (i.e. they have a different evolutionary history) would result in a UniFrac score of 1. Principal coordinates analysis (PCoA) can then visualize the UniFrac distances between samples in two-dimensional or three-dimensional space, allowing for the clustering of similar communities or separation of distinct communities to be easily distinguished visually.

UniFrac as a measure of beta diversity, coupled to PCoA, has the ability to distinguish differences between communities utilizing as little as 10 sequences per sample[23]. It is important to recognize that increased sequencing depth is not always necessary to recover biologically meaningful results when those results are obvious. Thus, by choosing diversity measurements that are appropriate for a study design, researchers utilizing modern sequencing methods are able to characterize differences between samples at relatively low sequence coverage. This enables researchers to assess fine-grained spatial and temporal patterns by characterizing hundreds to thousands of samples, such as timeseries across multiple patients or environments. The functionality of UniFrac, as well as a multitude of diversity measurements are available in QIIME and can be readily compared.

In general, pipelines for analyzing 16S rRNA and shotgun metagenomic data have separate workflows. Some initial steps, such as demultiplexing (removing barcodes from and separating pooled samples) and quality filtering, are common to both pipelines. However, for 16S rRNA data, sequences must be grouped into OTUs, chimeric sequences generated by incomplete template extension must be removed, and phylogenetic trees must be constructed. In contrast, in the metagenomic pipeline, sequences must be assigned to functions as well as to taxonomy (either as whole reads or after assembly). Once taxon or gene function tables are constructed, the pipelines begin to converge, at least conceptually: the interest is then in 1) the composition of each sample, 2) finding the taxa or functions that discriminate among groups of samples (e.g. according to clinical parameters), and 3) in asking whether the samples cluster according to any measured clinical states (or according to time). One exciting emerging direction is comparing metagenomic and 16S rRNA clustering directly using a technique called Procrustes analysis that allows the PCoA plots to be combined[24]. Another powerful tool is the use of machine learning and statistical techniques to build predictive models of taxa[25] or functions[26] that discriminate between groups of samples.

A unique advantage of QIIME relative to other pipelines is its ability to exploit “sample metadata”, e.g. clinical information about subjects, to produce visualizations that make the

main patterns in the data immediately apparent. Of particular interest, QIIME supports the MIMARKS (Minimum Information about a MARKer Sequence) standard[27] developed by the Genomic Standards Consortium[28], which is increasingly popular with other tools for microbial and community analysis such as MG-RAST[29], and has been adopted by the INSDC (International Nucleotide Sequence Database Consortium, which includes GenBank, EBI, and DDBJ) as the standard for metadata.

With these tools in hand, basic patterns of similarities and differences in the microbiota are now routine. The key challenge now is to extend analyses to include longitudinal studies and to understand the role of specific host and environmental factors in the development and maintenance of the microbiome.

## Dynamic interactions between human microbes and the environment

The gastrointestinal (GI) tract of a human infant provides a brand new environment for microbial colonization[30]. Indeed, the microbiota that an infant begins to acquire depends strongly on mode of delivery[31]. Twenty minutes after birth, the microbiota of vaginally delivered infants resembles the microbiota of their mother's vagina, while infants delivered via Cesarean section harbor microbial communities typically found on human skin[32]. The acquisition of microbiota continues over the first few years of life, as an infant's GI tract microbiome begins to resemble that of an adult as early as 1 year of life[33]. In one case-study following an infant's microbiota over the first 2.5 years of life, phylogenetic diversity increases significantly and linearly with time[34]. Additionally, significant changes in gut microbiota composition were apparent at five time points; starting a diet of breast milk, development of fever at day 92, introduction of rice cereal at day 134, introduction of formula and table foods at day 161, and antibiotic treatment and adult diet at day 371[34]. Interestingly, each dietary change was accompanied by changes in gut microbiota and the enrichment of corresponding genes. For example, as the infant began to receive a full adult diet, genes in the microbiome associated with vitamin biosynthesis and polysaccharide digestion became enriched[34].

The interaction between the human microbiota and the environment is dynamic, with human microbes flowing freely onto the surfaces we interact with everyday. Fierer et al. showed that human fingertips can transfer signature communities of microbes onto keyboards and these communities strongly differentiate individuals [35]. PCoA plots showed that it was possible to determine which fingers were typing on which keys, and which individuals were using which keyboards: it was even possible to link a person's hand to the computer mouse they use with up to 95% accuracy when compared to a database of other hands[35]. Overall, this study showed that microbial communities are constantly being transferred between surfaces, and that a dynamic interaction exists between environmental microbiota and different human body sites.

## Intrapersonal microbial diversity

Another interesting question that we are just beginning to answer is how stable the microbiome within an individual is over time. By defining what constitutes normal temporal variation in an individual over time, we will be better able to quantify and understand changes in microbial communities that result from dietary and pharmaceutical interventions. In the longest timeseries study to date, Caporaso et al. sampled two individual's microbial communities in the gut, oral cavity, and left and right palms over 396 time points spanning 15 months[36]. Communities at different body sites were readily distinguishable from one another using 3-D PCoA plots over a one year time span, even though the community structure within a given site was highly variable[36]. The level of diversity is also different between body sites, with the mouth and gut harboring the most diverse communities[37].

Taken together, these studies show that an individual's microbiota represents a highly variable and compartmentalized ecosystem.

Overall, it has yet to be conclusively proven that individuals, or even body sites, harbor a “core” set of specific bacterial taxa. For example, the Meta-HIT consortium defined a “core” set of lineages as those that were present in half of the subjects studied, although essentially no genes were present in all subjects studied[3]. Of course, it is important to recognize that sampling depth may be critical for distinguishing taxa that are absent from those that are merely very rare; the dynamic range of microbial abundance is also quite large, and even within the Meta-HIT “core” genes, 2000-fold ranges of abundance were not uncommon. Proving that a taxon is completely absent in the gut is not possible with these types of studies, so core calculations should always carry with them a caveat about sequencing depth. Another factor to consider when defining diversity and a core is that methodological artifacts can greatly increase the apparent numbers of OTUs in a sample (and hence reduce the apparent fraction that is shared). Both sequencing error[38, 39] and issues related to alignment, especially multiple sequence alignment[40-43], can inflate the number of OTUs immensely. It is important to ensure that the same methodological procedures were used when performing estimates of the core in terms of the fraction of individuals the core must be represented in, the minimum abundance, and the procedure for deciding which sequences count as “the same”. Finally, there is a key question about whether variation around a core is structured so that humans harbor only a few general types of microbiota profiles in a given body site: this is well established for the vagina[44] but more controversial in the gut[45]. In general, extreme caution must be applied when performing clustering procedures, as many will break up continuous variation into clusters where none exist[21]. Robust model selection procedures that incorporate the possibility that only continuous variation, not discrete clusters, exist remain to be developed within the context of microbial community analysis.

There is increasing evidence that individuals actually share a “core microbiome” rather than “core microbiota”. In a study of monozygotic and dizygotic twin pairs concordant for obesity or leanness, a subset of identifiable microbial genes, but not species, were shared between all individuals[15]. Remarkably, vastly different sets of microbial species yielded very similar functional KEGG pathways. However, deviations from this core microbiome were apparent in obese subjects, suggesting that it will be important to utilize metagenomic data in addition to determining microbial community composition with 16S marker gene studies when assessing differences between disease states. Understanding whether this principle holds true for other body sites will be fascinating; cross-biome metagenomic comparisons have been exceedingly rare to date[46, 47].

## **The gut microbiome plays an important role in digestion and nutrition**

The evidence is mounting for the inextricable link between a host's microbiota, digestion, and metabolism. In an analysis of humans and 59 additional mammalian species, 16S rRNA sequences clustered together carnivores, omnivores, and herbivores in principal coordinate spacing, showing that community structures differ depending on diets[48]. Dietary changes in mice can also lead to significant changes in bacterial metabolism, especially small chain fatty acids and amino acids, in as little as one week[49], and can lead to large changes after only one day[50]. Importantly, the genetic diversity found within our gut microbiota allows us to digest compounds via metabolic pathways not explicitly coded for in the mammalian genome, greatly increasing our ability to extract energy from our diverse diets[51, 52].

Gut microbiota also seem to play an important role in obesity. Germ-free mice that receive a transplant of gut microbiota from conventional mice have an increase in adiposity without



increasing food intake due to increased energy extraction from the diet and increased energy deposition into host adipocytes[53]. The two major microbial divisions, Firmicutes and Bacteroidetes, show different abundances depending on phenotype. Decreased Bacteroidetes and increased Firmicutes have been found in genetically obese mice (*ob/ob*) when compared to their lean counterparts[54], and the obesity phenotype can even be transferred to a germ-free but genetically wild-type mouse by way of the microbiota, and the phenotype is due to energy balance: bomb calorimetry of the fecal pellets reveal that the *ob/ob* mice extract more energy from their diet, and leave less behind in the feces[51]. Fascinatingly, the same effects hold true for another mouse model, the TLR5 knockout mice, which also become obese in some mouse facilities (but develop colitis in others, presumably due to differences in the background microbiota). The TLR5 knockout mice also produce a transmissible obesity phenotype, but no difference in the efficiency of energy harvest is involved. Instead, the altered microbiota somehow makes the mice hungrier, and their microbe-induced obesity can be cured by restricting the amount of food in their cages to that consumed by wild-type mice, as well as by antibiotics[55]. The correlation between microbes and obesity is perhaps best illustrated through weight loss. As different groups of human subjects were placed on either a fat-restricted or carbohydrate-restricted diet, their abundance of Bacteroidetes increased as their body weight decreased, transitioning from the signature 'obese' microbial community to a 'lean' community[56]. Thus, the modulation of a patient's microbiota might be a therapeutic option for promoting weight loss in obese patients or promoting weight gain in underweight children.

Surprisingly, the microbes that we ingest with our food might be providing our individual microbiome with new genes to digest new foods. Hehemann et al. found that a new class of glycoside hydrolases used to digest porphyran, a polysaccharide common in red algae, was also found in human stool samples as a gene in *Bacteriodes plebeius*. A closer examination of the stool metadata revealed that the stool samples containing the porphyran-digesting gene were only present in Japanese individuals; the gene was not found in the gut microbiome of the individuals of the United States. Why would a marine gene be found in human gut? The authors concluded that the seaweed common to the Japanese, but not American, diet contained the microorganism which transferred the genes to gut microbiome[57]. Thus, microbes have the ability to greatly increase the number of metabolic tools of the human gut, allowing us to digest an array of substrates.

## Plasticity of the Human Gut

Given the relative stability of the human gut microbiota, one key question is whether it is sufficiently plastic to allow well-defined interventions to improve health. As described above, the gut microbiota is fairly stable over time once established, at least compared to the differences between individuals. However, a number of studies demonstrate that external forces can alter the community of microbes located in the GI tract and antibiotics are an important example.

Antibiotics are mainly used to combat pathogenic bacterial species that reside within or have invaded a host, however the current generation of antibiotics are broad spectrum and target broad swaths of the normal microbiota as well. Thus, antibiotics significantly affect the host's innate gut microbiota. Three to four days after treatment with the broad-spectrum antibiotic ciprofloxacin the gut microbiota experience a decrease in taxonomic richness, diversity, and evenness[58, 59]. The large magnitude of changes in the gut microbiota demonstrated significant interpersonal variability. While the gut microbiota began to resemble its pre-treatment state a week after treatment, differences between individuals were seen with regards to how closely the post-treatment community resembled the pre-treatment community, and some taxa failed to return to the community[59, 60]. Indeed, the

reestablishment of some species can be affected for up to four years following antibiotic treatment[61]. Yet the overall recovery of the gut microbiota following antibiotic treatments suggests that there are factors within the community, biotic or abiotic, that promote community resilience, although these have yet to be elucidated.

Other antibiotics also tend to produce results that differ substantially between subjects[62, 63] and even body sites[64]. Because larger populations have not yet been studied, in part due to ethical issues with administration of antibiotics to healthy human subjects, the basis for these underlying differences has not yet been elucidated. Understanding the factors that determine the ability of a microbiota to resist and recover from perturbation, as well as understanding the factors that determine its current state, will be key to developing tools to assist in microbiome manipulation. For example, counter-intuitively, in rats the administration of antibiotics prior to cecal transplant actually reduces the chance that new microbes will establish[65].

One fascinating hint that the microbiota may be more plastic than imagined is the recent success of treatment of persistent *Clostridium difficile* infections via stool transplant, which has been successful in a number of studies[66-72], and in general the depauperate gut community produced during the *C. difficile* infection is replaced by the donor community[67, 73]. The success of this technique is remarkable, especially considering how little is known about the best community to supply. For example, is it better to receive the fecal community of a close relative or of a cohabiting individual, or perhaps to bank one's own stool before beginning antibiotic treatment so that it can be restored later? Is the same stool good for everyone, or do the vast differences in the microbiota imply that each person's microbes are specifically adapted relative to those they might receive from a donor? As with blood types, are there "universal donors" and "universal recipients"? These and many other questions remain to be answered.

## Conclusions and prospectus

As in every year since the initial sequencing of DNA, this year has resulted in an unprecedented growth in the amount of sequence data collected at an unprecedentedly low cost. Increasingly powerful tools used to extract meaningful patterns from this wealth of data have been developed or updated as well. Emerging technologies such as stool transplantation, 16S rRNA and whole-genome sequencing on the Illumina platform, the ability to transplant human microbial communities into mice with high efficiency even from frozen samples[50], and the creation of personalized culture collections[74] raises the prospect of a future in which therapies for individual humans are piloted in a battery of mice that are subjected to different treatments, and where leave-one-out experiments that reveal the effects of the deletion of individual species[74] or individual genes from within a species[75] allow insight into mechanism. Although the tools we have available are still imperfect (for example, the limited read length of today's high-throughput sequencing technologies limit the ability to detect bacterial species and strains, and analyses of viruses and eukaryotes are still very much an emerging frontier), the prospects for developing a mechanistic understanding of the factors that underlie the plasticity of the microbiome and then for manipulating the microbiome to improve health seem increasingly bright.

## Acknowledgments

Work cited in this review from the author's lab was supported in part by grants from the National Institutes of Health [HG4872 (RK)], the Crohn's and Colitis Foundation of America, and the Howard Hughes Medical Institute [RK].

## References

1. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. *Nature*. 2007; 449:804–810. [PubMed: 17943116]
2. Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, Deal C, et al. The NIH Human Microbiome Project. *Genome Research*. 2009; 19:2317–2323. [PubMed: 19819907]
3. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*. 2010; 464:59–65. [PubMed: 20203603]
4. Lederberg J, McCray A. Ome sweet 'omics: -- A genealogical treasury of words. *The Scientist*. 2001; 15:8.
5. Dobell C. The discovery of intestinal protozoa in man. *Proceedings of the Royal Society of Medicine*. 1920; 13:1–15. [PubMed: 19981292]
6. van Leeuwenhoek A. An abstract of a Letter from Antonie van Leeuwenhoek, Sep. 12, 1683. About Animals in the scrurf of the Teeth. *Philosophical Transactions of the Royal Society of London*. 1684; 14:568–574.
7. Pace NR. A molecular view of microbial diversity and the biosphere. *Science*. 1997; 276:734–740. [PubMed: 9115194]
8. Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nature Reviews Microbiology*. 2008; 6:776–788.
9. Lozupone CA, Knight R. Global patterns in bacterial diversity. *Proc Natl Acad Sci U S A*. 2007; 104:11436–11440. [PubMed: 17592124]
10. Marchesi JR. Prokaryotic and Eukaryotic Diversity of the Human Gut. *Advances in Applied Microbiology*, Vol 72. 2010; 72:43–62.
11. Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, Mahaffy JM, Mueller J, Nulton J, Rayhawk S, et al. Viral diversity and dynamics in an infant gut. *Research In Microbiology*. 2008; 159:367–373. [PubMed: 18541415]
12. Consortium IHGS. Finishing the euchromatic sequence of the human genome. *Nature*. 2004; 431:931–945. [PubMed: 15496913]
13. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature*. 2008; 452:872–876. [PubMed: 18421352]
14. Fierer N, Hamady M, Lauber CL, Knight R. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105:17994–17999. [PubMed: 19004758]
15. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, et al. A core gut microbiome in obese and lean twins. *Nature*. 2009; 457:480–484. [PubMed: 19043404]
16. Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC, Gevers D, Knight R. Experimental and analytical tools for studying the human microbiome. *Nature reviews Genetics*. 2012; 13:47–58.
17. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*. 2009; 75:7537–7541. [PubMed: 19801464]
18. Hartman AL, Riddle S, McPhillips T, Ludascher B, Eisen JA. Introducing W.A.T.E.R.S.: a workflow for the alignment, taxonomy, and ecology of ribosomal sequences. *BMC bioinformatics*. 2010; 11:317. [PubMed: 20540779]
19. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic acids research*. 2009; 37:D141–145. [PubMed: 19004872]



20. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, et al. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*. 2010; 7:335–336. [PubMed: 20383131]
21. Kuczynski J, Liu Z, Lozupone C, McDonald D, Fierer N, Knight R. Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. *Nat Methods*. 2010; 7:813–819. [PubMed: 20818378]
22. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and environmental microbiology*. 2005; 71:8228–8235. [PubMed: 16332807]
23. Kuczynski J, Costello EK, Nemergut DR, Zaneveld J, Lauber CL, Knights D, Koren O, Fierer N, Kelley ST, Ley RE, et al. Direct sequencing of the human microbiome readily reveals community differences. *Genome Biology*. 2010; 11
24. Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A, Fontana L, Henrissat B, Knight R, Gordon JI. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science*. 2011; 332:970–974. [PubMed: 21596990]
25. Knights D, Costello EK, Knight R. Supervised classification of human microbiota. *FEMS microbiology reviews*. 2011; 35:343–359. [PubMed: 21039646]
26. White JR, Nagarajan N, Pop M. Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS computational biology*. 2009; 5:e1000352. [PubMed: 19360128]
27. Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G, et al. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIXS) specifications. *Nature biotechnology*. 2011; 29:415–420.
28. Yilmaz P, Gilbert JA, Knight R, Amaral-Zettler L, Karsch-Mizrachi I, Cochrane G, Nakamura Y, Sansone SA, Glockner FO, Field D. The genomic standards consortium: bringing standards to life for microbial ecology. *The ISME journal*. 2011
29. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, et al. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*. 2008; 9:386. [PubMed: 18803844]
30. Mackie RI, Sghir A, Gaskins HR. Developmental microbial ecology of the neonatal gastrointestinal tract. *Am J Clin Nutr*. 1999; 69:1035S–1045S. [PubMed: 10232646]
31. Dominguez-Bello MG, Blaser MJ, Ley RE, Knight R. Development of the human gastrointestinal microbiota and insights from high-throughput sequencing. *Gastroenterology*. 2011; 140:1713–1719. [PubMed: 21530737]
32. Dominguez-Bello MG, Costello EK, 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. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:11971–11975. [PubMed: 20566857]
33. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. Development of the human infant intestinal microbiota. *PLoS Biol*. 2007; 5:e177. [PubMed: 17594176]
34. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE. Succession of microbial consortia in the developing infant gut microbiome. *Proceedings of the National Academy of Science, USA*. 2010
35. Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R. Forensic identification using skin bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:6477–6481. [PubMed: 20231444]
36. Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, Knights D, Gajer P, Ravel J, Fierer N, et al. Moving pictures of the human microbiome. *Genome Biology*. 2011; 12:R50. [PubMed: 21624126]
37. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial Community Variation in Human Body Habitats Across Space and Time. *Science*. 2009; 326:1694–1697. [PubMed: 19892944]

38. Kunin V, Engelbrekton A, Ochman H, Hugenholtz P. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology*. 2010; 12:118–123. [PubMed: 19725865]
39. Quince C, Lanzen A, Curtis TP, Davenport RJ, Hall N, Head IM, Read LF, Sloan WT. Accurate determination of microbial diversity from 454 pyrosequencing data. *Nature Methods*. 2009; 6:639–641. [PubMed: 19668203]
40. Sun Y, Cai Y, Huse SM, Knight R, Farmerie WG, Wang X, Mai V. A large-scale benchmark study of existing algorithms for taxonomy-independent microbial community analysis. *Briefings in bioinformatics*. 2011
41. Schloss PD, Westcott SL. Assessing and Improving Methods Used in Operational Taxonomic Unit-Based Approaches for 16S rRNA Gene Sequence Analysis. *Applied and Environmental Microbiology*. 2011; 77:3219–3226. [PubMed: 21421784]
42. Huse SM, Welch DM, Morrison HG, Sogin ML. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology*. 2010; 12:1889–1898. [PubMed: 20236171]
43. Sun Y, Cai Y, Liu L, Yu F, Farrell ML, McKendree W, Farmerie W. ESPRIT: estimating species richness using large collections of 16S rRNA pyrosequences. *Nucleic acids research*. 2009; 37:e76. [PubMed: 19417062]
44. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, et al. Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(1):4680–4687. [PubMed: 20534435]
45. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, et al. Enterotypes of the human gut microbiome. *Nature*. 2011; 473:174–180. [PubMed: 21508958]
46. Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, Furlan M, Desnues C, Haynes M, Li L, et al. Functional metagenomic profiling of nine biomes. *Nature*. 2008; 452:629–632. [PubMed: 18337718]
47. Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC, et al. Comparative metagenomics of microbial communities. *Science*. 2005; 308:554–557. [PubMed: 15845853]
48. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI. Evolution of mammals and their gut microbes. *Science*. 2008; 320:1647–1651. [PubMed: 18497261]
49. Martin FP, Sprenger N, Montoliu I, Rezzi S, Kochhar S, Nicholson JK. Dietary modulation of gut functional ecology studied by fecal metabonomics. *Journal of proteome research*. 2010; 9:5284–5295. [PubMed: 20806900]
50. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Science translational medicine*. 2009; 1:6ra14.
51. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*. 2006; 444:1027–1031. [PubMed: 17183312]
52. Gilmore MS, Ferretti JJ. *Microbiology: The thin line between gut commensal and pathogen*. *Science*. 2003; 299:1999–+. [PubMed: 12663906]
53. Backhed F, Ding H, Wang T, Hooper LV, Koh GY, Nagy A, Semenkovich CF, Gordon JI. The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S A*. 2004; 101:15718–15723. [PubMed: 15505215]
54. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:11070–11075. [PubMed: 16033867]
55. Vijay-Kumar M, Aitken JD, Carvalho FA, Cullender TC, Mwangi S, Srinivasan S, Sitaraman SV, Knight R, Ley RE, Gewirtz AT. Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science*. 2010; 328:228–231. [PubMed: 20203013]

56. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. *Nature*. 2006; 444:1022–1023. [PubMed: 17183309]
57. Hehemann J, Correc G, Barbeyron T, Helbert W, Czjzek M, Michel G. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature*. 2010; 464:908–912. [PubMed: 20376150]
58. Dethlefsen L, Huse S, Sogin ML, Relman DA. The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S rRNA Sequencing. *Plos Biology*. 2008; 6:2383–2400.
59. Dethlefsen L, Relman DA. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(1):4554–4561. [PubMed: 20847294]
60. Dethlefsen L, Huse S, Sogin ML, Relman DA. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *Plos Biology*. 2008; 6:e280. [PubMed: 19018661]
61. Jernberg C, Lofmark S, Edlund C, Jansson JK. Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *The ISME journal*. 2007; 1:56–66. [PubMed: 18043614]
62. Jernberg C, Lofmark S, Edlund C, Jansson JK. Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *Isme Journal*. 2007; 1:56–66. [PubMed: 18043614]
63. Jernberg C, Sullivan A, Edlund C, Jansson JK. Monitoring of antibiotic-induced alterations in the human intestinal microflora and detection of probiotic strains by use of terminal restriction fragment length polymorphism. *Applied and Environmental Microbiology*. 2005; 71:501–506. [PubMed: 15640226]
64. Jakobsson HE, Jernberg C, Andersson AF, Sjolund-Karlsson M, Jansson JK, Engstrand L. Short-Term Antibiotic Treatment Has Differing Long-Term Impacts on the Human Throat and Gut Microbiome. *PLoS ONE*. 2010; 5
65. Manichanh C, Reeder J, Gibert P, Varela E, Llopis M, Antolin M, Guigo R, Knight R, Guarner F. Reshaping the gut microbiome with bacterial transplantation and antibiotic intake. *Genome Research*. 2010; 20:1411–1419. [PubMed: 20736229]
66. Mellow MH, Kanatzar A. Colonoscopic fecal bacteriotherapy in the treatment of recurrent clostridium difficile infection--results and follow-up. *The Journal of the Oklahoma State Medical Association*. 2011; 104:89–91. [PubMed: 21608450]
67. Grehan MJ, Borody TJ, Leis SM, Campbell J, Mitchell H, Wettstein A. Durable alteration of the colonic microbiota by the administration of donor fecal flora. *Journal of Clinical Gastroenterology*. 2010; 44:551–561. [PubMed: 20716985]
68. Garborg K, Waagsbo B, Stallemo A, Matre J, Sundoy A. Results of faecal donor instillation therapy for recurrent Clostridium difficile-associated diarrhoea. *Scandinavian journal of infectious diseases*. 2010; 42:857–861. [PubMed: 20662620]
69. Silverman MS, Davis I, Pillai DR. Success of self-administered home fecal transplantation for chronic Clostridium difficile infection. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association*. 2010; 8:471–473. [PubMed: 20117243]
70. MacConnachie AA, Fox R, Kennedy DR, Seaton RA. Faecal transplant for recurrent Clostridium difficile-associated diarrhoea: a UK case series. *QJM : monthly journal of the Association of Physicians*. 2009; 102:781–784. [PubMed: 19726581]
71. You DM, Franzos MA, Holman RP. Successful treatment of fulminant Clostridium difficile infection with fecal bacteriotherapy. *Annals of internal medicine*. 2008; 148:632–633. [PubMed: 18413636]
72. Tvede M, Rask-Madsen J. Bacteriotherapy for chronic relapsing Clostridium difficile diarrhoea in six patients. *Lancet*. 1989; 1:1156–1160. [PubMed: 2566734]
73. Khoruts A, Dicksved J, Jansson JK, Sadowsky MJ. Changes in the composition of the human fecal microbiome after bacteriotherapy for recurrent Clostridium difficile-associated diarrhea. *Journal of clinical gastroenterology*. 2010; 44:354–360. [PubMed: 20048681]

74. Faith JJ, Rey FE, O'Donnell D, Karlsson M, McNulty NP, Kallstrom G, Goodman AL, Gordon JI. Creating and characterizing communities of human gut microbes in gnotobiotic mice. *The ISME journal*. 2010; 4:1094–1098. [PubMed: 20664551]
75. Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, Knight R, Gordon JI. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host & Microbe*. 2009; 6:279–289. [PubMed: 19748469]