DOCTORAL THESES IN FOOD SCIENCES AT THE UNIVERSITY OF TURKU

New Insights into Human Gut Microbiota Development in Early Infancy: Influence of Diet, Environment and Mother's Microbiota

JUHANI AAKKO

Food Chemistry and Food Development
Department of Biochemistry

DOCTORAL THESES IN FOOD SCIENCES AT THE UNIVERSITY OF TURKU
Food Development (tech.)

New Insights into Human Gut Microbiota Development in Early Infancy: Influence of Diet, Environment and Mother's Microbiota

JUHANI AAKKO

Functional Foods Forum
and
Food Chemistry and Food Development
Department of Biochemistry

TURKU, FINLAND – 2016

Food Chemistry and Food Development
Department of Biochemistry
University of Turku, Finland

Supervised by
Professor Seppo Salminen, Ph.D.
Functional Foods Forum
University of Turku
Turku, Finland

Akihito Endo, Ph.D.
Functional Foods Forum
University of Turku
Turku, Finland and
Department of Food and Cosmetic Science
Tokyo University of Agriculture
Hokkaido, Japan

Reviewed by
Konrad Domig, P.D. Dr.
Department of Food Science
and Technology
University of Natural Resources
and Life Sciences
Vienna, Austria

Karen Scott, Ph.D.
Rowett Institute of Nutrition and
Health
University of Aberdeen
Aberdeen, Scotland, UK

Opponent
Professor Cristophe Lacroix, Ph.D.
Department of Health Sciences and Technology
ETH Zürich
Zürich, Switzerland

The originality of this dissertation has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN: 978-952-93-8182-1 (paperback)
ISBN: 978-952-93-8183-8 (PDF)
ISSN: 2323-9395 (paperback)
ISSN: 2323-9409 (PDF)
Painosalama Oy – Turku, Finland 2016

TABLE OF CONTENTS

ABSTRACT

The human gut is home to an enormous number of microbes that play an important role in host physiology, metabolism, nutrition and immune function. Deviations in initial human gut microbiota development have been associated with a number of immune-mediated diseases including inflammatory bowel diseases, atopy and allergies, as well as obesity, which are reaching epidemic proportions in the westernized world. However, gut microbiota dysbiosis has also been linked to health problems common in the developing world including malnutrition. The aim of this thesis was to unravel the impact of specific factors, including environment, geographical location, diet and mother's microbiota, on the gut microbiota development in early infancy.

Analysis of gut microbiota composition of six-month-old infants from Malawi and Finland, representing developing and westernized countries, revealed that the Malawians had a distinct *Lactobacillus* microbiota that was richer and more diverse compared to their Finnish counterparts. Interestingly, we detected *L. rhamnosus* GG, a widely used probiotic in Finland, in a third of the Finnish infants even though the infants or their mothers did not receive any probiotic products. For the detection, we used a *L. rhamnosus* GG strain-specific detection system, whose specificity was confirmed in this thesis. In addition, we noted that lipid-based nutrient supplements do not have an impact on gut microbiota development in Malawian infants. These dietary supplements have been used in several clinical trials to promote growth of children from developing countries, but their impact on the gut microbiota has been previously unknown. In addition to environment and diet, a number of host factors may have an impact on the gut microbiota development. Here, we assessed how mother's microbiota influences the gut microbiota development in early infancy by comparing samples of maternal feces, placenta, amniotic fluid, colostrum, meconium and infant feces from mother-infant pairs. Our results suggest that the microbial colonization may begin already during fetal life. After birth, colostrum seems to influence the gut microbiota development of the neonate by serving as a source of microbes and microbiota modulating factors, such as human milk oligosaccharides. We also noted that human milk oligosaccharides and microbiota of the colostrum are interrelated.

Taken together, this thesis provides further knowledge on how different factors influence gut microbiota development in early life. Modification of gut microbiota composition during its development may be a viable strategy to prevent or treat inflammatory non-communicable diseases in the future.

TIIVISTELMÄ

Ihmisen suolistossa asuu valtava määrä mikrobeja, jotka vaikuttavat merkittävästi ihmisen fysiologiaan, metaboliaan, ravitsemukseen ja immunijärjestelmän toimintaan. Suoliston mikrobiston kehityksen häiriöiden on todettu olevan yhteydessä lukuisiin immuunivälitteisiin sairauksiin, kuten tulehduksellisiin suolistosairauksiin, atopiaan, allergioihin ja liikalihavuuteen. Nämä sairaudet ovat yleistyneet etenkin länsimaisissa hyvinvointivaltioissa viime vuosikymmeninä. Lisäksi viime aikoina on havaittu, että suoliston mikrobiston häiriöt ovat yhteydessä kehittyvissä maissa yleisiin sairauksiin kuten aliravitsemukseen. Tämän väitöskirjatyön tavoitteena oli tutkia eri tekijöiden kuten ympäristön, maantieteellisen sijainnin, ruokavalion ja äidin mikrobiston vaikutusta suoliston mikrobiston kehitykseen varhaislapsuudessa.

Malawilaisten ja suomalaisten kuuden kuukauden ikäisten lasten suolistomikrobiston analyysi osoitti, että malawilaisilla oli rikkaampi ja monipuolisempi *Lactobacillus* –mikrobisto, jonka koostumus oli selkeästi erilainen kuin suomalaisilla. Tutkimuksessa havaittiin yllättäen, että kolmasosalla suomalaisista lapsista oli *L. rhamnosus* GG probioottia suolistossa, vaikka lapset tai äidit eivät olleet nauttineet probioottituotteita. Havainto tehtiin käyttäen *L. rhamnosus* GG laji-spesifistä PCR-pohjaista menetelmää, jonka tarkkuus arvioitiin ja todettiin hyväksi tässä väitöskirjatyössä. Lisäksi osoitettiin, että rasvapohjaiset ravintolisät eivät vaikuta malawilaisten lasten suolistomikrobiston kehitykseen. Rasvapohjaisilla ravintolisillä on pyritty ehkäisemään lasten aliravitsemusta kehitysmaissa, mutta niiden vaikutusta suoliston mikrobistoon ei ole aikaisemmin tunnettu. Ympäristön ja ruokavalion lisäksi monet isännästä riippuvat tekijät voivat vaikuttaa suoliston mikrobiston kehitykseen. Tässä työssä selvitettiin äidin mikrobisto vaikutusta lapsen mikrobiston kehitykseen varhaislapsuudessa vertailemalla mikrobistonäytteitä äidin ulosteesta, istukasta, lapsivedestä, äidinmaidosta, mekoniumista, sekä lapsen ulosteesta. Tulokset osoittivat, suoliston kolonisaatio alkaa mahdollisesti jo ennen syntymää kohdussa. Tämän jälkeen lapsen mikrobiston kehitykseen vaikuttavat rintamaidon sisältämät mikrobit ja mikrobistoa muokkaavat tekijät, kuten äidinmaidon oligosakkaridit. Työssä havaittiin myös, että äidinmaidon oligosakkaridit ja mikrobisto ovat yhteyksissä toisiinsa.

Tämä väitöskirjatyö tuotti uutta tietoa siitä kuinka eri tekijät vaikuttavat suolistomikrobiston kehitykseen. Suoliston mikrobiston muokkaaminen kehityksen eri vaiheissa voi olla yksi toimiva vaihtoehto tulehduksellisten ei-tarttuvien sairauksien ehkäisemiseen ja parantamiseen tulevaisuudessa.

LIST OF ABBREVIATIONS

2'FL, 2'-fucosyllactose
3FL, 3-fucosyllactose
3'SL, 3'-sialyllactose
ANOVA, analysis of variance
ANCOVA, analysis of covariance
CI, confidence interval
DGGE, denaturing gradient gel electrophoresis
DFLNH, difucosyllacto-N-hexaose
DF-LNT, difucsyllactose-N-tetraose
DSLNH, disialyl-lacto-N-hexaose
DS-LNT, disialyllacto-N-tetraose
ERIC-PCR, enterobacterial repetitive intergenic consensus-PCR
FDSLNH, fucosyldisialyllacto-N-hexaose
FLNH, fucolacto-N-hexaose
FOS, fructo-oligosaccharide
FUT, fucosyltransferase
GEE, generalized estimating equations
GOS, galacto-oligosaccharide
HMO, human milk oligosaccharide
IBD, inflammatory bowel diseases
IgA, immunoglobulin A
IQR, interquartile range
LNFP, lacto-N-fucopentaose
LNT, lacto-N-tetraose
LNnT, lacto-N-neotetraose
LNS, lipid-based nutrient supplement
LST, sialyllactose-N-tetraose
MAMP, microbe-associated molecular pattern
PCA, Principal component analysis
PCoA, Principal coordinates analysis
PCR, polymerase chain reaction
PRR, pattern recognition receptor
PSA, polysaccharide A
qPCR, quantitative PCR
RAPD, randomly amplified polymorphic DNA
SCFA, short-chain fatty acid
SD, standard deviation
SFB, segmented filamentous bacteria
Spp., species

Subsp., subspecies
Th, helper T cell
TLR, toll like receptor
UPGMA, Unweighted Pair Group Method with Arithmetic Mean
16S rRNA, Bacterial ribosomal 16S RNA molecule

LIST OF ORIGINAL PUBLICATIONS

I. Endo A, Aakko J, Salminen S. Evaluation of strain-specific primers for identification of Lactobacillus rhamnosus GG. *FEMS Microbiol Lett* 2012, 337: 120–125.

II. Aakko J, Endo A, Mangani C, Maleta K, Ashorn P, Isolauri E, Salminen S. Distinctive Intestinal Lactobacillus Communities in 6-Month-Old Infants From Rural Malawi and Southwestern Finland. *J Pediatr Gastroenterol Nutr* 2015, 61: 641–648.

III. Aakko J, Grzeskowiak, L, Asukas, T, Päivänsäde, E, Lehto, KM, Fan YM, Mangani C, Maleta K, Ashorn P and Salminen S. Lipid-based nutrient supplements do not affect gut *Bifidobacterium* microbiota in Malawian infants: a randomised trial, *J Pediatr Gastroenterol Nutr.* (In press)

IV. Collado MC, Rautava S, Aakko J, Isolauri E, Salminen S. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. *Sci Rep* 2016, 6: 23129.

V. Aakko J, Kumar H, Rautava S, Wise, A, Autran, C, Bode, L, Isolauri E, Salminen S. The impact of human milk oligosaccharides on specific microbes in human breast milk. *Submitted*

1 INTRODUCTION

The colonization of the human gastrointestinal tract is a gradual process that is affected by host genotype, mother's microbiota, mode of delivery, gestational age, early feeding practices and the environment. Early microbial contact is essential for healthy intestinal and immunological development during infancy and thus can have long-term effects on the risk of gastrointestinal, immune-mediated and metabolic disease in later life. It has been hypothesized that a lack of microbial contact during infancy is one of the reasons for the increase in allergic and other chronic inflammatory diseases such as obesity and related disorders in the modern societies (Strachan, 1989; Isolauri *et al.*, 2009). In addition, deviations in gut microbiota composition have been also linked to diseases common in the developing world such as bacterial or viral diarrhea, and, more recently, malnutrition (Gordon *et al.*, 2012; Smith *et al.*, 2013).

During first years of life, the human microbiota is a dynamic ecosystem dominated by lactobacilli and bifidobacteria (Bäckhed *et al.*, 2015; Koenig *et al.*, 2011). Recent reports suggest that fetus may be exposed to microbes or microbial parts already *in utero*: microbial DNA in amniotic fluid and placenta has been noted by several studies (Satokari *et al.*, 2009; Rautava *et al.*, 2012b). However, current knowledge on the role of intrauterine microbes in human gut colonization is scarce. Following birth, breastfeeding is a major source of viable microbes and microbiota modulating factors, including human milk oligosaccharides, to the newborn infant (Jost *et al.*, 2015). Moreover, the gut microbiota may be affected by hygiene and sanitation and other environmental factors such as the presence of animals.

In the present study, we aimed to characterize the transfer of microbiota from mother to infant more thoroughly than before and to reveal other factors, such as geographic location and diet, that influence the gut microbiota development in early infancy. Comprehensive understanding of the mechanistic basis of microbiota development and the factors that have an impact on it may reveal new therapeutic approaches that could be used to modulate the gut microbiota to treat or prevent non-communicable diseases including allergic, autoimmune and inflammatory disorders.

2 REVIEW OF THE LITERATURE

2.1 Human gut microbiota composition

An average human body harbors approximately the same amount of microbial cells as eukaryotic cells and their genes outnumber the human genes by more than 100-fold (Sender *et al.*, 2016; Human Microbiome Project Consortium, 2012). The human microbiota is composed of all three domains of life: Bacteria, which predominate, Archaea, and Eukarya, and viruses.

The majority of the human microbiota resides in the gastrointestinal tract. Nevertheless, relatively few bacterial phyla reside in the human gut compared to communities in other body sites. The adult gut microbiota consists predominantly of two bacterial phyla, Firmicutes and Bacteroidetes, which are followed by Actinobacteria, Fusobacteria, Proteobacteria and Verrucomicrobia (Eckburg *et al.*, 2005). It has been estimated that the gut microbiota consists of an average of 500-1000 different microbial species with great variation between individuals (Jalanka-Tuovinen *et al.*, 2011). During the first 2-3 years of life the human gut microbiota is in a dynamic state progressively changing and becoming more diverse after which it stabilizes, and remains rather stable during adulthood. At later stages of life the microbiota composition becomes less diverse and again more dynamic (Ottman *et al.*, 2012).

The concentration and diversity of microbiota increases when moving from the esophagus distally to the rectum. In the esophagus and stomach, the number of bacteria can be around 10 per gram whereas in the colon and distal gut the bacterial cell count can reach 10^{12} per gram of contents (O'Hara and Shanahan, 2006). Different bacterial groups are known to colonize different parts and sites in the gastrointestinal (Figure 1). Stomach and small intestine are more harsh environments for the microbes compared to colon because of low pH and presence of bile salts, respectively. In addition to being bile tolerant the small intestinal microbiota reflects a community that can quickly respond to varying carbohydrate availability (Zoetendal *et al.*, 2012). Based on recent studies, the small intestine is mainly colonized by Bacilli (*Streptococcus* sp.), *Clostridium cluster* XIVa and coliforms (Zoetendal *et al.*, 2012). The microbiota in colon is more stable, but more diverse than in the small intestine. The colon is mainly populated by strict anaerobic microorganisms which form a complex and dynamic ecosystem (Eckburg *et al.*, 2005).

In addition to differences in microbiota between different longitudinal parts of gastrointestinal tract, there is also variation in microbiota at latitudinal level of the intestine. The intestinal epithelium is separated from the lumen by a thick layer of mucus, which serves as a barrier that protects the underlying

epithelial cells from pathogens as well as physical and chemical damage. The luminal and mucosal parts of the intestinal lining have been shown to differ in microbiota composition (Zoetendal *et al.*, 2002). The luminal microbiota tends to be dynamic and it is mainly affected by diet, while the mucosal microbiota is more stable. The luminal part has two specific layers: digesta and inter-fold regions. Species of the families Bacteroidaceae, Prevotellaceae and Rikenellaceae populate the digesta while the inter-fold regions are colonized by species of Lachnospiraceae and Ruminococcaceae families (Donaldson *et al.*, 2016). Also the mucus layer can be also divided into two distinct intersections: the loose outer layer and the dense inner layer close to the epithelium. The loose outer layer of mucus is populated by mucin-degrading bacteria, such as *Bacteroides acidifaciens, Bacteroides fragilis* and *Akkermansia muciniphila* (Salyers *et al.*, 1988; Derrien *et al.*, 2004; Donaldson *et al.*, 2016). The inner layer of mucus has been earlier described as devoid of bacteria (Johansson *et al.*, 2008), but recent studies suggest that low levels of *Bacteroides fragilis* and *Acinetobacter* spp. may reside closer to the epithelium (Donaldson *et al.*, 2016; Pédron *et al.*, 2012).

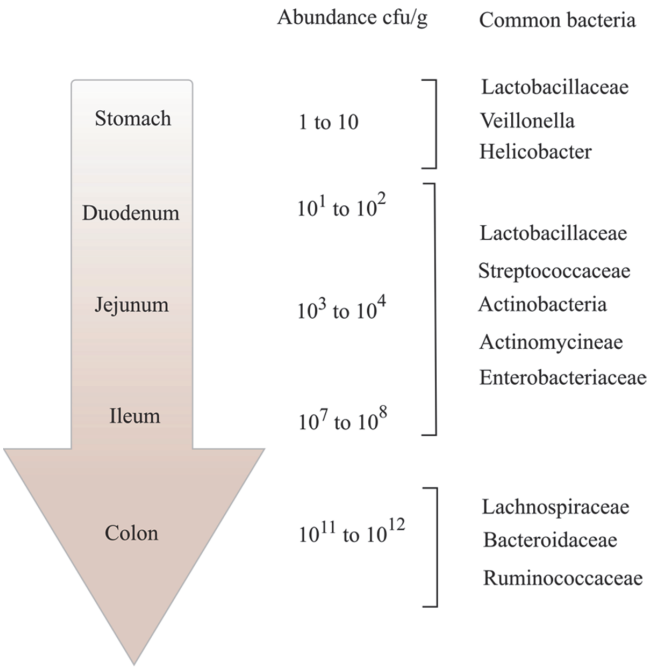


Figure 1. Variations in microbial numbers and composition across the human gastrointestinal tract

2.2 General functions of the human gut microbiota

The gut microbiota could be considered as an organ placed within an organ since it plays an essential role in the metabolic, nutritional, physiological and immunological processes of the human body (O'Hara and Shanahan, 2006). The microbiota extracts energy from otherwise indigestible dietary polysaccharides such as resistant starch and different dietary fibers. This leads to the production of important nutrients such as short-chain fatty acids, vitamins and amino acids, many of which humans are unable to produce themselves (LeBlanc *et al.*, 2013; Neis *et al.*, 2015; Cummings, 1984). The microbiota also participates in the defense against pathogens by producing antimicrobial compounds and by providing so called colonization resistance, which means that the microbiota protects itself directly or indirectly against incursion by new and often pathogenic microorganisms (Buffie and Pamer, 2013). Commensal microbes also have an impact on the normal development and maintenance of the gastrointestinal sensorimotor functions (Barbara *et al.*, 2005). Moreover, the gut microbiota has an important role in the normal development of the mucosal immune system, the induction of IgA production and epithelial barrier tightening (Ulluwishewa *et al.*, 2011; Macpherson *et al.*, 2008; Lepage *et al.*, 2013; Sutherland *et al.*, 2016).

2.3 Gut microbiota in health and disease - the extended hygiene-hypothesis

The so-called hygiene hypothesis suggested that a lack of exposure to infectious agents such as microorganisms and parasites in childhood may suppress immune system development and lead to increased prevalence of allergies and atopic disease (Strachan, 1989). Since the days of Strachan, we have learned that the gut microbiota is central to this phenomenon. The human gut microbiota and its host have co-evolved, but due to changes in our lifestyle and environment, we are losing key players of the microbiota as well as reducing overall diversity (Blaser and Falkow, 2009; Ley *et al.*, 2008). It also seems that some microbes traditionally considered as pathogens, such as *H. pylori*, may in fact be protective against allergic disease for instance, and thus their systematic eradication may have surprising consequences (Seiskari *et al.*, 2007; Ozerlat, 2011). Loss of specific microbial species that have a role in human health and decrease in overall diversity of the microbiota may carry over from generation to generation, possibly in an irreversible manner (Sonnenburg *et al.*, 2016).

Dysbiosis in the gut microbiota appears to be associated with more diseases than before, but the mechanisms are not well understood. In addition to allergic

disease, relatively strong evidence supports the role of gut microbiota dysbiosis in inflammatory bowel diseases, celiac disease, diabetes and obesity (Vos and Vos, 2012). Interestingly, associations between central nervous system diseases such as autism and Alzheimer's disease have also been proposed (Toh and Allen-Vercoe, 2015; Bhattacharjee and Lukiw, 2013). However, it is important to recognize that the so-called gut-brain axis is bidirectional and until now only associations have been reported (Collins and Bercik, 2009). Nevertheless, studies conducted on laboratory animals suggest that gut microbes may actually alter behaviour via influencing the vagus nerve (Bravo *et al.*, 2011).

Although it is clear that microbiota plays an important role in human health and disease, there is still no consensus on what is a "healthy" gut microbiota composition or if it even exists. More detailed studies are needed and also work reporting compositional differences of genetic origin or geographical location.

2.4 Host-microbe interactions

2.4.1 Lessons from germ-free animals

The gastrointestinal tract is the primary site of interaction between the host immune system and microorganisms, both symbiotic and pathogenic. The gut microbiota plays a key role in the maturation of the immune system and the structure and function of the gastrointestinal tract. Our understanding of early host-microbe interactions is mainly based on studies utilizing germ-free or gnotobiotic animals. Germ-free animals are more susceptible to infection by various pathogens because they have various developmental defects of the immune system. For instance, they show deficits in the development of gut-associated lymphoid tissues (GALT), numerous defects in the antibody production as well as fewer and smaller Peyer's patches and mesenteric lymph nodes (Round and Mazmanian, 2009). However, they can provide mechanistic answers to microbiota-host interactions. Nevertheless, one has to be careful when extrapolating results from studies conducted with mice to humans.

One famous study utilizing germ-free animals to demonstrate the role of gut microbiota in immune system development was conducted by Sudo et al. (1997). In this pioneering study the authors showed that monocolonizing a germ-free mouse with a *B. infantis* strain during the neonatal period could induce immune tolerance in an animal that normally lacks one. In the same study the authors demonstrated that colonization at an older age did not lead to immune tolerance. This indicates that an early microbial exposure is needed for the normal development of immune tolerance. In a similar vein, the impact of the gut microbiota on the function and structure of gastrointestinal tract has been demonstrated by monocolonizing germ-free mice with *Bacteroides*

thetaiotaomicron. The monocolonization had an impact on the expression of various host genes that influence nutrient uptake, metabolism, angiogenesis, mucosal barrier function and the development of the enteric nervous system (Xu and Gordon, 2003).

In addition to monocolonization of germ-free mice, whole microbiotas have been transplanted to see the impact of gut microbiota on host health. For instance, the causal relationship between gut microbiota dysbiosis and obesity was confirmed when gut microbiota of obese lep^{ob} mice and lean wild-type mice was transplanted to germ-free mice (Turnbaugh *et al.*, 2006). This study demonstrated that there was a greater increase in body fat for mice colonized with a gut microbiota associated with obesity. Later, similar findings were reported when microbiota was transplanted from obese humans and their lean twins to germ-free mice (Ridaura *et al.*, 2013).

2.4.2 Mechanisms of immunomodulation by gut microbiota

Various pattern recognition receptors (PRRs), such as Toll-like receptors and nucleotide-binding oligomerization domain- (NOD-) like receptors (NLRs), are key mediators in the communication between the host and the gut microbiota (Figure 2). The TLRs are mostly present on the membrane of immune and epithelial cells (Kawai and Akira, 2006) while NODs are present in the cytoplasm of enteric cells (Lavelle *et al.*, 2009). These receptors are evolutionary conserved molecules that promote immune responses by recognizing conserved molecular motives such as microbe-associated molecular patterns (MAMP) and pathogenic-associated molecular patterns (PAMP) (Kubinak and Round, 2012). The TLR family consists of eleven different transmembrane proteins, which are specialized in recognizing different microbial molecules. These molecules include lipopolysaccharides on the cell membrane of Gram-negative bacteria (TLR4), peptidogylcan (TLR2), bacterial flagellin (TLR5), lipoteichoic acid of Gram-positive bacteria (TLR1 and TLR6) and unmethylated CpG-motifs (TLR9) and muramyl dipeptide (NOD2).

Specific bacteria have been shown to shape the intestinal immune response via specific TLRs. For instance, TLR2 has been shown to be activated by a unique surface polysaccharide (polysaccharide A, PSA) of *Bacteroides fragilis* (Figure 2) (Round *et al.*, 2011). Before this mechanism was known, it was noticed that PSA can correct the systemic T cell deficiencies in germ-free mice (Mazmanian *et al.*, 2005). Later, it was shown that the PSA expressed by *B. fragilis* regulates the development of $Foxp3^+$ regulatory T (T_{reg}) cells while suppressing the T_H17 responses (Round and Mazmanian, 2010; Round *et al.*, 2011). The T_{reg} and T_H17 possess opposing actions and the imbalance between

these two has been suggested to be a contributing factor in various immune-mediated diseases (Eisenstein and Williams, 2009).

The gut microbiota derives various compounds from food that the host consumes. For instance, microbes produce short-chain fatty acids, such as acetate, propionate and butyrate, from non-digestible dietary carbohydrates and thus their levels are highly dependent on the diet of the host as well as the community structure of the microbiota (David *et al.*, 2014; Flint *et al.*, 2015; Scott *et al.*, 2013). After their production, SCFA are absorbed and used for different biosynthetic purposes by the host. For instance, epithelial cells in the colon obtain most of their energy from these bacterial metabolites (Macfarlane and Macfarlane, 2012). SCFAs, butyrate in particular, also play an important role in the maintenance of gut barrier function: both by increasing mucin production which may result in changes on bacterial adhesion (Jung *et al.*, 2015) and improving tight-junction integrity (Peng *et al.*, 2009). Moreover, butyrate is known to regulate the proliferation and activation of regulatory T cells in the colon (Figure 2) (Furusawa *et al.*, 2013). Interestingly, SCFAs may also have a critical role in appetite regulation and energy homeostasis (Frost *et al.*, 2014; Lin *et al.*, 2012).

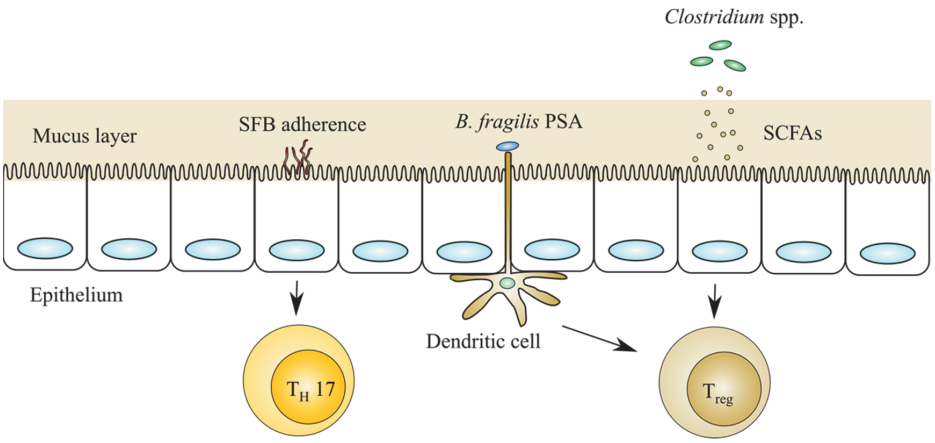


Figure 2. Immunomodulation by gut microbes. PSA, polysaccharide A; SCFA, short chain fatty acid; SFB, segmented filamentous bacteria

2.5 Development of gut microbiota

The microbiota of a healthy, breastfed, vaginally delivered, full-term infant who remains healthy for a longer follow-up time is considered as the gold standard for microbiota development (Collado *et al.*, 2015). Microbiota during infancy has been described as less diverse and more dynamic than an adult's microbiota and it is dominated by bifidobacteria and lactobacilli (Yatsunenko

et al., 2012; Ottman *et al.*, 2012). As the infant ages, the diversity of the microbiota increases with the emergence and dominance of the phyla Firmicutes and Bacteroidetes (Rodríguez *et al.*, 2015; Ottman *et al.*, 2012).

Since the microbes living in human gut are not found in other environments, it seems logical that the human infant inherits a majority of its microbiota from its mother (Rautava *et al.*, 2012b; Ley *et al.*, 2006). It has been proposed that the mother's gut microbiota changes towards a composition that is beneficial for both the mother and the infant (Koren *et al.*, 2012). During pregnancy, the gut microbiota of a pregnant mother has been shown to change particularly during the thid trimester and it is characterized by a decreased diversity with a decrease in the abundance of *Faecalibacterium* spp. and an increase in Proteobacteria, *Streptococcus* spp. and Actinobacteria (Koren *et al.*, 2012). Apart from mother's microbiota, type of feeding, environment and the use of antibiotics, probiotics and prebiotics may influence the microbiota development of the infant.

2.5.1 Is there microbial contact *in utero*?

Since the studies conducted by Tissier (Tissier, 1900) in early 20th century concerning the acquisition of gut microbiota during infancy, the idea that the fetus is sterile and that microbial colonization starts after birth was widely accepted. When thinking from an evolutionary point of view, this dogma of the microbiologically and immunologically naïve fetal life seems bizarre, since humans among other mammals have evolved in a world already inhabited by a colossal variety of different microorganisms (Rautava *et al.*, 2012b). Currently, there is increasing evidence, although still lacking, that the fetus may be exposed to microbes or microbial parts already *in utero*. Thus, it has been suggested, that the intrauterine environment provides the embryo and fetus a controlled environment in which there is limited and selective exposure to microbes during the vulnerable periods of early development (Rautava *et al.*, 2012b).

Microorganisms in the amniotic fluid and placenta have been mainly associated with intrauterine infections and preterm deliveries (Goldenberg *et al.*, 2008). Anyhow, bacterial DNA has been reported also in placenta and amniotic fluid in normal conditions (Satokari *et al.*, 2009; Rautava *et al.*, 2012a; Aagaard *et al.*, 2014). In addition, intact bacteria have been detected in the basal plate of placenta in both term and preterm gestations (Stout *et al.*, 2013) and bacterial cells have been obtained by culturing from amniotic fluid (DiGiulio *et al.*, 2008).

In addition to amniotic fluid and placenta, meconium, which was also thought to be sterile in the past, seems to have a complex, albeit low abundance,

microbiota (Dominguez-Bello *et al.*, 2010; Gosalbes *et al.*, 2013; Mshvildadze *et al.*, 2010). In fact, the meconium microbiota has been reported to reflect maternal health and also to be associated with the development of disease in offspring (Gosalbes *et al.*, 2013). However, often it is not clear how long after membrane rupture or birth the meconium has been passed, and thus it cannot be ruled out that the bacteria in the meconium origin from elsewhere than the intrauterine environment. Moreover, it must be kept in mind that DNA-based methods are prone to contamination and possible background DNA when assessing microbiota in low-biomass sites such as placenta or meconium (Salter *et al.*, 2014). Therefore proper negative controls should be included and the results interpreted with caution since they may provide a distorted image of the real bacterial population in the low-biomass samples.

If assumed that the intrauterine environment is not sterile, the route that bacteria may use to enter the amniotic cavity remains to be elucidated. It has been suggested that microbes might ascend from the vagina and the cervix, haematogenously disseminate through the placenta, or spread through the fallopian tubes from abdominal cavity (Goldenberg *et al.*, 2008). It seems that bacterial DNA is frequently detected in the amniotic fluid already during the second trimester and is not always associated with adverse outcome or inflammation (Markenson *et al.*, 2003). It is also known that non-pregnant women have a microbiota in the uterus with a distinct composition compared to the vaginal microbiota (Andrews *et al.*, 2005; Verstraelen *et al.*, 2016).

Interestingly, bacterial efflux from mother to fetus has been demonstrated in mice. Jimenez and coworkers inoculated a genetically labeled *Enterococcus faecium* strain orally to pregnant mice and later isolated it from the amniotic fluid of the inoculated animals (Jimenez *et al.*, 2005, 2008). Also, specific bacteria have been isolated from the cord blood from healthy neonates born by cesarean section (Jimenez *et al.*, 2005) suggesting the possible mother-to-infant efflux of commensal microbes via blood in humans also.

One of the main arguments against the *in utero* colonization hypothesis is that we are able to derive “germ-free” animals by hysterectomy (Arvidsson *et al.*, 2011) or embryo transfer (Inzunza *et al.*, 2005), and thus *in utero* colonization would not happen. However, this does not exclude the possibility that, during fetal life, humans or other animals may be exposed to microbial parts or non-proliferating microbes, which may have an impact on the development of the immune system for instance. As indicated by a study by Rautava and coworkers (2012a), the microbe-associated molecular patterns (MAMPs), such as bacterial DNA, can induce immuno-stimulatory effects via the stimulation of TLRs without the need for microbial cells to enter the amniotic cavity (Rautava *et al.*, 2012a). Data also suggests that fetal intestine

innate immune gene expression can be modulated by maternal supplementation with specific probiotic bacteria.

Also, the maternal exposure to a variety of microbes during pregnancy can have an impact on the infant's health. Maternal exposure to farm animals appears to protect the infant from asthma, hay fever and eczema, although continued exposure may be required for optimal protection (Douwes *et al.*, 2008; Roduit *et al.*, 2011). In addition, antibiotic exposure during pregnancy has been associated with a risk of developing asthma later in childhood (Stensballe *et al.*, 2013). In contrast, probiotic interventions aiming to reduce the risk of atopic dermatitis appear to be more effective when conducted perinatally compared to prenatal or postnatal interventions only (Doege *et al.*, 2012). However, the exact mechanisms remain unknown.

2.5.2 Mode of delivery

The major colonization process of the gastrointestinal tract starts at birth when maternal vaginal and fecal microbiota serve as inocula to the newborn child. During cesarean section, this direct contact is absent and the infant is exposed to environmental microbes instead, which can lead to disturbances in the gut microbiota development. Epidemiological and experimental studies have shown that these disturbances caused by cesarean section may be associated with diseases such as asthma (Sevelsted *et al.*, 2015), inflammatory bowel diseases (IBD) (Sevelsted *et al.*, 2015) and obesity (Darmasseelane *et al.*, 2014).

During the first day of life, infants born vaginally have been reported to be predominantly colonized by bacteria associated with vaginal environment such as *Lactobacillus*, *Prevotella* or *Sneathia* spp., whereas cesarean section delivered infants are colonized by microbes typically found on the skin surface, such as *Staphylococcus*, *Corynebacterium*, and *Propionibacterium* spp. (Dominguez-Bello *et al.*, 2010). Moreover, colonization by *Lactobacillus*, *Bifidobacterium* and *Bacteroides* has been noted to be delayed in infants delivered by cesarean section (Adlerberth *et al.*, 2006; Grönlund *et al.*, 1999). A recent study demonstrated that especially colonization by lactate-producing and -utilizing bacteria is delayed in infants born by cesarean section (Pham *et al.*, 2016). Disturbances in the microbiota caused by cesarean section have been reported to last up to 6 months after birth (Grönlund *et al.*, 1999). Another study using culture based techniques reported differences in gut microbiota between children born by vaginal delivery or cesarean section still at seven years of age (Salminen *et al.*, 2004).

2.5.3 Breastfeeding

After birth, the transfer of microbiota from mother to infant continues through breastfeeding. In addition to its nutritional value, breast milk contains various bioactive components that are thought to be protective and to supplement the innate immunity and have an impact on the infant gut microbiota composition and health (Petherick, 2010; Walker, 2010). These compounds include regulatory cytokines and growth factors, human milk oligosaccharides (HMO) and viable microbes as well as regulatory cytokines and growth factors that support and guide the development and maturation of the immune system. The gut microbiota of a breast-fed infant is typically dominated by species of *Bifidobacterium*, while formula fed infants harbour a more diverse microbiota containing *Escherichia coli, Clostridium difficile*, *Bacteroides* spp. and *Lactobacillus* spp. (Penders *et al.*, 2005; Cooke *et al.*, 2005; Bezirtzoglou *et al.*, 2011).

2.5.3.1 Breast milk microbiome

First reports of microbes in human milk date back to 1970's, but the focus was on pathogenic bacteria as possible sources of infection (Carroll *et al.*, 1979; Eidelman and Szilagyi, 1979). The idea about human milk as a possible source of commensal microbes emerged in the beginning of 2000's (Martin *et al.*, 2003; Heikkilä and Saris, 2003).

The breast milk microbiota has been described to be dominated by the phyla Proteobacteria and Firmicutes and the taxa *Staphylococcus*, *Pseudomonas*, *Enterobacteriaceae*, *Streptococcus*, *Lactobacillus, Weissella* and *Leuconostoc* (Cabrera-Rubio *et al.*, 2012; Hunt *et al.*, 2011; Urbaniak *et al.*, 2016). In addition, a number of studies also suggest that breast milk contains a high number of bifidobacteria (Martín *et al.*, 2009; Makino *et al.*, 2011). Interestingly, the milk microbiota varies between individuals and it is affected by mode of delivery, possibly due to hormonal changes produced during labor, and maternal weight (Cabrera-Rubio *et al.*, 2012, 2016). In addition, the milk microbiota seems to reflect the mother's health status also. For instance, atopic mothers have been described to have a lower abundance of *Bifidobacterium* sp. in their milk compared to healthy mothers (Grönlund *et al.*, 2007).

The origin of bacteria in the breast milk is not entirely clear yet. At first it was thought that the bacteria in breast milk originate from the skin of the breast area. Further investigations revealed that breast milk contains bacterial species such as lactobacilli and bifidobacteria that are not found on skin, but rather in the gastrointestinal tract of the mother (Gueimonde *et al.*, 2007; Martin *et al.*, 2003; Sinkiewicz and Nordström, 2005). Current belief is that bacteria may be

transferred from the gastrointestinal tract to the mammary gland via the so-called entero-mammary pathway. It has been proposed that dendritic cells may serve as transport vehicles for the bacteria. These cells are able to penetrate the intestinal wall (Rescigno *et al.*, 2001) and possibly take up non-pathogenic bacteria from inside the gut and feed them in the lymphatic and blood circulation through which they finally reach the milk producing mammary gland (Perez *et al.*, 2007). Interestingly, bacteria have also been detected in breast tissue in non-lactating women, which could suggest that at least part of the microbiota resides naturally in the mammary gland (Urbaniak *et al.*, 2014).

2.5.3.2 Human milk oligosaccharides

Maybe the most important factors influencing the infant's gut microbiota development are the HMOs (Figure 3). The major HMO components in human milk are mono- and difucosyllactose, lact-N-tetraose (LNT), and their mono- and difucosylated derivatives as well as fucosylated lacto-N-hexaoses and lacto-N-octaoses (Kunz *et al.*, 2000). However, the HMO profile of the human milk is unique to each mother and so far more than a hundred different HMOs have been identified (Kobata, 2010; Kunz *et al.*, 2000). The genotype of an individual affects HMO profile which is largely dependent on the expression of glycosyltransferases and the secretor type, defined by the ability to encode the alpha1,2-fucosyltransferase FUT2 (Kobata, 2010; Kunz *et al.*, 2000). This enzyme affects the Lewis blood group and controls the secretion of ABO histo-blood group antigens at mucosal surfaces (Kelly *et al.*, 1995; Oriol *et al.*, 1981). For instance, FUT2 non-secretors, who don't have an active FUT2 allele, lack the oligosaccharide 2'-fucosyllactose which is the most abundant oligosaccharide in breast milk of secretors. Oligosaccharide amount and composition in milk also varies over the course of lactation. Colostrum has been noted to have the highest concentration with as much as 20-25 g/l of HMOs, whereas after maturation the HMO concentrations in breast milk are typically around 5-20 g/l (Chaturvedi *et al.*, 2001; Coppa *et al.*, 1999; Thurl *et al.*, 2010). Interestingly, the oligosaccharide concentration in bovine milk has been reported to be very low (0.05 g/l) in comparison to human milk (Gopal and Gill, 2000).

HMOs are known to act in the infant gut as substrates for specific bacteria, such as specific bifidobacterial species (Bode, 2012), but HMOs can also act as decoy molecules for pathogens (Zivkovic *et al.*, 2011; Zopf and Roth, 1996). Moreover, HMOs are known to directly affect intestinal epithelial cells and modulate their gene expression as well as influence lymphocyte cytokine production (Bode, 2012). It has been also suggested that sialic acid, a

component of some HMOs, is an essential nutrient for brain development and cognition (Wang, 2009).

The impact of HMOs on the growth of bifidobacteria has been studied extensively *in vitro*. It seems that the ability to consume HMOs is a species or even strain specific trait for the bifidobacteria (Ruiz-Moyano *et al.*, 2013). Thus it is incorrect to refer to HMOs as "bifidogenic". For instance, *B. longum* subsp. *infantis* is known to grow well in HMOs when they are offered as the sole carbohydrate source, whereas *B. bifidum* grows slower in HMOs, and *B. longum* subsp. *longum* and *B. breve* can only utilize lacto-N-tetraose (LNT), but not lacto-N-neotetraose (LNnT) (Asakuma *et al.*, 2011; LoCascio *et al.*, 2007; Xiao *et al.*, 2010). The genome of *B. longum* subsp. *infantis* has been sequenced and it revealed entire gene clusters that control the expression of specific glycosidases, sugar transporters and glycan-binding proteins dedicated to HMO utilization (Sela *et al.*, 2008). In addition to bifidobacteria, two species of *Bacteroides, Bacteroides thetaiotaomicron* and *Bacteroides fragilis*, have been identified as consumers of HMOs (Marcobal *et al.*, 2011).

Despite a number of *in vitro* studies conducted, only a limited number of studies have looked at the impact of HMOs on infant gut microbiota composition in humans so far. It seems that the maternal FUT2 secretor status has an impact on the infant's gut bifidobacterial communities with bifidobacteria being established earlier and more often in infants fed by secretor mothers (Lewis *et al.*, 2015). A study conducted by Wang and coworkers (2015) showed that it is possible to some extent to predict the infant's gut microbiota composition based on the mother's milk's oligosaccharide composition. Later, other groups have reported similar results (De Leoz *et al.*, 2015). Interestingly, no reports have been available on the impact of HMOs on the breast milk microbiota composition itself.

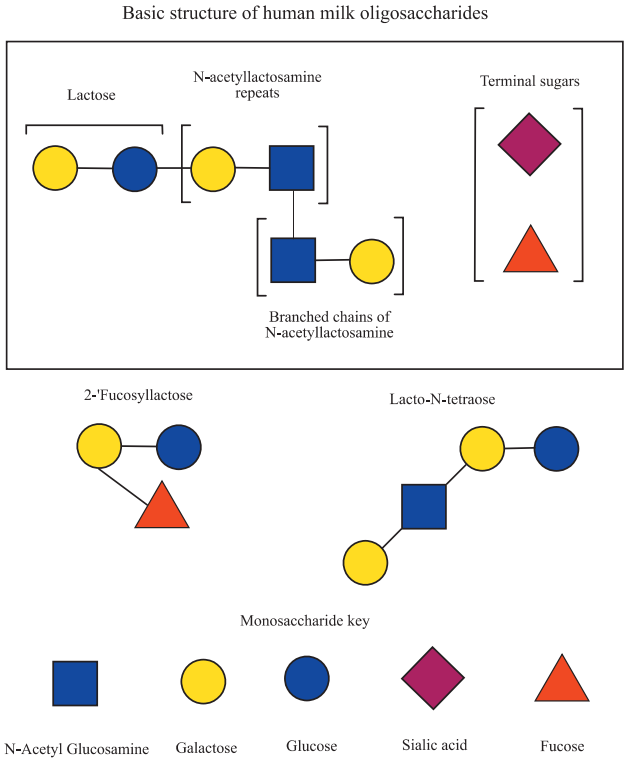


Figure 3. Basic structure of human milk oligosaccharides.

2.5.4 Introduction of complementary foods and weaning

The introduction of complementary foods is also an important step affecting the development of the gut microbiota shifting it towards a more adult-like microbiota. After the introduction of solid foods, the diversity of the gut microbiota increases and typically there is an increase in the proportion bacteria that are specialized in breaking down plant polysaccharides, such as members of the phyla Bacteroidetes and some species of Firmicutes, which may increase the production of SCFAs (Koenig *et al.*, 2011). Another study conducted by Fallani and coworkers (2011) reported that bifidobacteria, enterobacteria and *C. difficile* and *C. leptum* species decrease while *C. coccoides* and C. *leptum* increase after weaning. However, it has been also suggested that actually the cessation of breast-feeding is the major driver in the development of an adult type microbiota instead of the introduction of solid foods (Bäckhed *et al.*, 2015). Nevertheless, at 2-3 years of age, the gut microbiota and its functions start to resemble an adult's microbiota and its activity (Koenig *et al.*, 2011).

2.5.5 Geographic location

A number of studies indicate that infants in different geographical locations have differences in the gut microbiota composition. It has been hypothesized that humans among other animals have adapted to colder climates by increasing their body mass (Bergmann's rule) (Holliday and Hilton, 2010). Thus, some have suggested that people in the colder climates would have a microbiota that is associated with weight gain with an increase in the proportion of Firmicutes and a decrease in the proportion of Bacteroidetes (Suzuki and Worobey, 2014).

Data indicates that gut microbiota composition of infants is different between European countries. Fallani and coworkers (2010) suggested that there is a north-south gradient in the microbiota composition across young Europeans infants: Northern European infants had higher levels of *Bifidobacterium*, *Atopobium*, *Clostridium perfringens* and *Clostridium difficile*, and lower levels of *Bacteroides*, *Eubacterium* and *Lactobacillus* than the Southern European infants.

Differences between European and African infants have been reported also. Infants in Burkina Faso have been reported to have higher proportions of Bacteroidetes while Firmicutes and Enterobacteriaceae were under-represented compared to Italian infants (De Filippo *et al.*, 2010). Another study conducted by Grześkowiak and coworkers (2012a) compared Malawian and Finnish infants. In this study, the Malawians had higher levels of bifidobacteria, *Bacteroides-Prevotella* and *Clostridium histolyticum* while the species *Bifidobacterium adolescentis*, *Clostridium perfringens* and *Staphylococcus aureus* were only detected in Finnish infants (Grześkowiak *et al.*, 2012a).

In both of the aforementioned studies the authors speculated that diet is most likely the most important factor influencing the observed differences in gut microbiota composition between the European and African infants. In the African countries, infants are typically breastfed for a longer period of time, but complementary foods introduced earlier. The complementary foods are often rich in fiber and complex carbohydrates, which the European diet often lacks, and these may modify the gut microbiota composition (De Filippo *et al.*, 2010; Grześkowiak *et al.*, 2012a).

2.5.6 Probiotics, prebiotics and dietary supplements

Gut microbiota development may be modulated by probiotics, prebiotics or dietary supplements. According to the definition, probiotics are live microorganisms that, when administered in adequate amounts, confer a health benefit on the host (Hill *et al.*, 2014; FAO/WHO, 2002). For prebiotics, the following definition has been proposed recently: "a prebiotic is a nondigestible

compound that, through its metabolization by microorganisms in the gut, modulates composition and/or activity of the gut microbiota, thus conferring a beneficial physiological effect on the host" (Bindels *et al.*, 2015). The idea behind the probiotic/prebiotic therapy is to modulate an unbalanced microbiota towards a more balanced state and thus lower the risk of disease.

The majority of the probiotics currently used are strains from the genera *Lactobacillus* and *Bifidobacterium*, such as *Lactobacillus rhamnosus* GG and *Bifidobacterium animalis* subsp. *lactis* Bb12, which are maybe the most studied probiotics in the world. As mentioned earlier, probiotics seem to lower the risk of disease better when administered to the mother during pregnancy. It seems that the infant's microbiota may be modulated better this way instead of providing the probiotic to the infant. For instance, increased levels of bifidobacteria have been reported in the feces of infants whose mothers received *Lactobacillus rhamnosus* GG during pregnancy (Gueimonde *et al.*, 2006b; Lahtinen *et al.*, 2009). Prenatal probiotics have also been reported to influence the milk microbiota composition. In addition, probiotic lactobacilli consumed by mothers have been recovered later from milk and infant feces (Abrahamsson *et al.*, 2009).

It has been proposed that specific prebiotics, such as galacto-oligosaccharides (GOS) and fructo-oligosaccharides (FOS), could have similar functionalities as HMOs, and thus be beneficial to infants who are not breastfed for one reason or another (Oozeer *et al.*, 2013). However, data about prebiotic interventions in early life and during pregnancy is scarce compared to probiotic interventions. It has been reported that supplementation of infant formula with a mixture of GOS and FOS may protect against infections during the first six months of life in formula fed infants (Arslanoglu *et al.*, 2007). There is also some evidence that the addition of a mixture of GOS and FOS in infant feed may prevent eczema in infants up to two years of age (Osborn and Sinn, 2013).

Diet often has a poor nutritional value in developing countries, and malnutrition is a problem. It has been shown that the gut microbiota has a role in weight regulation and thus also in malnutrition (Gordon *et al.*, 2012). More recent reports support the hypothesis that the source of protein in dietary supplements may in fact have an impact on recovery rates in children with (moderate acute) malnutrition (Stobaugh *et al.*, 2016). In developing countries, different dietary supplements are used to fight malnourishment. However, there is limited information on whether these supplements impact the gut microbiota development. Optimally, the gut microbiota composition would be modulated by the supplements to promote linear growth in infants at risk of malnutrition.

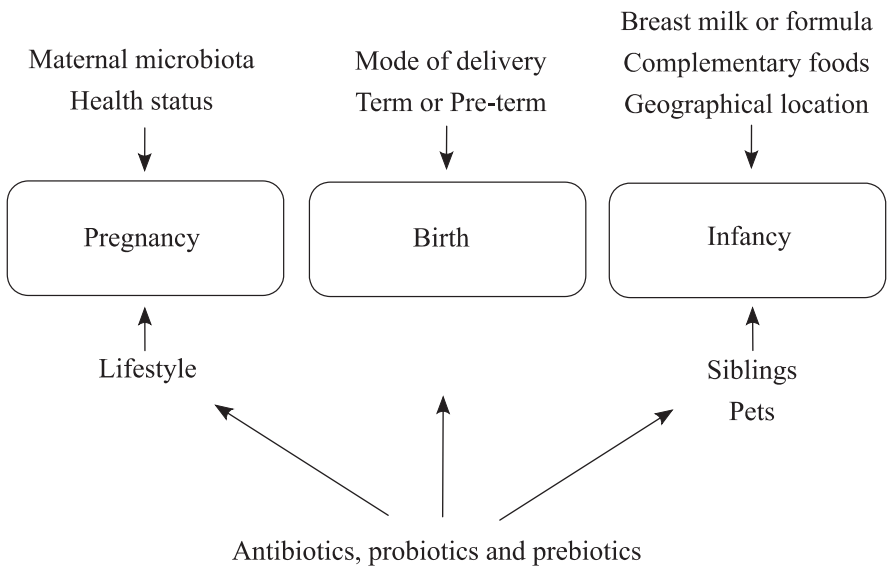


Figure 4. Factors that have an impact on gut microbiota development during early life.

2.6 Methods to study the human microbiota

2.6.1 Culture-dependent methods

Human gut microbiota has been studied since the beginning of 1900 when Henry Tissier isolated bifidobacteria from the stool of breast-fed infants (Tissier, 1900). Until the 1980s, the analysis of the microbial community in human gut was largely dependent on the ability to grow strict anaerobes. The cultured isolates were then identified and characterized by different phenotypic methods. However, this allowed only a limited view of the human gut microbiota since only a minority (30-40%) of gut bacteria can be cultured even by the most advanced culturing techniques today (Rettedal *et al.*, 2014; Lagier *et al.*, 2012). When the culture-independent molecular methods revolutionized the gut microbiota studies in 1980s and 1990s, it became clear that the taxonomic and phylogenetic diversity of the human gut microbiota had been enormously underestimated (Namsolleck *et al.*, 2004). Nevertheless, it is important to remember that the culture method gives important information on the viability and culturability of various microorganisms and their complex relations in gut microbiota assessment.

2.6.2 Polymerase chain reaction, the dawn of molecular methods

Almost all molecular methods used in microbiota studies today are based on the amplification of DNA by polymerase chain reaction (PCR). PCR, as we know it today, was invented in the 1980s by Kary Mullis and colleagues at Cetus Corporation in California (Bartlett and Stirling, 2003). PCR allows the amplification of a single or a few copies of a piece of DNA exponentially generating millions of copies of the desired DNA sequence, which can be then analyzed using various methods. In microbiology typically the 16S rRNA gene is amplified and sequenced because it is highly conserved within prokaryotes, but the sequence differs between organisms of other genera and species (Woese *et al.*, 1990).

Today, DNA-based methods can be divided into targeted amplicon studies and the more recent shotgun metagenomic studies. The former mentioned studies focus on specific known marker genes, such as the 16S rRNA gene, and primarily gives taxonomically informative data. The metagenomic studies are indirect and take into account the whole microbial community, which can give information about the functional potential of the human microbiome.

However, molecular methods based on PCR do not come without limitations. PCR is prone to the so-called PCR induced bias (Polz and Cavanaugh, 1998). Especially in microbiota studies, this can cause noticeable error. Relative abundancies can increase or decrease due to differences in the reaction efficiency. For instance, reaction efficiency is typically reduced if the template DNA has a high GC content. Sequence-based techniques can also totally miss minor but clinically relevant populations, including pathogenic bacteria such as *Salmonella* Typhi, *Tropheryma whipplei* and *Yersinia enterocolitica* due to the so called depth bias (Lagier *et al.*, 2012).

Prior to PCR, the DNA has to be extracted from the microbes. This step can also influence the final results. Due to the cell-wall structure of gram-positive microorganisms, poor yield of DNA is common from these cells if they are not lysed properly. Better yield can be achieved by including a step where the cells are physically disrupted with bead-beating for instance (Nylund *et al.*, 2010). However, this may cause DNA shearing and thus the method has to be properly optimized (Miller *et al.*, 1999; Yuan *et al.*, 2012).

In microbiota studies, the results in PCR-based methods may also be strongly influenced by PCR primer choice. For example, it has been shown that some widely utilized primers targeting the 16S rRNA may not be able to amplify some bacterial DNA, such as that from Actinobacteria, and therefore the primers should be modified to overcome this problem (Walker *et al.*, 2015).

2.6.3 Polymerase chain reaction denaturing gradient gel electrophoresis (PCR-DGGE)

In the beginning of 1990s, denaturing gradient gel electrophoresis (DGGE) was introduced for the analysis of the diversity of complex microbial ecosystems (Muyzer *et al.*, 1993). Together with temperature gradient gel electrophoresis (TGGE), this method remained the golden standard in the gut microbiota community analysis until the next-generation sequencing methods emerged.

DGGE is based on the mobility of partially melted DNA molecules, for instance the 16S rRNA gene, in polyacrylamide gels (Muyzer and Smalla, 1998). Partially melted DNA molecules have a decreased mobility compared to completely helical forms of the molecule. The gel has an increasing denaturing gradient in which the melted strands fragment completely into single strands and then come to a stop. Mobility of the DNA fragment in DGGE is dependent on its GC content: DNA molecules with higher GC content are more stable and thus denature slower and travel faster and further in the polyacrylamide gel. This makes it possible to separate DNA fragments of the same length but with varied sequences providing a semi-quantitative community fingerprint of the bacterial species present in the sample. These fragments can be also isolated from the DGGE gels for identification by sequencing.

One of the strongest points of DGGE is that it can be used to analyze multiple samples simultaneously. Due to this, DGGE can be used for comparative purposes; for instance, comparing the gut microbiota of a person with a disease to a healthy one. Although some think that high-throughput methods have made this technique obsolete, it is still utilized in gut microbiota research due to its low cost, reliability, reproducibility and ease of use.

2.6.4 Quantitative PCR

While PCR can be used to detect bacterial DNA in given samples, used alone it is unreliable for quantification. qPCR is similar to standard PCR, but the reaction mixture includes a compound that fluoresceces when bound to double-stranded DNA. When the fluorescence is measured in each PCR cycle, it can be used to estimate the amount of DNA present in the test sample when compared to a standard curve obtained from dilution series of reference DNA (Carey *et al.*, 2007). qPCR can be used to quantitate bacterial load accurately up to strain level. However, if the copy numbers of 16S rRNA gene are counted for instance, it must be kept in mind that some bacterial genomes contain multiple copies of this gene, which may affect the end result (Farrelly *et al.*, 1995; Nadkarni *et al.*, 2002).

For gut microbiota research, quantitative PCR (qPCR) was combined with primers targeting different bacterial groups (Gueimonde *et al.*, 2004; Rinttila *et*

al., 2004). This method is still one of the most accurate methods available for microbiota analysis. It also allows the analysis of bacteria whose relative abundance is low and that are often lost in high-throughput methods such as next-generation sequencing. However, qPCR can be only used to target already known microbial species and is therefore often complemented with some other method. Due to the high accuracy of this method, it still has a wide variety of applications in human microbiota studies.

2.6.5 Next-generation sequencing of 16S rRNA gene

Before the introduction of the so-called next-generation sequencing technologies, direct sequencing of the 16S rRNA genes from complex microbial communities was very cumbersome and costly. The 16S rRNA gene sequences had to be amplified with universal primers, cloned into vectors, transformed into *E. coli*, then colonies were picked for sequencing, purification of plasmid and bidirectional sequencing (Tringe and Rubin, 2005).

The next-generation sequencing technologies have allowed the sequencing of massive numbers of DNA templates at the same time in the same reaction set-up. The 454 sequencing (Roche Diagnostics GMBH Ltd, Mannheim, Germany) was one of the first technologies that allowed the sequencing of amplified DNA fast and at a lower cost (Margulies *et al.*, 2005). This technique was quickly adapted for microbial community analysis by several groups (Andersson *et al.*, 2008; Sogin *et al.*, 2006; McKenna *et al.*, 2008) allowing a high resolution profiling of microbial communities in complex environments such as the gut microbiota. Since then, other commercially available technologies have emerged from other companies as well, the major ones being Illumina (Illumina, San Diego, CA, USA) and SOLiD (Life Technologies, Carlsbad, CA, USA).

Next-generation sequencing has various advantages such as phylogenetic identification, detection of unknown bacteria, the ability to collect quantitative data and speed. Unfortunately, sequencing of the whole 16S rRNA gene is not widely available yet and thus these technologies are limited to genus level analysis in microbiome studies. The fragment length, PCR bias and choice of 16S rRNA gene region influences the community diversity estimates by next-generation sequencing (Youssef *et al.*, 2009). For instance, shorter amplicon lengths may result in relatively higher diversity estimates compared to longer fragments (Engelbrektson *et al.*, 2010). Thus, the choice of the hypervariable region of the 16S rRNA gene is highly important (Kim *et al.*, 2011). Due to these methodological differences, it may often be difficult to compare community diversity results across different studies.

The next-generation sequencing technologies have become a staple in microbiota research. Nevertheless, the studies based on the 16S rRNA gene can only give us information about phylogenetics, not what the bacteria actually do in the gut. For this purpose, more thorough analysis of bacterial genomes, such as metagenomics, is needed.

2.6.6 Metagenomics

While the research based on 16S rRNA gene sequence attempts to answer the question "who's there?", metagenomics may be used to answer the additional questions of "what can they do?", "what are they doing?", and "what is the genetic potential of the non-cultured bacterial fraction of the gut microbiota?". Metagenomics has been described as a genomic analysis applied to all microorganisms of an ecosystem without previous identification (Lepage *et al.*, 2013). Metagenomics was first described by Handelsman and Rodon in 1998 and used in the analysis of aquatic and soil ecosystems (Handelsman *et al.*, 1998; Rondon *et al.*, 1999).

The emergence of next-generation technologies has made metagenomics more cost-effective. Nowadays it is possible to directly shotgun sequence the metagenomic DNA content of an ecosystem. This means that the DNA of a mixed DNA sample, for instance from gut microbiota, is randomly fragmented and then these fragments are sequenced and the overlapping sequences are reconstructed to assemble them into a continuous sequence (Fraher *et al.*, 2012). This data then allows the assessment of genetic diversity and functions of the gut microbiota.

Differences in sequencing platforms, DNA preparation methods and the complexity of the samples studied can cause different or biased observations (Poretsky *et al.*, 2014). Another fallback of metagenomics is that it does not permit monitoring the activity or gene expression of the microbiota (Lepage *et al.*, 2013). For example, it has been shown in marine ecosystems that microbial presence and activity are not linked together since the most expressed genes are not the most represented ones within the metagenome (Frias-Lopez *et al.*, 2008). To overcome this fallback, the messenger RNA (mRNA) sequences from the microbial community should be analyzed. This method is known as metatranscriptomics, which again has its own challenges and limitations.

Metagenomics can be combined with other "omics" technologies. In addition to the aforementioned, metatranscriptomics, metaproteomics (protein mapping profile) and metabolomics (metabolic profile) may provide additional useful data. For instance, metagenomics combined with metabolomics has been used to identify the role of microbiota in dietary phospholipid metabolism, contributing to atherosclerosis (Wang *et al.*, 2011).

Table 1. Advantages and disadvantages of common methods used in gut microbiota analysis

Method	Description	Advantages	Disadvantages
Culturing	Isolation of bacteria on selective media	Cheap, semi-quantitative, bacteria can be isolated and their properties studied further	Labor intensive, most of the human gut microbes cannot be cultured by today's techniques
DGGE/TGGE	Gel separation of 16S rRNA amplicons using denaturing or temperature gradient	Fast, semi-quantitative, bands can be excised for identification	PCR bias, only semi-quantitative, no phylogenetic identification
qPCR	Amplification and quantification 16S rRNA. Reaction mixture contains a compound that fluorescenses when it binds to double-stranded DNA	Phylogenetic identification, fast, accurate	PCR bias, laborious if many different bacteria/bacterial groups are targeted
Sequencing of 16S rRNA amplicons	Massive parallel sequencing of partial 16S rRNA amplicons with 454 Pyrosequencing for instance	Phylogenetic identification, quantitative	PCR bias, expensive
Microbiome shotgun sequencing	Massive parallel sequencing of the whole genome with 454 pyrosequencing or Illumina	Phylogenetic identificaiton, quantitative	PCR bias, expensive, computationally intensive

2.7 Future perspectives

During the past decade, the development of high-throughput technologies has caused an avalanche of metagenomic data. The cost effectiveness has increased significantly during recent years and will continue to increase in the future. Together with exponential increase in computational power a more thorough analysis of the gut microbiota at a lower cost will be available to almost everyone. Moreover, the research will most likely expand into studying the metatranscriptome and metaproteome of the microbiota. Therefore, data analysis and interpretation will have an even bigger role than today.

Despite the advances in techniques used in microbiota research, a lot remains to be uncovered about the role of human microbiota in human health. The information we have today about this subject is mainly based on association studies and a shift to studying mechanisms and causality is needed for true understanding of the role of gut microbiota in health and disease

(Waldor *et al.*, 2015). Moreover, since microbes interact with each other in the gut, it will be important to identify bacterial genomes that work as functional groups (guilds) in the gut (Zhang *et al.*, 2015). Understanding the mechanistic bases for interaction between microbial communities and their hosts will most likely reveal new therapeutic approaches aiming to modulate the gut microbiota to treat non-communicable diseases, such as allergic, autoimmune and anti-inflammatory disorders, as well as obesity.

Finally, it must not be forgotten that there are also other microorganisms in the intestine in addition to bacteria that may have a role in human health. In fact, it has been noted that bacteriophages outnumber bacterial cells by tenfold in the human gut (Dutilh *et al.*, 2014). Similar to gut bacteria, gut viromes have been shown to be specific for individual (Reyes *et al.*, 2010) and they seem to respond to diet also (Minot *et al.*, 2011). However, the impact of the virome on host health remains to be elucidated. In addition to viruses, the role of parasites in human health has been discussed recently and it has been suggested that lack of some specific helminths for instance might be associated with risk developing of IBD (Ramanan *et al.*, 2016).

3 AIMS OF THE STUDY

The primary aim in the present study was to characterize the potential microbiota transfer from mother to infant and uncover factors, such as geographic location and diet, that affect microbiota development in early infancy.

The specific aims were:

1. To evaluate the specificity of two *Lactobacillus rhamosus* GG specific primer sets (**I**).

2. To analyze gut *Lactobacillus* microbiota development in a developing country and to compare it to a westernized country (**II**).

3. To study the impact of lipid-based nutrient supplements with different sources of protein on the development of *Bifidobacterium* microbiota in Malawian infants (**III**).

4. To characterize the microbial populations in placenta, amniotic fluid and colostrum and to elucidate their role as the initial inoculum for the intestinal microbiota (**IV**).

5. To assess whether human milk oligosaccharides affect the human milk microbiota composition (**V**).

4 MATERIALS AND METHODS

4.1 Materials, subjects and study design

4.1.1 Evaluation of *L. rhamnosus* GG strain specific primers (I)

In study **I**, 41 strains of *L. rhamnosus*, including *L. rhamnosus* GG and the *L. rhamnosus* type strain, were used to evaluate two primer sets reported to be *L. rhamnosus* GG strain specific (Ahlroos and Tynkkynen, 2009; Brandt and Alatossava, 2003). *L. rhamnosus* GG (= ATCC53103) was used as a positive control and obtained from the American Type Culture Collection (Manassas, Virginia). The type strain of *L. rhamnosus*, DSM 20021^T, obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany), was used as a negative control. The rest of the *L. rhamnosus* strains used in study **I** were obtained from the Belgian Coordinated Collection of Microorganisms/LMG (Gent, Belgium). The LMG strains were dairy and human clinical isolates identified as *L. rhamnosus* originating from different countries. All *L. rhamnosus* strains were routinely grown in MRS broth (Oxoid Ltd., Basingstoke, Hampshire, England) at 37 °C overnight.

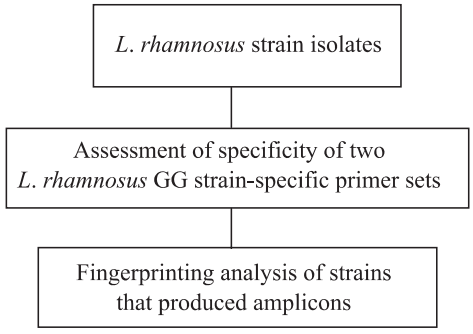


Figure 5. Study **I** design

4.1.2 *Lactobacillus* microbiota of Malawian and Finnish infants (II)

In study **II**, the populations consisted of healthy 6-month-old infants living in rural Malawi (n = 44) and Southwestern Finland (n = 31) (Figure 6). The Malawian infants were enrolled for a clinical trial assessing the impact of selected dietary interventions on early childhood growth (identifier: NCT00524446). In Finland, the infants were participants of a prospective randomized study in Turku and nearby areas in Southwestern Finland (identifier: NCT00167700). Infants and their mothers included in the present

study did not receive probiotics or prebiotics. Malawian and Finnish infants represented typical infants consuming typical diets for each area.

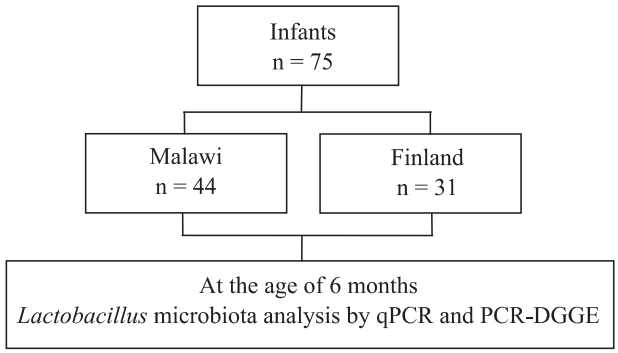


Figure 6. Study **II** design

4.1.3 Impact of dietary intervention on bifidobacterial populations of Malawian infants (III)

Subjects of the study **III** represent a sub-population of the Lungwena Child Nutrition Intervention Study (LCNI-5). The study participants were healthy infants who met the following inclusion criteria: age 5.5 – 6.5 months, residence in the study area, and available informed consent from at least one guardian. An additional inclusion criterion was the availability of a fecal sample at six and twelve months of age. The exclusion criteria were as follows: weight-for-length (WFL) 80 % of the World Health Organization (WHO) reference median, presence of edema, severe illness warranting hospitalization on the enrollment day, history of peanut allergy, concurrent participation in another clinical trial and any symptoms of food intolerance within 30 min of ingesting a 5 g test dose of lipid-based nutrient supplement (either milk or soy-based) used in the trial. Of the total 840 participants in the main trial, 160 were selected to the current sub-study (Figure 7). The study **III** infants were randomly assigned to one of four intervention schemes for a 6-month period (Figure 7). Infants either received 71g/day of micronutrient-fortified corn-soy blend (CSB), 54 g/day of micronutrient-fortified lipid-based nutrient supplements (LNS) with milk protein base (milk-LNS) or 54 g/day of micronutrient-fortified LNS with soy protein base (soy-LNS) between 6 and 12 months of age.

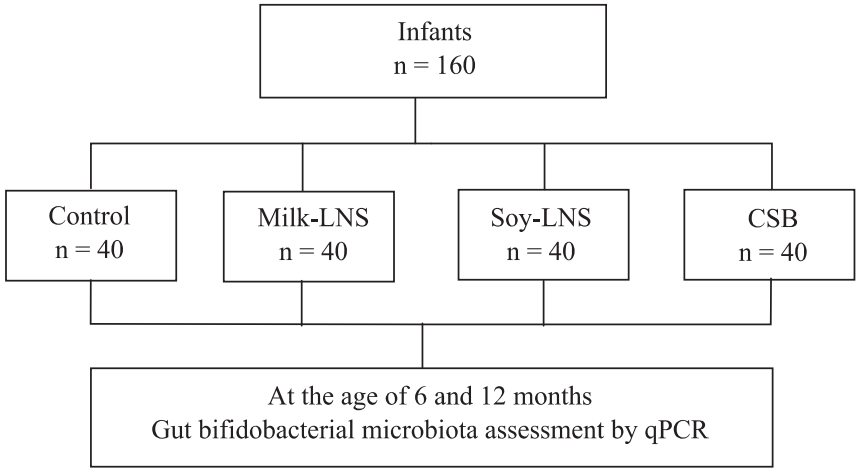


Figure 7. Study **III** design

4.1.4 Microbial contact *in utero* and in early infancy (IV)

The study **IV** samples were collected from subjects participating in a randomized double-blind placebo-controlled clinical trial (identifier: NCT00167700), which aimed to assess the impact of probiotic modulation on maternal microbial contact in late pregnancy. Pregnant women, who were scheduled to undergo cesarean section after 37 weeks of gestation, were recruited to obtain placenta samples without risk of contamination taking place during delivery. Mothers with conditions, which might affect placental and fetal physiology (e.g. pre-eclampsia, intrauterine growth retardation, fetal anomalies, onset of labour, asphyxia) or contaminate the placenta (rupture of membranes, vaginal delivery, infection), were excluded from the study. Amniotic fluid, placenta, meconium, colostrum, infant feces and maternal feces samples were available from 15 mother-infant pairs and included in the study **IV** (Figure 8).

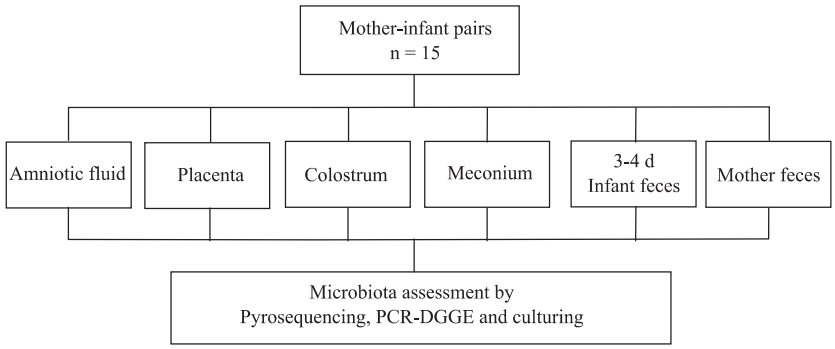


Figure 8. Study **IV** design

4.1.5 Impact of human milk oligosaccharides on human milk microbiota (V)

The study **V** participants were from the same randomized, double-blind placebo-controlled trial as the study **IV**. For study **V**, colostrum samples from eleven mothers after full-term cesarean section were available for analysis (Figure 8).

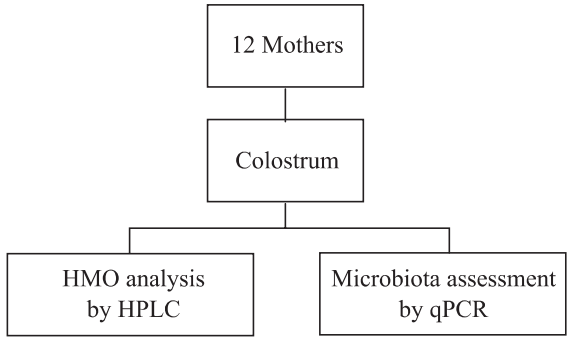


Figure 9. Study **V** design

4.2 Methods

4.2.1 Sample collection (II-V)

In study **II,** the Finnish fecal samples were collected, stored frozen and transported to the laboratory in the Functional Foods Forum, University of Turku, Finland, and stored at -70 °C until analysis. For studies **II** and **III**, fecal samples from Malawian infants were collected at home by parents, then brought to the clinic and collected by lab technicians, which took 2-4 hours in total. The samples were stored at -20 °C in the study site and later at a central

laboratory in Mangochi at -40 °C until being transported from Malawi to Finland and stored at -70 °C until analysis. In study **IV** amniotic fluid and placenta samples were obtained during sterile cesarean section and stored at -80 °C. Meconium (0-2 d) and infant fecal samples (3-4 d) were collected from diapers after they had been passed. Meconium, infant and maternal fecal samples were collected in sterile plastic containers, refrigerated and processed without further delay. In study **IV** and **V**, colostrum samples were collected in the maternity hospital using milk produced within 24 hours of delivery. The mothers were given written instructions for standardized collection of samples in the mornings and the samples were stored at -20 °C until analyzed.

4.2.2 DNA extraction (I-V)

4.2.2.1 DNA extraction from reference strains (**I**)

In study **I**, the DNA was extracted from 1 ml of the cultured cells using the method described by Endo and coworkers (2007). In short, the cells were washed twice with 1 ml of TE buffer (10mM Tris-HCl, 1 mM EDTA; pH 8.0) and then resuspended in a solution containing 250 µl of extraction buffer (100 mM Tris-HCl, 40mM EDTA; pH 9.0), 50 µl of 10 % sodium dodecyl sulfate, and 150 µl of benzyl chloride. In addition, 100 mg of glass beads (0.1 mm in diameter) were added to the suspension and the mixture was beaten three times for 30 s with FastPrep-24 (FP120-230, Bio 101 ThermoSavant, Holbrook, NY). Next, 150 µl of 3 M sodium acetate was added to the beaten sample and then the sample was kept on ice for 15 min. Then, the sample was centrifuged at 15000×g for 10 min after which the supernatant was transferred to a new tube and DNA was precipitated with isopropanol. The precipitated DNA was suspended in 50 µl of TE buffer. DNA yield was measured with NanoDrop ND-1000 spectrophotometer (NanoDrop® Technologies, Wilmington, DE, US) and the DNA was diluted to 10 µg/µl prior to use.

4.2.2.2 Manual DNA extraction utilizing the Qiagen column (**II, IV and V**)

In study **II** DNA was extracted from the fecal samples using QIAmp DNA Stool Mini Kit (Qiagen, Venlo, The Netherlands) as described by Grzeskowiak and coworkers (2012). Briefly, fecal samples were weighed (0.5 g), diluted 1:10 (w/v) with PBS buffer (130mM sodium chloride, 10 mM sodium phosphate, 0.05 % cysteine, pH 7.2) and homogenization was performed by thorough agitation in a vortex. Next, aliquots of these dilutions were used for DNA extraction with the QIAmp DNA Stool Mini Kit following the

manufacturer's instructions. In studies **IV** and **V**, instead of vortexing, the samples were homogenized using repeated bead-beating with FastPrep-24 (FP120-230, Bio 101 ThermoSavant, Holbrook, NY).

4.2.2.3 Semi-automated DNA extraction using KingFisher (**III**)

In study **III**, the samples were pre-treated with repeated bead-beating and DNA was extracted using InviMag Stool DNA kit (Stratec Molecular, Berlin, Germany) and automated KingFisher DNA extraction system (Thermo Fisher Scientific Oy, Vantaa, Finland), as previously described (Nylund *et al.*, 2010). Briefly, fecal samples were weighted (0.1-0.125 g) and mixed with 0.5 ml of lysis buffer P, glass-beads and glass-sand (0.25g of glass-sand, 0.1 mm and 4 pieces of glass-beads, 1.5 mm) in 2 ml screw cap tube. Homogenization and mechanical cell disruption was performed by two rounds of bead-beating (treatment time 3 min) with FastPrep-24 (FP120-230, Bio 101 ThermoSavant, Holbrook, NY) and heat treatment (95 C, 15 min).

4.2.3 Lactobacillus rhamnosus GG strain-specific PCR (I and II)

Two different *L. rhamnosus* GG specific primer sets were used in study **I**: the first set targeted a putative transposase gene (Ahlroos and Tynkkynen, 2009) and the second a phage-related gene of *L. rhamnosus* GG (Brandt and Alatossava, 2003). For each primer set, the reaction mixture and the amplification of DNA was done as previously described (Ahlroos and Tynkkynen, 2009; Brandt and Alatossava, 2003). Successful amplification was verified by subjecting the PCR products to gel electrophoresis in 1.0% agarose, followed by ethidium bromide staining.

4.2.4 Fingerprinting analysis (I)

Rep-PCR, RAPD and ERIC PCR fingerprinting were carried out to distinguish the different *L. rhamnosus* strains from each other in study **I**. The rep-PCR was carried out using the primers $(GTG)_5$ and a primer set of REP1R-I and REP2-I. The preparation of reaction mixture and amplification of DNA was performed as previously described (Gevers *et al.*, 2001). For RAPD, six different primers (C0540, 1251, OPA-03, D, E, and F) were used, and the reaction mixture and amplification of DNA was performed as previously described by Endo & Okada (Endo and Okada, 2005). Finally, ERIC-1 and ERIC-2 were used for the ERIC PCR, for which the preparation of the reaction mixture and amplification of DNA were performed using the method described by Ventura et al. (2003). The amplification products were subjected to gel electrophoresis in 1.0%

agarose, followed by ethidium bromide staining. Electrophoresis images were imported to Bionumerics software version 6.6 (Applied Maths, Belgium) for further analysis.

4.2.5 Denaturing gradient gel electrophoresis (DGGE) (II and IV)

PCR-DGGE was used to characterize the microbiota in fecal samples in studies **II** and **IV**, and in breast milk, placenta and amniotic fluid samples in study **IV**. The primer set Lac1 and Lac-GC (Walter *et al.*, 2001) was used in study **II** to analyze the *Lactobacillus* group community in the fecal samples. The reaction mixture was prepared as previously described (Endo and Okada, 2005) and the thermal cycling was conducted according to Walter et al. (2001). In study **IV**, total bacteria primers F357-GC and R518 (Muyzer *et al.*, 1993) were used and the reaction mixture was prepared and the thermal cycling conducted as described by Endo et al. (2008). Successful amplification was confirmed by gel electrophoresis in 1.0% agarose.

The DGGE analysis was performed using the DCode system (Bio-Rad Laboratories, Hercules, CA) according to the method described by Endo et al. (2007) (Endo *et al.*, 2007). For lactobacilli, a denaturing gradient of 35% to 50% was used and the electrophoresis was performed at a constant voltage of 23 mA for 16 hours. For total bacteria, a denaturing gradient of 35% to 75% was used instead and the electrophoresis was conducted at a constant voltage of 28 mA for 16 h. The gels were stained, unknown bands were excised and sequenced according to the method described by Endo et al. (2007). Digitalized DGGE images were imported to Bionumerics software version 6.6 (Applied Maths, Belgium) for further analysis.

4.2.6 DNA fingerprinting and DGGE images processing (I, II, IV)

In studies **I**, **II** and **IV**, electrophoresis images were imported to Bionumerics software version 6.6 (Applied Maths, Belgium) for normalization and band detection. Band search and band matching were conducted using a band tolerance of 1% as implemented in Bionumerics. Bands and band matching were manually checked and corrected when deemed necessary.

In studies **II** and **IV**, the band presence/absence and intensity matrices were exported from Bionumerics and imported to R version 3.1.2 (R Core Team, 2015) for further analysis. Number of bands present in the DGGE profile was used as a measure of richness for the samples, and the Shannon diversity index, H', was calculated using the equation $H' = -\Sigma P_i \ln (P_i)$. P_i is the importance

probability of the bands in a gel lane is calculated as $P_i = n_i/N$, in which n_i, is the intensity of band I and N is the sum of intensities of all bands in the densitometric profile.

4.2.7 Quantitative polymerase chain reaction (qPCR) analysis of microbiota (II, III and V)

Quantitative PCR was used to characterize the fecal and breastmilk microbiota in studies **II**, **III**, and **V** using group- and species-specific primers. PCR amplification and detection were performed with an ABI PRISM 7300-PCR sequence detection system (Applied Biosystems, Foster City, CA). The 25 µl reaction mixture composed of Power SYBR Green PCR Master Mix (Applied Biosystems), 1 µL of each of the primers at a concentration of 0.2 mol/l and 1 µl of template DNA. The fluorescent products were detected in the last step of each cycle and a melting curve analysis was performed after the amplification to distinguish the targeted PCR product from the nontargeted ones. Bacterial concentration in each sample was determined by comparing the Ct values obtained from standard curves. The standard curves were made from serial dilutions of DNA isolated from a known amount of each pure culture of the different reference strains.

The following reference strains were used: *B. longum* DSM 20219 (also for *Bifidobacterium* genus), B. *catenulatum* DSM 20224, *B. bifidum* DSM 20456, *B. lactis* DSM 20606, *B. infantis* DSM 20090, *B. adolescentis* DSM 20083, *B. breve* DSM 20213, *Akkermansia muciniphila* ATCC BAA-835, *S. aureus* DSM 20231, *C. coccoides* DSM 935^T, *C. leptum* DSM753^T, *L. rhamnosus* DSM20021^T, *Streptococcus thermophilus* CECT 986, *Enterococcus faecalis* CECT 481, and *Escherichia coli* K-12 (for total bacteria). *Bacteroides fragilis* DSM 2151. The sequences and annealing temperatures of the primers used for qPCR in studies **II**, **III** and **V** are presented in Table 2.

Table 2. Primers used in qPCR analysis in studies **II**, **III** and **V**

Target bacterial group/species	Sequence (5' to 3')	Annealing temp.	Reference
Total bacteria	AACGCGAAGAACCTTAC CGGTGTGTACAAGACCC	58	Muyzer et al. 1993
Bifidobacterium spp.	CTCCTGGAAACGGGTGG GGTGTTCTTCCCGATATCTACA	60	Matsuki et al. 2002
B. adolescentis	GGATCGGCTGGAGCTTGCTCCG CCCCGAAGGCTTGCTCCCAGT	63	Rinne et al. 2005
B. longum group	TTCCAGTTGATCGCATGGTCTTCT GGCTACCCGTCGAAGCCACG	65	Rinne et al. 2005
B. breve	AATGCCGGATGCTCCATCACAC GCCTTGCTCCCTAACAAAAGAGG	62	Rinne et al. 2005
B. bifidum	TGACCGACCTGCCCCATGCT CCCATCCCACGCCGATAGAAT	61	Rinne et al. 2005
B. catenulatum	GCCGGATGCTCCGACTCCT ACCCGAAGGCTTGCTCCCGAT	64	Rinne et al. 2005
Staphylococcus spp.	GGCCGTGTTGAACGTGGTCAAATCA TIACCATTTCAGTACCTTCTGGTAA	58	Martineau et al. 2001
Staphylococcus aureus NUC	GCGATTGATGGTGATACGGTT AGCCAAGCCTTGACGAACTAAAGC	55	Fang & Hedin 2002
Streptococcus group	GAAGAATTGCTTGAATTGGTTGAA GGACGGTAGTTGTTGAAGAATGG	62	Collado et al. 2009
Lactobacillus group (Study **II**)	AGCAGTAGGGAATCTTCCA CACCGCTACACATGGAG	58	Rinttila et al. 2004
Lactobacillus group (Study **V**)	GGAAACAG(A/G)TGCTAATACCG CACCGCTACACATGGAG	61	Heilig et al. 2002
Enterococcus spp.	ACTCGTTGTACTTCCCATTGT CCTTATTGTTAGTTGCCATCATT	62	Rinttilä et al. 2004
Akkermansia muciniphila	CAGCACGTGAAGGTGGGGAC CCTTGCGGTTGGCTTCAGAT	60	Collado et al. 2007
Bacteroides-Prevotella group	ATAGCCTTTCGAAAGRAAGAT CCAGTATCAACTGCAATTTTA	50	Matsuki et al. 2002
Clostridium cluster XIVa - XIVb	AAATGACGGTACCTGACTAA CTTTGAGTTTCATTCTTGCGAA	53	Matsuki et al. 2002
Clostridium cluster IV	GCACAGCAGTGGAGT CTTCCTCCGTTTTGTCAA	60	Matsuki et al. 2002

4.2.8 Bacterial culture from placenta and amniotic fluid samples and identification of bacterial isolates (IV)

Placenta samples were kept under anaerobic conditions (AnaeroGen, Oxoid, Hampshire, United Kingdom) and analyzed in less than 2 hours to avoid alterations in bacterial viability. The placenta tissue specimens were homogenized in 200 µl of a phosphate-buffered saline (PBS) solution by pipetting and thorough agitation in a vortex. Each homogenized sample was randomly plated on different culture media (100 µl). The PBS solution used in the homogenization was plated as a negative control. The following media were used: Gifu anaerobic medium (GAM agar, Nissui Pharmaceutica Co., LTD., Tokyo, Japan) and LB medium (MP Biomedicals, LLC, France). The plates were incubated under anaerobic conditions at 37 °C for 48-72 h.

All viable and cultivable bacteria recovered were isolated and re-streaked onto the same agar media. For preliminary identification, traditional microbiological methods were used, including analysis of colony and cellular

morphology and Gram staining. Next, the isolates were grown in the same isolation broth media and harvested at the late log growth phase and DNA was extracted using the QIAmp DNA Stool Mini Kit (QIAgen) following manufacturer's instructions. DNA from each isolate was amplified with 16S rRNA gene targeting primers (968f and 1401r). The primer sequences and annealing temperatures have been published elsewhere (Nübel *et al.*, 1996). The amplification products were subjected to gel electrophoresis in 1.0% agarose gels, purified using Gel Band DNA purification kit (GE Healthcare, Buckinghamshire, UK), and sequenced in an ABI Prism-3130XL genetic analyzer (Applied Biosystems). Search analyses to determine the closest relatives of the partial 16S rRNA gene sequences retrieved were conducted in GenBank using Basic Local Alignment Search Tool (BLAST) algorithm, and sequences with more than 99% similarity were considered to be of the same species.

4.2.9 454 Pyrosequencing (IV)

Pyrosequencing was used in study **IV** to characterize the bacterial communities in fecal, meconium, colostrum, placenta and amniotic fluid samples. A barcoded primer set based on universal primers 27F and 533R was used to amplify a 500bps long part of the 16S rRNA gene covering the hyper-variable regions V1 to V3 (V1-V3). The thermal cycling was carried out using a KAPA-HiFi polymerase (Kapa Biosystems, INC., Wilmington, MA) with an annealing temperature of 52 °C and 30 cycles. Purified PCR products were pooled in equimolar amounts according to manufacturer's instructions, and submitted for pyrosequencing using the Genome Sequencer FLX Titanium Series (454 Life Science, Branford, USA). All of the procedures followed 454 Roche protocols (454 Life Science).

From the resulting raw data set provided by pyrosequencing, low quality sequences and sequences having a length shorter than 100 nucleotides were filtered out. Moreover, chimeric sequences were removed using UCHIME software (Edgar *et al.*, 2011). A dereplicate request on the QIIME pipeline (v. 1.8) commands with default parameters was used for identifying representative sequences for each operational taxonomic unit (OUT) generated from the complete linkage clustering with a 97% similarity and aligned to fully-sequenced microbial genomes (Green Genes 13_8 database).

To estimate diversity conservatively and reduce noise in patterns of beta diversity, singleton OTUs were removed prior to community analysis. Alpha diversity indices were determined from rarefied tables using the Shannon-Wiener index for diversity, the Chao1 index for richness and Observed Species

(number of unique OTUs) and Phylogenetic Distance (PD_whole) were also determined. Most of the samples reached a plateau suggesting sufficient coverage of the microbes present in the samples. Microbial Beta-diversity between samples was evaluated and computed from the previously constructed OTU table using UniFrac, a phylogenetic distance metric that measures community similarity based on the degree to which pairs of communities share branch length in a common phylogenetic tree (Lozupone *et al.*, 2011).

4.2.10 High-performance liquid chromatography analysis of human milk oligosaccharides (V)

Human milk oligosaccharide analysis was performed as previously described at the University of California, San Diego, CA, USA (Alderete *et al.*, 2015). Raffinose was added to each milk sample as an internal standard for absolute quantification. The following HMOs were detected based on retention time comparison with commercial standard oligosaccharides and mass spectrometry analysis: 2'-fucosyllactose (2'-FL), 3-fucosyllactose (3'-FL), 3'-sialyllactose (3-'SL), lacto-N-tetraose (LNT), lacto-N-neotetraose (LNnT), lacto-N-fucopentaose (LNFP) I, LNFP II, and LNFP III, sialyl-LNT b (LST b) and LST c, difucosyl-LNT (DF-LNT), disialyl-LNT (DS-LNT), fucosyl-lacto-N-hexaose (FLNH), difucosyl-lacto-N-hexaose (DFLNH), fucosyl-disialyl-lacto-N-hexaose (FDSLNH) and disialyl-lacto-N-hexaose (DSLNH). Secretor status was defined by the presence of 2'-FL or LNFP.

4.2.11 Statistics

Statistical analysis was performed using IBM SPSS version 21.0 (Studies **II** and **IV**) and R versions 3.1.2 (Study **II** and **IV**) and 3.2.2 (Studies **III** and **V**) (R Core Team, 2015).

Study I:
In study **I**, all fingerprinting data were combined to make a compositional data set. Next, a dendrogram was constructed from this composite data using Dice coefficients with the unweighted pair-group method using arithmetic averages (UPGMA) clustering method as implemented in Bionumerics software version 6.6 (Applied Maths, Belgium).

Study II:
The *Lactobacillus* group bacterial counts (log 16S rRNA gene copies g^{-1}) assessed by qPCR were expressed as means with 95% confidence interval. Independent samples *t* test was used to assess differences in means of the

bacterial counts between Malawian and Finnish infants. Richness in the samples was not normally distributed and is thus presented as medians with interquartile range (IQR) and analyzed using the Mann-Whitney U test. The Shannon diversity indices were normally distributed and are expressed as means with 95% confidence intervals (CI) and analyzed by independent samples t test. Fisher exact test was used to determine statistical difference in the prevalence of each band class between the Malawian and Finnish infants. Band classes identified as the same species were combined.

Cluster analysis of the DGGE patterns was performed by the UPGMA based on the Pearson correlation similarity coefficient as implemented in Bionumerics software version 6.6 (Applied Maths, Belgium). Principal component analysis based on Euclidean distances was performed based on the DGGE fingerprints using the R function prcomp. The first 2 principal components were used to generate a 2-dimensional scatter plot. PERMANOVA analysis was performed using the adonis function of the vegan R software package (Oksanen *et al.*, 2015), where a euclidean distance matrix was calculated using 999 permutations.

Study III:

Bacterial counts were not normally distributed and therefore non-parametric methods were used for the analysis. Bacterial detection rates are presented as percentages (%) and counts (log gene copies g^{-1}) as medians with interquartile ranges (IQR). Samples in which bacterial DNA was not detected by qPCR were excluded from the statistical analysis.

McNemar's test was applied in the comparison of bacterial colonization rates between time points in the whole study population. Similarly, Wilcoxon signed-rank test was used in the assessment of differences in bacterial counts between the time points in the whole study population.

Geeglm function from the R package geepack (Højsgaard *et al.*, 2014) was used to fit a generalized estimating equations (GEE) model with a compound symmetry structure to determine the impact of the diet intervention on the bacterial detection rates. In this model, intervention type and time were added as independent variables, the primary variable measured at baseline was included as a covariate and individual was included as a random factor. A non-parametric analysis of covariance (ANCOVA) was performed utilizing the sm.ancova function of the R package sm (Azzalini and Bowman, 2014) to assess statistically significant differences in the bacterial counts between the intervention groups after the intervention when controlling for the baseline bacterial counts.

Study IV:

Cluster analysis of the DGGE patterns was performed by the UPGMA based on the Pearson correlation similarity coefficient as implemented in Bionumerics software version 6.6 (Applied Maths, Belgium). Analysis of variance (ANOVA) followed by Tukey HSD test was used to assess differences in means of richness and diversity calculated from the DGGE fingerprints between the different sample types.

Unweighted and weighted UniFrac distances and sample metadata comprised the data matrices used as inputs for principal coordinate analysis (PCoA). Unweighted UniFrac distances compare microbial communities based on presence/absence of members (community membership: presence/absence matrix), and weighted UniFrac also incorporates relative abundance information (community structure:presence/absence/abundance matrix). PCoA plots were used to assess the variation in the composition of microbial communities between samples and to visualize potential clustering of samples by metadata. Samples were also hierarchically clustered based on their inter-sample UniFrac distances using UPGMA. Biplots were generated as part of the beta diversity analyses, using genus level OTU tables showing principle coordinate sample clustering alongside weighted taxonomic group data. All beta-diversity measures were performed on OTU tables rarefied to 500 sequences per sample for all samples to account for variations in sequencing depth. Pyrosequencing data on assigned sequences at family level shared between samples were used to generate a Venn diagram.

After taxonomical assignment, relative frequencies of different taxonomic categories obtained were calculated using the Statistical Analysis of Metagenomic Profiles program (STAMP v.2.0.0). Statistical differences between experimental groups were estimated by ANOVA analysis with the Games-Howell post-hoc test and the multiple test correction of Benjamini-Hochberg as implemented in STAMP. Multivariate statistical analysis and clustering were performed using METAGENassist software (Arndt *et al.*, 2012). Data-filtering was performed by the interquantile range method followed by quantile normalization within replicates after log transformation. Hierarchical clustering of OTUs were done using core microbiota defined as the OTUs that are present in at least 50% of the samples and Cluster analysis was performed using the Euclidean distance. Principal component analyses and identification of significant features were performed for all sample groups together.

Study V:

A spearman’s correlation matrix was calculated for the HMO groups and different bacteria as well as individual HMO structures and the bacteria.

Heatmaps of the correlation coefficients were constructed with the use of microbiome R package (Lahti and Salojärvi, 2012).

4.2.12 Ethics

The protocols of study **II** and **III** were approved by the College of Medicine research and ethics committee (University of Malawi) and the ethics committee of the Pirkanmaa Hospital District, Finland. The protocols of studies **IV** and **V** were approved by the Ethics committee of the Intermunicipal Hospital District of Southwest Finland.

5 RESULTS

5.1 Specificity of *L. rhamnosus* GG specific primers (I)

The *L. rhamnosus* GG strain-specific PCR system targeting the putative transposase gene (Ahlroos and Tynkkynen, 2009) produced an amplicon of the expected size from eight of the tested 41 *L. rhamnosus* strains. Sequence analysis revealed that the eight strains shared completely identical sequences of the putative transposase gene, including *L. rhamnosus* GG. In contrast, the second *L. rhamnosus* GG specific PCR system targeting a phage-related gene produced an amplicon of an expected size from five of the 41 *L. rhamnosus* strains tested. The five strains were included in the eight strains that produced amplicons with the PCR system targeting the putative transposase gene. The strains LMG 18025, LMG 18030, and LMG 18038, originating from Egyptian fermented milk products, produced an amplicon with the first system but not with the second (Table 3).

Table 3. *L. rhamnosus* strains used in study **I** and their PCR reactions with two LGG strain-specific primer sets.

			PCR reaction with	
Strain	**Source**	**Country (year)**	**Transposase***	**Phage†**
GG (ATCC 53103)	Infant feces	USA	+	+
DSM 20021^T			-	-
LMG 8153	Healthy adult urethra	Canada	-	-
LMG 10768	Human blood	Sweden (1988)	-	-
LMG 10770	Human bowel drain	Sweden (1989)	-	-
LMG 10772	Wine	Sweden (1989)	-	-
LMG 10773	Human sputum	Sweden (1989)	-	-
LMG 10775	Hip punction	Sweden (1990)	-	-
LMG 10776	Human pleura	Sweden (1990)	-	-
LMG 12166	Homemade cheese	Yugoslavia (1968)	-	-
LMG 18020	Mish cheese	Egypt (1995)	-	-
LMG 18025	Domiatti cheese	Egypt (1992)	+	-
LMG 18028	Laban Rayeb	Egypt (1994)	-	-
LMG 18030	Zabady (yogurt)	Egypt (1992)	+	-
LMG 18038	Domiatti cheese	Egypt (1992)	+	-
LMG 19716	Human blood	Denmark	-	-
LMG 19717	Human blood	Denmark	-	-
LMG 19720	Human blood	Denmark	-	-
LMG 23277	Human blood	Belgium (1993)	-	-
LMG 23278	Human pleural fluid	Belgium (1994)	-	-
LMG 23280	Human blood	Belgium (1995)	-	-
LMG 23282	Human blood	Belgium (1997)	-	-
LMG 23285	Human pus diverticulitis	Belgium (1995)	-	-
LMG 23288	Human throat	Belgium	-	-
LMG 23294	Human blood	Belgium (1996)	-	-
LMG 23296	Human bronchoalveolar lavage fluid	Belgium (1999)	-	-
LMG 23299	Human upper respiratory tract	Belgium (1993)	-	-
LMG 23302	Human bronchoalveolar lavage fluid	Belgium (1997)	-	-
LMG 23304	Human feces	Belgium	-	-
LMG 23320	Human blood	Finland	+	+
LMG 23325	Human blood	Finland	+	+
LMG 23522	Cheese	Finland	-	-
LMG 23524	Human feces	Italy	-	-
LMG 23525	Cheese	Italy	-	-
LMG 23527	Human feces	Finland (1999)	-	-
LMG 23534	Human feces	Finland (1999)	+	+
LMG 23550	Human blood	UK (1990)	-	-
LMG 23553	Human blood	UK (1990)	-	-
LMG 23667	Infant feces	Sweden	-	-
LMG 25859	Starter culture of probiotic yogurt drink	Netherlands (2010)	+	+
LMG 25881	Dairy product	China (2010)	-	-

*Putative transposase gene-targeting system

† Phage-related system

The fingerprinting analysis by Rep-PCR, RAPD and ERIC PCR was carried out to distinguish the *L. rhamnosus* strains which produced amplicons with either one of the *L. rhamnosus* GG specific PCR systems. *L. rhamnosus* DSM 20021^T was included as a negative control. This analysis revealed that the strains originating from Egyptian fermented milk products (LMG 18025, LMG 18030 and LMG 18038) and DSM 20021^T were genotypically distinct from *L. rhamnosus* GG, while strains LMG 23320 and LMG 23325 originating from human blood in Finland, LMG 23534 originating from human feces in Finland, and LMG 25859, a dairy starter strain, produced very similar profiles to *L. rhamnosus* GG. Clustering analysis of the fingerprinting data indicated that the tested strains fell into two clusters (Figure 10). The first cluster contained the strains originating from the Egyptian fermented milk products, which were highly similar between each other, but showed less than 60% similarity with *L. rhamnosus* GG. The second cluster contained the strains *L. rhamnosus* GG, LMG 23520, LMG 23525, LMG 23534, and LMG 25859 and they shared over 90% similarities. The *L. rhamnosus* type strain, DSM 20021^T, was distinct from the other strains tested.

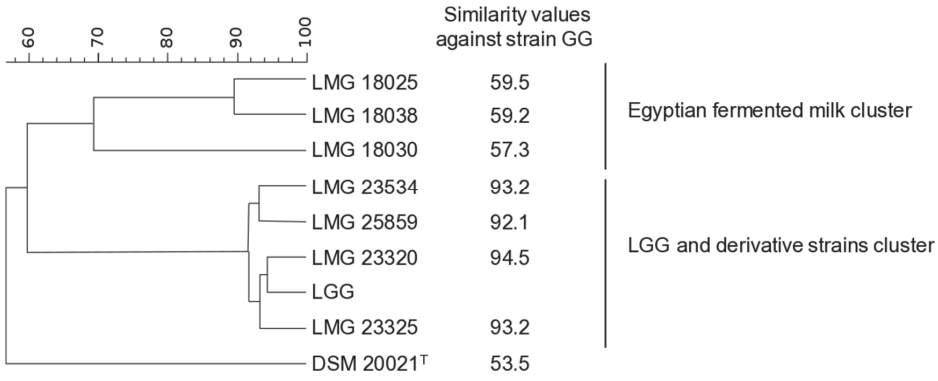


Figure 10. Dendrogram based on cluster analysis of the composite data set of the ERIC PCR, RAPD and rep-PCR fingerprinting

5.2 *Lactobacillus* microbiota of Finnish and Malawian infants (II)

Clinical characteristics of the study **II** infants are presented in Table 4. In Malawi, all infants were breastfed and received local complementary foods including liquids such as tea, water and rice water, and solid foods such as thin maize porridge and bean and fish soups. In Finland, three of the 31 infants were exclusively breastfed, while 18 infants received breast milk and complementary foods such as pureed vegetables (e.g. potato, carrot), fruits (e.g.

peach, banana and blueberry), chicken or fish meat, and cereals (e.g. oats and rice). Ten of the Finnish infants received infant formula instead of breast milk and complementary foods (Table 4). The mean age when infant formula was introduced to these infants was 2.2 (SD 2.2) months.

Table 4. Clinical characteristics of the infants (Grześkowiak *et al.*, 2012a)

Clinical characteristics	Malawian infants (n = 44)	Finnish infants (n = 31)	P-value
Sex			
Male	20/44 (45.5%)	14/31 (45.2%)	1.000
Female	24/44 (54.5%)	18/31 (54.8%)	1.000
Age (months)	6 (5.6-6.0)	6.1 (5.9-6.4)	0.007
Weight (kg)	6.6 (6.2-7.2)	8.1 (7.5-8.7)	< 0.001
Mode of delivery (% vaginal)	Not available	24/31 (77.4%)	n/a
Breastfeeding	44/44 (100.0%)	21/31 (67.7%)	< 0.001
Breast milk + complementary food	44/44 (100.0%)	18/31 (58.1%)	< 0.001
Breast milk + complementary food + formula	Non-existent	3/31 (9.7%)	0.067
Formula + complementary food	Non-existent	10/31 (32.2%)	< 0.001
Breast milk + complementary food + animal milk*	18/44 (40.9%)	None	<0.001

*animal milk includes cow's and goat's milk
The P-value indicates the differences between the two groups as analyzed by the Mann-Whitney U or Fisher's exact tests.

Lactobacilli were detected in all study **II** infants in Malawi and Finland, but the Malawian infants harboured higher counts of lactobacilli than the Finnish infants (7.45 log gene copies g^{-1} [95% CI 7.32–7.57] vs. 6.86 log gene copies g^{-1} [95% CI 6.57–7.16]; P < 0.001, respectively) based on qPCR analysis. Based on PCR-DGGE analysis, bacterial richness determined based on the number of DNA fragments per sample was higher in Malawian compared with the Finnish infants (5 bands [IQR 3.25–6.00] vs. 4 bands [IQR 2.00–4.00]; P = 0.001, respectively). In addition, diversity (Shannon diversity index, H') of the *Lactobacillus*-group microbiota was higher in Malawian infants than in the Finnish infants (1.31 [95% CI 1.21–1.41] vs. 0.98 [95% CI 0.83 – 1.13]; P < 0.001, respectively).

We also observed differences in the species composition of the *Lactobacillus* microbiota between the Malawian and Finnish infants. The Malawian infants had the species *Leuconostoc citreum*, *Weissella confusa, L. ruminis*, *L. gasseri*, *L. acidophilus* and *L. mucosae* more often than the Finnish infants 100.0% vs 74.2%, $P < 0.001$; 95.5% vs 41.9%, $P < 0.001$; 59.1% vs 0.0%, $P < 0.001$; 38.6% vs 9.7%, $P = 0.005$; and 29.5% vs 0%, $P < 0.001$; 22.7% vs 3.2%, $P < 0.017$, respectively) (Table 5). However, *L. casei* group, *L. delbrueckii* and *L. parabuchneri* were detected more often in the Finnish infants compared with Malawians (0.0% vs. 58.1%, $P < 0.001$; 0.0% vs. 19.4%, $P < 0.004$; 0% vs. 12.9%, $P < 0.026$, respectively) (Table 5).

Table 5. Identification of the PCR-DGGE bands by sequencing and their presence in the fecal samples of six-month-old Malawian and Finnish infants.

Best Blast hit (best cultured hit, similarity)	**Band position**	**Detected bands**	**% of Malawi (n = 44)**	**% of Finland (n = 31)**	**p (Fisher's exact test)**
Leuconostoc citreum	54.9	67	100.0	74.2	< 0.001
Weissella confusa	45.9	55	95.5	41.9	< 0.001
Uncultured (*L. ruminis* 99%)	58.8	26	59.1	0.0	< 0.001
L. gasseri	40.1	20	38.6	9.7	0.004
L. casei group	80.0	18	0.0	58.1	< 0.001
Uncultured (*L. fermentum* 98%)	43.3 and 47.3	16	22.7	19.4	0.478
L. acidophilus	42.3	13	29.5	0.0	< 0.001
Uncultured (*L. mucosae* 98%)	76.7 and 82.2	11	22.7	3.2	0.017
Not sequenced	41.4	8	13.6	6.5	0.275
Uncultured (*L. oris* 99%)	64.4	8	11.4	9.7	0.565
L. delbrueckii	50.4	6	0.0	19.4	0.004
L. paracasei	81.4	4	4.5	6.5	0.551
L. parabuchneri	61.0	4	0.0	12.9	0.026

A total of 18 Finnish samples (58.1% in total), which produced *L. casei* group bands in DGGE profiles, were further analyzed by species-specific PCR with *L. rhamnosus*-, *L. paracasei*-, and *L. casei* specific primer sets and an *L. rhamnosus* GG strain-specific primer set. All of the 18 samples presented *L. rhamnosus*, 8 of them presented *L. paracasei*, but none had *L. casei*. Furthermore, the strain *L. rhamnosus* GG was detected in 10 of the Finnish samples. Interestingly, *L. rhamnosus* group was more common in breastfed infants than in the formula fed infants in Finland (76.2% vs. 20.0%, $p = 0.005$).

UPGMA analysis of the DGGE fingerprints based on the Pearson correlation coefficients revealed two main clusters (Figure 11). Cluster 1 composed of infants from both countries and was characterized by bright bands identified as *Leuconostoc citreum* and *Weissella confusa*. Moreover, Cluster 1 included a Malawian sub-cluster that was distinguished by a strong *L. ruminis* band. Cluster 2 composed of only Finnish infants, which produced a bright *L. casei* group band and a weak or no *Leuconostoc* or *Weissella* bands. Similar clustering was seen in principal components analysis (Figure 12), which was confirmed by PERMANOVA analysis ($P < 0.001$; $F = 9.96$).

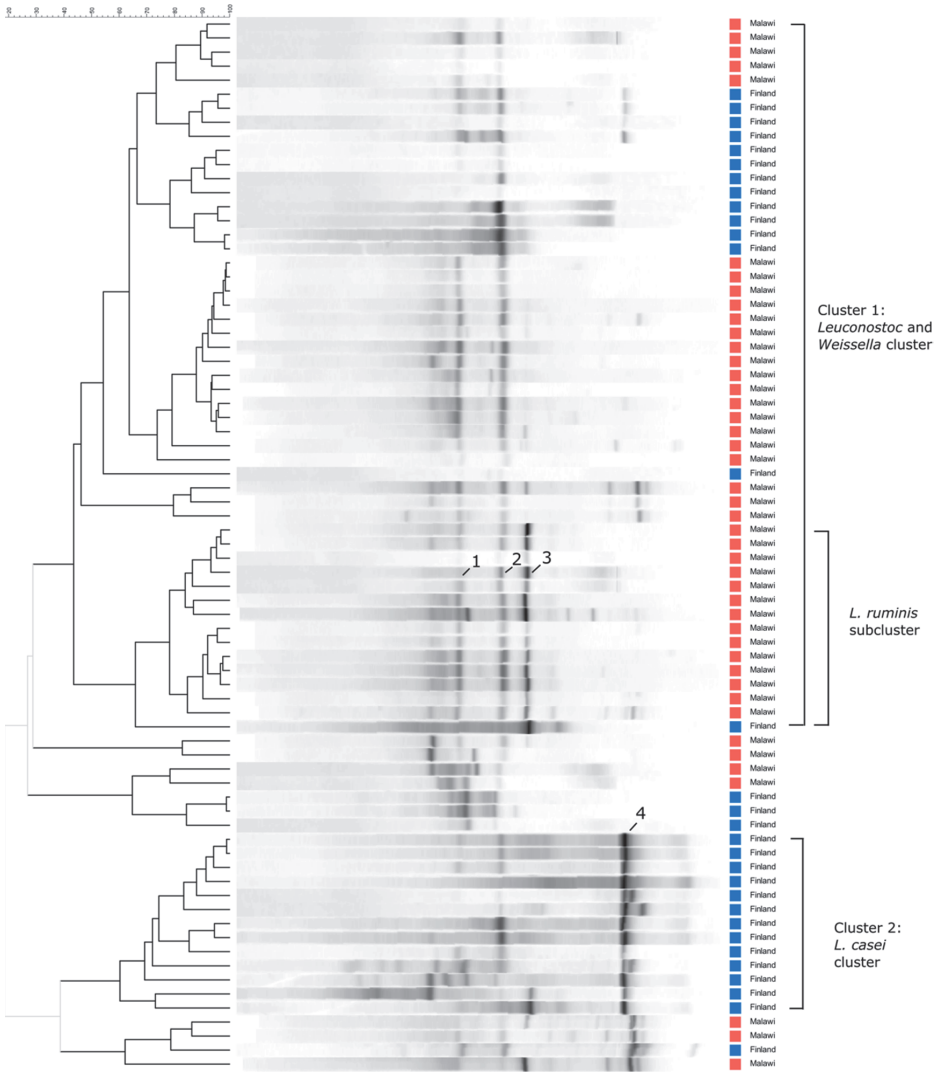


Figure 11. UPGMA dendrogram illustrating the correlation between the DGGE-fingerprints of Malawian and Finnish samples. The bands that had major influence in the clustering analysis are numbered and identified as follows: 1. *Weissella confusa*, 2. *Leuconostoc citreum*, 3. *Lactobacillus ruminis* and, 4. *Lactobacillus casei* group.

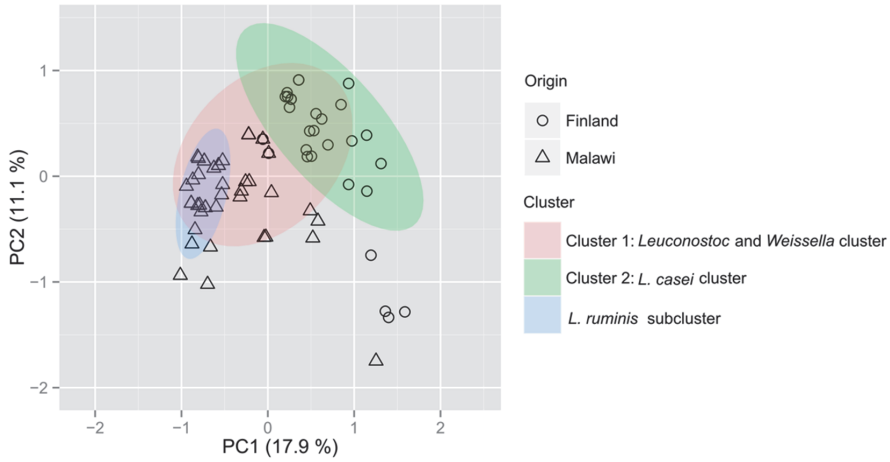


Figure 12. Principal component analysis score plot based on the *Lactobacillus* DGGE-fingerprints of Malawian and Finnish infants' faecal samples. Principal components 1 and 2 accounted for 17.9% and 11.1% of the variance, respectively. The cluster names correspond to the ones shown in Figure 11.

5.3 Impact of lipid-based nutrient supplements on Malawian infants' microbiota (III)

Clinical characteristics of the study **III** infants are presented in Table 6. No clinically significant differences were detected between the intervention groups at baseline. Only two of the study infants had been exclusively breastfed, while the others had received complementary local foods (Table 6). Five infants in study **III** were hospitalized during the follow up due to pneumonia, anemia or malaria, but only two received antibiotics during the study period (Table 6)

Table 6. Clinical characteristics of the study infants

Clinical characteristics	Control (n = 40)	Milk-LNS (n = 40)	Soy-LNS (n = 40)	CSB (n = 40)
Male sex	19/40 (47.5 %)	16/40 (40.0 %)	18/40 (45.0 %)	21/40 (52.5 %)
Age (months)	5.94 (0.29)	5.93 (0.29)	5.94 (0.28)	5.98 (0.30)
Weight (kg)	6.87 (0.69)	7.05 (0.86)	6.86 (0.89)	7.03 (0.93)
Length (cm)	62.6 (1.9)	63.3 (2.5)	62.5 (2.9)	62.8 (2.0)
Weight-for-age z score	-0.89 (0.88)	-0.62 (0.96)	-0.87 (1.22)	-0.78 (1.00)
Length-for-age z score	-1.79 (0.91)	-1.38 (1.17)	-1.78 (1.17)	-1.78 (0.78)
Weight-for-length z score	0.43 (0.98)	0.45 (0.82)	0.45 (1.33)	0.56 (1.06)
Breastfeeding	40/40 (100.0%)	40/40 (100.0%)	40/40 (100.0%)	40/40 (100.0%)
Breastfeeding + complementary food (6 m)	40/40 (100.0%)	38/39 (97.4%)	39/40 (97.5%)	40/40 (100.0%)
Breastfeeding + complementary food (12 m)	38/38 (100.0%)	39/39 (100.0%)	39/39 (100.0%)	39/39 (100.0%)
Breastfeeding + complementary food + animal milk* (6 m)	11/40 (27.5%)	15/40 (37.5%)	12/40 (30%)	15/40 (37.5%)
Breastfeeding + complementary food + animal milk* (12 m)	23/38 (57.5%)	20/39 (50%)	22/39 (55.0%)	22/39 (55.0%)
No. treated with antibiotics from 6 to 12 m	0/40 (0.0%)	1/40 (2.5%)	0/40 (0.0%)	1/40 (2.5%)
No. hospitalized during study period	1/40 (2.5%)	2/40 (5.0%)	2/40 (5%)	0/40 (0.0%)

*animal milk includes cow's and goat's milk

Data are presented as a mean with standard deviation (SD), unless otherwise stated. CSB, corn-soy blend; milk-LNS, milk powder containing lipid-based nutrient supplement; soy-LNS, soy-flour containing lipid-based nutrient supplement.

Bifidobacteria were detected in all study **III** infants and the highest colonization rates were noted for the species *B. longum*, *B. bifidum, B. infantis* and *B. breve*, which were all detected in at least 90% of the infants at baseline (Table 7). In general, bifidobacteria seemed to form a major part of the gut microbiota of the study **III** infants with bacterial counts reaching 10^{10} cells g^{-1} of feces and higher for the species *B. longum* and *B. infantis* across all intervention groups at baseline samples from the 6 month old infants (Table 8).

The dietary supplements did not have an impact on the colonization rates or bacterial counts when the analysis was controlled for baseline values ($p > 0.05$). However, changes in the microbiota composition were observed in the whole population over time. At twelve months, all study **III** infants still harboured bifidobacteria in their feces, but the colonization rates of *B. longum*, *B. bifidum*, *B. infantis*, *B. breve* and *B. lactis* were lower at twelve months compared to six months of age (100% vs. 91.7%, $p = 0.001$; 94.5% vs. 82.8%, $p = 0.001$; 94.5% vs. 84.1%, $p = 0.012$; 95.9% vs. 83.4%, $p = 0.001$; and 78.6% vs. 60.0%, $p < 0.001$, respectively) (Table 7). In addition, *S. aureus* was detected at a lower abundance at twelve than at six months of age (31.0% vs. 13.8%, $p < 0.001$) (Table 7).

We also found shifts in the total bacterial counts over time. The total bacterial count was slightly higher at twelve months of age (11.1 log gene copies g^{-1}, IQR 10.2 – 11.9 vs. 11.4 log gene copies g^{-1}, IQR 10.8 – 11.8, $p = 0.028$). Also the *Bifidobacterium* genus counts were slightly higher (10.5 log gene copies g^{-1}, IQR 9.8 – 11.0 vs. 10.7 log gene copies g^{-1}, IQR 10.2 – 11.1, $p = 0.027$) even though the colonization rates of different *Bifidobacterium* species were generally lower after the intervention (Table 8). However, counts of *B. infantis*, *B. lactis* and *B. longum* decreased (10.5 log gene copies g^{-1}, IQR 9.1 – 11.5 vs. 9.4 log gene copies g^{-1}, IQR 7.6 – 10.7, $p < 0.001$; 4.0 log gene copies g^{-1}, IQR 3.3 – 4.7 vs. 3.3 log gene copies g^{-1}, IQR 0.0 – 4.4, $p < 0.001$; and 10.1 log gene copies g^{-1}, 9.5 – 10.7 vs. 9.8 log gene copies g^{-1}, 8.9 – 10.3, $p < 0.001$, respectively) while *B. catenulatum* increased (0.0 log gene copies g^{-1}, IQR 0.00 – 2.9 vs. 4.0 log gene copies g^{-1}, IQR 0.0 - 7.9, $p = 0.031$) at twelve months compared to six months of age (Table 8). Moreover, *S. aureus* counts were lower (0.0 log gene copies g^{-1}, IQR 0.0 – 3.1 vs. 0.0 log gene copies g^{-1}, IQR 0.0 – 0.0, $p < 0.001$) at twelve months of age (Table 8).

Table 7. Detection rates of the measured bacteria in the different intervention groups presented as percentages at baseline (six-month-old) and after dietary intervention (twelve-month-old), analyzed by quantitative polymerase chain reaction

Bacterial group	Time point	Control (n = 35)	Milk-LNS (n = 38)	Soy-LNS (n = 34)	CSB (n = 38)	p-value*
Total bacteria	6 months	100.0	100.0	100.0	100.0	
	12 months	100.0	100.0	100.0	100.0	1.000
Bifidobacterium genus	6 months	100.0	100.0	100.0	100.0	
	12 months	100.0	100.0	100.0	100.0	1.000
B. longum	6 months	100.0	100.0	100.0	100.0	
	12 months	91.4	97.4	82.4	94.9	0.194
B. catenulatum	6 months	48.6	36.8	38.2	53.8	
	12 months	65.7	60.5	35.3	48.7	0.202
B. bifidum	6 months	91.4	94.7	94.1	94.9	
	12 months	77.1	89.5	82.4	82.1	0.811
B. infantis	6 months	94.3	97.4	91.2	94.9	
	12 months	88.6	92.1	73.5	79.5	0.955
B. adolescentis	6 months	8.6	5.3	0.0	0.0	
	12 months	0.0	10.5	2.9	0.0	1.000
B. breve	6 months	97.1	94.7	94.1	94.9	
	12 months	82.9	84.2	85.3	82.1	0.825
B. lactis	6 months	74.3	71.1	79.4	87.2	
	12 months	42.9	63.2	61.8	69.2	0.520
Staphylococcus aureus	6 months	40.0	34.2	20.6	28.2	
	12 months	14.3	7.9	8.8	23.1	0.295

* The p-value (generalized estimating equations, GEE) indicates a difference between groups in change of bacterial prevalence between time points. CSB, corn-soy blend; milk-LNS, milk powder containing lipid-based nutrient supplement; soy-LNS, soy-flour containing lipid-based nutrient supplement

Table 8. Bacterial counts (log gene copies g^{-1}) presented as medians with interquartile range in fecal samples at baseline (six-month-old) and after dietary intervention (twelve-month-old), analyzed by quantitative polymerase chain reaction

Bacterial group	Time point	Control (n = 35)	Milk-LNS (n = 38)	Soy-LNS (n = 34)	CSB (n = 38)	p-value*
Total bacteria	6 months	11.1 (10.0 - 11.5)	10.7 (10.1 - 11.7)	10.7 (10.1 - 11.5)	11.7 (11.0 - 12.4)	
	12 months	11.4 (11.1 - 11.9)	11.4 (10.5 - 11.8)	11.2 (10.8- 11.7)	11.3 (10.8 - 11.9)	0.165
Bifidobacterium genus	6 months	10.6 (9.8 - 11.0)	10.5 (9.8 - 10.9)	10.2 (9.8 - 10.9)	10.6 (9.9 - 11.4)	
	12 months	10.9 (10.3 - 11.3)	10.6 (10.1 - 11.2)	10.7 (10.0 - 11.0)	10.7 (10.2 - 11.1)	0.460
B. longum	6 months	10.1 (9.6 - 10.6)	10.0 (9.5 - 10.8)	9.9 (9.4 - 10.5)	10.5 (9.7 - 10.9)	
	12 months	9.6 (8.9 - 10.3)	9.8 (9.0 - 10.5)	9.6 (8.3 - 10.0)	9.7 (9.1- 10.2)	0.527
B. catenulatum	6 months	0 (0.0 - 5.6)	0 (0.0 - 4.9)	0 (0.0 - 5.3)	4.9 (0.0 - 8.1)	
	12 months	4.6 (0.0 - 8.6)	5.0 (0.0 - 8.2)	0 (0.0 - 4.1)	2.0 (0.0 - 7.2)	0.710
B. bifidum	6 months	7.2 (5.5 - 8.5)	7.2 (5.7 - 8.4)	7.3 (6.2 - 8.2)	7.5 (6.8 - 8.5)	
	12 months	7.0 (4.6 - 8.1)	8.2 (6.6 - 8.9)	6.9 (4.1 - 8.5)	7.1 (6.0 - 9.0)	0.418
B. infantis	6 months	10.5 (8.9 - 11.5)	10.5 (9.4 - 11.6)	10.2 (9.1 - 11.0)	11.2 (10.0 - 12.1)	
	12 months	10.0 (7.6 - 10.7)	9.7 (7.9 - 10.8)	9.4 (0.0 - 10.7)	9.2 (7.6 - 10.1)	0.094
B. adolescentis	6 months	0 (0.0 - 0.0)	0 (0.0 - 0.0)	ND	ND	
	12 months	ND	0 (0.0 - 0.0)	0 (0.0 - 0.0)	ND	-
B. breve	6 months	7.9 (7.3 - 8.9)	8.3 (7.0 - 8.8)	8.7 (7.3 - 9.4)	8.7 (7.8 - 9.6)	
	12 months	8.5 (4.0 - 9.1)	9.0 (7.6 - 9.6)	8.5 (6.2 - 9.2)	8.9 (5.9 - 9.8)	0.850
B. lactis	6 months	3.6 (0.0 - 4.4)	3.5 (0.0 - 4.6)	4.3 (3.1 - 4.9)	4.3 (3.3 - 4.8)	
	12 months	0 (0.0 - 3.4)	3.4 (0.0 - 4.4)	3.3 (0.0 - 4.4)	3.8 (0.0 - 4.6)	0.074
Staphylococcus aureus	6 months	0 (0.0 - 3.8)	0 (0.0 - 3.4)	0 (0.0 - 0.0)	0 (0.0 - 2.9)	
	12 months	0 (0.0 - 0.0)	0 (0.0 - 0.0)	0 (0.0 - 0.0)	0 (0.0 - 0.0)	0.532

* The p-value (non-parametric ANCOVA) indicates a difference in the bacterial counts at twelve months of age between intervention groups when controlling for baseline. CSB, corn-soy blend; milk-LNS, milk powder containing lipid-based nutrient supplement; ND, not detected; soy-LNS, soy-flour containing lipid-based nutrient supplement

5.4 Microbial contact *in utero* and in early infancy (IV)

Altogether 15 mother-infant pairs from whom maternal feces, placenta, amniotic fluid, colostrum, meconium and infant feces samples were available for analysis were included in study **IV**. All study neonates were born by elective cesarean section at full term with no symptoms or signs of intrauterine infection. One mother received antenatal antibiotic therapy with clindamycin for a bacterial infection of the skin whilst the remaining 14 mothers received no antibiotics prior to or during the cesarean section. None of the neonates were administered antibiotics during the test period. Clinical details of the study **IV** mother-infant pairs are presented in Table 9.

Table 9. Clinical characteristics of mother-infant pairs of study **IV**

Mother/Infant	Reason for cesarean section	Prenatal antibiotics	Gestational age (weeks)	Gender	Birthweight (grams)	Apgar score at 5 minutes
1	breach position	clindamycin	39 1/7	male	3400	9
2	hemorrhoids	none	39 3/7	male	3750	9
3	pelvic diameter	none	39 1/7	female	3230	9
4	breach position	none	39 3/7	male	3470	9
5	pelvic diameter	none	38 6/7	female	3100	9
6	previous section	none	39 3/7	female	4325	7
7	pelvic diameter	none	39	female	3720	9
8	fear of childbirth	none	39 1/7	female	3595	9
9	breach position	none	39 1/7	female	3980	8
10	breach position	none	39	male	3490	10
11	breach position	none	39	female	3000	10
12	previous section	none	39	female	3250	10
13	placenta praevia	none	38 1/7	female	2770	9
14	previous section	none	38 6/7	male	3200	9
15	fear of childbirth	none	39 1/7	male	3670	9

Based on 16S rRNA gene pyrosequencing and PCR-DGGE analysis, we found a unique microbial community in the placenta and amniotic fluid that was clearly distinct from the microbial community in the maternal feces (Figure 13, 14 and 15). We could not obtain PCR products from the negative controls for further analysis. The microbial populations in the amniotic fluid and placenta samples were highly consistent across individuals with low abundance, low richness and low diversity (Figure 14, 15 and 16).

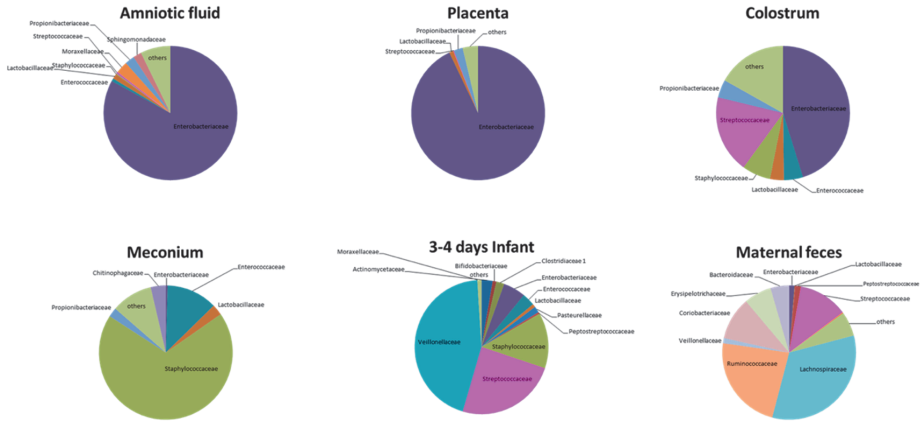


Figure 13. Diversity of the most abundant bacterial taxa identified at family level.

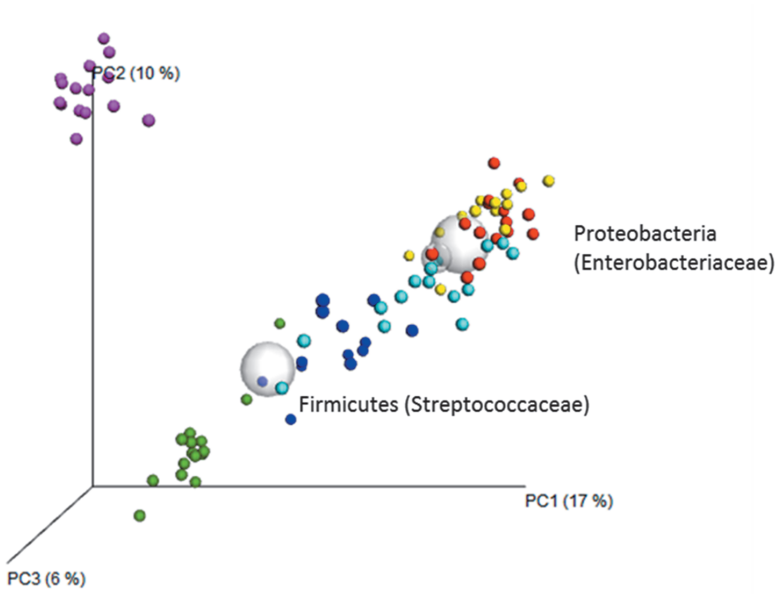


Figure 14. Principal Coordinate Analysis (PCoA) plot based on the unweighted UniFrac distance matrix. Individual samples of maternal feces (pink), placenta (yellow), amniotic fluid (red), colostrum (light blue), meconium (dark blue), and infant feces (green) are presented as individual points.

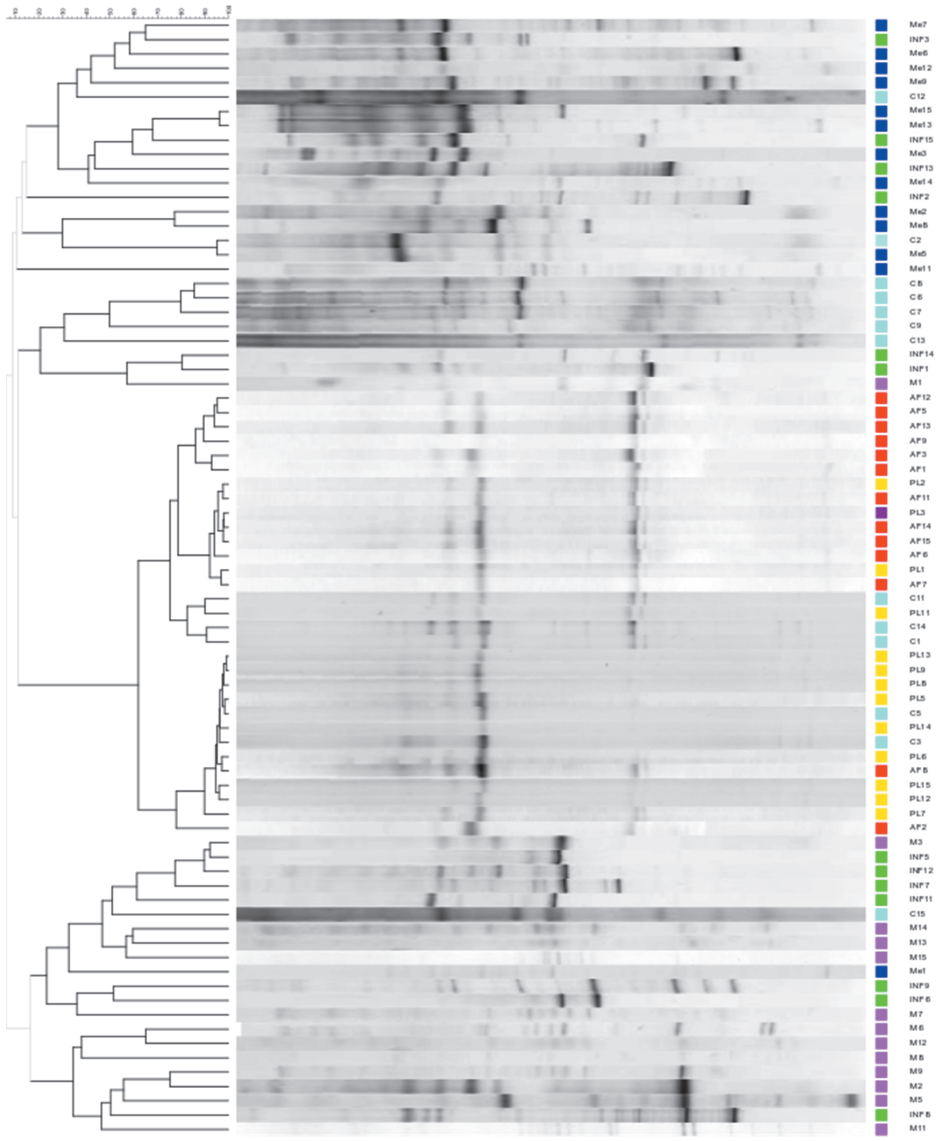


Figure 15. UPGMA Cluster analysis using Pearson correlation coefficient performed on PCR-DGGE analysis fingerprints from mother feces (M), placenta (PL), amniotic fluid (AF), colostrum (C), meconium (Me) and infant feces (INF) samples. UPGMA = unweighted pair-group method using arithmetic averages.

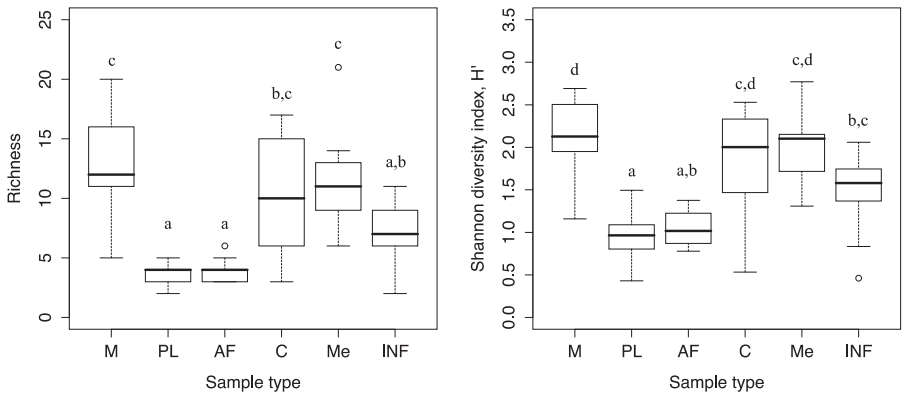


Figure 16. Richness and diversity of the gut microbiota in mother feces (M), placenta (PL), amniotic fluid (AF), colostrum (C), meconium (Me) and infant feces (INF) samples. The letters above the box-plots indicate results of Tukey HSD test following a significant 1-way ANOVA. Mean values not sharing the same letters are significantly different from each other ($p < 0.05$)

Proteobacteria was the predominant phylum in the amniotic fluid and placenta samples with a high abundance of species belonging to the Enterobacteriaceae family. At genus level, *Enterobacter* and *Escherichia*/*Shigella* were the most predominant Proteobacteria in these samples (Figure 13). These genera were also detected in colostrum, meconium and infant feces, but their levels were relatively low compared to other genera. *Propionibacterium* was the second most predominant genus detected in amniotic fluid and placenta, and was also detected in meconium. *Streptococcus* genus was present in the amniotic fluid, placenta and meconium samples, but in low abundance (< 1%). However, this genus was present in higher abundance in colostrum and infant feces (12% and 24%, respectively). *Staphylococcus* genus was also present in amniotic fluid (< 1%) and placenta (< 1%), but in lower abundance when compared to meconium samples (20%). Moreover, *Lactobacillus* genus was present in amniotic fluid, placenta, colostrum and meconium samples. Interestingly, we successfully retrieved staphylococci and propionibacteria by culturing from the placenta and amniotic fluid samples (Table 10).

In colostrum samples, the microbiota composition was more varied with higher richness and diversity (Figure 16) compared to other sample types. Similar to amniotic fluid and placenta samples, Proteobacteria and Enterobacteriaceae in particular, were frequent in the colostrum samples (Figure 13).

Table 10. Identification of cultured microbial isolates from amniotic fluid and placenta by sequencing of a partial 16S rRNA gene.

Source	Identification Ribosomal Database	SeqMatch	Seqmatch score
Amniotic fluid	*Propionibacterium* (uncultured)	S000565825	0.930
Amniotic fluid	*Propionibacterium granulosum*	S001331259	0.987
Amniotic fluid	*Propionibacterium acnes*	S001549679	0.983
Amniotic fluid	*Propionibacterium* (uncultured)	S001028822	0.994
Amniotic fluid	*Streptomyces puniceus*	S000581716	0.579
Amniotic fluid	*Propionibacterium granulosum*	S001331269	0.985
Amniotic fluid	*Propionibacterium acnes*	S001549679	0.964
Amniotic fluid	*Staphylococcus* (uncultured)	S001245567	0.491
Amniotic fluid	*Staphylococcus epidermis*	S00345154	1.000
Amniotic fluid	*Propionibacterium acnes*	S001381916	0.974
Amniotic fluid	*Lachnospiraceae* (uncultured)	S001245230	0.800
Amniotic fluid	*Propionibacterium* (uncultured)	S001245230	0.952
Amniotic fluid	*Propionibacterium* (uncultured)	S001245230	0.974
Amniotic fluid	*Propionibacterium* (uncultured)	S000381988	0.969
Amniotic fluid	*Staphylococcus lugdunensis*	S000381988	0.978
Amniotic fluid	*Propionibacterium* (uncultured)	S001089380	0.986
Placenta	*Propionibacterium acnes*	S001549675	0.955
Placenta	*Propionibacterium acnes*	S001549679	0.971
Placenta	*Propionibacterium* (uncultured)	S001247911	1.000
Placenta	*Propionibacterium* (uncultured)	S001245230	9.966
Placenta	*Staphylococcus pasteuri*	S000995943	1.000
Placenta	*Staphylococcus warneri*	S001794135	0.991
Placenta	*Propionibacterium* (uncultured)	S001248017	0.958
Placenta	*Propionibacterium* (uncultured)	S001245230	0.978

The meconium microbiota shared features with the placenta, amniotic fluid and colostrum microbiota (Figure 17). Forty-one of the 75 bacterial family phylotypes at the family level detected in meconium samples were also present in amniotic fluid and placenta samples. Additional 15 phylotypes at family level were shared between meconium and amniotic fluid samples, but not with placenta samples. In some mother-infant pairs, the meconium microbiota was similar to colostrum microbiota. Of the 75 bacterial family phylotypes detected in meconium, 54 were also observed in colostrum (Figure 17). Also, colostrum, amniotic fluid and placenta samples shared features in their microbial composition to some extent (Figure 17). The microbial composition in infant fecal samples collected later during the first week of life was clearly distinct from the meconium microbiota and resembled more that in colostrum (Figures 15 and 17).

Taken together, a core microbiome consisting of OTUs detected in 50% of each sample type was found to include *Streptococcus*, unclassified Enterobacteriaceae, *Propionibacterium*, *Lactobacillus* and unclassified *Bacillales*. Bacteria belonging to *Escherichia*/*Shigella*, *Lactobacillus* and

Propionibacterium were present in all placenta, amniotic fluid and meconium samples.

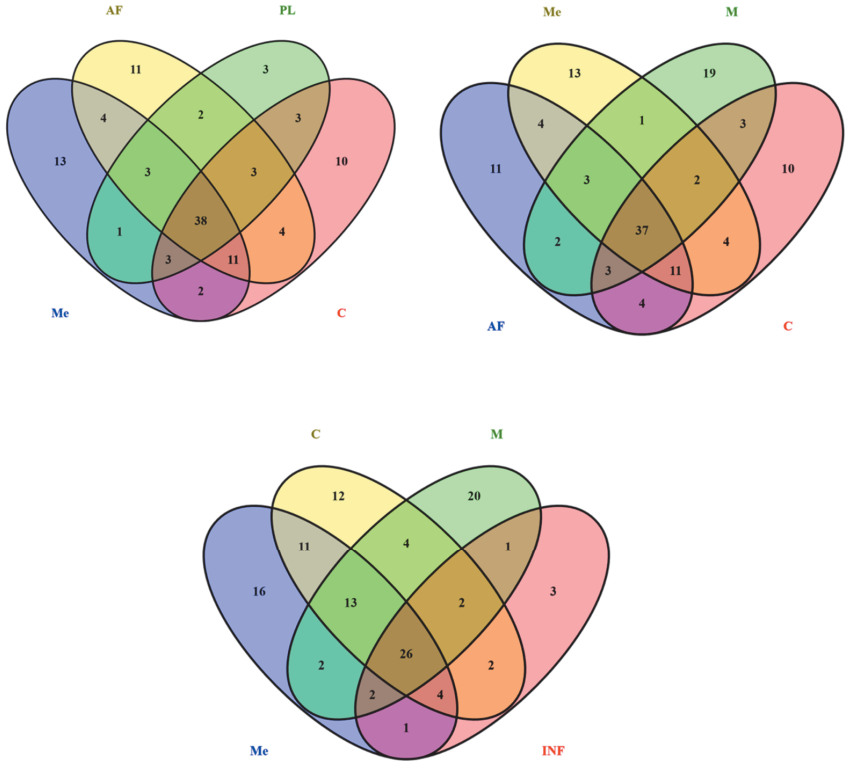


Figure 17. Venn diagrams of exclusive and shared bacterial phylotypes at family level in mother feces (M), placenta (PL), amniotic fluid (AF), colostrum (C), meconium (Me) and infant feces (INF) samples based on 16S rRNA gene pyrosequencing.

5.5 Influence of HMOs on breast milk microbiota composition (V)

In study **V**, eleven colostrum samples (from mothers engaged in study **IV**) were available for microbial and HMO analysis. Clinical characteristics of the study **V** individuals are presented in Table 11.

Table 11. Clinical characteristics of the mother-infant pairs of study **V.**

Mother/ Infant	Gestational age (weeks)	Gender	Birthweight (grams)	Apgar score at 5 minutes
1	39 1/7	female	3230	9
2	39 3/7	male	3470	9
3	38 6/7	female	3100	9
4	39 3/7	female	4325	7
5	39	female	3720	9
6	39 1/7	female	3595	9
7	39 1/7	female	3980	8
8	37	female	3850	9
9	39	male	3490	10
10	37 4/7	male	2740	9
11	39	female	3000	10

The mean concentrations of the HMOs detected in the colostrum samples are presented in Table 12. The most abundant HMOs in the samples were 2'-fucosyllactose (2'-FL), lacto-N-fucopentaose (LNFP I), difucosyl-lacto-N-tetraose (DF-LNT) and lacto-N-tetraose (LNT) with mean concentrations (μg/ml) of 2 733 (SD 512), 1832 (SD 639) and 958 (SD 218), respectively (Table 12). The oligosaccharide profile of the human milk is dependant on the expression of glycosyltransferases and the secretor type which is defined by the ability to encode alpha1,2-fucosyltransferase FUT2 (Kobata, 2010). In study **V**, all mothers were identified as secretors as indicated by the presence of 2'-FL or LNPF I in the sample.

Table 12. Mean concentrations (μg/ml) with standard deviations of human milk oligosaccharides in colostrum samples (n = 11), analyzed by high-pressure liquid chromatography

HMO	**Mean (SD)**
2'FL	2733 (512)
3FL	50 (34)
LNnT	409 (82)
3'SL	413 (90)
LNT	958 (219)
LNFP I	1832 (639)
LNFP II	466 (146)
LNFP III	65 (21)
LSTb	61 (19)
LSTc	475 (248)
DF-LNT	1140 (512)
DS-LNT	285 (88)
FLNH	93 (41)
DFLNH	113 (41)
FDSLNH	86 (40)
DSLNH	96 (63)

HMO, human milk oligosaccharide; SD, standard deviation

Bacterial counts (log gene copies g^{-1}) determined by qPCR are presented in Table 13. *Bifidobacterium* spp., *Streptococcus* spp., *Bacteroides-Prevotella* group and *Lactobacillus* group were detected in all of the colostrum samples (Table 13). *Bifidobacterium* spp. was the most abundant bacterial group followed by the *Lactobacillus* group (4.6, IQR 4.5 – 4.8; and 3.4, IQR 4.5 – 4.8, respectively). *B. longum* group was the most abundant bifidobacterial group followed by *B. bifidum* (Table 13). In contrast, *B. breve* was less common and detected in less than half of the samples (Table 13).

Table 13. Median bacterial counts with interquartile range (IQR) (log gene copies g^{-1}) and detection rates in colostrum samples (n=11), analyzed by qPCR

Bacterium	Median (IQR)	Detection rate (%)
Total bacteria	5.1 (5.0 – 5.8)	100.0
Bifidobacterium spp.	4.6 (4.5 – 4.8)	100.0
Bifidobacterium longum	3.9 (3.9 – 4.0)	100.0
Bifidobacterium breve	0.0 (0.0 – 3.9)	45.5
Bifidobacterium bifidum	2.7 (2.6 – 2.8)	100.0
Staphylococcus spp.	4.5 (3.7 – 4.8)	81.8
Staphylococcus aureus NUC	0.0 (0.0 – 2.0)	45.5
Streptococcus group	2.0 (1.8 – 3.0)	100.0
Lactobacillus group	3.4 (3.2 – 3.6)	100.0
Enterococcus spp.	1.7 (1.3 – 1.8)	81.8
Akkermansia muciniphila	0.9 (0.0 – 1.5)	63.6
Bacteroides-Prevotella group	2.8 (2.7 – 3.2)	100.0
Clostridium cluster XIVa - XIVb	4.6 (4.0 – 4.7)	81.8
Clostridium cluster IV	2.8 (2.5 – 3.2)	81.8

We discovered associations between different groups of HMOs (Figure 18a) and microbiota as well as individual HMO structures and microbes (Figure 18b). The total HMO concentration correlated positively with *Bifidobacterium* spp. and *B. breve* (ρ = 0.63, p = 0.036 and ρ = 0.66, p = 0.027, respectively). Moreover, fucosylated HMOs correlated positively with *Bifidobacterium* spp., and *Akkermansia muciniphila* (ρ = 0.67, p = 0.024 and ρ = 0.70, p = 0.017, respectively) while the sialylated HMOs correlated with *B. breve* (ρ = 0.84 , p = 0.001). There was also a positive correlation between HMOs that were both fucosylated and sialylated and *S. aureus* (ρ = 0.75, p = 0.007) and HMOs that were neither fucosylated nor sialylated and *B. longum* group (*B. longum* subsp. *infantis* and *B. longum* subsp. *longum*) (ρ = 0.65, p = 0.030). Between individual HMOs and the microbiota, the strongest correlations were observed for LNT and *B. longum* group, LNFP III and *B. breve*, fucosyldisialyllacto-N-hexaose (FDSLNH) and *S. aureus*, LNFP I and *Akkermansia muciniphila*, and LST c and *B. breve* (ρ = 0.82, p = 0.002; ρ = 0.79, p = 0.004; ρ = 0.75, p = 0.007 and ρ = 0.74, p = 0.010, respectively) (Figure 18b).

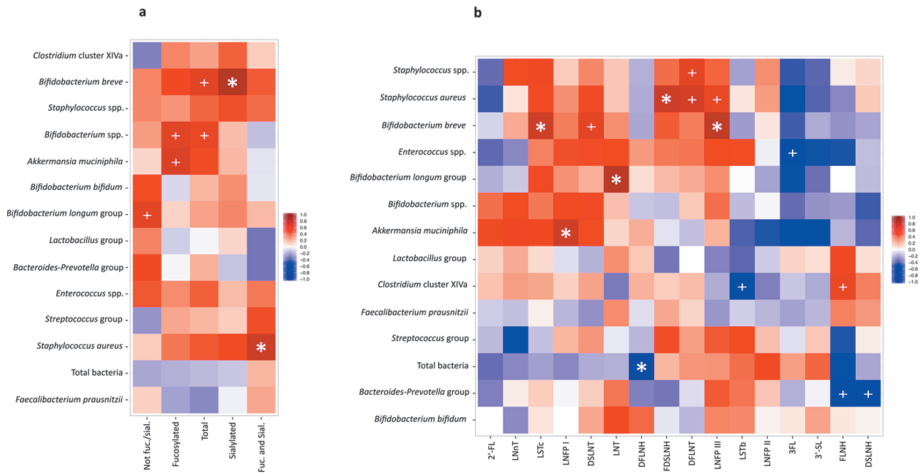


Figure 18. Heat map of spearman rank correlations between human milk oligosaccharide groups and bacterial counts (a) and between individual human milk oligosaccharides and bacterial counts (b). Negative correlations are colored in shades of blue and positive correlations are red. +, $p < 0.05$; *, $p \leq 0.01$.

6 DISCUSSION

6.1 Evaluation of *L. rhamnosus* GG specific primers (I and II)

PCR-based strain-specific identification using strain-specific primers is a valuable tool for identifying probiotics in commercial products as well as human specimens in intervention studies. The specificity of the primers used must be ensured to avoid false positive results. In study **I** we evaluated the specificity two previously published *L. rhamnosus* GG specific primer sets.

The specific primer set targeting a transposase gene gave false positive results from three strains (LMG 18025, LMG 18030 and LMG 18038) which were clearly distinguishable from *L. rhamnosus* GG based on the results of fingerprinting analysis. On the other hand, the *L. rhamnosus* GG detection system targeting the phage-related gene described by Brandt & Alatossava (Brandt and Alatossava, 2003) detected *L. rhamnosus* GG and *L. rhamnosus* GG-like strains, suggesting that this system is a more reliable tool for the detection of *L. rhamnosus* GG. However, this system still produced amplicons from *L. rhamnosus* strains that were closely related to or nearly indistinguishable from *L. rhamnosus* GG using these primers. Interestingly, all of these isolates originated from human clinical samples from Finland. Finland has a long history of using *L. rhamnosus* GG in several food matrices and thus it would be tempting to speculate that *L. rhamnosus* GG might colonize and produce derivative strains in the human gut. Moreover, one of the positive strains (LMG 25859) is used as a starter strain for the probiotic yogurt drink Yoba. The strain was reported to originate from infant feces, suggesting that the strain is also a *L. rhamnosus* GG derivative. This was confirmed by a recent comparative genomics study, which suggested that this particular strain is a *L. rhamnosus* GG derivative which is missing the genomic islands that contain genes encoding pili subunits (Sybesma *et al.*, 2013).

In the present series (**II**), we detected *L. rhamnosus* GG in half of the *L. rhamnosus* positive fecal samples of Finnish infants, although neither the infants nor the mothers received any probiotic products, which further strengthens the hypothesis that *L. rhamnosus* GG may colonize the human gut and produce derivative strains. Similar findings have been reported earlier from Finland, where almost half of the fecal samples in a placebo group contained *L. rhamnosus* GG like strains (Gueimonde *et al.*, 2006a). Nevertheless, it remains to be elucidated whether or not this hypothesis is true.

6.2 Impact of geographical location and diet on gut microbiota development during infancy (II and III)

Diet, hygiene and geographical location have been reported to have an impact on the gut microbiota development in infancy. In the present series (**II**), we compared the gut microbiota composition between Malawian and Finnish infants, representing low-income and high-income countries respectively. These two countries offer completely different environments and dietary conditions. These are also reflected in the different prevalence of microbiota associated diseases including allergies, obesity and gastrointestinal disorders in Finland, and bacterial (or viral) diarrhea and malnutrition in Malawi. Moreover, we also studied whether lipid-based nutrient supplements affect the gut microbiota composition in Malawian infants over time (**III**).

Earlier studies have indicated that Malawian infants have greater proportions of bifidobacteria, *Bacteroides-Prevotella*, and *Clostridium histolyticum* compared to Finnish infants, while the species *Bifidobacterium adolescentis*, *Clostridium perfringens*, and *Staphylococcus aureus* were only present in Finnish infants (Grześkowiak *et al.*, 2012a). In the present series (**II**), we expanded the microbiota studies to species composition of lactobacilli which have been also identified among the first microbes to colonize the human gastrointestinal tract, possibly already in utero (Rautava *et al.*, 2012b; Satokari *et al.*, 2009). Lactobacilli also form the majority of the microbiota of the birth canal (Hummelen *et al.*, 2010) and breast milk (Cabrera-Rubio *et al.*, 2012). Lactobacilli are believed to be part of a healthy gut microbiota and thus members of this group are commonly used as probiotics as well.

In study **II**, we observed that at 6 months of age, Malawian infants had a more abundant, richer and more diverse *Lactobacillus* microbiota compared to Finnish infants of the same age. This information may be useful in understanding why some gut-related diseases are more common in high-income or low-income countries. Typically, higher levels of lactobacilli have been reported in the feces of infants from areas where allergic and immune-mediated diseases are less common. For instance, higher levels of lactobacilli have been detected in Estonian versus Swedish infants (Sepp *et al.*, 1997) and also in rural Thai infants versus Singaporean infants (Mah *et al.*, 2008). However, it has been demonstrated that early colonization with specific lactobacilli is associated with the risk of developing allergy (Johansson *et al.*, 2011) and atopic dermatitis (Penders *et al.*, 2010). The mechanisms behind the protective effect of early colonization by lactobacilli against these diseases are not well understood.

Leuconostoc citreum and *Weissella confusa* were the most common lactobacilli in both the Finnish and Malawian infants, but more often detected

in the Malawians. These species most likely originate from breast milk, since *Leuconostoc* spp. and *Weissella* spp. have been reported to be among the predominant bacteria in breast milk (Cabrera-Rubio *et al.*, 2012). Surprisingly, these species have not, to our knowledge, been reported in infant fecal samples before. One possible reason for the non-detection of these two species in the former studies could be the fact that these organisms are sensitive to acids. The *Lactobacillus* microbiota was originally studied by culture-dependent methods utilizing acidified media, but the acidified media are not suitable for these acid-sensitive organisms (Dal Bello *et al.*, 2003). *W. confusa* has been shown to be less widely distributed in the feces of irritable bowel syndrome patients compared to healthy controls (Ponnusamy *et al.*, 2011), but otherwise the role of *Weissella* and *Leuconostoc* in health and disease is not well known. Both of them have been considered as opportunistic pathogens associated with neonatal bacteremia (Svec *et al.*, 2007), but our results suggest that these species are common inhabitants in the infant gut regardless of geographic location.

A species identified as *L. ruminis* was only detected in Malawian infants (**II**). This bacterium has been described as one of the predominant *Lactobacillus* species in adults (Reuter, 2001), but only a few studies reported it as a part of the microbiota in infants. Another study has reported *L. ruminis* in Malawian infants as well (Yatsunenko *et al.*, 2012), suggesting that this species could be a characteristic of infants from this area. This could mean that the Malawians develop a more adult-like *Lactobacillus* microbiota earlier than the Finnish infants. However, it is also possible that this species is of animal origin in the Malawian infants. The Malawians live in close contact with animals and *L. ruminis* has been identified as an indigenous species in several animals including pigs (Yin and Zheng, 2005), cows (Krause *et al.*, 2003) and chickens (Kovalenko *et al.*, 1989).

L. casei group, mainly the species *L. rhamnosus*, was a characteristic of Finnish infants and encountered more often in the breastfed compared to formula-fed infants (**II**). Similarly, *L. rhamnosus* has been reported as a predominant *Lactobacillus* species in Sweden (Ahrne *et al.*, 2005) and Greece (Mitsou *et al.*, 2008). Interestingly, infants colonized by *L. casei* group were less often colonized by *Leuconostoc* spp. and *Weissella* spp. in our study (**II**). It is possible that these species share adhesion sites and may also compete for the same nutrients. Nevertheless, species of the *L. casei* group have been linked to a decreased risk of developing allergies (Johansson *et al.*, 2011) and atopic dermatitis (Penders *et al.*, 2010) and therefore this group of species could be considered as a healthy parameter in infants from Western countries.

In study **II**, the DNA extraction process did not include mechanical disruption of the cells by repeated bead-beating that was used in the other studies in the present series. This may have led to poor recovery of DNA from

some bacterial species (Nylund *et al.*, 2010; Walker *et al.*, 2015), but all samples were treated similarly and thus it should not affect the comparison between Finland and Malawi. However, it must be noted that the varying storage temperatures for the Malawian samples may have had an impact on the sample quality.

We believe that the environment, including diet and hygiene, may be among the factors associated with the differences in bacterial profiles. In Malawi, local foods naturally rich in complex carbohydrates are introduced soon after birth whereas in Finland infant formula and complementary foods are introduced when breast-feeding is not sufficient. Plant polysaccharides have been reported as an important gut microbiota modulating factor in another African country, Burkina Faso (De Filippo *et al.*, 2010). In addition, hygiene and the presence of animals in the living areas of the infants might also explain the observed differences in the microbiota between Malawi and Finland (**II**). In Malawi, hygiene standards are generally lower compared to Finland, and people live in close contact with animals. However, even indirect contact with animals might be enough for gut microbiota modulation. It has been reported that even the dust from a house which has pets, can increase the number of lactobacilli and protect against airway allergens in mice (Fujimura *et al.*, 2013).

Different strains of lactobacilli are often used as probiotics in clinical intervention studies, including cases of atopic eczema and the results have been promising for some probiotic strains (Elazab *et al.*, 2013; Isolauri *et al.*, 2000; Kalliomaki *et al.*, 2003). Also, studies on probiotic use in one Norwegian birth cohort suggest that consumption of specific *Lactobacillus* probiotics during the perinatal period may promote a reduction in the risk of atopic eczema in infants (Bertelsen *et al.*, 2014). As indicated in this study (**II**), *Lactobacillus* microbiota composition differs between countries during infancy and thus we could ask the question: could such differences have an impact on the efficacy of a probiotic? *L. casei* group species have been widely used for the same purposes in several countries around the globe, regardless of whether the species are indigenous to the population or not. Actually, *L. rhamnosus* Yoba, a derivative strain of LGG, is used as a "generic probiotic" in Southern Africa (Kort and Sybesma, 2012; Mpofu *et al.*, 2014). The concept of generic probiotics refers to probiotic bacteria introduced under a novel brand name after intellectual property rights have expired, or when there is no patent coverage, or the use is restricted to certain countries, methods of use or clinical applications (Kort and Sybesma, 2012). The goal of this is to create access to probiotics for people in the developing world.

As mentioned earlier, we also studied the impact of lipid-based nutrient supplements with different sources of protein on the gut microbiota composition of the Malawian infants over time (**III**). These supplements have

been used in several clinical trials to improve growth in children from developing countries (Ashorn *et al.*, 2015a, 2015b; Ciliberto *et al.*, 2005; Maleta *et al.*, 2015; Phuka *et al.*, 2008; Sandige *et al.*, 2004), but their impact on the gut microbiota has been unknown. Gut microbiota is closely linked to malnutrition (Gordon *et al.*, 2012; Smith *et al.*, 2013) and recent reports suggest that an immature microbiota in early infancy may reduce or even prevent successful nutritional interventions that aim to promote growth (Subramanian *et al.*, 2014).

Similar findings on Malawian infants from the same study volunteers as in our study (**III**) suggest that these supplements do not have an impact on the gut microbiota composition (Cheung *et al.*, 2016). These findings were based on next-generation sequencing, and to complement them, we performed a targeted qPCR approach to monitor possible changes at species level that might have been lost in the high-throughput analysis. We targeted bifidobacteria due to their health-promoting properties and possible role in malnutrition (Collado *et al.*, 2010; Dinh *et al.*, 2016; Kalliomaki *et al.*, 2001; Rautava *et al.*, 2012b; Subramanian *et al.*, 2014). For instance, a decline in the number of bifidobacteria, especially *B. longum*, has been noted to be associated with malnutrition in infants and young children from Bangladesh and South India (Dinh *et al.*, 2016; Million *et al.*; Subramanian *et al.*, 2014).

We included *S. aureus* as an additional target due to its clinical relevancy and possible impact on body weight (Kalliomaki *et al.*, 2008; Song *et al.*, 2014; von Mutius and Radon, 2008). Previous studies have also demonstrated that *S. aureus* is a characteristic of infants from developed world while remaining rare in infants from developing countries such as Malawi (Grześkowiak *et al.*, 2012a, 2012b). Interestingly, it has been shown that *S. aureus* may trigger the occurrence of sCD14 in plasma and therefore it has been hypothesized that perinatal colonization by *S. aureus* may lower the risk of allergic manifestations (Lundell *et al.*, 2007; Salminen *et al.*, 2016).

In this series (**III**), the lipid-based nutrient supplements or the corn-soy blend did not have an impact on the measured bacterial populations. However, some dietary components including micronutrients, such as iron, have been shown to have a negative impact on the developing gut microbiota by decreasing the numbers of bifidobacteria and promoting a more pathogenic microbial profile with increased levels of enterobacteria in children from both low- and high-income countries (Zimmermann *et al.*, 2010; Krebs *et al.*, 2013). Nevertheless, when looking at the study population (**III**) as a whole, we saw shifts in bacterial colonization rates and counts in the Malawian infants over time. In general, the *Bifidobacterium* microbiota seemed to change towards a more of an adult type profile with lower prevalence of *B. infantis*, *B. lactis*, *B. longum*, *B. bifidum* and *B. breve*, which have been reported as typical

bifidobacteria in infants, and higher levels of *B. catenulatum*, a bifidobacterium more often encountered in adults (Junick and Blaut, 2012; Matsuki *et al.*, 1999, 2004).

Deviations in the microbiota composition are critical to early immune programming and influence the health of the infant and risk of subsequent disorders. Therefore, dietary interventions aiming to promote growth should also promote normal gut microbiota development to improve the impact of nutrition supplementation in young infants at risk of malnutrition.

6.3 Microbial colonization in early life (IV and V)

In addition to a variety of environmental factors, a number of host factors may influence the gut microbiota development. In the present study (**IV**), we assessed the impact of maternal microbiota on infant gut microbiota development from a nontraditional point of view by studying the possible microbial contact *in utero* and during the first days of life by comparing microbiota in amniotic fluid, placenta, meconium, colostrum, infant feces and maternal feces. In addition, we assessed how colostrum human milk oligosaccharide (HMO) content interacts with the colostrum microbiota itself (**V**).

Contrary to earlier beliefs, recent reports suggest that fetus might be exposed to microbes and/or microbial DNA already in the placenta without adverse pregnancy outcomes (Aagaard *et al.*, 2014; Rautava *et al.*, 2012b). In our study (**IV)**, we also found a unique microbial community in the placenta and amniotic fluid, which may provide the initial inoculum for gut colonization. Our findings were consistent with previous reports on placenta microbiota indicating a predominance of Proteobacteria, especially species belonging to the Enterobacteriaceae family (Aagaard *et al.*, 2014; Antony *et al.*, 2015; Zheng *et al.*, 2015). These findings, including ours (**IV**), suggest that the microbiota in the placenta is consistent between healthy individuals, which could mean that there are selective mechanisms to introduce specific maternal microbes or microbial parts to the fetus.

The origin of the intrauterine microbiota remains to be fully uncovered. Similarities between human oral and placental microbial communities have been reported (Aagaard *et al.*, 2014). Nevertheless, the study relied on oral microbiota data from a previous report based on non-pregnant individuals and the mechanism of possible bacterial transportation between these maternal compartments remains unknown. However, it has been demonstrated in laboratory animals that intestinal bacterial translocation is increased during pregnancy (Koren *et al.*, 2012; Perez *et al.*, 2007). Moreover, specific labeled bacteria that were fed to pregnant mice have been isolated from the placenta as

well as the meconium of the neonate after sterile cesarean section (Jimenez *et al.*, 2008). A more recent study by Agüero and coworkers (2016) utilizing germ-free experimental animals demonstrated that nonviable bacterial fragments are transferred to placenta where they stimulate the immune system development of the fetus. However, the authors were not able to culture the described microbes from the placenta or the neonate.

In the present series (**IV**), we were able to culture bacteria from some of the placenta and amniotic fluid samples. However, most of the other studies investigating placenta or amniotic fluid microbiota in humans have only used culture-independent methods. Based on them, it cannot be concluded whether there are viable, intact bacteria in placenta or amniotic fluid. The bacterial DNA that is detected using PCR-based techniques could be present in the intracellular space or, for instance, within immune cells. It is also important to keep in mind that sequencing of low biomass samples is prone to PCR bias and thus the results may present a distorted image of the real bacterial population in such samples. Whether or not viable microbes exist in the amniotic fluid or placenta, it has been noted that already the bacterial DNA or other lifeless microbial parts may influence the immune system maturation of the fetus *in utero* (Agüero *et al.*, 2016; Rautava *et al.*, 2012a).

In our study (**IV**), placenta samples were collected during elective cesarean section to avoid contamination by microbes from the birth canal or maternal feces. Moreover, mothers who presented with rupture of membranes, labour or signs of infection were excluded from the study. Contamination of the DNA extraction and PCR reagents has been recognized as a major concern when studying samples with low bacterial population (Laurence *et al.*, 2014; Salter *et al.*, 2014; Spangler *et al.*, 2009). Therefore, we used controls containing no template within the DNA extraction and PCR to rule out the possibility of contamination, and we could not obtain amplicons of 16S rRNA gene sequences for further analysis.

In contrast to our findings, a recently published study demonstrated that placental samples could not be reliably distinguished from background contamination controls using qPCR and 16S rRNA gene sequencing combined with two different methods for DNA extraction (Lauder *et al.*, 2016). Nevertheless, the authors found some extremely low abundant OTUs from the maternal side (basal plate) of the placental samples that could represent sporadic, and potentially rather rare placental microbiota (Lauder *et al.*, 2016). If true, this would be in accordance with earlier reports on intact intracellular microbes in the basal plate (Stout *et al.*, 2013). This study demonstrates that possibly new optimized DNA extraction methods and pipelines with lower detection limits are needed for reliable analysis of low biomass samples. Such work has been started already with laboratory animals, in which low copy

numbers of bacterial DNA could be reproducibly detected with no contaminations in brain, muscle, adipose tissue, liver and heart (Lluch *et al.*, 2015).

Several reports suggest that microbes are also present in meconium (Dominguez-Bello *et al.*, 2010; Gosalbes *et al.*, 2013; Mshvildadze *et al.*, 2010), but their source has been unknown. It has been speculated that the microbes or microbial DNA may be derived from swallowing amniotic fluid (Ardissone *et al.*, 2014). In the present series (**IV**), the meconium microbiota shared some similarities with amniotic fluid and placenta, but also that in the colostrum. The meconium samples (**IV**) were collected during the first two days of life, during which all infants had been breastfed from the first hours of life. Therefore it is possible that microbes of colostrum may have already had an impact on the meconium microbiota in some individuals. It is also possible that these microbes originated from other postnatal environmental sources because the meconium samples were collected non-invasively from diapers. Though, it is also possible that meconium and colostrum microbiota share a common maternal source if microbes or microbial parts are actively transferred from the maternal gut to mammary gland and placenta, as suggested by some research groups (Agüero *et al.*, 2016; Jimenez *et al.*, 2008; Perez *et al.*, 2007).

Although the transfer of microbes or microbial compounds to placenta has been demonstrated in experimental animals (Agüero *et al.*, 2016; Jimenez *et al.*, 2005, 2008), production of germ-free animals is still possible by hysterectomy (Arvidsson *et al.*, 2011) and embryo transfer (Inzunza *et al.*, 2005). This suggests that maybe permanent colonization does not happen *in utero* or it may only happen during the final stages of the pregnancy. In the future, this phenomenom should be studied more carefully when deriving germ-free animals to demonstrate whether there are microbes or microbial parts in the placenta or the fetus. To our knowledge, there have been no studies on the bacterial load of germ-free derived animals at their fetal stage utilizing modern technologies. It has been suggested that microbes or microbial parts may also be washed out or sterilized during the derivation procedure of germ-free animals since the concentration of microbes in placenta and amniotic fluid appears to be very low (Reid *et al.*, 2015). It also remains unknown how active the microbes in these sites are in the first place.

The role of placenta or meconium microbiota in health and disease is not well understood. However, some recent reports suggest that both the placenta and meconium microbiota may reflect the health status of the mother. Altered placenta and meconium microbiota has been associated with preterm birth (Ardissone *et al.*, 2014; Prince *et al.*, 2016). The placenta microbiota was reported to be altered with and without chrioamnionitis compared to healthy individuals (Prince *et al.*, 2016). Also, the placenta microbiota has been

associated with low birth weight in full term neonates (Amarasekara *et al.*, 2015). In addition to reflecting the health status of the mother (Hu *et al.*, 2013) to the meconium microbiota has also been associated with development of diseases in the offspring (Gosalbes *et al.*, 2013).

The microbial colonization process continues after birth via breastfeeding. After birth, breast milk serves both as the most important source of microbes and of other microbiota modulating factors, such as HMOs (Jost *et al.*, 2015). In the studies **IV** and **V** presented here, all mothers delivered by cesarean section, which may affect the colostrum microbiota composition possibly as a result of hormonal changes produced during labor (Cabrera-Rubio *et al.*, 2012, 2016). The infant fecal microbiota at 3-4 days was notably similar to the colostrum microbiota. However, based on the current sample material (**IV**), we can only hypothesize that the breast milk microbes contribute to the gut colonization during early infancy. To make robust conclusions on the role of breast milk microbes in gut microbiota development, formula fed infants and breastfed infants should be compared in future studies.

The impact of HMOs on infant microbiota has been reported (De Leoz *et al.*, 2015; Wang *et al.*, 2015), but it has been unknown whether the HMOs influence the human milk microbiota composition itself. Our results suggest that the HMO content of human milk and the human milk microbiota are interrelated (**V**).

Specific species of bifidobacteria are known for their ability to consume HMOs (Bode, 2012), which may explain some of our findings in study **V**. For instance, we saw an association between *B. longum* group bacteria and LNT, one of the core HMOs of both secretors and non-secretors, and subspecies of this bacterial group are known for their ability to metabolize this specific HMO by utilizing lacto-N-biosidase (LoCascio *et al.*, 2010; Sakurama *et al.*, 2013). In addition to bifidobacteria, HMO content was associated with counts of *S. aureus* and *Akkermansia muciniphila*. HMOs have been reported to promote the growth of *S. aureus* (Hunt *et al.*, 2012), but for *Akkermansia muciniphila* HMO metabolism has not been documented. However, the genome of *Akkermansia* includes a variety of genes, such as sialidases and fucosidases, which suggest that this bacterium could also metabolize HMOs (Derrien *et al.*, 2004). These results suggest that the impact of HMOs is not limited to the infant, but also to the microbial composition of human milk.

In conclusion, our results suggest that the fetus may be exposed to microbes already *in utero*, and the gut colonization continues in a step-wise manner following birth via microbes in breast milk, which is directed by host factors such as HMO composition of the milk. The final two parts of the study series (**IV** and **V**) are based on a relatively small cohort and thus future research should include multi-centre studies to confirm the presence of microbes in both

the placenta and the amniotic fluid samples and to elucidate their potential role for infant health. Bacterial transfer from mother to infant may provide a new target for clinical interventions aiming at improving the pregnant mother's nutrition and microbiota for further impact on reducing the inflammatory non-communicable disease risk by modulating the early host-microbe interactions and later programming of infant health. This will require new technologies and methods to assess the microbiota and its deviations. This would form the basis for counteracting potential deviations by nutrition, probiotics and prebiotics.

7 SUMMARY

Exposure to microbes and microbial structures in early life is crucial for the normal development of the gut microbiota and immune system maturation. Human gut microbiota development is a step-wise process, which possibly starts already during the fetal stages of life, based on our findings. Following birth, the microbiota development continues via the introduction of microbes and microbiota modulating factors in human milk, whose composition varies between individuals. Moreover, the environment plays a role in introducing microbes to the infant and the microbiota development is affected by environmental factors such as hygiene, sanitation and geographical location as demonstrated by the results in the present series.

It may be possible to modulate the gut microbiota composition at different stages of development to reduce the risk of non-communicable diseases, including allergic, autoimmune and anti-inflammatory diseases, such as obesity. In addition, gut microbiota modulation during infancy may provide a new complementary approach to promote growth of infants at risk of malnutrition. Nevertheless, further knowledge on the mechanisms of host-microbe interactions as well as functional capability of the human microbiota is needed, since the microbiota composition itself does not necessarily reflect the function profile that it has. This can be achieved by utilizing metagenomic, metatranscriptomic and metaproteomic approaches among others, but the cost of these methods remains high today and the data analysis methods are still under development, In addition, the role of archaeae, viruses and parasites in human health remains to be elucidated.

It is also plausible that microbiota research expands from the gastrointestinal tract to other niches as well. Microbes or microbial parts have been discovered in various niches previously thought to be sterile, at least in healthy individuals, including blood and several tissues, but their role in health and disease remains to be elucidated. For this purpose, novel optimized methods for analysing low biomass samples with minimal risk of contamination should be developed.

ACKNOWLEDGEMENTS

The work for this thesis was carried out at Functional Foods Forum and the Food chemistry and Food development unit of the Department of Biochemistry, University of Turku. The Doctoral Programme in Molecular Life Sciences of the University of Turku Graduate School is acknowledged for its financial support. I am also grateful for all the travel grants provided by the Turku University Foundation, which allowed me to participate in numerous conferences around the world

I wish to express my gratitude to my supervisors, Professor Seppo Salminen and Dr. Akihito Endo for excellent guidance and support through the years of my master's and doctoral studies.

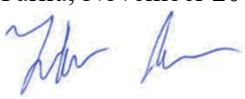
I also want to acknowledge the official reviewers, Dr. Karen Scott and Dr. Konrad Domig, for their constructive criticism and valuable comments on the manuscript.

My steering group members Baoru Yang and Maria Carmen Collado are thanked for their advice and support. I also wish to express gratitude to all my co-authors involved in the articles of this thesis for their contributions. I warmly thank Professor Erika Isolauri for providing samples for the studies as well as providing clinical expertise and insights in the clinical field together with Samuli Rautava.

In addition, all colleagues at FFF and Food Chemistry and Food Development unit are acknowledged for creating a friendly and supportive working environment. I wish to acknowledge Satu Tölkkö for her excellent technical assistance and making sure that everything is in order in the laboratory.

Finally, I wish to express my deepest gratitude to my dear family and close friends supporting and standing by me. Special warm thanks go to Sofia who keeps it in my mind that life is not that bad after all. I am truly grateful for your existence!

Turku, November 2016



Juhani Aakko

REFERENCES

Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. (2014). The placenta harbors a unique microbiome. *Sci Transl Med* **6**: 237ra65.

Abrahamsson TR, Sinkiewicz G, Jakobsson T, Fredrikson M, Björkstén B. (2009). Probiotic Lactobacilli in Breast Milk and Infant Stool in Relation to Oral Intake During the First Year of Life: *J Pediatr Gastroenterol Nutr* **49**: 349–354.

Adlerberth I, Lindberg E, Aberg N, Hesselmar B, Saalman R, Strannegård I-L, *et al.* (2006). Reduced enterobacterial and increased staphylococcal colonization of the infantile bowel: an effect of hygienic lifestyle? *Pediatr Res* **59**: 96–101.

Agüero MG de, Ganal-Vonarburg SC, Fuhrer T, Rupp S, Uchimura Y, Li H, *et al.* (2016). The maternal microbiota drives early postnatal innate immune development. *Science* **351**: 1296–1302.

Ahlroos T, Tynkkynen S. (2009). Quantitative strain-specific detection of Lactobacillus rhamnosus GG in human faecal samples by real-time PCR. *J Appl Microbiol* **106**: 506–514.

Ahrne S, Lonnermark E, Wold AE, Aberg N, Hesselmar B, Saalman R, *et al.* (2005). Lactobacilli in the intestinal microbiota of Swedish infants. *Microbes Infect Inst Pasteur* **7**: 1256–1262.

Alderete TL, Autran C, Brekke BE, Knight R, Bode L, Goran MI, *et al.* (2015). Associations between human milk oligosaccharides and infant body composition in the first 6 mo of life. *Am J Clin Nutr* **102**: 1381–1388.

Amarasekara R, Jayasekara RW, Senanayake H, Dissanayake VHW. (2015). Microbiome of the placenta in pre-eclampsia supports the role of bacteria in the multifactorial cause of pre-eclampsia. *J Obstet Gynaecol Res* **41**: 662–669.

Andersson AF, Lindberg M, Jakobsson H, Bäckhed F, Nyrén P, Engstrand L. (2008). Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PloS One* **3**: e2836.

Andrews WW, Goldenberg RL, Hauth JC, Cliver SP, Conner M, Goepfert AR. (2005). Endometrial microbial colonization and plasma cell endometritis after spontaneous or indicated preterm versus term delivery. *Am J Obstet Gynecol* **193**: 739–745.

Antony KM, Ma J, Mitchell KB, Racusin DA, Versalovic J, Aagaard K. (2015). The preterm placental microbiome varies in association with excess maternal gestational weight gain. *Am J Obstet Gynecol* **212**: 653.e1-653.16.

Ardissone AN, Cruz DM de la, Davis-Richardson AG, Rechcigl KT, Li N, Drew JC, *et al.* (2014). Meconium Microbiome Analysis Identifies Bacteria Correlated with Premature Birth. *PLOS ONE* **9**: e90784.

Arndt D, Xia J, Liu Y, Zhou Y, Guo AC, Cruz JA, *et al.* (2012). METAGENassist: a comprehensive web server for comparative metagenomics. *Nucleic Acids Res* **40**: W88-95.

Arslanoglu S, Moro GE, Boehm G. (2007). Early Supplementation of Prebiotic Oligosaccharides Protects Formula-Fed Infants against Infections during the First 6 Months of Life. *J Nutr* **137**: 2420–2424.

Arvidsson C, Hallén A, Bäckhed F. (2011). Generating and Analyzing Germ-Free Mice. In: *Current Protocols in Mouse Biology*. John Wiley & Sons, Inc. http://onlinelibrary.wiley.com/doi/10.1002/9780470942390.mo120064/abstract (Accessed November 8, 2016).

Asakuma S, Hatakeyama E, Urashima T, Yoshida E, Katayama T, Yamamoto K, *et al.* (2011). Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria. *J Biol Chem* **286**: 34583–34592.

Ashorn P, Alho L, Ashorn U, Cheung YB, Dewey KG, Gondwe A, *et al.* (2015a). Supplementation of Maternal Diets during Pregnancy and for 6 Months Postpartum and Infant Diets Thereafter with Small-Quantity Lipid-Based Nutrient Supplements Does Not Promote Child Growth by 18 Months of Age in Rural Malawi: A Randomized Controlled Trial. *J Nutr* **145**: 1345–1353.

Ashorn P, Alho L, Ashorn U, Cheung YB, Dewey KG, Harjunmaa U, *et al.* (2015b). The impact of lipid-based nutrient supplement provision to pregnant women on newborn size in rural Malawi: a randomized controlled trial. *Am J Clin Nutr* **101**: 387–397.

Azzalini A, Bowman A. (2014). sm: Smoothing methods for nonparametric regression and density estimation.

Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, *et al.* (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**: 690–703.

Barbara G, Stanghellini V, Brandi G, Cremon C, Nardo GD, De Giorgio R, *et al.* (2005). Interactions Between Commensal Bacteria and Gut Sensorimotor Function in Health and Disease. *Am J Gastroenterol* **100**: 2560–2568.

Bartlett JS, Stirling D. (2003). A Short History of the Polymerase Chain Reaction. In: Bartlett JS, Stirling D (eds) Methods in Molecular Biology™. *PCR Protocols*. Humana Press, pp 3–6.

Bertelsen RJ, Brantsaeter AL, Magnus MC, Haugen M, Myhre R, Jacobsson B, *et al.* (2014). Probiotic milk consumption in pregnancy and infancy and subsequent childhood allergic diseases. *J Allergy Clin Immunol* **133**: 165-71–8.

Bezirtzoglou E, Tsiotsias A, Welling GW. (2011). Microbiota profile in feces of breast- and formula-fed newborns by using fluorescence in situ hybridization (FISH). *Anaerobe* **17**: 478–482.

Bhattacharjee S, Lukiw WJ. (2013). Alzheimer's disease and the microbiome. *Front Cell Neurosci* **7**. e-pub ahead of print, doi: 10.3389/fncel.2013.00153.

Bindels LB, Delzenne NM, Cani PD, Walter J. (2015). Towards a more comprehensive concept for prebiotics. *Nat Rev Gastroenterol Hepatol* **12**: 303–310.

Blaser MJ, Falkow S. (2009). What are the consequences of the disappearing human microbiota? *Nat Rev Microbiol* **7**: 887–894.

Bode L. (2012). Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology* **22**: 1147–1162.

Brandt K, Alatossava T. (2003). Specific identification of certain probiotic Lactobacillus rhamnosus strains with PCR primers based on phage-related sequences. *Int J Food Microbiol* **84**: 189–196.

Bravo JA, Forsythe P, Chew MV, Escaravage E, Savignac HM, Dinan TG, *et al.* (2011). Ingestion of Lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc Natl Acad Sci* **108**: 16050–16055.

Buffie CG, Pamer EG. (2013). Microbiota-mediated colonization resistance against intestinal pathogens. *Nat Rev Immunol* **13**: 790–801.

Cabrera-Rubio R, Collado MC, Laitinen K, Salminen S, Isolauri E, Mira A. (2012). The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. *Am J Clin Nutr* **96**: 544–551.

Cabrera-Rubio R, Mira-Pascual L, Mira A, Collado MC. (2016). Impact of mode of delivery on the milk microbiota composition of healthy women. *J Dev Orig Health Dis* **7**: 54–60.

Carey CM, Kirk JL, Ojha S, Kostrzynska M. (2007). Current and future uses of real-time polymerase chain reaction and microarrays in the study of intestinal microbiota, and probiotic use and effectiveness. *Can J Microbiol* **53**: 537–550.

Carroll L, Osman M, Davies DP, McNeish AS. (1979). Bacteriological criteria for feeding raw breast-milk to babies on neonatal units. *Lancet Lond Engl* **2**: 732–733.

Chaturvedi P, Warren CD, Altaye M, Morrow AL, Ruiz-Palacios G, Pickering LK, *et al.* (2001). Fucosylated human milk oligosaccharides vary between individuals and over the course of lactation. *Glycobiology* **11**: 365–372.

Cheung YB, Xu Y, Mangani C, Fan Y-M, Dewey KG, Salminen SJ, *et al.* (2016). Gut microbiota in Malawian infants in a nutritional supplementation trial. *Trop Med Int Health TM IH* **21**: 283–290.

Ciliberto MA, Sandige H, Ndekha MJ, Ashorn P, Briend A, Ciliberto HM, *et al.* (2005). Comparison of home-based therapy with ready-to-use therapeutic food with standard therapy in the treatment of malnourished Malawian children: a controlled, clinical effectiveness trial. *Am J Clin Nutr* **81**: 864–870.

Collado MC, Isolauri E, Laitinen K, Salminen S. (2010). Effect of mother's weight on infant's microbiota acquisition, composition, and activity during early infancy: a prospective follow-up study initiated in early pregnancy. *Am J Clin Nutr* **92**: 1023–1030.

Collado MC, Rautava S, Isolauri E, Salminen S. (2015). Gut microbiota: a source of novel tools to reduce the risk of human disease? *Pediatr Res* **77**: 182–188.

Collins SM, Bercik P. (2009). The Relationship Between Intestinal Microbiota and the Central Nervous System in Normal Gastrointestinal Function and Disease. *Gastroenterology* **136**: 2003–2014.

Cooke G, Behan J, Clarke N, Gorman W, Costello M. (2005). Comparing the gut flora of Irish breastfed and formula-fed neonates aged between birth and 6 weeks old. *Microb Ecol Health Dis* **17**: 163–168.

Coppa GV, Pierani P, Zampini L, Carloni I, Carlucci A, Gabrielli O. (1999). Oligosaccharides in human milk during different phases of lactation. *Acta Paediatr Oslo Nor 1992Supplement* **88**: 89–94.

Cummings JH. (1984). Microbial digestion of complex carbohydrates in man. *Proc Nutr Soc* **43**: 35–44.

Dal Bello F, Walter J, Hammes WP, Hertel C. (2003). Increased complexity of the species composition of lactic acid bacteria in human feces revealed by alternative incubation condition. *Microb Ecol* **45**: 455–463.

Darmasseelane K, Hyde MJ, Santhakumaran S, Gale C, Modi N. (2014). Mode of delivery and offspring body mass index, overweight and obesity in adult life: a systematic review and meta-analysis. *PloS One* **9**: e87896.

David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, *et al.* (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**: 559–563.

De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, *et al.* (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A* **107**: 14691–14696.

De Leoz ML, Kalanetra KM, Bokulich NA, Strum JS, Underwood MA, German JB, *et al.* (2015). Human milk glycomics and gut microbial genomics in infant feces show a correlation between human milk oligosaccharides and gut microbiota: a proof-of-concept study. *J Proteome Res* **14**: 491–502.

Derrien M, Vaughan EE, Plugge CM, de Vos WM. (2004). Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int J Syst Evol Microbiol* **54**: 1469–1476.

DiGiulio DB, Romero R, Amogan HP, Kusanovic JP, Bik EM, Gotsch F, *et al.* (2008). Microbial prevalence, diversity and abundance in amniotic fluid during preterm labor: a molecular and culture-based investigation. *PloS One* **3**: e3056.

Dinh DM, Ramadass B, Kattula D, Sarkar R, Braunstein P, Tai A, *et al.* (2016). Longitudinal Analysis of the Intestinal Microbiota in Persistently Stunted Young Children in South India. *PloS One* **11**: e0155405.

Doege K, Grajecki D, Zyriax B-C, Detinkina E, Zu Eulenburg C, Buhling KJ. (2012). Impact of maternal supplementation with probiotics during pregnancy on atopic eczema in childhood--a meta-analysis. *Br J Nutr* **107**: 1–6.

Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, *et al.* (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A* **107**: 11971–11975.

Donaldson GP, Lee SM, Mazmanian SK. (2016). Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol* **14**: 20–32.

Douwes J, Cheng S, Travier N, Cohet C, Niesink A, McKenzie J, *et al.* (2008). Farm exposure in utero may protect against asthma, hay fever and eczema. *Eur Respir J* **32**: 603–611.

Dutilh BE, Cassman N, McNair K, Sanchez SE, Silva GGZ, Boling L, *et al.* (2014). A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat Commun* **5**: 4498.

Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, *et al.* (2005). Diversity of the human intestinal microbial flora. *Science* **308**: 1635–1638.

Eidelman AI, Szilagyi G. (1979). Patterns of bacterial colonization of human milk. *Obstet Gynecol* **53**: 550–552.

Eisenstein EM, Williams CB. (2009). The Treg/Th17 Cell Balance: A New Paradigm for Autoimmunity. *Pediatr Res* **65**: 26R–31R.

Elazab N, Mendy A, Gasana J, Vieira ER, Quizon A, Forno E. (2013). Probiotic administration in early life, atopy, and asthma: a meta-analysis of clinical trials. *Pediatrics* **132**: e666-76.

Endo A, Mizuno H, Okada S. (2008). Monitoring the bacterial community during fermentation of sunki, an unsalted, fermented vegetable traditional to the Kiso area of Japan. *Lett Appl Microbiol* **47**: 221–226.

Endo A, Okada S. (2005). Lactobacillus satsumensis sp. nov., isolated from mashes of shochu, a traditional Japanese distilled spirit made from fermented rice and other starchy materials. *Int J Syst Evol Microbiol* **55**: 83–85.

Endo A, Okada S, Morita H. (2007). Molecular profiling of Lactobacillus, Streptococcus, and Bifidobacterium species in feces of active racehorses. *J Gen Appl Microbiol* **53**: 191–200.

Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H, *et al.* (2010). Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J* **4**: 642–647.

Fallani M, Amarri S, Uusijarvi A, Adam R, Khanna S, Aguilera M, *et al.* (2011). Determinants of the human infant intestinal microbiota after the introduction of first complementary foods in infant samples from five European centres. *Microbiology* **157**: 1385–1392.

Fallani M, Young D, Scott J, Norin E, Amarri S, Adam R, *et al.* (2010). Intestinal microbiota of 6-week-old infants across Europe: geographic influence beyond delivery mode, breast-feeding, and antibiotics. *J Pediatr Gastroenterol Nutr* **51**: 77–84.

FAO/WHO. (2002). Report of a Joint FAO/WHO Working Group on Drafting Guidelines for the Evaluation of Probiotics in Food. ftp://ftp.fao.org/es/esn/food/wgreport2.pdf.

Farrelly V, Rainey FA, Stackebrandt E. (1995). Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Appl Environ Microbiol* **61**: 2798–2801.

Flint HJ, Duncan SH, Scott KP, Louis P. (2015). Links between diet, gut microbiota composition and gut metabolism. *Proc Nutr Soc* **74**: 13–22.

Fraher MH, O'Toole PW, Quigley EMM. (2012). Techniques used to characterize the gut microbiota: a guide for the clinician. *Nat Rev Gastroenterol Hepatol* **9**: 312–322.

Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW, *et al.* (2008). Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci U S A* **105**: 3805–3810.

Frost G, Sleeth ML, Sahuri-Arisoylu M, Lizarbe B, Cerdan S, Brody L, *et al.* (2014). The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism. *Nat Commun* **5**: 3611.

Fujimura KE, Demoor T, Rauch M, Faruqi AA, Jang S, C Johnson C, *et al.* (2013). House dust exposure mediates gut microbiome Lactobacillus enrichment and airway immune defense against allergens and virus infection. *Proc Natl Acad Sci U S A*. e-pub ahead of print, doi: 10.1073/pnas.1310750111.

Furusawa Y, Obata Y, Fukuda S, Endo TA, Nakato G, Takahashi D, *et al.* (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**: 446–450.

Gevers D, Huys G, Swings J. (2001). Applicability of rep-PCR fingerprinting for identification of Lactobacillus species. *FEMS Microbiol Lett* **205**: 31–36.

Goldenberg RL, Culhane JF, Iams JD, Romero R. (2008). Epidemiology and causes of preterm birth. *Lancet Lond Engl* **371**: 75–84.

Gopal PK, Gill HS. (2000). Oligosaccharides and glycoconjugates in bovine milk and colostrum. *Br J Nutr* **84 Suppl 1**: S69-74.

Gordon JI, Dewey KG, Mills DA, Medzhitov RM. (2012). The human gut microbiota and undernutrition. *Sci Transl Med* **4**: 137ps12.

Gosalbes MJ, Llop S, Valles Y, Moya A, Ballester F, Francino MP. (2013). Meconium microbiota types dominated by lactic acid or enteric bacteria are differentially associated with maternal eczema and respiratory problems in infants. *Clin Exp Allergy J Br Soc Allergy Clin Immunol* **43**: 198–211.

Grönlund M-M, Gueimonde M, Laitinen K, Kociubinski G, Grönroos T, Salminen S, *et al.* (2007). Maternal breast-milk and intestinal bifidobacteria guide the compositional development of the Bifidobacterium microbiota in infants at risk of allergic disease. *Clin Exp Allergy J Br Soc Allergy Clin Immunol* **37**: 1764–1772.

Grönlund MM, Lehtonen OP, Eerola E, Kero P. (1999). Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery. *J Pediatr Gastroenterol Nutr* **28**: 19–25.

Grześkowiak Ł, Collado MC, Mangani C, Maleta K, Laitinen K, Ashorn P, *et al.* (2012a). Distinct gut microbiota in southeastern African and northern European infants. *J Pediatr Gastroenterol Nutr* **54**: 812–816.

Grześkowiak Ł, Grönlund M-M, Beckmann C, Salminen S, von Berg A, Isolauri E. (2012b). The impact of perinatal probiotic intervention on gut microbiota: double-blind placebo-controlled trials in Finland and Germany. *Anaerobe* **18**: 7–13.

Gueimonde M, Kalliomaki M, Isolauri E, Salminen S. (2006a). Probiotic intervention in neonates--will permanent colonization ensue? *J Pediatr Gastroenterol Nutr* **42**: 604–606.

Gueimonde M, Laitinen K, Salminen S, Isolauri E. (2007). Breast milk: a source of bifidobacteria for infant gut development and maturation? *Neonatology* **92**: 64–66.

Gueimonde M, Sakata S, Kalliomäki M, Isolauri E, Benno Y, Salminen S. (2006b). Effect of maternal consumption of lactobacillus GG on transfer and establishment of fecal bifidobacterial microbiota in neonates. *J Pediatr Gastroenterol Nutr* **42**: 166–170.

Gueimonde M, Tölkkö S, Korpimäki T, Salminen S. (2004). New Real-Time Quantitative PCR Procedure for Quantification of Bifidobacteria in Human Fecal Samples. *Appl Environ Microbiol* **70**: 4165–4169.

Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* **5**: R245-249.

Heikkilä M p., Saris P e. j. (2003). Inhibition of Staphylococcus aureus by the commensal bacteria of human milk. *J Appl Microbiol* **95**: 471–478.

Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, *et al.* (2014). Expert consensus document: The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol* **11**: 506–514.

Højsgaard S, Halekoh U, Yan J. (2014). geepack: Generalized Estimating Equation Package.

Holliday TW, Hilton CE. (2010). Body proportions of circumpolar peoples as evidenced from skeletal data: Ipiutak and Tigara (Point Hope) versus Kodiak Island Inuit. *Am J Phys Anthropol* **142**: 287–302.

Hu J, Nomura Y, Bashir A, Fernandez-Hernandez H, Itzkowitz S, Pei Z, *et al.* (2013). Diversified microbiota of meconium is affected by maternal diabetes status. *PloS One* **8**: e78257.

Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* **486**: 207–214.

Hummelen R, Fernandes AD, Macklaim JM, Dickson RJ, Changalucha J, Gloor GB, *et al.* (2010). Deep sequencing of the vaginal microbiota of women with HIV. *PloS One* **5**: e12078.

Hunt KM, Foster JA, Forney LJ, Schütte UME, Beck DL, Abdo Z, *et al.* (2011). Characterization of the Diversity and Temporal Stability of Bacterial Communities in Human Milk. *PLOS ONE* **6**: e21313.

Hunt KM, Preuss J, Nissan C, Davlin CA, Williams JE, Shafii B, *et al.* (2012). Human milk oligosaccharides promote the growth of staphylococci. *Appl Environ Microbiol* **78**: 4763–4770.

Inzunza J, Midtvedt T, Fartoo M, Norin E, Osterlund E, Persson A-K, *et al.* (2005). Germfree status of mice obtained by embryo transfer in an isolator environment. *Lab Anim* **39**: 421–427.

Isolauri E, Arvola T, Sutas Y, Moilanen E, Salminen S. (2000). Probiotics in the management of atopic eczema. *Clin Exp Allergy J Br Soc Allergy Clin Immunol* **30**: 1604–1610.

Isolauri E, Kalliomaki M, Rautava S, Salminen S, Laitinen K. (2009). Obesity - extending the hygiene hypothesis. *Nestle Nutr Workshop SeriesPaediatric Programme* **64**: 75-85–9, 251–7.

Jalanka-Tuovinen J, Salonen A, Nikkilä J, Immonen O, Kekkonen R, Lahti L, *et al.* (2011). Intestinal Microbiota in Healthy Adults: Temporal Analysis Reveals Individual and Common Core and Relation to Intestinal Symptoms. *PLOS ONE* **6**: e23035.

Jimenez E, Fernandez L, Marin ML, Martin R, Odriozola JM, Nueno-Palop C, *et al.* (2005). Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section. *Curr Microbiol* **51**: 270–274.

Jimenez E, Marin ML, Martin R, Odriozola JM, Olivares M, Xaus J, *et al.* (2008). Is meconium from healthy newborns actually sterile? *Res Microbiol* **159**: 187–193.

Johansson MA, Sjogren YM, Persson JO, Nilsson C, Sverremark-Ekstrom E. (2011). Early colonization with a group of Lactobacilli decreases the risk for allergy at five years of age despite allergic heredity. *PloS One* **6**: e23031.

Johansson MEV, Phillipson M, Petersson J, Velcich A, Holm L, Hansson GC. (2008). The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci* **105**: 15064–15069.

Jost T, Lacroix C, Braegger C, Chassard C. (2015). Impact of human milk bacteria and oligosaccharides on neonatal gut microbiota establishment and gut health. *Nutr Rev* **73**: 426–437.

Jung T-H, Park JH, Jeon W-M, Han K-S. (2015). Butyrate modulates bacterial adherence on LS174T human colorectal cells by stimulating mucin secretion and MAPK signaling pathway. *Nutr Res Pract* **9**: 343–349.

Junick J, Blaut M. (2012). Quantification of human fecal bifidobacterium species by use of quantitative real-time PCR analysis targeting the groEL gene. *Appl Environ Microbiol* **78**: 2613–2622.

Kalliomaki M, Collado MC, Salminen S, Isolauri E. (2008). Early differences in fecal microbiota composition in children may predict overweight. *Am J Clin Nutr* **87**: 534–538.

Kalliomaki M, Kirjavainen P, Eerola E, Kero P, Salminen S, Isolauri E. (2001). Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *J Allergy Clin Immunol* **107**: 129–134.

Kalliomaki M, Salminen S, Poussa T, Arvilommi H, Isolauri E. (2003). Probiotics and prevention of atopic disease: 4-year follow-up of a randomised placebo-controlled trial. *Lancet* **361**: 1869–1871.

Kawai T, Akira S. (2006). TLR signaling. *Cell Death Differ* **13**: 816–825.

Kelly RJ, Rouquier S, Giorgi D, Lennon GG, Lowe JB. (1995). Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *J Biol Chem* **270**: 4640–4649.

Kim M, Morrison M, Yu Z. (2011). Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J Microbiol Methods* **84**: 81–87.

Kobata A. (2010). Structures and application of oligosaccharides in human milk. *Proc Jpn Acad B Phys Biol Sci* **86**: 731–747.

Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, *et al.* (2011). Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci U S A* **108**: 4578–4585.

Koren O, Goodrich JK, Cullender TC, Spor A, Laitinen K, Kling Bäckhed H, *et al.* (2012). Host Remodeling of the Gut Microbiome and Metabolic Changes during Pregnancy. *Cell* **150**: 470–480.

Kort R, Sybesma W. (2012). Probiotics for every body. *Trends Biotechnol* **30**: 613–615.

Kovalenko NK, Golovach TN, Kvasnikov EI. (1989). Lactic bacteria in the digestive tract of poultry. *Mikrobiologiia* **58**: 137–143.

Krause DO, Smith WJ, Conlan LL, Gough JM, Williamson MA, McSweeney CS. (2003). Diet influences the ecology of lactic acid bacteria and Escherichia coli along the digestive tract of cattle: neural networks and 16S rDNA. *Microbiol Read Engl* **149**: 57–65.

Krebs NF, Sherlock LG, Westcott J, Culbertson D, Hambidge KM, Feazel LM, *et al.* (2013). Effects of different complementary feeding regimens on iron status and enteric microbiota in breastfed infants. *J Pediatr* **163**: 416–423.

Kubinak JL, Round JL. (2012). Toll-Like Receptors Promote Mutually Beneficial Commensal-Host Interactions. *PLoS Pathog* **8**. e-pub ahead of print, doi: 10.1371/journal.ppat.1002785.

Kunz C, Rudloff S, Baier W, Klein N, Strobel S. (2000). Oligosaccharides in human milk: structural, functional, and metabolic aspects. *Annu Rev Nutr* **20**: 699–722.

Lagier J-C, Armougom F, Million M, Hugon P, Pagnier I, Robert C, *et al.* (2012). Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* **18**: 1185–1193.

Lahti L, Salojärvi J. (2012). microbiome R package.

Lahtinen SJ, Boyle RJ, Kivivuori S, Oppedisano F, Smith KR, Robins-Browne R, *et al.* (2009). Prenatal probiotic administration can influence Bifidobacterium microbiota development in infants at high risk of allergy. *J Allergy Clin Immunol* **123**: 499–501.

Lauder AP, Roche AM, Sherrill-Mix S, Bailey A, Laughlin AL, Bittinger K, *et al.* (2016). Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. *Microbiome* **4**: 29.

Laurence M, Hatzis C, Brash DE. (2014). Common Contaminants in Next-Generation Sequencing That Hinder Discovery of Low-Abundance Microbes. *PLOS ONE* **9**: e97876.

Lavelle EC, Murphy C, O'Neill L a. J, Creagh EM. (2009). The role of TLRs, NLRs, and RLRs in mucosal innate immunity and homeostasis. *Mucosal Immunol* **3**: 17–28.

LeBlanc JG, Milani C, de Giori GS, Sesma F, van Sinderen D, Ventura M. (2013). Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Curr Opin Biotechnol* **24**: 160–168.

Lepage P, Leclerc MC, Joossens M, Mondot S, Blottière HM, Raes J, *et al.* (2013). A metagenomic insight into our gut's microbiome. *Gut* **62**: 146–158.

Lewis ZT, Totten SM, Smilowitz JT, Popovic M, Parker E, Lemay DG, *et al.* (2015). Maternal fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants. *Microbiome* **3**: 13–015–0071–eCollection 2015.

Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. (2008). Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* **6**: 776–788.

Ley RE, Peterson DA, Gordon JI. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**: 837–848.

Lin HV, Frassetto A, Kowalik EJ, Nawrocki AR, Lu MM, Kosinski JR, *et al.* (2012). Butyrate and propionate protect against diet-induced obesity and regulate gut hormones via free fatty acid receptor 3-independent mechanisms. *PloS One* **7**: e35240.

Lluch J, Servant F, Païssé S, Valle C, Valière S, Kuchly C, *et al.* (2015). The Characterization of Novel Tissue Microbiota Using an Optimized 16S Metagenomic Sequencing Pipeline. *PLoS ONE* **10**. e-pub ahead of print, doi: 10.1371/journal.pone.0142334.

LoCascio RG, Desai P, Sela DA, Weimer B, Mills DA. (2010). Broad Conservation of Milk Utilization Genes in Bifidobacterium longum subsp. infantis as Revealed by Comparative Genomic Hybridization. *Appl Environ Microbiol* **76**: 7373–7381.

LoCascio RG, Ninonuevo MR, Freeman SL, Sela DA, Grimm R, Lebrilla CB, *et al.* (2007). Glycoprofiling of Bifidobacterial Consumption of Human Milk Oligosaccharides Demonstrates Strain Specific, Preferential Consumption of Small Chain Glycans Secreted in Early Human Lactation. *J Agric Food Chem* **55**: 8914–8919.

Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. (2011). UniFrac: an effective distance metric for microbial community comparison. *ISME J* **5**: 169–172.

Lundell A-C, Adlerberth I, Lindberg E, Karlsson H, Ekberg S, Aberg N, *et al.* (2007). Increased levels of circulating soluble CD14 but not CD83 in infants are associated with early intestinal colonization with Staphylococcus aureus. *Clin Exp Allergy J Br Soc Allergy Clin Immunol* **37**: 62–71.

Macfarlane GT, Macfarlane S. (2012). Bacteria, Colonic Fermentation, and Gastrointestinal Health. *J AOAC Int* **95**: 50–60.

Macpherson AJ, McCoy KD, Johansen F-E, Brandtzaeg P. (2008). The immune geography of IgA induction and function. *Mucosal Immunol* **1**: 11–22.

Mah KW, Sangsupawanich P, Tunyapanit W, van Bever H, Shek LP, Chua KY, *et al.* (2008). Gut microbiota of children living in rural south Thailand and urban Singapore. *Allergol Int Off J Jpn Soc Allergol* **57**: 65–71.

Makino H, Kushiro A, Ishikawa E, Muylaert D, Kubota H, Sakai T, *et al.* (2011). Transmission of Intestinal Bifidobacterium longum subsp. longum Strains from Mother to Infant, Determined by Multilocus Sequencing Typing and Amplified Fragment Length Polymorphism. *Appl Environ Microbiol* **77**: 6788–6793.

Maleta KM, Phuka J, Alho L, Cheung YB, Dewey KG, Ashorn U, *et al.* (2015). Provision of 10-40 g/d Lipid-Based Nutrient Supplements from 6 to 18 Months of Age Does Not Prevent Linear Growth Faltering in Malawi. *J Nutr* **145**: 1909–1915.

Marcobal A, Barboza M, Sonnenburg ED, Pudlo N, Martens EC, Desai P, *et al.* (2011). Bacteroides in the Infant Gut Consume Milk Oligosaccharides via Mucus-Utilization Pathways. *Cell Host Microbe* **10**: 507–514.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, *et al.* (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376–380.

Markenson GR, Adams LA, Hoffman DE, Reece MT. (2003). Prevalence of Mycoplasma bacteria in amniotic fluid at the time of genetic amniocentesis using the polymerase chain reaction. *J Reprod Med* **48**: 775–779.

Martín R, Jiménez E, Heilig H, Fernández L, Marín ML, Zoetendal EG, *et al.* (2009). Isolation of Bifidobacteria from Breast Milk and Assessment of the Bifidobacterial Population by PCR-Denaturing Gradient Gel Electrophoresis and Quantitative Real-Time PCR. *Appl Environ Microbiol* **75**: 965–969.

Martin R, Langa S, Reviriego C, Jiminez E, Marin ML, Xaus J, *et al.* (2003). Human milk is a source of lactic acid bacteria for the infant gut. *J Pediatr* **143**: 754–758.

Matsuki T, Watanabe K, Fujimoto J, Kado Y, Takada T, Matsumoto K, *et al.* (2004). Quantitative PCR with 16S rRNA-gene-targeted species-specific primers for analysis of human intestinal bifidobacteria. *Appl Environ Microbiol* **70**: 167–173.

Matsuki T, Watanabe K, Tanaka R, Fukuda M, Oyaizu H. (1999). Distribution of bifidobacterial species in human intestinal microflora examined with 16S rRNA-gene-targeted species-specific primers. *Appl Environ Microbiol* **65**: 4506–4512.

Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL. (2005). An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122**: 107–118.

McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, Liu Z, *et al.* (2008). The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathog* **4**: e20.

Miller DN, Bryant JE, Madsen EL, Ghiorse WC. (1999). Evaluation and Optimization of DNA Extraction and Purification Procedures for Soil and Sediment Samples. *Appl Environ Microbiol* **65**: 4715–4724.

Million M, Diallo A, Raoult D. Gut microbiota and malnutrition. *Microb Pathog*. e-pub ahead of print, doi: 10.1016/j.micpath.2016.02.003.

Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, *et al.* (2011). The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res* **21**: 1616–1625.

Mitsou EK, Kirtzalidou E, Oikonomou I, Liosis G, Kyriacou A. (2008). Fecal microflora of Greek healthy neonates. *Anaerobe* **14**: 94–101.

Mpofu A, Linnemann AR, Sybesma W, Kort R, Nout MJR, Smid EJ. (2014). Development of a locally sustainable functional food based on mutandabota, a traditional food in southern Africa. *J Dairy Sci* **97**: 2591–2599.

Mshvildadze M, Neu J, Shuster J, Theriaque D, Li N, Mai V. (2010). Intestinal microbial ecology in premature infants assessed with non-culture-based techniques. *J Pediatr* **156**: 20–25.

von Mutius E, Radon K. (2008). Living on a farm: impact on asthma induction and clinical course. *Immunol Allergy Clin North Am* **28**: 631–647, ix–x.

Muyzer G, Smalla K. (1998). Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie Van Leeuwenhoek* **73**: 127–141.

Muyzer G, Waal EC de, Uitterlinden AG. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695–700.

Nadkarni MA, Martin FE, Jacques NA, Hunter N. (2002). Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiol Read Engl* **148**: 257–266.

Namsolleck P, Thiel R, Lawson P, Holmstrøm K, Rajilic M, Vaughan EE, *et al.* (2004). Molecular Methods for the Analysis of Gut Microbiota. *Microb Ecol Health Dis* **16**. e-pub ahead of print, doi: 10.3402/mehd.v16i2-3.7939.

Neis EPJG, Dejong CHC, Rensen SS. (2015). The Role of Microbial Amino Acid Metabolism in Host Metabolism. *Nutrients* **7**: 2930–2946.

Nübel U, Engelen B, Felske A, Snaidr J, Wieshuber A, Amann RI, *et al.* (1996). Sequence heterogeneities of genes encoding 16S rRNAs in Paenibacillus polymyxa detected by temperature gradient gel electrophoresis. *J Bacteriol* **178**: 5636–5643.

Nylund L, Heilig HGHJ, Salminen S, de Vos WM, Satokari R. (2010). Semi-automated extraction of microbial DNA from feces for qPCR and phylogenetic microarray analysis. *J Microbiol Methods* **83**: 231–235.

O'Hara AM, Shanahan F. (2006). The gut flora as a forgotten organ. *EMBO Rep* **7**: 688–693.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, *et al.* (2015). vegan: Community Ecology Package.

Oozeer R, van Limpt K, Ludwig T, Ben Amor K, Martin R, Wind RD, *et al.* (2013). Intestinal microbiology in early life: specific prebiotics can have similar functionalities as human-milk oligosaccharides. *Am J Clin Nutr* **98**: 561S–71S.

Oriol R, Danilovs J, Hawkins BR. (1981). A new genetic model proposing that the Se gene is a structural gene closely linked to the H gene. *Am J Hum Genet* **33**: 421–431.

Osborn DA, Sinn JKH. (2013). Prebiotics in infants for prevention of allergy. *Cochrane Database Syst Rev* CD006474.

Ottman N, Smidt H, de Vos WM, Belzer C. (2012). The function of our microbiota: who is out there and what do they do? *Front Cell Infect Microbiol* **2**. e-pub ahead of print, doi: 10.3389/fcimb.2012.00104.

Ozerlat I. (2011). Helicobacter pylori: H. pylori infection might suppress allergy in children. *Nat Rev Gastroenterol Hepatol* **8**: 418.

Pédron T, Mulet C, Dauga C, Frangeul L, Chervaux C, Grompone G, *et al.* (2012). A Crypt-Specific Core Microbiota Resides in the Mouse Colon. *mBio* **3**. e-pub ahead of print, doi: 10.1128/mBio.00116-12.

Penders J, Thijs C, Mommers M, Stobberingh EE, Dompeling E, Reijmerink NE, *et al.* (2010). Intestinal lactobacilli and the DC-SIGN gene for their recognition by dendritic cells play a role in the aetiology of allergic manifestations. *Microbiol Read Engl* **156**: 3298–3305.

Penders J, Vink C, Driessen C, London N, Thijs C, Stobberingh EE. (2005). Quantification of Bifidobacterium spp., Escherichia coli and Clostridium difficile in faecal samples of breast-fed and formula-fed infants by real-time PCR. *FEMS Microbiol Lett* **243**: 141–147.

Peng L, Li Z-R, Green RS, Holzman IR, Lin J. (2009). Butyrate enhances the intestinal barrier by facilitating tight junction assembly via activation of AMP-activated protein kinase in Caco-2 cell monolayers. *J Nutr* **139**: 1619–1625.

Perez PF, Dore J, Leclerc M, Levenez F, Benyacoub J, Serrant P, *et al.* (2007). Bacterial imprinting of the neonatal immune system: lessons from maternal cells? *Pediatrics* **119**: e724-32.

Petherick A. (2010). Development: Mother's milk: A rich opportunity. *Nature* **468**: S5–S7.

Pham VT, Lacroix C, Braegger CP, Chassard C. (2016). Early colonization of functional groups of microbes in the infant gut. *Environ Microbiol* n/a-n/a.

Phuka JC, Maleta K, Thakwalakwa C, Cheung YB, Briend A, Manary MJ, *et al.* (2008). Complementary feeding with fortified spread and incidence of severe stunting in 6- to 18-month-old rural Malawians. *Arch Pediatr Adolesc Med* **162**: 619–626.

Polz MF, Cavanaugh CM. (1998). Bias in template-to-product ratios in multitemplate PCR. *Appl Environ Microbiol* **64**: 3724–3730.

Ponnusamy K, Choi JN, Kim J, Lee S-Y, Lee CH. (2011). Microbial community and metabolomic comparison of irritable bowel syndrome faeces. *J Med Microbiol* **60**: 817–827.

Poretsky R, Rodriguez-R LM, Luo C, Tsementzi D, Konstantinidis KT. (2014). Strengths and Limitations of 16S rRNA Gene Amplicon Sequencing in Revealing Temporal Microbial Community Dynamics. *PLoS ONE* **9**. e-pub ahead of print, doi: 10.1371/journal.pone.0093827.

Prince AL, Ma J, Kannan PS, Alvarez M, Gisslen T, Harris RA, *et al.* (2016). The placental microbiome is altered among subjects with spontaneous preterm birth with and without chorioamnionitis. *Am J Obstet Gynecol*. e-pub ahead of print, doi: 10.1016/j.ajog.2016.01.193.

R Core Team. (2015). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria.

Ramanan D, Bowcutt R, Lee SC, Tang MS, Kurtz ZD, Ding Y, *et al.* (2016). Helminth infection promotes colonization resistance via type 2 immunity. *Science* **352**: 608–612.

Rautava S, Collado MC, Salminen S, Isolauri E. (2012a). Probiotics modulate host-microbe interaction in the placenta and fetal gut: a randomized, double-blind, placebo-controlled trial. *Neonatology* **102**: 178–184.

Rautava S, Luoto R, Salminen S, Isolauri E. (2012b). Microbial contact during pregnancy, intestinal colonization and human disease. *Nat Rev Hepatol* **9**: 565–576.

Reid G, Brigidi P, Burton JP, Contractor N, Duncan S, Fargier E, *et al.* (2015). Microbes Central to Human Reproduction. *Am J Reprod Immunol* **73**: 1–11.

Rescigno M, Urbano M, Valzasina B, Francolini M, Rotta G, Bonasio R, *et al.* (2001). Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nat Immunol* **2**: 361–367.

Rettedal EA, Gumpert H, Sommer MOA. (2014). Cultivation-based multiplex phenotyping of human gut microbiota allows targeted recovery of previously uncultured bacteria. *Nat Commun* **5**: 4714.

Reuter G. (2001). The Lactobacillus and Bifidobacterium microflora of the human intestine: composition and succession. *Curr Issues Intest Microbiol* **2**: 43–53.

Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, *et al.* (2010). Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* **466**: 334–338.

Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, *et al.* (2013). Gut Microbiota from Twins Discordant for Obesity Modulate Metabolism in Mice. *Science* **341**: 1241214.

Rinttila T, Kassinen A, Malinen E, Krogius L, Palva A. (2004). Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. *J Appl Microbiol* **97**: 1166–1177.

Rodríguez JM, Murphy K, Stanton C, Ross RP, Kober OI, Juge N, *et al.* (2015). The composition of the gut microbiota throughout life, with an emphasis on early life. *Microb Ecol Health Dis* **26**. e-pub ahead of print, doi: 10.3402/mehd.v26.26050.

Roduit C, Wohlgensinger J, Frei R, Bitter S, Bieli C, Loeliger S, *et al.* (2011). Prenatal animal contact and gene expression of innate immunity receptors at birth are associated with atopic dermatitis. *J Allergy Clin Immunol* **127**: 179–85, 185.e1.

Rondon MR, Raffel SJ, Goodman RM, Handelsman J. (1999). Toward functional genomics in bacteria: analysis of gene expression in Escherichia coli from a bacterial artificial chromosome library of Bacillus cereus. *Proc Natl Acad Sci U S A* **96**: 6451–6455.

Round JL, Lee SM, Li J, Tran G, Jabri B, Chatila TA, *et al.* (2011). The Toll-Like Receptor 2 Pathway Establishes Colonization by a Commensal of the Human Microbiota. *Science* **332**: 974–977.

Round JL, Mazmanian SK. (2010). Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc Natl Acad Sci* **107**: 12204–12209.

Round JL, Mazmanian SK. (2009). The gut microbiome shapes intestinal immune responses during health and disease. *Nat Rev Immunol* **9**: 313–323.

Ruiz-Moyano S, Totten SM, Garrido DA, Smilowitz JT, German JB, Lebrilla CB, *et al.* (2013). Variation in consumption of human milk oligosaccharides by infant gut-associated strains of Bifidobacterium breve. *Appl Environ Microbiol* **79**: 6040–6049.

Sakurama H, Kiyohara M, Wada J, Honda Y, Yamaguchi M, Fukiya S, *et al.* (2013). Lacto-N-biosidase Encoded by a Novel Gene of Bifidobacterium longum Subspecies longum Shows Unique Substrate Specificity and Requires a Designated Chaperone for Its Active Expression. *J Biol Chem* **288**: 25194–25206.

Salminen S, Endo A, Isolauri E, Scalabrin D. (2016). Early Gut Colonization With Lactobacilli and Staphylococcus in Infants: The Hygiene Hypothesis Extended. *J Pediatr Gastroenterol Nutr* **62**: 80–86.

Salminen S, Gibson GR, McCartney AL, Isolauri E. (2004). Influence of mode of delivery on gut microbiota composition in seven year old children. *Gut* **53**: 1388–1389.

Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, *et al.* (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* **12**: 87.

Salyers AA, Pajeau M, McCarthy RE. (1988). Importance of mucopolysaccharides as substrates for Bacteroides thetaiotaomicron growing in intestinal tracts of exgermfree mice. *Appl Environ Microbiol* **54**: 1970–1976.

Sandige H, Ndekha MJ, Briend A, Ashorn P, Manary MJ. (2004). Home-based treatment of malnourished Malawian children with locally produced or imported ready-to-use food. *J Pediatr Gastroenterol Nutr* **39**: 141–146.

Satokari R, Gronroos T, Laitinen K, Salminen S, Isolauri E. (2009). Bifidobacterium and Lactobacillus DNA in the human placenta. *Lett Appl Microbiol* **48**: 8–12.

Scott KP, Gratz SW, Sheridan PO, Flint HJ, Duncan SH. (2013). The influence of diet on the gut microbiota. *Pharmacol Res* **69**: 52–60.

Seiskari T, Kondrashova A, Viskari H, Kaila M, Haapala A-M, Aittoniemi J, *et al.* (2007). Allergic sensitization and microbial load − a comparison between Finland and Russian Karelia. *Clin Exp Immunol* **148**: 47–52.

Sela DA, Chapman J, Adeuya A, Kim JH, Chen F, Whitehead TR, *et al.* (2008). The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci* **105**: 18964–18969.

Sender R, Fuchs S, Milo R. (2016). Revised estimates for the number of human and bacteria cells in the body. *bioRxiv* 36103.

Sepp E, Julge K, Vasar M, Naaber P, Bjorksten B, Mikelsaar M. (1997). Intestinal microflora of Estonian and Swedish infants. *Acta Paediatr Oslo Nor 1992* **86**: 956–961.

Sevelsted A, Stokholm J, Bonnelykke K, Bisgaard H. (2015). Cesarean section and chronic immune disorders. *Pediatrics* **135**: e92-8.

Sinkiewicz G, Nordström EA. (2005). 353 Occurrence of Lactobacillus Reuteri, Lactobacilli and Bifidobacteria in Human Breast Milk. *Pediatr Res* **58**: 415–415.

Smith MI, Yatsunenko T, Manary MJ, Trehan I, Mkakosya R, Cheng J, *et al.* (2013). Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* **339**: 548–554.

Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR, *et al.* (2006). Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci U S A* **103**: 12115–12120.

Song W-J, Chang Y-S, Lim M-K, Yun E-H, Kim S-H, Kang H-R, *et al.* (2014). Staphylococcal enterotoxin sensitization in a community-based population: a potential role in adult-onset asthma. *Clin Exp Allergy J Br Soc Allergy Clin Immunol* **44**: 553–562.

Sonnenburg ED, Smits SA, Tikhonov M, Higginbottom SK, Wingreen NS, Sonnenburg JL. (2016). Diet-induced extinctions in the gut microbiota compound over generations. *Nature* **529**: 212–215.

Spangler R, Goddard NL, Thaler DS. (2009). Optimizing Taq Polymerase Concentration for Improved Signal-to-Noise in the Broad Range Detection of Low Abundance Bacteria. *PLOS ONE* **4**: e7010.

Stensballe LG, Simonsen J, Jensen SM, Bønnelykke K, Bisgaard H. (2013). Use of antibiotics during pregnancy increases the risk of asthma in early childhood. *J Pediatr* **162**: 832–838.e3.

Stobaugh HC, Ryan KN, Kennedy JA, Grise JB, Crocker AH, Thakwalakwa C, *et al.* (2016). Including whey protein and whey permeate in ready-to-use supplementary food improves recovery rates in children with moderate acute malnutrition: a randomized, double-blind clinical trial. *Am J Clin Nutr* **103**: 926–933.

Stout MJ, Conlon B, Landeau M, Lee I, Bower C, Zhao Q, *et al.* (2013). Identification of intracellular bacteria in the basal plate of the human placenta in term and preterm gestations. *Am J Obstet Gynecol* **208**: 226.e1-226.e7.

Strachan DP. (1989). Hay fever, hygiene and household size. *BMJ* **299**: 1259–1260.

Subramanian S, Huq S, Yatsunenko T, Haque R, Mahfuz M, Alam MA, *et al.* (2014). Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**: 417–421.

Sudo N, Sawamura S, Tanaka K, Aiba Y, Kubo C, Koga Y. (1997). The requirement of intestinal bacterial flora for the development of an IgE production system fully susceptible to oral tolerance induction. *J Immunol Baltim Md 1950* **159**: 1739–1745.

Sutherland DB, Suzuki K, Fagarasan S. (2016). Fostering of advanced mutualism with gut microbiota by Immunoglobulin A. *Immunol Rev* **270**: 20–31.

Suzuki TA, Worobey M. (2014). Geographical variation of human gut microbial composition. *Biol Lett* **10**: 20131037.

Svec P, Sevcikova A, Sedlacek I, Bednarova J, Snauwaert C, Lefebvre K, *et al.* (2007). Identification of lactic acid bacteria isolated from human blood cultures. *FEMS Immunol Med Microbiol* **49**: 192–196.

Sybesma W, Molenaar D, van IJcken W, Venema K, Kort R. (2013). Genome instability in Lactobacillus rhamnosus GG. *Appl Environ Microbiol* **79**: 2233–2239.

Thurl S, Munzert M, Henker J, Boehm G, Müller-Werner B, Jelinek J, *et al.* (2010). Variation of human milk oligosaccharides in relation to milk groups and lactational periods. *Br J Nutr* **104**: 1261–1271.

Tissier H. (1900). Recherches sur la Flore Intestinale des Nourrissons (etat Normal et Pathologique). University of Paris.

Toh MC, Allen-Vercoe E. (2015). The human gut microbiota with reference to autism spectrum disorder: considering the whole as more than a sum of its parts. *Microb Ecol Health Dis* **26**. e-pub ahead of print, doi: 10.3402/mehd.v26.26309.

Tringe SG, Rubin EM. (2005). Metagenomics: DNA sequencing of environmental samples. *Nat Rev Genet* **6**: 805–814.

Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**: 1027–1031.

Ulluwishewa D, Anderson RC, McNabb WC, Moughan PJ, Wells JM, Roy NC. (2011). Regulation of Tight Junction Permeability by Intestinal Bacteria and Dietary Components. *J Nutr* **141**: 769–776.

Urbaniak C, Angelini M, Gloor GB, Reid G. (2016). Human milk microbiota profiles in relation to birthing method, gestation and infant gender. *Microbiome* **4**: 1.

Urbaniak C, Cummins J, Brackstone M, Macklaim JM, Gloor GB, Baban CK, *et al.* (2014). Microbiota of human breast tissue. *Appl Environ Microbiol* **80**: 3007–3014.

Ventura M, Meylan V, Zink R. (2003). Identification and Tracing of Bifidobacterium Species by Use of Enterobacterial Repetitive Intergenic Consensus Sequences. *Appl Environ Microbiol* **69**: 4296–4301.

Verstraelen H, Vilchez-Vargas R, Desimpel F, Jauregui R, Vankeirsbilck N, Weyers S, *et al.* (2016). Characterisation of the human uterine microbiome in non-pregnant women through deep sequencing of the V1-2 region of the 16S rRNA gene. *PeerJ* **4**: e1602.

Vos WM de, Vos EA de. (2012). Role of the intestinal microbiome in health and disease: from correlation to causation. *Nutr Rev* **70**: S45–S56.

Waldor MK, Tyson G, Borenstein E, Ochman H, Moeller A, Finlay BB, *et al.* (2015). Where Next for Microbiome Research? *PLOS Biol* **13**: e1002050.

Walker A. (2010). Breast Milk as the Gold Standard for Protective Nutrients. *J Pediatr* **156**: S3–S7.

Walker AW, Martin JC, Scott P, Parkhill J, Flint HJ, Scott KP. (2015). 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. *Microbiome* **3**: 26–015–0087–4. eCollection 2015.

Walter J, Hertel C, Tannock GW, Lis CM, Munro K, Hammes WP. (2001). Detection of Lactobacillus, Pediococcus, Leuconostoc, and Weissella species in human feces by using group-specific PCR primers and denaturing gradient gel electrophoresis. *Appl Environ Microbiol* **67**: 2578–2585.

Wang B. (2009). Sialic Acid Is an Essential Nutrient for Brain Development and Cognition. *Annu Rev Nutr* **29**: 177–222.

Wang M, Li M, Wu S, Lebrilla CB, Chapkin RS, Ivanov I, *et al.* (2015). Fecal Microbiota Composition of Breast-fed Infants is Correlated with Human Milk Oligosaccharides Consumed. *J Pediatr Gastroenterol Nutr*. e-pub ahead of print, doi: 10.1097/MPG.0000000000000752.

Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, DuGar B, *et al.* (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**: 57–63.

Woese CR, Kandler O, Wheelis ML. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* **87**: 4576–4579.

Xiao J, Takahashi S, Nishimoto M, Odamaki T, Yaeshima T, Iwatsuki K, *et al.* (2010). Distribution of in vitro fermentation ability of lacto-N-biose I, a major building block of human milk oligosaccharides, in bifidobacterial strains. *Appl Environ Microbiol* **76**: 54–59.

Xu J, Gordon JI. (2003). Honor thy symbionts. *Proc Natl Acad Sci U S A* **100**: 10452–10459.

Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, *et al.* (2012). Human gut microbiome viewed across age and geography. *Nature* **486**: 222–227.

Yin Q, Zheng Q. (2005). Isolation and identification of the dominant Lactobacillus in gut and faeces of pigs using carbohydrate fermentation and 16S rDNA analysis. *J Biosci Bioeng* **99**: 68–71.

Youssef N, Sheik CS, Krumholz LR, Najar FZ, Roe BA, Elshahed MS. (2009). Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Appl Environ Microbiol* **75**: 5227–5236.

Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ. (2012). Evaluation of Methods for the Extraction and Purification of DNA from the Human Microbiome. *PLoS ONE* **7**. e-pub ahead of print, doi: 10.1371/journal.pone.0033865.

Zhang C, Yin A, Li H, Wang R, Wu G, Shen J, *et al.* (2015). Dietary Modulation of Gut Microbiota Contributes to Alleviation of Both Genetic and Simple Obesity in Children. *EBioMedicine* **2**: 968–984.

Zheng J, Xiao X, Zhang Q, Mao L, Yu M, Xu J. (2015). The Placental Microbiome Varies in Association with Low Birth Weight in Full-Term Neonates. *Nutrients* **7**: 6924–6937.

Zimmermann MB, Chassard C, Rohner F, N'goran EK, Nindjin C, Dostal A, *et al.* (2010). The effects of iron fortification on the gut microbiota in African children: a randomized controlled trial in Cote d'Ivoire. *Am J Clin Nutr* **92**: 1406–1415.

Zivkovic AM, German JB, Lebrilla CB, Mills DA. (2011). Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proc Natl Acad Sci U S A* **108 Suppl 1**: 4653–4658.

Zoetendal EG, Raes J, van den Bogert B, Arumugam M, Booijink CC, Troost FJ, *et al.* (2012). The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J* **6**: 1415–1426.

Zoetendal EG, Wright A von, Vilpponen-Salmela T, Ben-Amor K, Akkermans ADL, Vos WM de. (2002). Mucosa-Associated Bacteria in the Human Gastrointestinal Tract Are Uniformly Distributed along the Colon and Differ from the Community Recovered from Feces. *Appl Environ Microbiol* **68**: 3401–3407.

Zopf D, Roth S. (1996). Oligosaccharide anti-infective agents. *Lancet Lond Engl* **347**: 1017–1021.

Turun yliopisto
University of Turku