Role of the intestinal microbiome in health and disease: from correlation to causation

Willem M de Vos and Elisabeth AJ de Vos

Recorded observations indicating an association between intestinal microbes and health are long-standing in terms of specific diseases, but emerging high-throughput technologies that characterize microbial communities in the intestinal tract are suggesting new roles for the supposedly normal microbiome. This review considers the nature of the evidence supporting a relationship between the microbiota and the predisposition to disease as associative, correlative, or causal. Altogether, indirect or associative support currently dominates the evidence base, which now suggests that the intestinal microbiome can be linked to a growing number of over 25 diseases or syndromes. While only a handful of cause-and-effect studies have been performed, this form of evidence is increasing. The results of such studies are expected to be useful in monitoring disease development, in providing a basis for personalized treatments, and in indicating future therapeutic avenues.

INTRODUCTION

Virtually every day we are all confronted with the activity of our intestine, and it is no surprise that at least some of us have developed a fascination for our intestinal condition and its relation to health and disease. Following his discovery of microbial life, Antonie van Leeuwenhoek reported in 1681 the first observation relating a disturbed microbial composition to the diarrhea he experienced, possibly after drinking polluted Amsterdam canal water.¹ The account that his watery excrements contained more and different “little animals,” as he called the bacteria, is largely correct. It reflects the association between some forms of diarrhea and a sudden shedding of intestinal microbes, including those that inhabit the mucosa. As we now know, the mucosal microbial communities may differ in composition and abundance from those present in the colon.² This hallmark discovery of several centuries ago was not followed up by further well-documented work until the end of the last century, when interest in the intestinal microbes experienced a real renaissance.³ It led to the notion that intestinal microbes can be considered a personalized human organ with a metabolic activity second only to that of the liver.⁴ It also resulted in a change in terminology: what was first known as “microflora” – a term still found in some publications and medical textbooks – has now been renamed “microbiota” on the basis of the diversity of microorganisms revealed mainly by microbial ecologists, who used molecular systematics and positioned microbes in ancestral evolutionary terms far away from plants.⁵ With the implementation of genomics-based approaches, the exploration of the human intestinal microbiome, defined here as all microbiota in the intestinal tract, has begun. The collective genomes within the microbiome have been found to contain more than 3 million unique genes.⁶

It is known that the intestinal microbiota shows a specific spatial organization,² but as the vast majority of microbes are found in the colon, virtually all present-day studies focus on the microbiome that is recovered from fecal samples. An understanding of the human intestinal microbiome is now rapidly developing, and in many cases the relationship between the microbiome and health and disease is being explored. In most instances, however, this
has simply meant an analysis of associations with disease or functional disturbances, and only in special cases are specific correlations described in which specific microbial groups relate to a healthy or a diseased state in a manner that implies a linear relationship. Finally, there are only a handful of examples in which the cause-and-effect relations satisfying Koch’s postulates apply, but even these relate mainly to studies in animal models, thereby providing hypotheses for human disease and human intervention tests. While one may argue that probiotic interventions can be seen as providing causal evidence for their roles, these interventions involve the use of specific bacteria that may have a direct effect on the host rather than on the intestinal microbiota per se. Probiotics are therefore not considered here, particularly as they have been reviewed exhaustively elsewhere.

Following a short overview of the present knowledge of the human microbiota and the available high-throughput analytical approaches along with their promise and pitfalls, the most noteworthy studies that relate to the human microbiome in health and its predisposition to disease are summarized. Where possible, evidence suggesting simply correlations is differentiated from that suggesting actual causal relationships. It is expected that this information will 1) provide a conceptual basis for how the human intestinal microbiome affects disease, 2) contribute to the development of techniques for monitoring disease development, 3) provide a basis for personalized treatment, and 4) indicate future therapeutic avenues.

A main driver for the increased understanding of the intestinal microbiome has been the development of molecular and high-throughput tools that obviated the need for culturing and permitted the analysis of microbial function. The application of these tools reinforced the conclusions of a decade ago that humans are colonized from birth by a developing intestinal microbiota that, in adult life, is highly individual, temporally stable, and similar in monozygotic twins and other genetically related subjects. Moreover, metagenomic developments with next-generation technology (NGT) sequencing approaches have now provided a catalog of over 3 million genes, which, in terms of the average microbiota composition, are derived mainly from prokaryotic Bacteria and, to a lesser extent, Archaea, with only a few fungal genes encountered. This and other analyses based on quantitative analysis showed the human intestinal communities to be highly complex, predicted to contain more than 1,000 different prokaryotic species belonging to a limited set of a dozen taxa and dominated by gram-positive anaerobes (Figure 1). Further computational analysis has led to the notion that the apparently diverse microbial communities can be grouped into three so-called enterotypes, consisting of networks between different microbial groups that are robust and evident in subjects from different continents. It has been suggested that these enterotypes may well affect the response of

---

**Figure 1** Phylogenetic tree of the most important microbial taxa in the human intestinal tract, along with their relative contributions.
subjects to dietary and pharmaceutical interventions, and hence it is of interest that a biased distribution was observed in irritable bowel syndrome (IBS) patients and that the enterotypes were reportedly affected by diet.

The intestinal microbial communities are highly complex; therefore, the analytical coverage of this complexity, its reproducibility, and its accuracy are important factors that determine the quality of the analytical assessment. Yet, despite the rapid developments of NGT sequencing systems, deep metagenomic sequence analysis is still time-consuming, costly, and rather challenging from a bioinformatics and data storage point of view. Hence, most studies addressing the intestinal Bacteria and Archaea have focused on the 1.5-kb bacterial 16S rRNA gene sequences that are well-established phylogenetic markers. There is already a growing number of over a million entries in accessible databases. Only a fraction derive from the human intestinal microbiota, and these have been curated some years ago, leading to the conclusion that only approximately 400 species have been cultured from human intestine. Since only microbes that have been cultured can be characterized taxonomically, this leads to the sometimes difficult situation that a specific 16S rRNA sequence or signature is now found to be associated with disease or a healthy state, but this rRNA sequence cannot be assigned to a species. These may then be termed species-level phylotypes or operational taxonomic units that are identified only by their complete 16S rRNA sequence and usually defined as sharing 97% or less sequence identity with other entries in the ribosomal databases. Applying this strict criterion has led to a database of around 1,200 different phylotypes that provide a systematic framework of the microbial diversity in the human gastrointestinal tract. This information is of great value in defining a healthy microbiota and comparing it with that of diseased subjects. Moreover, this database has been instrumental in the design of high-throughput approaches such as the Human Intestinal Tract Chip (HITChip), a phylogenetic DNA microarray for the comprehensive analysis of gastrointestinal tract microbiota at multiple levels of taxonomic resolution.

A wide range of high-throughput approaches, mainly based on microarray hybridization, polymerase chain reaction (PCR), and NGT sequence analysis as well as combinations thereof, have been applied successfully to monitor the human intestinal microbiota, and these have been reviewed extensively. While instrumental in providing deep insight, they all suffer from inherent biases that vary from sequence errors, i.e., PCR chimeras or cloning artifacts, to cross-hybridization. However, several of these approaches have been compared to each other, and, overall, highly similar results were obtained by HITChip analysis and NGT sequencing of diagnostic regions of 16S rRNA amplicons. A discriminating factor has been the depth of the analysis that can be seen in terms of the rRNA sequences that can be reliably quantified. Phylogenetic microarrays such as the HITChip may quantify the microbiota in a highly reproducible way, representing a depth of $10^4$ to $10^5$, which is comparable to over 200,000 reads (approximately 100 Mb of sequence information) on an NGT sequencer. Since the fraction of single microbial groups may vary 100- to 1,000-fold, it is critical to have a detection limit as low as possible to obtain good discrimination in the microbiota analysis.

This is exemplified in a simple experiment in which the microbiotas of 65 healthy subjects of different nationalities were analyzed using the HITChip (Figure 2). In a single subject, around 900 phylotype-like species could be detected, which is slightly higher than the number found by ultra-deep metagenomic analysis, which emphasizes the depth of the HITChip analysis. Similar studies have been reported with different-sized groups of subjects, and shared phylotype-like sequences were identified as what is known as the "common core." It is evident from the displayed plot of sequences that this common core is dependent not only on the number of subjects but, notably, on the depth of the analysis (Figure 2A). In these healthy subjects, a common core of over 450 species-like taxa could be defined. Detailed analysis of this common core showed it to consist of a series of well-known species or genera, including those belonging to Bifidobacterium, Clostridium, Collinsella, Dorea, Eubacterium, Parabacteroides, Prevotella, Ruminococcus, and Streptococcus spp. (Figure 2B). The vast majority of this common core (a total of 387, representing 85%), however, included microbes that have not yet been cultured and hence are phylotypes that show similarity to 16S rRNA sequences from bacteria that have not yet been cultured.

In spite of the many technological advances in the last decade, there are still considerable challenges that call for caution. Several problems are related to the inappropriate use of technologies, the interpretation of results, or other systematic errors. DNA isolation and PCR amplification seem trivial but have been shown to provide biased views that may explain retrospectively the early observations that adult samples lack Bacteroides spp. or that samples from babies do not contain bifidobacteria. Moreover, the interpretation of 16S rRNA sequence information should be based on the appropriate knowledge and use of this phylogenetic biomarker. A recent careful analysis of mock mixtures of microbial DNA by NGT sequencing methods revealed a large range of inaccuracies and indicated that many new taxa are incorrectly identified due to chimer formation. This can now be avoided by new algorithms, but these are still imperfect, and hence low-abundance taxa should be treated with caution and as potential artifacts. Another issue is that many prokaryotes have multiple copies of 16S rRNA in their genomes.
This should be taken into consideration when addressing quantitative effects. Moreover, in some cases, significant sequence heterogeneity exists in the copies of the 16S rRNA genes, such as in *Bifidobacterium adolescentis*, an important inhabitant of the adult colon. Similarly, metagenomic studies also suffer from biases, albeit of a different kind, such as inaccurate quantifications of bacterial populations, limitations of databases or computing capacity, and incorrect assembly of short reads or repeated sequences, including rRNA operons. However, the most important factors that may explain most of the differences in the present literature include the limited number of samples analyzed, the great variety of analysis platforms, and the differences between experimental procedures. While this emphasizes the need for standardization, it also implies that caution should be used when comparing various studies unless rigorous tests for robustness and reproducibility have been performed.

**MICROBIOTA IN HEALTH: COMPOSITION, ACTIVITY, AND STABILITY**

When addressing the intestinal microbiota in disease, it is essential to know the baseline in healthy subjects. Defining health, however, is much more complex than defining disease, as has been recently emphasized. The World Health Organization (WHO) has defined health as a state of complete physical, mental, and social well-being and not merely as the absence of disease or infirmity. It is of interest to note that, in addition to the physical state, the mental and social aspects of well-being are also included. These are of relevance to the intestinal tract, which is composed of a single layer of epithelial cells surrounded by the enteric nervous system, the largest reservoir of nerve cells in the human body apart from that in the brain. Hence, the intestinal tract is part of the brain-gut axis and has also been recognized as the second brain. In recent years, it has become evident that the intestinal microbes communicate with the epithelial cells, as shown by the extension of early observations in monoassociated mouse models to new studies in healthy human volunteers. It has also now emerged that specific intestinal microbes also interact with the enteric nervous system. So far, this has only been reported for animal models, but the impact on animal behavior and anxiety is striking. It is therefore appropriate to consider the intestinal microbiota in the context of the WHO definition of health by including mental and social aspects. The recent conclu-
sion from experimental animals that their social behavior changes in relation to the microbiota also allows new experimental approaches to test the nature of microbiota/behavioral relationships.35

In operational terms, it is not easy to apply the complete WHO definition of health. In most cases, comparative studies are performed with control groups of healthy subjects who have been selected on the basis of the absence of disease and infirmity. In some cases, however, extensive and validated questionnaires have been used that address the quality of life (QoL) and incorporate the physical, social, and mental aspects in some way. These are particularly useful in differentiating healthy subjects from those who suffer from IBS, a frequently occurring aberration that involves the brain-gut axis.36 In the data presented above (Figure 2), QoL questionnaires were administered to the subjects, so the common core defined in this group may represent a healthy microbiota. This work was recently extended in a similar analysis of over 100 healthy subjects using the same HITChip platform.37 When the same approach was applied to the intestinal microbiota of patients with ulcerative colitis (UC), however, one of the forms of inflammatory bowel disease (IBD; see below), a much smaller general core was found. Moreover, the common core found in UC patients was markedly different from that in the healthy subjects, illustrating the power of this comparative approach.38 Apart from detecting compositional differences in health and disease, the diversity of the intestinal microbiota is also often addressed. Diversity is here defined in the ecological terms of species richness and evenness, reflecting the number of phylotypes and their relative abundance. This is particularly important because there is a strong correlation between species diversity and resilience in many ecosystems, and there is no reason to assume this is different for the healthy intestinal microbiota. While the microbial diversity increases rapidly in early life, it has been found to stabilize during adulthood and is maintained stably throughout all later phases, though it may decrease slightly in subjects over 100 years of age.37 In several diseases, there is a marked effect on the microbial diversity, the most prominent being that observed in patients with recurrent *Clostridium difficile* infection (CDI).39

Apart from the microbial composition, the activity of the intestinal microbiota is a major contributing factor to health and disease. Various ways to define the activity of the intestinal microbiota have been described.39 By using cell sorting combined with specific dyes, it was found that fecal samples on average contain approximately 30% dead and 20% injured cells that show a nonstochastic phylogenetic distribution, probably because some bacterial groups are more easily damaged than others.40 More global approaches capitalize on functional metagenomics, such as metatranscriptomics, metaproteomics, and metabolomics.7 While metatranscriptomics is a high-throughput method that exploits NGT sequence analysis, the recovery of mRNA from the intestinal tract is a great challenge because the half-life of mRNA in prokaryotes is in the order of minutes. So far, metatranscriptomic approaches have been successfully applied to intestinal systems with a sufficiently high flux that permits rapid sampling and processing. These include the intestinal tracts of babies and ileostoma patients, in whom vitamin production and sugar metabolism by bifidobacteria and streptococci, respectively, were found to be among the most abundant functions.41,42 Metaproteomics capitalizes on the fast and global mass spectrometry analysis of proteins that are generally much more stable than transcripts. While exploiting the rapidly growing metagenome databases, metaproteomics has developed into an established tool to assess microbial function in the complex ecosystem of the intestinal tract.43–46 The potential of a new metaproteomics annotation approach has been illustrated by revealing the in situ activity of mucus-degrading *Akkermansia* spp., a member of the Verrucomicrobia (Figure 1), in healthy volunteers.44 Last, but not least, there is the metabolomics approach, which is a powerful tool that has been used in a great variety of studies addressing the impact of the intestinal microbiota on human health.47 Notably, urine and blood metabolomics provided new insight into microbiota function, leading to the recent discovery of the involvement of intestinal microbes in promoting suitable diets for patients with atherosclerosis.48 One of the limitations of the present-day metabolomics, however, lies in identifying the observed metabolites and, in some cases, reliably determining their concentration.

While these functional metagenomic tools will be instrumental in analyzing the relation between microbiota function and health, they have not yet been applied in large-scale studies comparing healthy and diseased subjects. Hence, in the section below, the main focus is on global microbiota analysis based on high-throughput approaches. It is expected, however, that in the near future such studies will be paralleled by functional approaches that expand beyond only the composition of the microbiota.

Apart from the composition and function of the microbiota, a third factor needs to be addressed, and that is time. The temporal variation of the microbiota composition in healthy subjects has been addressed in various time windows, varying from weeks to months to years.17,20,49 In all cases, a high temporal stability was observed that resulted in the maintenance of a recognizable individual microbiota composition for periods of over 10 years. However, the immediate effects of antibiotic use were observed and affected the temporal
stability, confirming model experiments with a small number of volunteers. Moreover, in a weekly follow-up study of QoL-controlled healthy subjects, it was observed that traveling across time zones may affect the temporal stability of the microbiota. In addition, by linking intestinal health to the microbiota, it could be established that abdominal pain was inversely correlated with the amounts of bifidobacteria. These unexpected observations testify to the power of these global approaches and provide a basis for further prospective studies to establish cause-and-effect relationships. Moreover, they underline the need to incorporate time as an additional factor to take account of the temporal stability of the intestinal microbiota.

**STRONG ASSOCIATIONS OF MICROBIOTA AND DISEASE: CAUSE-AND-EFFECT STUDIES**

In recent years, associations with varying degrees of support have been established between human intestinal microbiota and an increasing number of over 25 diseases, syndromes, or functional aberrations. The support for these associations can vary from anecdotal indications, such as those described below, to much firmer evidence obtained from large cohorts. Here, the focus is on 10 of the strongest associations that are supported by multiple studies (Table 1). Specific correlations between function or disease and intestinal microbes and, where possible, causation are also described and, in some cases, are supported by studies in animal models.

The majority of the studies relating microbiota and disease concern only a few aberrations that have a prominent effect on health; these include IBD, IBS, and CDI (Table 1). The two main IBD conditions, Crohn’s disease (CD) and UC, have been associated with genetic predispositions, and several dozens of host genes have been described, reflecting the complexity of these diseases. The intestinal microbiota associations in IBD have been studied by comparing healthy and compromised subjects. To correct for the genetic impact, however, monozygotic and dizygotic twins discordant for the disease are often studied. Both CD and UC are associated with a reduced diversity of the intestinal microbiota. There are, however, marked differences between the two diseases that reflect their very different nature. In CD, reduced numbers of *Faecalibacterium prausnitzii* have been repeatedly observed, and this anaerobic butyrate producer is reported to have anti-inflammatory properties in a mouse model. Similarly, in a comparative study of IBD-discordant twins, an increased level of *Faecalibacterium prausnitzii* was also found in the CD patients, but the microbiota of the UC patients were similar to those of their healthy twin siblings. This latter finding may be attributed to the limited depth of the analysis or to other technical factors, since other studies showed marked differences in the microbiota of UC patients. Moreover, in a recent study that specifically addressed the mucosal bacteria in IBD patients, the level of *Akkermansia muciniphila* was reported to be 10-fold reduced in CD patients and 100-fold reduced in UC patients, and it was

<table>
<thead>
<tr>
<th>Aberration</th>
<th>Most relevant observations and potential correlation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crohn’s disease</td>
<td>Diversity decrease – reduced <em>F. prausnitzii</em></td>
<td>Kaser et al. 2010(^51); Sokol et al. 2009(^52); Willing et al. 2010(^53)</td>
</tr>
<tr>
<td>Ulcerative colitis</td>
<td>Diversity decrease – reduced <em>A. muciniphila</em></td>
<td>Png et al. 2010(^54); Käser et al. 2010(^55); Lepage et al. 2011(^56)</td>
</tr>
<tr>
<td>Irritable bowel syndrome</td>
<td>Global signatures – increased <em>Dorea</em> and <em>Ruminococcus</em></td>
<td>Salonen et al. 2010(^57); Saulnier et al. 2011(^58); Rajilić-Stojanović et al. 2011(^59)</td>
</tr>
<tr>
<td><em>Clostridium difficile</em> infection</td>
<td>Strong diversity decrease – presence of <em>C. difficile</em></td>
<td>Grehan et al. 2010(^60); Khoruts et al. 2010(^61)</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Variation in <em>Bacteroides</em> spp. – increased fusobacteria</td>
<td>Sobhani et al. 2011(^62); Wang et al. 2012(^63); Marchesi et al. 2011(^64)</td>
</tr>
<tr>
<td>Allergy/atopy</td>
<td>Altered diversity – specific signatures</td>
<td>Stsepova et al. 2007(^65); Bisgaard et al. 2011(^66); Storro et al. 2011(^67)</td>
</tr>
<tr>
<td>Celiac disease</td>
<td>Altered composition, notably in small intestine</td>
<td>Nistal et al. 2012(^68); Di Cagno et al. 2011(^69); Kalliomäki et al. 2012(^70)</td>
</tr>
<tr>
<td>Type 1 diabetes</td>
<td>Signature differences</td>
<td>Vaarela 2011(^71); Giongo et al. 2011(^72); Brown et al. 2011(^73)</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>Signature differences</td>
<td>Larssen et al. 2010(^74); Wu et al. 2010(^75); Koote et al. 2012(^76)</td>
</tr>
<tr>
<td>Obesity</td>
<td>Specific bacterial ratios (<em>Bacteroidetes/Firmicutes</em>)</td>
<td>Ley et al. 2006(^77); Turnbaugh et al. 2009(^78); Musso et al. 2011(^79)</td>
</tr>
</tbody>
</table>
suggested that this bacterium could be a health biomarker. Akkermansia muciniphila is a mucus-degrading and propionate-producing bacterium belonging to the Verrucomicrobia (Figure 1) that not only stimulates the immune system but also the barrier function in a mouse model. Very recently, it was shown in a mouse model of IBD with human-relevant disease-susceptibility mutations that Koch’s postulates were fulfilled by common commensal Bacteroides spp. but not by members of the Enterobacteriaceae. Remarkably, the latter were > 100-fold enriched in the IBD model but were not associated with disease, so this elegant experiment stresses the need for cause-and-effect rather than association studies.

While IBD affects only a fraction of the population, IBS is a highly prevalent aberration that may affect over 5% of the population and can be differentiated into several different types that relate to bowel habits. Two recent studies in adults and children have produced a series of global signatures that differentiate IBS subjects from healthy individuals. In spite of the differences in series of global signatures that differentiate IBS subjects from healthy individuals. In spite of the differences in the fecal microbiota for up to 24 weeks was described. In a more systematic study, a dozen patients were monitored for up to 2 months, and a consistent pattern of increased diversity indistinguishable from the donor microbiota was observed in the cured patients (van Nood et al., unpublished data, 2012). These cases all represent clear examples of a cause-and-effect relationship in which a diverse donor microbiota is stably established in low-diversity recipients with CDI who were thereby cured of the disease. There are some single case reports of successful fecal transplantation in patients with UC and other IBD, even IBS. While these observations can also be considered as indicating that the altered intestinal microbiota in IBD and IBS are a cause rather than an effect, larger sets of fecal transplantation patients, criteria that define an optimal donor microbiota, and better descriptions of the medical conditions, the efficacy, and the changes in the microbiota are needed to support that conclusion.

Another important disease for which a series of recent studies support an association with the microbiota is colorectal cancer (CRC), a life-threatening disease that in some cases is linked to colitis, the so-called colitis-associated cancer (Table 1). It is assumed that dietary components, such as nitrate, which is a precursor for carcinogenic nitrosamines, can be converted into (pro)carcinogens by enzymes of the intestinal microbiota and hence can promote the onset of CRC, as has been recently reviewed. Testifying to the interest in this area is the very recent insight gained in a series of studies in which the fecal and, in some cases, the mucosal microbiota of CRC patients were compared with those of healthy subjects. In the fecal microbiota of over several hundreds of CRC patients and healthy controls, an increased level of bacteria belonging to the Bacteroides/Prevotella was observed in France using a quantitative PCR approach, while in a global study in China, a more complex pattern was observed that was characterized by a reduction of potentially pathogenic gram-negative bacteria in CRC patients, an unequal distribution of some Bacteroides spp., and a reduced number of butyrate-producing bacteria. More detailed differences were observed by comparing the microbiota on tumor biopsies with those of the neighboring healthy tissues, and this revealed a set of global differences. Similar approaches with fewer patients but more powerful genomic approaches revealed a variety of differences that systematically included an overrepresentation of Fusobacterium spp. in the tumor sites. A strain of Fusobacterium nucleatum was isolated, genomically characterized, and found to be invasive in a human cell line. Remarkably, close inspection of the earlier reported study with biopsies of CRC patients also showed an increase in fusobacteria in the tumor samples. Whether the fusobacteria really are involved in the onset of CRC, however, remains to be established, and it is well possible that these gram-positive bacteria, which are rather common intestinal colonizers, just prefer affected tissue. Inflamed appendices removed after appendicitis were also found to contain much more fusobacteria, and microscopic evidence of invasion of fusobacteria into the enterocytes was provided; moreover, the level of bacteria related to the mucus-degrading Akkermansia muciniphila was greatly reduced, as in IBD (see above).

Similarly, it was found that there was a relation between the level and invasiveness of Fusobacterium nucleatum and the severity of...
inflammation in mucosal samples of IBD patients. This supports the hypothesis that invasion precedes inflammation and, in some cases, like colitis-associated cancer, may lead to the development of cancer.

In several of the diseases described so far, notably IBD and CDI, the recurrent use of antibiotics and a high level of hygiene also have been implicated in their prevalence. This so-called “hygiene hypothesis” also applies to the development of allergy, which is often already developing in early life and manifested in various forms of atopy. In a series of comparisons of babies and young children with atopy compared with age-matched healthy controls, correlations were found between specific microbiota differences and the manifestation of allergy. In some cases, specific groups of bacteria, including Bifidobacterium spp., have been implicated, but the development of the diversity also appears to be involved. A complicating factor is that, in early life, both the infant and the microbiota are still developing and are subject to large variations, so their proper study will require substantial numbers in the cohorts. In a recent study, over 1,000 babies were monitored over time. It was reported that colonization by Clostridium difficile at the age of 1 month was associated with wheeze, eczema, and asthma at a later age. Although these cohort study results are promising, the analysis of the microbiota was performed by quantitative PCR. Analysis by the new all-encompassing techniques has not yet been performed, which thereby limits the significance of the study. Similar aspects of power apply to the analysis of celiac disease, a chronic inflammation of the mucosa in the small intestine that develops in genetically susceptible persons in response to dietary gluten, present in barley, wheat, and rye. A set of over 1,000 genes have been involved in celiac disease, testifying to its complexity. Another major factor, however, appears to be the intestinal microbiota, and it has been proposed that aberrations in early-life colonization result in inappropriate host immune responses that lead to enteric inflammation. There is, however, no clear picture yet of the link between the microbiota and celiac disease. A variety of associations between various different microbial groups have been made, especially in small intestinal biopsies. Furthermore, connections to the expression of ileal genes, notably those involved in immune regulation, have been made. A complicating factor is that the current knowledge of early-life colonization is still limited and not supported by large cohort studies. Moreover, with the involvement of host factors, the microbiota, and diet, celiac disease is particularly complex, as noted in several recent reviews.

Lifestyle diseases are a last group of highly prevalent aberrations that increasingly impact human life. These include diabetes, obesity, and metabolic syndrome. The associations between these diseases and the microbiota have been researched extensively following the first description in human obesity that linked this condition with microbial ecology (Table 1). Because of its potentially major significance and, possibly, because some of the initial observations could not be reproduced by other or the same authors, the relation between microbiota, diabetes, and obesity is one of the most extensively reviewed research domains. While both type 1 diabetes (T1D) and type 2 diabetes (T2D) are characterized by high levels of blood glucose, their mechanistic basis differs: T1D patients have a defect in the production of insulin, while T2D patients are generally insulin resistant. T1D is not a lifestyle disease but is considered to result from autoimmune destruction of the insulin-producing beta cells of the pancreas; currently, there are few reports on an association with the microbiome, and these deal with only a handful of patients. Hence, there is a need for much more robust analysis. Animal experiments, however, have provided clear indications for an important role of the microbiota that relates to its impact on immune signaling in T1D.

Similarly, elegant transplantation studies in mice showed a clear causal link between the microbiota and T2D, implying the involvement of Toll-like receptor 5 signaling. The relation between T2D and the microbiota has been further explored in human subjects; a series of microbiota signatures have been associated with T2D, but they differ considerably from study to study. A common theme is a low-grade inflammation that, in animal experiments, has been associated with high levels of lipopolysaccharide and an aberrant microbiota. T2D is linked to obesity, which is usually defined as a body mass index (BMI) of over 30 and can be easily assessed without complex phenotyping.

The first hallmark study addressing the association between the human microbiota and obesity suggested an inverse relationship between obesity and the Bacteroides/Firmicutes ratio and showed that this ratio increased following weight loss. This study was accompanied by a seminal experiment showing an increased energy-harvesting capacity of the microbiota of obese mice with a low Bacteroides/Firmicutes ratio, a property that can be transplanted to lean, germ-free mice, providing a causal relationship between microbiota and obesity in experimental animals. As both the Bacteroides/Firmicutes ratio and BMI can be easily measured, many studies with human subjects have subsequently attempted to address the relationship between these parameters. Often, however, a relationship could not be established, resulting in inconsistent results (see Diamant et al. for a recent extensive review). This may be due to the use of different analysis platforms, the reproducibility issues of NGT analyses in particular, and other biases that call for caution (see above). Another explanation, however, could
lie in the fact that the original study monitored severe weight loss but did not specify the initial BMI of the obese subjects, instead reporting only their weight reduction.\textsuperscript{74} It is possible that very obese subjects were selected to maximize the effect. In a recent study using the HITChip phylogenetic microarray, an ultra-deep analysis of the intestinal microbiota of morbidly obese subjects with an average BMI of >45 was performed. The \textit{Bacteroides}/\textit{Firmicutes} ratio was found to be significantly lower compared with that in lean subjects (BMI of 25) (Verdam et al., unpublished data, 2012). This study also indicated that a high BMI correlated significantly with a high level of proteobacteria, known to produce lipopolysaccharide that was associated with low-grade inflammation in animal experiments.\textsuperscript{96} This requires deep analysis of the microbiota as well as advanced phenotyping and genotyping of the human subjects. Moreover, causal relations need to be established; for example, further mechanistic insight can be gained from ongoing studies with fecal transplantations of lean donors, which have shown that the insulin resistance of T2D patients can be corrected.\textsuperscript{102}

**IMPLICATIONS OF THE MICROBIOTA IN OTHER DISEASES AND ABERRATIONS**

As the number of studies on the intestinal microbiota increases, the number of diseases and aberrations analyzed for associations with the intestinal microbiota likewise increases. In addition, many models of mice and other experimental animals are used in disease models to study the associations with the intestinal microbiota. Needless to say, these are very useful in providing mechanistic insight and do stimulate follow-up studies, but the results cannot be extrapolated to humans. It should be noted that humans and mice have considerably different microbiota, as their intestinal architectures, physical-chemical environments, and biological environments are very dissimilar. A summary of some of the presently anticipated associations that are less strong than those discussed above is provided in Table 2. Some of the support derives from analyses in human trials and observational studies of single patients, and some from a careful analysis of the literature. Other support relates exclusively to trials in animals, often wild-type, mutant, or germ-free or antibiotic-treated mice. In many cases, the evidence for the associations is rather premature and relies on single case reports. Several diseases, however, are so important that they are worth including, e.g., the data on nonalcoholic fatty liver disease, which affects over 30\% of the US population.\textsuperscript{101} Other diseases, such as Alzheimer’s disease, multiple sclerosis, and Parkinson’s disease, are almost untreatable diseases. Hence, sufficient care should be taken not to overinterpret the indications in the summary presented in Table 2. In some cases, other circumstantial evidence supports the relation to intestinal microbiota. This is the case with Alzheimer’s disease, multiple sclerosis, and Parkinson’s disease, for which non-peer-reviewed but promising results have been reported.

<table>
<thead>
<tr>
<th>Disease or aberration</th>
<th>Type of support</th>
<th>Reference*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s disease</td>
<td>Microbiota in a mouse model of Alzheimer’s disease</td>
<td>Karri et al. 2010\textsuperscript{103}</td>
</tr>
<tr>
<td>Atherosclerosis</td>
<td>Analysis of plaques in humans</td>
<td>Koren et al. 2011\textsuperscript{104}</td>
</tr>
<tr>
<td>Autistic spectrum disorders</td>
<td>Analysis of mucosa in children with autism spectrum disorders</td>
<td>Williams et al. 2011\textsuperscript{105}</td>
</tr>
<tr>
<td>Chronic fatigue syndrome</td>
<td>Cultured microbiota in patients with chronic fatigue syndrome</td>
<td>Sheedy et al. 2009\textsuperscript{106}</td>
</tr>
<tr>
<td>Colic babies</td>
<td>Longitudinal analysis of colic babies cohort</td>
<td>de Weerth et al. 2012 unpublished data</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>Cardiovascular-diseased mice and microbial metabolism</td>
<td>Wang et al. 2011\textsuperscript{101}</td>
</tr>
<tr>
<td>Depression and anxiety</td>
<td>Probiotic intervention in stressed mice</td>
<td>Bravo et al. 2011\textsuperscript{104}</td>
</tr>
<tr>
<td>Frailty</td>
<td>Analysis of elderly and high frailty scores</td>
<td>van Tongeren et al. 2005\textsuperscript{107}</td>
</tr>
<tr>
<td>Graft-vs-host disease</td>
<td>Review of human data on graft-vs-host disease</td>
<td>Murphy et al. 2011\textsuperscript{108}</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>Involvement of microbiota in mice with multiple sclerosis</td>
<td>Berer et al. 2011\textsuperscript{109}</td>
</tr>
<tr>
<td>Nonalcoholic fatty liver disease</td>
<td>Effect of choline depletion in humans</td>
<td>Spencer et al. 2011\textsuperscript{101}</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>Role of enteric nervous system and review of Parkinson’s disease development</td>
<td>Braak et al. 2003\textsuperscript{110}</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>Microbiota as predisposing factor in rheumatoid arthritis</td>
<td>Scher and Abramson 2011\textsuperscript{111}</td>
</tr>
<tr>
<td>Retrovirus infection</td>
<td>Mouse retrovirus infection relies on microbiota</td>
<td>Kane et al. 2011\textsuperscript{112}</td>
</tr>
<tr>
<td>Poliovirus infection</td>
<td>Mouse microbiota promotes poliovirus infection</td>
<td>Kuss et al. 2011\textsuperscript{113}</td>
</tr>
</tbody>
</table>

\* The most recent single reference is given.
in single cases by using fecal transplantations. In addition, it has been reported that the levels of Bifidobacterium spp. and Akkermansia muciniphila, previously identified as a health-related bacteria (see above), were decreased in children with autism spectrum disorders compared with healthy controls.

A set of other studies provide completely new mechanistic insight that can be applied to the two recent studies involving the role of intestinal microbiota in virus replication in mouse models. These breakthrough results provide a radically new vision of the development and progression of viral infections, the involvement of the immune system, and the contribution of the intestinal microbiota. It will of great interest to see the impact of the intestinal microbiota on viral infections in compromised and healthy humans. While these studies may provide support for the efficacy of antibiotic treatments in viral infections, it should be remembered that repeated antibiotic treatment disturbs the intestinal microbiota.

CONCLUSION

A total of over 25 diseases, syndromes, or other aberrations have now been associated with the intestinal microbiota. The most robust ones discussed here are supported by multiple and robust studies in humans (Table 1). In addition, a series of other less convincing or first reports are listed as well (Table 2). In several cases, specific correlations have been made with specific potentially health-associated bacteria, such as some Bifidobacterium spp., Faecalibacterium prausnitzii, and Akkermansia muciniphila. As expected, these all belong to the common core of QoL-controlled healthy subjects (Figure 2). Cause-and-effect relationships, however, are scarce, but it is of interest to see a Koch’s postulate applied to Bacteroides strains in a mouse IBD model. Moreover, there is a large body of literature on the use of single or multiple strains of lactic acid bacteria or yeasts that are marketed as probiotics and provide a health benefit. While these probiotic strains in general do not colonize and hence are not considered to constitute the intestinal microbiota, the experimental approaches used to show the efficacy of these strains may well serve as guidance for further mechanistic studies of the intestinal microbiota. Finally, in recent years, fecal transplantations of healthy microbiota have increased in both frequency and sophistication and can be seen as demonstrating cause-and-effect relationships, provided the studies are carefully controlled and documented. A first example of a double-blind trial involving fecal transplantation has recently been documented. Now, however, it is necessary to characterize the microbiota changes over longer periods of time, link these changes to health status, and provide insight into the attributes of the donor microbiota. Moreover, as

specified recently, continuing support from gastroenterologists is needed.

Once an association with the microbiota has been established, there are various obvious next steps to be taken that include biomarker development, early diagnostics, and monitoring of disease development. Moreover, specific segmentation of the patients may be possible using biomarkers, such as those based on the enterotypes, which may lead to personalized treatments. Finally, especially when causal effects have been established, specific therapeutic avenues can be developed on the basis of health-promoting microbes or their biomolecules. Given the importance of the diseases described here, it is imperative to 1) build on the established associations between intestinal microbiota and the predisposition to disease, 2) elucidate new associations, and 3) establish causality.

Acknowledgments

The authors are grateful to Dr. Mirjana Rajilić-Stojanović and Jonna Jalanka-Tuovinen for assistance with Figures 1 and 2, respectively.

Funding. This work was partially supported by project 40274/06 of TEKES, Finland, project PSG 93609004 and the unrestricted Spinoza Award of the Netherlands Organization of Scientific Research, grant 25.0.1.7 from the European Research Council (MicrobesInside) and research grants 118165, 137389 and 141130 from the Academy of Finland.

Declaration of interest. The authors have no relevant interests to declare.

REFERENCES


