REVIEW ARTICLE

Infant gut microbial colonization and health: recent findings from metagenomics studies

Adrien Fischer*,1, Katrine Whiteson1, Vladimir Lazarevic1, Jonathan Hibbs1, Patrice Francois1, Jacques Schrenzel1.

1Genomic Research Laboratory, Department of Internal Medicine, Geneva University Hospitals, Rue Gabrielle-Perret-Gentil 4, 1211 Geneva 14, Switzerland.

Received: 24 October 2011 Accepted: 11 April 2012 Available Online: 13 April 2012

Abstract

New DNA sequencing technologies which have emerged in the last decade enable in-depth study of human gut microbiota. The bacterial communities inhabiting our gut influence our immune system development and maturation as well as our general health. However, the balance between host and bacterial community is affected by changes in lifestyle. Increasing rates of caesarean delivery, formula-feeding, antibiotic treatments, high fat diet, urbanization and hygiene may create important changes in the colonization of the gut microbiota. Emergent diseases and conditions like asthma, allergies, necrotizing enterocolitis (NEC), obesity and diabetes may be related to modifications in the microbiota. In this review we focus on studies related to early bacterial colonization of the gut, and how the evolution of gut microbiota during the first years of life may lead to new perspectives on the treatment of these diseases. Diet complementation with pre- or probiotics in formula or replacement of a disease associated-microbiota with a healthy one are currently the most studied approaches in the treatment of microbiota-related disorders. Bacteriophages may provide an alternative means for manipulating gut bacterial communities. However, the question is whether we can affect infants gut microbiota without any risk to health. High-throughput sequencing (HTS) techniques not only allow evaluation of the composition of the gut microbiome, but also follow-up of its evolution in real-time, either naturally or altered by environmental influences. This review discusses these techniques, evaluates the impact of microbiome composition on infant development and outlines possible improvements in health care based on this knowledge.

Keywords: Microbiota; Infant; Gut; Delivery; Feeding; Probiotic; Prebiotic.

1. Introduction

The last decades have given rise to new genomic approaches for the study of uncultivated cells and viruses. Micro-organism community structure and diversity from inert surfaces (soil or sea) [1-3] or associated with humans and other eukaryotic organisms can now be studied [4-6]. Jo Handelsman originally described this as metagenomics, the study of all the microbial genomes in a community (metagenome) including different species and kingdoms (i.e. bacteria and viruses).

This culture-independent approach is now widely used to study the composition of human microbiome and viriome and evaluate their contribution to health. Bacterial cells in the human gut are ten times more abundant than the host’s somatic cells [7,8] and are in a confined space. Studying the evolution of bacterial populations in this context with high resolution methods will supplement classical techniques which are probably able to identify less than 50% of bacterial species present [9]. The greatest part of the human microbiome correspond to the bacterial community in the distal part of the gut, while the bacterial load at other anatomical sites (i.e. skin and mucosae) is relatively low [10]. The microbiota can shape the development of human gut epithelial cells and contribute to our digestion by their ability to synthesize enzymes to hydrolyze otherwise indigestible oligosaccharides [11,12]. Gut microbiota play an important role in the maturation of the intestine [13-17] and act as a barrier against pathogens. They also influence host metabolism (drug, amino-acid, lipid and anti-oxidant),

*Corresponding author: Adrien FISCHER. Genomic Research Laboratory, Department of Internal Medicine, Geneva University Hospitals, 1211 Geneva 14, Switzerland. Phone (+4122) 372 9818, Fax (+4122) 372 9830, Email Address: adrien.fischer@genomic.ch
nutrient absorption as well as immune development and response [18,19,15]. There is a complex relationship between gut microbiome, environment, diet and even genetic predisposition to specific diseases, especially auto-immune disorders [20].

In the 20th century, advances in agriculture, transportation, urbanization and medical care, among other factors, have changed our diet and lifestyle significantly. These modifications may have a deep impact on our gut microbiota composition [21] that may be related to the emergence of new diseases like asthma, obesity, autoimmune disorders or inflammatory bowel diseases in adults and children [22-27]. Changes in microbial community structure have been observed in many human diseases including diabetes and inflammatory bowel diseases [23,28-34].

Human gut colonization begins at birth with microbes and viruses found in the environment and present in maternal skin or vagina (depending on the delivery mode). During the first two years of life, our microbiome and viriome change and evolve to a community different from that found at birth [35]. A long list of important questions is being addressed in this quickly growing field by researchers around the world. Understanding bacterial colonization of the infant gut as a function of environment, genetic composition or phage community may help us avoid and manage diseases not only in infancy, but throughout life. Will this new genomic approach lead to new treatments? Are gut microbiota associated with disease as cause, consequence, or both? Does the virus population in an infant’s gut play a role in the emergence of new diseases? Could manipulation of the gut microbiota prevent, treat or worsen diseases of infants, older children or adults?

2. Tools for physical characterization of microbial diversity

Studying variation in the human gut microbiota is currently possible with a range of techniques which are decreasing in cost. Many of these techniques have been used by ecologists studying soil or ocean microbial communities before the explosion of interest in human microbial communities.

2.1 Metabolites based method

Most methods use nucleic acids as their substrates. It is possible, however, to sample microbial populations using other molecular substrates.

Mass spectrometry

Mass spectrometry (MS) can directly detect the chemical products of microbial metabolism. For example, MS shows that the gut microbiome has an impact on mammalian blood metabolites [36], highlighting the influence of the microbiome on the drug processing capacity of the host. This method is not yet widely used to study commensal bacteria. Introducing MS complements the phenotypic information derived from nucleic acid-based methods, and may help us to learn more about bacterial metabolism. Because it requires no prior assumptions about what molecular components will be present, mass spectrometry is an open-ended technique and allows potential detection of previously unknown organisms.

2.2 Nucleic acid based methods

Variations in the 16S rRNA sequence allow phylogenetic classification of bacterial and archeal species [37]. The 16S rRNA gene is ubiquitous among prokaryotes, and has become the main genetic and phylogenetic tag to characterize and compare bacterial communities using classic techniques or high-throughput sequencing (HTS).

Polymerase chain reaction (PCR)-based techniques have been used to characterize prokaryotic phylogeny since 1980’s [38,39]. More recently, PCR has been used to complement other methods like fluorescence in situ hybridization [40] for the study of microorganism variation. In addition to 16S based approach, bacterial communities may be investigated analyzing random metagenomic fragments [41].

Restriction Fragment Length Polymorphism

In the 1990’s PCR with fluorescent primers was used to amplify ribosomal DNA [42]. The resulting amplicons were digested with restriction enzymes and fragments obtained were separated electrophoretically. This method, called Terminal Restriction Fragment Length Polymorphism (T-RFLP) has been used to determine the microbial composition of samples from ocean crust [43], cystic-fibrosis affected lungs [44], and many other sources. The resolution of T-RFLP is rather moderate, since the number of restriction sites in the 16S rRNA genes is limited.

Automated ribosomal intergenic spacer analysis

A PCR based method relying on the length heterogeneity in the intergenic region spacer of the 16S-23S ribosomal RNA was developed by Fisher and Triplett in 1999 [45]. This method called ARISA (Automated Ribosomal Intergenic Spacer Analysis) allows for the comparison of bacterial communities but also provides an estimation of microbial richness and diversity from many kinds of samples including soil, water (ocean) or human gut [46-50]. ARISA was used in concert with other classical analysis tools to assess a correlation between gut microbiome and diabetes in rats [51]. This method is cheaper than most of the others and allows a rapid estimation of bacterial composition, but has some limits because bacterial 16S and 23S rRNA genes are not always organized as an operon; in silico modeling shows a loss of linearity when the species richness increases [52].
The utilization of complementary methods, such as statistics [53] or sequencing [54] could help to correct this bias.

**Fluorescence in situ hybridization**

Fluorescence in situ hybridization (FISH) allows the localization of specific DNA sequences after hybridization with specific fluorescently labeled oligonucleotide probes [55,56]. It was used to compare different bacterial communities e.g. the gut microbiota of 6 month old formula-fed or breast-fed babies. This method allowed Rinne and colleagues to identify differences related to a decrease of *Bifidobacterium* in formula-fed infants gut microbiota [40]. FISH has the advantage of observing specific members of the bacterial community directly in the sample, for instance, within buccal epithelial cells [57,58]. The basic limitation of the method is that it only detects taxa covered by the probes. FISH is an example of a closed-ended diagnostic technique, like T-RFLP or ARISA, while metagenomics includes the study of unknown bacterial communities, which can only be revealed through more open-ended methods, assuming little prior knowledge of the bacteria under study.

**Microarray**

A microarray consists of a huge number of single-strand DNA molecules spotted on glass slides. Hybridization with labeled cDNA of interest allows identification of bacterial species, differences between strains and the presence or absence of genes of interest. Microarray assays may be used within high-throughput sequencing. For example microarray analysis revealed a dramatic shift in the gut viral population between 1 and 2 weeks of life [59], while high-throughput sequencing allowed the detection of some stable members in viral community until 3 months of age [59]. Analysis of bacterial communities in patients suffering noma disease was conducted using microarrays [60,61].

Using microarrays alone may miss important unpredicted or unknown members of a bacterial community. This is because microarrays in general are another closed-ended technique, designed in advance with known sequences. This approach, which assumes cognition of all sequences of interest before the experiment, is considered to be more efficient than using random sequences on the array. It introduces a bias against previously unknown organisms, but it does allow screening of many known sequences, enough to identify most Operational Taxonomic Units (OTUs) known at the time the experiment is done. Samples obtained from two separate samples, processed similarly with identical microarrays, may provide information regarding the differences in their microbiota profiles.

**High-Throughput sequencing methods**

Recently, the Roche/454 and Illumina/Solexa HTS proved their potential in analyzing microbial communities [62,63,41]. Currently, 454 pyrosequencing produces reads close to $10^5$ bases, whereas the length of Illumina reads is limited to $10^3$ bases. However, the number of sequence reads generated by Illumina is, for the same cost, about 150x that produced using the Roche/454 platform. HTS-based study of microbiota may include random sequencing of metagenomic fragments or sequencing of 16S rRNA amplicons libraries [62,63,41]. 16S rRNA contains very stable as well as highly variable regions, allowing for phylogenetic comparison. The HTS of partial 16S rRNA genes of a microbial community generates a large amount of classifiable sequences. All target sequences, however, are not amplified with the same efficiency using “universal” 16S primers, the copy number is not stable across organisms (which complicates abundance analysis), and the classification of 16S rDNA sequences is typically limited to the genus level [64]. When genomic DNA from the community is randomly sequenced, a higher fraction of sequences remain taxonomically unassigned because of the lack of homology in sequence databases. Nevertheless, the taxonomy of sequences corresponding to different groups of organisms, including bacteria, archaea and fungi as well as viruses may be assessed simultaneously.

While 10 sequences per sample may be enough to identify differences between bacterial communities [28], it is necessary to increase the number of sequences per sample to identify underrepresented taxa. In some cases, even after obtaining one million sequences from a single sample, further sequencing will still lead to the discovery of new phylotypes [65]. HTS has aspects of a closed-ended system since it assumes all the target organisms have conserved elements (of the 16s gene, for example). It allows considerable variability within those limits, however (including the detection of new 16s variable sequences pertaining to a previously unknown species, such as the discovery of *Trophyrema whippelii* [66]). In this sense it is a partially open-ended system, allowing for some discovery of unexpected organisms that fit fairly broad pre-experimental assumptions. The possibility of sequencing the complete genome of bacteria without any prior knowledge of their sequences aside from the 16s primers confirmed the open-ended aspects of HTS.

Although the HTS culture-independent approach could out-compete other existing diagnostic and typing methods in microbiology, questions remain about the cost and best strategy for analysis of the massive datasets generated by this technology.

**Analytical strategies for HTS data**

Data analysis is one of the most labor intensive aspects of metagenomic projects. Computing advances do not keep pace with constant huge increase in the data generated by high-throughput sequencing. This leads to problems with storage space and computer memory as well as with the selection of consistent and reproducible analytic algorithms [67].
A typical HTS data set must be cleaned following predefined parameters for the study. For example, sequences which are too short for a given technology contain ambiguous characters, or those that do not carry the primer sequence are removed. Sequences containing putative amplification and sequencing errors are also removed. Remaining sequences are then clustered into Operational Taxonomic Units (OTU or phylotypes) with varying percentages of sequence identity [68]. Currently 97% identity is often used to represent the species level cut-off. In many cases a representative sequence is chosen from each OTU, by largest abundance or even a randomly chosen sequence, and this is used to build a phylogenetic tree. The phylogenetic diversity in the tree can be confirmed by BLASTing all the reads in a parallel analysis. Finally, the taxonomy can be assigned to each OTU, after using RDP classifier [69,70]. Much can be learned about the composition of the microbial community in a particular sample by simply looking at the abundance and identity of the top ten or more OTUs; doing this at different identity cut-offs is also revealing.

Within sample (alpha) and between sample (beta) diversity may be estimated from the abundance of 16s rDNA sequences [71]. Although different studies with different approaches will each have their own unique bias, comparisons of the communities found in different samples with the same methods are likely to reveal biologically relevant differences. Branch length-based phylogenetic diversity comparisons using UniFrac [72], or nearest taxon index (NTI) [73] are widely used. UniFrac metric, that calculates the fraction of the branch length shared by two communities in their common phylogenetic tree, can be used to estimate beta diversity. For more than two communities, a distance matrix that relates each pair of communities may be subject to hierarchical clustering and Principal Component Analysis [74,75]. ANOSIM (Analysis of Similarity) [76] is a powerful way of testing the similarity of community composition of samples with a shared condition against a null hypothesis that they are randomly distributed. PERMANOVA [77,76] determines which species in a community are driving the differences. Powerful software is emerging to accommodate these types of analysis, including MG-RAST [78], QIIME [79], mothur [80], RDP [69], CARMA [81], CAMERA [82], IMG/M [83], PRIMER [76] and PC-ORD (MJM Software Design, Gleneden Beach, OR).

3. Using metagenomic tools to study development of the human gut microbiome during infancy

Colonization of an infant’s gut begins just after birth. This initial colonization may be important as it could establish long term microbial communities. Gut microbiota are highly variable and poorly diversified during the first two years of life before the more stable and diversified adult microbiota is established, often around the time of weaning [84,85]. Recent research suggests the importance of the first colonization (bacterial or viral), which is linked to the delivery mode, the environment, food source and genetic background in infant’s health [86-95].

3.1 Initial colonization

During the first month of life, *Bifidobacterium* are predominant in gut of almost all infants [96]. A recent study on gut microbiota from vaginally delivered and at least partially breast-fed infants that were never exposed to surgical intervention, antibiotics, pre- or probiotics, provided insight into “normal” bacterial colonization and evolution of the gut microbiota during the first 4 months of life [97]. The study, partially represented in table 1, identified a large number of bacteria and revealed the presence of *Staphylococcus*, *Bifidobacterium* and *γ*-Proteobacteria in most 4 days old infants’ gut. The authors highlighted that the evolution of gut microbiota includes shifts in both composition and structure. From day 4 to 120, *Lactobacillus 2a*, *Veillonella*, *Lachnospiraceae2* and *Bifidobacterium 1* populations increased whereas *Staphylococcus*, *Streptococcus* and an uncultured bacterium population decreased. Some strains could cause severe diseases only in particular contexts. For example, if low-level toxin-producing staphylococcal strains are supplanted by a high-level toxin producer, the gut microbial community might not appear to change but the baby may become ill. Moreover, the use of antibiotics may have a drastic effect on development of gut microbiota and health especially if infants are not breast-fed [88-98-100].

3.2 Delivery Mode

3.2.1 Vaginal vs. C-section delivery

The most important first source of inocula after birth is the mother’s vaginal and fecal microbiota [92,93] from the birth canal. This ecosystem contains a limited number of bacterial taxa [101,102]. Depending on the delivery mode, these first inocula could be completely different (Table 1). Indeed, if babies are delivered vaginally, their gut microbiota just after birth (before 24 hours of life) is similar to their mother’s vaginal (and fecal) microbiota (sampled 1 hour before delivery) dominated by *Lactobacillus, Prevotella* or *Sneathia* spp [94]. In contrast, children delivered by C-section are colonized by species similar to those from their mother’s skin microbiota such as *Staphylococcus*, *Corynebacterium* and *Propionibacterium* spp [94] or from the environment (equipment, air, other infants, nurses) [95]. *Staphylococcus aureus* (*S. aureus*) colonization is more common in the gut of infants delivered by C-section likely because of the presence of this species on their mothers’ skin [103,104]. However, Eggesbo and colleagues [97] demonstrated *Staphylococcus aureus* in 95% of healthy 4 days old vaginally delivered and breast-fed infants decreasing after 120 days to...
a prevalence around 60% [97]. Further studies over longer time periods and larger samples are required to establish whether delivery mode affects S. aureus colonization.

Babies born by C-section have a microbiome similar to vaginally born infants whose mothers received antibiotics before normal delivery or during breast-feeding [95]. This may be a consequence of a shift in mother’s vaginal microbiota due to antibiotic treatment or a direct effect of antibiotics from mother’s milk on the vaginally-delivered babies gut microbiota. Common use of antibiotics before C-section may also contribute to this observed similarity. The delivery mode has a major role in first colonization of the infant gut. How can we determine, however, whether delivery mode and its effect on microbial community composition affect the development of increasingly common disorders including asthma, allergy, auto-immune diseases and colitis?

3.2.2 Clostridium difficile and C-section

Clostridium difficile (C. difficile) is an anaerobic spore-forming gram positive rod which under some circumstances elaborates an exotoxin that kills epithelial cells. In extreme cases, enteric infection with C. difficile causes inflammatory diarrhea, toxic megacolon and death. The organism can be recovered from the stool of healthy individuals [106], however, exposure to antibiotics alters gut flora in a way that favors overgrowth of C. difficile, elaboration of its toxin, and disease [107]. Clindamycin, an anti-anaerobic agent to which Clostridium difficile is resistant, is most strongly associated with C. difficile diarrhea, but risk also increases even with short courses of many other antibiotics, such as pre-surgical prophylaxis with cefazolin [108,109].

A correlation between C. difficile colonization and delivery mode was not found in a 1986 study in Italy [110]. Thirteen percent of the newborns were colonized with C. difficile and the colonization rate was higher but insignificant in caesarean than in vaginally delivered infants [110]. However, according to Penders et al, infants born by caesarean are more likely to be colonized with C. difficile [88]. One study involving 1032 one month old infants found that the prevalence of C. difficile in infants’ gut was 42% for C-section delivered, 26% for vaginally-delivered in the hospital, and only 19% for home delivered babies [88]. A recent paper from the Chicago area highlighted the increased prevalence of C. difficile infections in C-section delivered infants with an incidence of 2.2 / 1000 births while only 0.2 / 1000 of

Table 1. Infant gut bacterial population depends on the delivery mode. *: After 8 days of life; **: Babies vaginally delivered, breast-fed and without any treatments (babies and mother) before the study; +: predominant. This Table is based on the results presented in the following studies: Dominguez-Bello et al [94]; Fallani et al [95]; Eggesbo et al [97]; Morowitz et al [105]; Hyman et al [101]; Penders et al [88].
vaginally delivered newborns were infected [111]. It is possible that the different findings reflect relatively higher antibiotic use with caesarean delivery in the U.S. and Italy; higher antibiotic use would be expected to promote more *Clostridium difficile*. However, 60% to 70% of infants guts (vaginally or C-section delivered in Boston) are asymptptomatically colonized with *Clostridium difficile* [112]. Asymptomatic *Clostridium difficile* colonization correlates with high levels of Firmicutes (also seen in C-section born infants, except for *Staphylococcus* or *Bacteroides* but it is clearly not correlated with the presence of *Bifidobacterium* [95,113,107]. *Clostridium difficile*-associated disease correlates with the absence of Bacteroidetes [114]. Usually, *Clostridium difficile* disappears in children between 1 and 2 years old. The gut microbiota of infants, characterized by a low diversity, might not provide an efficient protection against *Clostridium difficile* infection. The absence of some bacteria may allow unusual *Clostridium difficile* colonization and overgrowth leading to disease. Table 2 and Khoruts’ study (described below) support the idea that Bacteroides colonization decreases the chances of *Clostridium difficile* pathogenicity. Therefore, decreased *Bacteroides* and *Bifidobacterium* populations in infant gut, due to the use of antibiotics [88] may increase the risk of *Clostridium difficile* infection. In older patients *Clostridium difficile*-associated disease is related to a decrease in overall bacterial populations [34]. Rousseau et al therefore suggested that *Clostridium difficile* colonization results in a microbial community disorder of the gut. Alternatively, disturbed microbial populations may predispose to overgrowth of *Clostridium difficile*. Varma and colleagues observed that the supernatant from *Lactobacillus fermentum* culture alone is enough to inhibit *S. aureus* and *Pseudomonas aeruginosa* growth in vitro, suggesting that compounds released by *Lactobacillus* inhibit pathogen growth and may be used to treat hospital-acquired infections [115]. Whether competition from particular bacteria or by inhibitory compounds released during their growth affects bacterial colonization remains to be clarified.

The association between disease-causing *Clostridium difficile* and C-section, found in some studies [88,111] was not confirmed in the others [110,112]. One study pointed to a shift or a decrease in overall bacterial populations in *Clostridium difficile*-associated disease. Therefore, it appears that *Clostridium difficile* may be a member of the gut microbiota of healthy infants and that illness might be triggered by factor which causes a shift in the gut microbiome.

Initial microbial colonization in infants may influence health in later life. The gut microbiome undergoes several drastic shifts in composition during the first two years of life, and later becomes more stable [116,117,85]. Bacterial taxa differ in their ability to persist in the infant gut [97]. The forces that shape the formation of bacterial communities in the gut, including the immune system and the consequences of random environmental variables, are the subject of intense study.

### 3.3 Feeding mode

**3.3.1 Lifestyle and geography**

During the last decades, our life style has changed drastically. Nowadays, more people live in cities than in rural areas. We use antibiotics, eat heavily processed food and pay more attention to hygiene. All these changes have an impact on our microbiota and thus on our health. Carlotta de Filippo and colleagues have compared the gut microbiome of African children having a life style close to the time of the birth of agriculture to the gut microbiome from European children [22]. They highlighted a diet, richer in fiber from complex grains for African children and, noted an enrichment of Bacteroidetes and a depletion of Firmicutes among African compared with European children. The effect of dietary fibers on the relative abundance of Bacteroidetes and Firmicutes was recently confirmed in the ob/ob mouse model [118]. Gut microbiota of African children are enriched in bacteria containing genes for cellulose or xylose hydrolysis, as well as for short-chain fatty acid metabolism.

**3.3.2 Breast-feeding**

Human milk and thus breast-feeding affects growth, immune system development, cognitive development, as well as susceptibility to toxins and immune diseases like asthma and allergy [119-125]. At least some of these effects may be mediated by the effect of breast-feeding on gut microbiome.

Differences in breast-feeding practice influence the observed microbiota differences in the gut between human subpopulations [22,95]. In the De Filippo’s study [22], the microbial communities of breast-fed African and Italian children were quite similar, but diverged when breast-feeding was discontinued in the Italian infants. A study of infants in Granada and Stockholm found that the Granada infants had lower carriage of *Bifidobacteria* (19 vs. 60%) and higher carriage of *Bacteroides* (21% vs. 6%), which the authors attributed to the lower rate of breast-feeding in Granada (43% vs. 76%) [95]. On the other hand, limited evidence suggests that feeding mode is not required for *S. aureus* colonization. Sequencing the gut metageneome of an 8-day-old premature breast-fed neonate revealed a limited population of *S. aureus* [105]. *S. aureus* population grew significantly at 10 days of age, suggesting *S. aureus* reservoir was independent of feeding type [105]. In another case, contamination by a methicillin resistant *S. aureus* (MRSA) through mother’s breast milk was reported [126], suggesting that the colonization by *S. aureus* may be in some cases related to the feeding mode. While term-delivered infant gut microbiota is less diverse than that of adults, it remains more diversified than premature infants gut microbiota [85,127,33]. Among low weight infants colonized by coagulase negative *staphylococci* (CoNS), breast-feeding decreases duration of persistent CoNS bacteremia [128]. Breast-feeding might be a good way to protect against *S. aureus* diseases. It must be conceded, however, that the
apparent ecological relationship between breast-feeding and gut microbiota variations is highly susceptible to confounding, since children living in Western cities are not exposed to the same physical or social environment as children living in developing countries or in farms.

In addition to ecologic and demographic demonstrations, microbiologic studies of individuals also tend to support the influence of breast-feeding on the infant gut microbiota. Breast-feeding gives rise to a less diversified microbiota mainly composed of *Bifidobacterium* and *Lactobacillus* [40,129,130], while formula-feeding promotes *Bacteroides* and *Clostridium coccoides* gut colonization [95,131], resulting in a microbiota profile closer to those of adults. One obvious explanation for such differences, illustrated in the Figure 1, is the presence of abundant and complex oligosaccharides in human milk absent in formula. These oligosaccharides, indigestible by humans, are a major substrate for *Bifidobacterium* (Bifidobacterium bifidum and *Bifidobacterium longum* subsp. *infantis*) [132] and are necessary for the “normal” infant gut microbiota development [132]. They promote the growth of *Bifidobacterium* and in some cases of *Lactobacillus* [133]. As oligosaccharides from human milk are similar to some human cell surface carbohydrates, they play an important role in the protection against pathogenic bacteria by competing with their target ligands [134,135]. Oral supplementation with *Bifidobacteria* or *Lactobacillus* (probiotics) may have beneficial effects. For instance, the addition of *Lactobacillus* in the diet of infants, aged between 2 weeks and 13 years, during the three month preceding antimicrobial therapy, decreased the incidence of diarrhea 2 weeks after the introduction of the antimicrobial treatment [136]. Administration of Lactobacillus in children from 2 months to 6 years old was associated with lower rotavirus diarrhea duration and, improved effect of parenteral rehydration [137]. Co-administration of *Lactobacillus* and *Bifidobacterium* was associated with reduced incidence of NEC in low-birth weight infants [138], and reduced diarrhea duration [139] as well as stool frequency [139]. However in some instances, the benefit from administration of *Lactobacillus* or *Bifidobacterium* was not observed, e.g. on the duration of non-rotaviral diarrhea [137], and the incidence of death, NEC or nosocomial infections in low birth-weight infants [140,141].

These observations raise the question of whether prebiotics should be added to formula. Prebiotics are selectively fermented ingredients that promote growth and/or activity of the gut bacteria and confer benefits to the host [142]. Several studies demonstrated that introducing prebiotics like galacto-oligosaccharides or fructo-oligosaccharides into formula shifted the gut microbiota to be more similar to breast-fed infants [129,130,40,143,131]. Moro and colleagues [143] observed a dose-related positive association between prebiotics in formula and lactobacilli number in infants gut but Fallani *et al* [95] found a similar abundance of lactobacilli in formula-fed without prebiotic supplementation and in breast-fed infants. The distribution of *Bifidobacterium* species in the fecal samples from standard and prebiotic-supplemented formula-fed infants resembled those of breast-fed infants and adults, respectively [129]. The infant’s gut microbial community is thus highly malleable and sensitive to changes in the composition of the diet. The use of prebiotics should be approached with caution, because their long-term effects are currently unknown.

4. Role of intestinal flora in the emergence of new diseases

Intestinal colonization is the host’s earliest contact with micro-organisms [147], with important consequences on maturation of the immune system and on metabolism [148]. In this section, we examine mechanisms by which intestinal flora might influence pathophysiology of diseases mediated by nutrient absorption and by the host immune system. This is a dynamic literature, and every week new studies demonstrate associations between altered microbiota and disease states. We note, however, that at this stage most published observations lack the resolution to disentangle cause and effect. In most cases we cannot yet determine whether altered microbiota cause disease or are actually a marker of underlying problems.

4.1 Obesity and nutrition

Gut microbiota may modulate ability to access nutrients and therefore body weight and composition. Studies comparing germ free (GF) mice raised in sterile conditions with normal mice have found that they respond differently to the same high fat diet and suggested that GF mice can resist obesity. Compared to mice with normal gut flora, GF mice consumed fewer calories, increased lipid excretion and enhanced insulin sensitivity [149]. No differences were observed during a low fat diet between GF and conventional mice [150,151], suggesting a primary impact of the gut flora on absorption of fats.

Intestinal microbiota may also influence nutrition by their effect on gastrointestinal hormone production. *Helicobacter pylori* (*H. pylori*), an ancient and dominant commensal microbial inhabitant, regulates ghrelin and leptin production, two hormones involved in body weight regulation [152-155]. As *H. pylori* has become less prevalent in infants’ commensal flora since the beginning of the early 1900s, these two hormones persist and may cause infants’ obesity and type-2-diabetes.

Immunologically active receptors may be another means by which gut microbiota influence obesity. High levels of FIAF (fasting-induced adipose factor), AMPK (AMP-activated protein kinase regulating fatty acid oxidation) activation and a lack of ileal TNF-α induction have been observed in GF mice that remain lean despite a high fat diet [156,150,151]. Lean GF mice treated with gut flora from obese mice become obese [157]. Whatever the mechanism, it seems clear that obesity is
epidemiologically associated with the gut microbiome, and in particular with a predominance of Firmicutes over Bacteroidetes [23]. Among discordant twins, bacterial communities in the obese twins are less diverse [65].

4.2 The enteric immune system

Healthy development of the immune system has been associated with infants’ gut microbiota colonization [158,159]. Complex interactions with commensal microbiota modulate the response of the immune system to self molecules, harmless and pathogenic microbes. Commensal microbiota is not essential for animal life [160,161], and its presence may even appear disadvantageous in some respects. For instance, GF mice have decreased pro-inflammatory T cell response (and therefore less pathology) in an induced...
encephalomyelitis model [162]. However, GF mice also show altered antibody production, lymph node structure and gut capillary and lymphoid tissue development [163-166]. Changes in interactions between the immune system and the microbiota may lead to the emergence of allergy, asthma, diabetes, obesity or gastric disorders [158,29,167,168,146]. Immune system response is involved in this process as this disease is TNF-α and INF-γ dependent [20]. Pathology in apparently “autoimmune” diseases may require interaction between the host genome and specific microbiological exposure, as with the cytokine-mediated induction of a Crohn’s disease-like illness in mice by a combination of a host mutation with a specific norovirus infection [20].

Migration of Mast cells from blood to the intestine might be promoted by the commensal microbiota as GF mice carry less Mast cells in their intestine [169]. Mast cells or Mastocytes are involved in allergic disease, but also in wound healing and protection against pathogens [170]. GF mice underexpress Keratinocyte-derived cytokine, lipopolysaccharide-induced cytokine, and macrophage inflammatory protein 2, ligands for the receptor CXCR2 required for intestinal localization of Mast cells [169]. Oligosaccharides have a big impact on infant gut microbiota, and their similarity to pathogenic ligands allows them to compete with pathogenic bacteria to protect infants from infection [134,135]. Kunii et al [169] hypothesize that commensal bacteria in mice stimulate production of a ligand necessary for the intestinal recruitment of Mast cells.

4.3 Are treatments available to regulate gut microbial community disorder?

Even without a sophisticated understanding of how gut microbiota is assembled, acceptance of their importance allows for novel medical treatments of infectious diseases. C. difficile diarrhea is a paradigmatic example. Antibiotic treatments before surgery often lead to loss of “healthy bacteria” and susceptibility to opportunistic pathogens, including C. difficile. One such patient suffered chronic intestinal C. difficile associated disease, with diarrhea every 15 minutes, dramatic weight loss and confinement to a wheelchair [114,171]. The patient received a fecal transplant from her husband, and had a normal bowel movement the next day. Fecal transplants have been used successfully for this kind of infection in dozens of patients since the 1950s, but the treatment is far from accepted routinely. In this case, rapid and drastic changes in the patient’s gut microbiota was followed using metagenomics, showing how the patient’s gut bacteria came to resemble the healthy donor community, with a predominance of Bacteroidetes spp. strains, and the end of the patient’s symptoms. These results showed that in the absence of healthy gut microbiota, pathogenic bacteria overgrew and caused life-threatening diarrhea. If the opportunistic strain is exterminated and its niches are quickly populated by “healthy microbiota”, the pathogen will not be able to colonize the patient.

4.3.1 Prebiotics/Probiotics as a solution?

Many studies report the benefit of using pre or probiotics to treat infants’ disease linked to gut microbiota modifications. Studies underway in Europe, the US (At the UC Davis Children’s hospital http://www.ucdmuc.ucdavis.edu/children/pediatric_research) and South America (https://clinicaltrials.gov/clinical-trials/show/NCT00727363) all seek to understand initial infant colonization, and whether probiotics protect against NEC morepowerfully than antibiotics as previously proposed [172-177]. In some hospitals, notably in Sweden, all premature babies are given probiotics to prevent gut disorders such as NEC, despite the lack of certainty described by Millar et al [178]. Figure 2 summarizes the involvement of some compounds in infants’ intestinal diseases and makes a link with inflammatory immune reactions in the intestine. Although -biotics help to prevent or cure some diseases, like NEC or allergic rhinitis [179,180], their role remains to be clarified. Indeed, adding prebiotic to breast-fed infants lead to a different gut microbiota than those who were breast-fed exclusively [181].

4.3.1.1 Necrotizing Enterocolitis

It is not clear whether NEC has a specific agent or agents, or if so what those agents might be. Enterobacter sakazakii, found in infants’ formula powder, is one candidate [193,194] while Morowitz and colleagues proposed Citrobacter strains [105]. However, they observed that healthy children also carry Citrobacter strains. In other studies NEC was related to C. difficile, γ-proteobacteria or decreased bacterial diversity in infants’ gut [33,127]. This suggests that NEC is not caused by one particular bacterium. Moreover, Morowitz and colleagues observed increased Citrobacter phage counts, while others did not look at the phage population. The largest study of gut phage and viruses in humans did not find a tight link between viral population and regulation of bacterial population [195]. However, by analogy with aquatic environments [196,197] it is likely phages act as regulators of microbiota structure in the gut. Treatments may need to target not one, but several possible pathogens, and do so without adversely affecting beneficial microbiota.

Lactoferrin is an antimicrobial (reviewed by Jenssen et al [198]) produced by our immune system [199,200] that stimulates gut epithelial cell growth, proliferation and differentiation [183,184], and also displays antiviral properties [201]. The use of lactoferrin may prevent infant NEC [182] by inducing proliferation and differentiation of intestinal cells but cannot cure this illness [202]. A recent study of NEC showed a link between the use of the vegetal antimicrobial trans-cinnamaldehyde and the inactivation of the pathogen Enterobacter sakazakii [185]. This could prevent NEC in infants. However in preterm infants, NEC may be related to an allergic reaction to cow milk [186,187]. The question of whether this chain of events can be prevented or disrupted by appropriate microbiota remains
4.3.1.2 Allergic Colitis

Allergic colitis is often due to antigens against cow’s milk [203,204]. Lactobacillus is thought to preserve the infant intestine by favoring antigen degradation. Lactobacillus preserves the intestinal mucosal barrier, competing with pathogenic bacteria. It also enhances the production of cytokines IL-10 and TGF-β [205,206], both involved in the process of immune tolerance during inflammation or allergy [207,208]. Savino and colleagues observed improved gut motility and decreased pain in 24 of 25 infants suffering colitis after 21 days of Lactobacillus diet complementation, but only 15 of 21 children treated with placebo [181,188,189]. Infants given probiotic containing Lactobacillus were better protected against rectal bleeding and allergic colitis. The importance of Lactobacillus in immune protection against inflammatory colitis was documented by Nermes and colleagues [190]. They added Lactobacillus to the diet of 19 infants between 3 and 13 months of age during three months exclusively formula-fed. Compared with 20 controls, the Lactobacillus infants had decreases in the number of cells secreting IgA and IgM (by 7% and 20% respectively) while the controls had increases in the same cell populations (by 22% and 31% respectively) between the beginning and the end of the study. At the same time the Lactobacillus-enriched diet correlated with an increase of CD19+ and CD27+ B memory cells compared with controls. The authors observed no difference in Bifidobacterium populations between the treated and the control groups, both of which were breast-fed before the study. Lactobacillus supplementation reduced the level of calprotectin, a marker of cow’s milk allergic colitis, and allowed a better recovery of the intestinal mucosa inflammation in infants suffering from hematochezia [189]. Arvola and colleagues reported a group of infants suffering rectal bleeding who were largely breast-fed [192]. The cow’s milk elimination diet in a subset of infants did not affect the duration of rectal bleeding. Compared to the control group, infants suffering from rectal bleeding had lower bacteria counts and their populations of Bifidobacterium and Lactobacillus were around ten times lower than in healthy infants. The authors thus suggested the possibility of probiotic intervention aimed at normalizing the level of bifidobacteria and lactobacilli.

4.3.2 Dangers of antibiotherapy

The use of antibiotics continues to increase in human medicine [209]. In addition, low dose antibiotics are used routinely to increase the weight and growth rate of livestock.
The widespread use of antibiotics resulted in the emergence of multidrug resistant bacteria and possibly in changes in the persistent human-associated microbial communities. At present, immune development is influenced by a lower diversity of bacteria than it was the case in previous generations [210]. Animal models suggest the microbiota shift associated with increased antibiotic use may increase the incidence of obesity [210] which has reached pandemic proportions among humans. The use of antibiotics affects the colonization of the infant gut by a potential pathogen S. aureus [103,211,212]. Lindberg and colleagues found a higher prevalence of S. aureus colonization in the gut of Swedish than in Italian infants until the age of one year. This difference was not related to the delivery mode or the diet and was attributed to a common use of antibiotics effective against S. aureus in Italy, while penicillin V, inefficient against S. aureus, is mostly used on Swedish infants. Not only the use of antibiotics may create new niches for resistant bacteria but may also eradicate useful bacteria like Lactobacillus, or favor the overgrowth of pathogenic bacteria like C. difficile [213,98,107]. Therefore, the use of antibiotics in preterm infants or babies delivered by C-section may bias studies on C. difficile and S. aureus carriage and their relationship with disease. While several studies suggest that changes in the overall bacterial community are associated with S. aureus pathogenicity, we do not know if the changes in bacterial community are causative, as they are thought to be with C. difficile.

Treating pneumonia with ceftriaxone for 5 days in babies < 6 months old results in extirpation of Lactobacillus from infants’ gut microbiota, as well as a decreased population of other commensal bacteria like Enterobacteriaceae [98]. However, 20 days after the end of the treatment, the gut bacterial communities were greatly recovered and included Lactobacillus that had disappeared during the treatment. Recent research on adult gut microbiota modifications following 4 days of ciprofloxacin treatment repeated twice over 2 months reported a likely permanent shift in the rare gut microbiota species [214]. However, some of the newly acquired species were close to those present before the treatment. Recently, Hvid and colleagues pointed out the relationship between the number of courses of antibiotics and risk of inflammatory bowel disease in infants [215]. Difficulties of the gut microbiota to recover after antibiotic therapy in early life highlights the importance of the initial gut microbiota establishment and a potential danger of antibiotic (over)use [216].

5. Viral gut community

5.1 Who are they?

The human microbiota is associated with an abundant and diverse community of viruses, in particular phages [217-220,59]. These phage populations are likely to play an important role in fast developing infant gut microbiome diversity [59]. However, a relative stability of both, phage and bacterial populations in adult feces samples led to the hypothesis that a “predatory viral microbial dynamic” does not exist in the adult human gut [195]. The interaction of phages and bacteria in the gut remains relatively poorly understood. One of the main issues in identification of phages, and viruses in general, is the lack of sequence similarity between metagenomic sequence reads and known viral genomes [195].

5.2 Relationship between viruses and bacteria in the gut

Phage communities in Western infants’ gut present low diversity but they might influence the abundance of the microbial community [59]. Breitbart and colleagues suggested that fecal phages do not originate from a dietary source as they are different from the phages present in mothers’ milk or in formulas [59]. They hypothesize that the first viruses in the gut originate from prophages induction from the first colonizing bacteria rather than from an environmental source. Persistence of the phage population in infants’ gut over extended periods raised the hypothesis of a completely different mode of colonization than that of bacteria. Indeed no significant clustering of viral population could be observed between co-twins or between twins and their mothers [195] while their bacterial populations were closer compared to those of unrelated people [146]. However, with a different approach, using microarrays, Breitbart et al reported a dramatic shift in the gut viral population between 1 and 2 weeks of life [59], related to an important modification in bacterial population [221]. Therefore, the results obtained in different studies should be considered in the light of the methodology used. Combining the data generated using different methodologies will improve our understanding of the gut viral communities.

6. Outlook

Metagenomic studies of infants’ gut microbiota provide important insights into understanding bacterial colonization of the gut and its role in human health. However, differences in sampling procedures and methodologies used as well as the lack of data on previous use of antibiotics can complicate comparisons of the results from different studies. While metagenomics allows to reveal changes in bacterial communities as a function of geography, diet and medical treatments, in many instances the causal relation between altered microbiota and a disease remains unclear. To deal with this issue, long-term studies with large cohort are needed. Given the complexity of the interactions between the host and environmental factors, and members of the gut bacterial community, one of the challenges is to define normal and abnormal gut microbiota. In that light, the exposure of germ-free mice to defined bacterial consortia under controlled conditions may provide insights into the
impact of early bacterial colonization on the host physiology. As shown for prebiotics and probiotics, bacteriophage supplementation in infant’s diet might have health-promoting effect. However, possible benefits and risk of phage-, prebiotic- and probiotic supplements on the health in later life remain to be determined.

The prevalence and incidence of allergies and autoimmune diseases have been increasing over recent decades. Metagenomics provides tools to relate these conditions to the gut microbiota composition and structure which may contribute to the development of novel prophylactic and therapeutic approaches.

Acknowledgments

This work was supported by funding from the Swiss National Science Foundation (Grant 3100A0-112370/1 to J.S.) and from bioMérieux.

References
