Infant microbiota and short chain fatty acids

Relation to lifestyle and effect on health

Monica Gio-Batta

Department of Infectious Diseases

Institute of Biomedicine

Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2023

monica.gio-batta@gu.se

ISBN 978-91-8069-369-1 (PRINT)
ISBN 978-91-8069-370-7 (PDF)

Printed in Borås, Sweden 2023
Printed by Stema Specialtryck AB

Infant microbiota and short chain fatty acids

Relation to lifestyle and effect on health

Monica Gio-Batta

Department of Infectious Diseases, Institute of Biomedicine
Sahlgrenska Academy, University of Gothenburg
Gothenburg, Sweden

ABSTRACT

The commensal microbiota has both beneficial and negative effects on the host. It is a source of infectious and inflammatory microorganisms, but may also provide immune stimulation during a critical period during early infancy that possibly offers protection against allergy and other immune mediated diseases. Here, we examine four birth cohorts to investigate the interplay between lifestyle factors, microbiota acquisition and maturation, production by the microbiota of immunomodulatory short chain fatty acids (SCFA) and allergy development. The FARMFLORA birth cohort consists of children raised on small family-owned dairy farms and non-farming controls from the same rural areas, the BAS cohort followed infants born in Jämtland until 13 years of age, the NICE cohort includes infants born in Norrbotten, and the ÖFLORA cohort consists of extremely preterm infants cared for at a neonatal intensive care unit in Göteborg. Our hypotheses were that preterm neonates acquire a microbiota that may predispose to sepsis and the inflammatory condition necrotizing enterocolitis, while a microbiota dominated by anaerobic bacteria yields protection from allergy development, in part through production of SCFAs. Using 16S rRNA gene sequencing, we demonstrate that growing up on a farm, with pets, or with elder siblings are all associated with enrichment of core anaerobic commensals, which in turn is reflected in more complex SCFAs in feces and protection from allergy at 3 and 8 years of age. The SCFA valeric

acid was associated with protection from allergy in later childhood, a finding confirmed in the BAS cohort. Using newly developed highly sensitive liquid chromatography-mass spectroscopy, we measured SCFAs in infants' blood plasma at 4 months of age, as well as corresponding samples of mother's milk and blood plasma in the NICE cohort. Plasma levels of several SCFAs were lower in infants who later became allergic and/or sensitized. Interestingly, the SCFAs butyrate and caproate were enriched 100-fold in mother's breast milk compared to her blood plasma, suggesting selective transport into the milk to serve the demands of the infant. Further, iso-butyrate and valerate were also enriched (4-fold) in milk, while other SCFAs were less prevalent in milk than in maternal blood plasma. Lastly, characterization of the gut, oral and skin microbiota of extremely preterm infants revealed a microbiota with a profound lack of anaerobes, dominated by Gram-positive facultatives, especially coagulase-negative staphylococci, and with a high prevalence of yeasts. Treatment with antibiotics affected microbiota development, as also did the degree of prematurity. The characteristics of the preterm microbiota may increase the risk of infectious and inflammatory conditions. Taken together, our results suggest that early life exposures impact on the composition of the infant gut microbiota and that SCFAs might be mediators contributing to immune maturation and protection against allergy.

Keywords: microbiota, gut, oral, skin, short chain fatty acids, allergy, siblings, farm, pets, extremely preterm, quantitative culture, 16S rRNA gene sequencing

ISBN 978-91-8069-369-1 (PRINT)
ISBN 978-91-8069-370-7 (PDF)

Sammanfattning på svenska

Vår hud och många av våra slemhinnor täcks av en normalflora som etableras under nyföddhetsperioden och blir alltmer komplex under tidig barndom. Normalfloran har både positiva och negativa effekter. Den kan ge oss infektioner och orsaka inflammation, men också stimulera immunsystemet så att det utvecklas på korrekt sätt och inte överreagerar mot ofarliga ämnen; enligt den så kallade ”hygienhypotesen” ger en riklig exponering för mikrober under nyföddhetsperioden skydd mot utveckling av allergi och andra immunregleringssjukdomar. Normalfloran i tjocktarmen tillverkar också korta fettsyror så som ättikssyra, smörsyra och valieransyra som utgör näring för tarrmslemhinnan och påverkar immunsystemets celler. I avhandlingen studeras barn från fyra födelsekohorter, alltifrån barn som växer upp på bondgård till för tidigt födda barn på neonatal intensivvårdsavdelning. Vi studerar betydelsen av olika livsstilsfaktorer för hur normalfloran etableras och mognar, dess produktion av olika kortkedjiga fettsyror och relation till senare allergiutveckling. Vi visar att barn som växer upp på bondgård, med husdjur eller med äldre syskon, tidigt får en komplex anaerob tarmflora och mer av vissa fettsyror i avföringen. Den korta fettsyran valeriansyra, liksom en tarmflora tidigt dominerad av anaerober, var förknippad med minskad allergiutveckling. Valeriansyra var också förknippad med minskad allergiutveckling i en annan födelsekohort (Barn Allergistudien). I födelsekohorten NICE sågs en koppling mellan koncentration av flera olika korta fettsyror i spädbarnets blod och minskad risk för senare allergiutveckling och att smörsyra och kapronsyra selektivt anrikades i mammans bröstmjölk där de förekom i 100 gånger högre koncentration än i hennes blod; kanske utgör de näring för barnets slemhinnor. Normalfloran hos extremt för tidigt födda barn var fattig på anaeroba bakterier och både mun-, hud- och tarmflora dominerades av grampositiva fakultativa anaerober; jästsvampar förekom rikligt. Prematurfloran var framför allt kopplad till den myckna antibiotikabehandlingen och kan vara en riskfaktor för utveckling av sepsis och nekrotiserande enterokolit som ofta drabbar dessa barn. Sammantaget tyder våra resultat på att många faktorer påverkar sammansättningen av spädbarnets normalflora och att tarmfloran, liksom de korta fettsyror den bildar, kan bidra till immunsystemets mognad och skydd mot allergiutveckling.

List of papers

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. **Gio-Batta M**, Sjöberg F, Yazdanshenas S, Nookaew I, Hesselmar B, Saalman R, Adlerberth I, Wold A. E. Delayed acquisition of typical commensal anaerobes is associated with various early life exposures and allergy development. Manuscript.

II. **Gio-Batta M**, Sjöberg F, Jonsson K, Barman M, Lundell A-C, Adlerberth I, Hesselmar B, Sandberg A-S, Wold A. E. Fecal Short Chain Fatty Acids in Children Living on Farms and a Link between Valeric Acid and Protection from Eczema. Scientific Reports 2020; 10: 22449.

III. **Gio-Batta M**, Spetz K, Barman M, Bråbäck L, Norin E, Björkstén B, Wold A. E, Sandin A. Low concentration of fecal valeric acid at one year of age is linked with eczema and food allergy at 13 years of age – findings from a Swedish birth cohort. International Archives of Allergy and Immunology 2022; 183(4):398-408.

IV. Barman M, **Gio-Batta M**, Andrieux L, Stråvik M, Saalman R, Fristedt R, Rabe H, Sandin A, Wold A. E and Sandberg A-S. Short-chain fatty acids (SCFA) in infant plasma – relation to SCFA levels in breast milk and plasma and to subsequent food allergy, atopic eczema or sensitization in a prospective birth-cohort study. Manuscript (submitted)

V. **Gio-Batta M**, Thordarson T, Sjöberg F, Bry K, Wold A. E, Elfvin A, Adlerberth I. Establishment of early gut, oral and skin microbiota in extremely preterm infants – a comprehensive culture-based study. Manuscript.

Contents

Abbreviations

Acetyl CoA	Acetyl coenzyme A
APGAR	Score of **A**ppearance, **P**ulse, **G**rimace, **A**ctivity and **R**espiration
BCRP	Breast cancer resistance protein
CFU	Colony-forming unit
CoNS	Coagulase-negative staphylococci
CPAP	Continuous positive airway pressure
FDR	False discovery rate
GPR	G protein-coupled receptor
IgA, IgG, IgE	Immunoglobulin A, G and E
LOS	Late-onset sepsis
MRSA	Methicillin-resistant *Staphylococcus aureus*
NEC	Necrotizing enterocolitis
NICU	Neonatal intensive care unit
OPLS	Orthogonal Partial Least Squares
SCFA	Short chain fatty acid
SIgA	Secretory immunoglobulin A
Treg	Regulatory T cell
Th2	T helper 2 cell

Introduction

The human body is home to an immense diversity of microbial life known as the microbiota, which is made up of trillions of individual bacteria, as well as fungi and sometimes archaea. Unlike pathogens infecting us, or transient bacteria that are swallowed in food or contaminate the skin from outside sources, the microbiota is an ecosystem of organisms which live and replicate on and in the body. Normally, these organisms exert only limiting harm on their hosts. For this reason, they are referred to as the body's normal flora. However, it should be acknowledged that most bacterial infections are caused by bacteria that have been commensals for a shorter or longer period of time, before they enter a normally sterile site and cause infection. The vast amount of inflammatogenic substances present in and on bacterial cells may also be a source of chronic inflammation, paving way for diseases such as colon cancer.

The microbiota is acquired during infancy in a gradual process that is influenced by the array of bacteria present in the environment and modified by factors such as diet and exposure to antibiotics. This thesis concerns the composition and evolution of the microbiota in relation to some external exposures, such as having elder siblings or pets, and being raised on a farm, but also being born much too early and being exposed to the very particular environment of a neonatal intensive care unit.

The composition of the infant's microbiota may affect several aspects of future health. The microbiota contains a vast pool of foreign antigens as well as microbial danger signals and therefore has the potential of influencing the development of the immune system during the critical period of early infancy, including the capacity to either actively tolerize environmental antigens, or to develop sensitization and allergy. However, the mechanisms by which the microbiota affects maturation of the immune system in early infancy are unknown. Here, we focus on short-chain fatty acids, metabolites produced by the colonic microbiota that have known immune-modulating properties.

Conversely, many organisms residing in the microbiota may cause infections, if reaching too high population counts, or if spread into the blood, tissues or normally sterile mucosae. Neonates, especially if born prematurely, are at increased risk of such infections.

Microbial classification and characteristics

All cellular life can be classified into three domains: bacteria, archaea and eukaryotes, and all of these may be represented in the human microbiota. Bacteria and archaea are both single-celled organisms with no cell nucleus, i.e. procaryotes. Fungi are eukaryotes but contain a cell wall. Although they are usually complex multicellular organisms, some of them can also exist in unicellular form that may colonize animals, including humans, and sometimes cause disease.

All living organisms can be classified according to a taxonomic system of eight levels – domain, kingdom, phylum, class, order, family, genus and species (**Figure 1**). At each level, taxonomic groups become smaller and include organisms with increasingly similar characteristics. Taking humans as an example, we belong to the domain 'eukaryota', kingdom 'animalia', phylum 'chordata', class 'mammalia', order 'primates', family 'hominidae', genus 'homo' and species 'homo sapiens'. For microorganisms, each species may also contain different strains, representing different genetic subtypes. If a colony of microbial cells all descend from a single ancestral cell and are therefore genetically identical, these cells are referred to as clones.

In addition to the formal naming system, certain key characteristics are used to describe different groups of bacteria. Facultatively anaerobic bacteria grow in the presence of oxygen but are also able to survive and replicate in anaerobic conditions, while anaerobic bacteria cannot use oxygen for growth and are unable to derive energy by using oxygen as terminal electron acceptor. Among anaerobes, tolerance for oxygen varies and some are aerotolerant, while obligate anaerobes die within minutes or hours of being in an oxygenated atmosphere. Anaerobic bacteria carry out many different metabolic processes, including fermentation, but also oxidation/reduction of

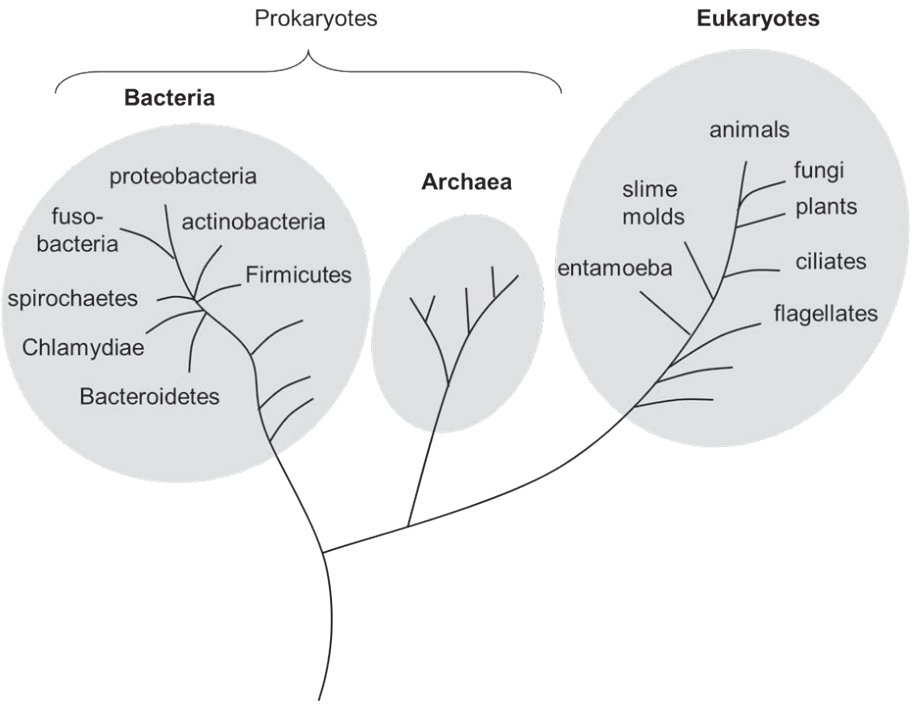

Figure 1: The tree of life. Branches show the classification of living organisms into different groups. Prokaryotes are single-celled organisms with no cell nucleus belonging to the domains Bacteria and Archaea. Eukaryotes include all organisms with a cell nucleus and may be single-celled, such as protozoa (amoeba, ciliates and flagellates) and some fungi, or multicellular, such as animals, plants and many fungi.

small compounds such as sulphate and nitrate. They may produce compounds that are partially toxic to facultative bacteria and encroach on their metabolism and living space. Therefore, an abundance of different anaerobic bacteria will reduce the capacity of anaerobes to proliferate and their populations will dwindle in the presence of a varied anaerobic microbiota. Facultative bacteria are able to translocate over epithelia, causing invasive infections. A higher abundance of intestinal facultative bacteria increases the risk of their translocation to mesenteric lymph nodes, and antibiotic-induced depletion of anaerobes in the gut also facilitates translocation of facultatives (1). Meanwhile, anaerobes seldom give rise to infections, probably because they cannot survive the high oxygen tension in live tissue (2). Metabolism of anaerobes and facultatives also differs. Anaerobes only have anaerobic metabolism and produce short-chain fatty

acids (SCFAs) through partial oxidation of metabolic substrates, a process termed fermentation. Under anaerobic conditions fermentation occurs also in facultatives, but when oxygen levels suffice facultative anaerobes use aerobic respiration to fully oxidise substrates to carbon dioxide and water. This yields much more energy than fermentation.

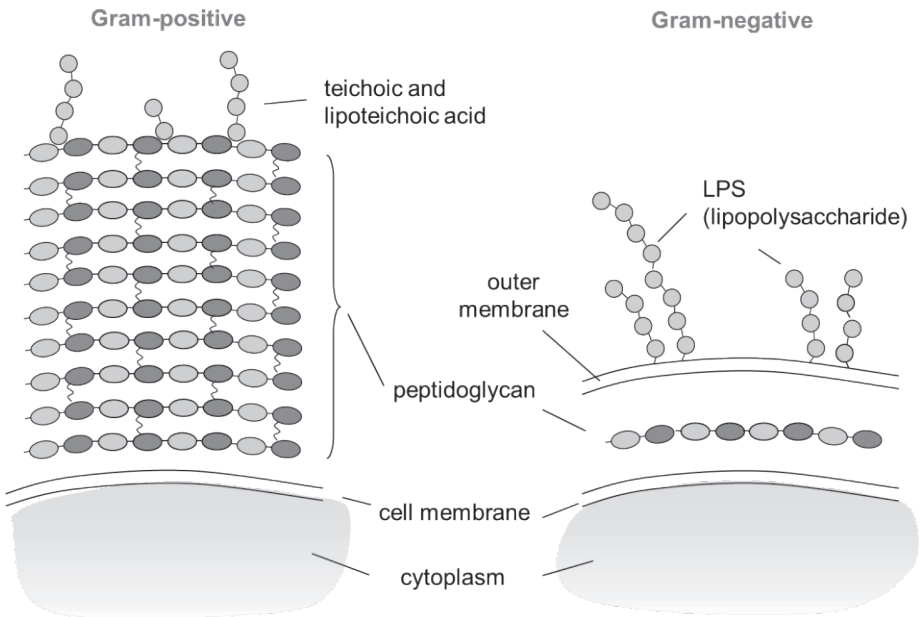

Figure 2: Gram-positive and Gram-negative bacterial cell walls. In Gram-positive bacteria (left) the cell wall is made up of many layers of peptidoglycan. In Gram-negative bacteria (right) the cell wall consists of a much thinner layer of peptidoglycan, which is itself surrounded by an outer membrane. This outer membrane contains lipopolysaccharide (LPS), which is recognized by our immune defences and activates an inflammatory response.

Another fundamental grouping of bacteria is between Gram-positive species, which have a thick peptidoglycan cell wall, and Gram-negative species which have a thin peptidoglycan cell wall and also an outer membrane (**Figure 2**). Lipopolysaccharide (LPS) is a major component of the outer membrane of Gram-negative bacteria, and teichoic and lipoteichoic acid are found in the

cell wall of Gram-positive bacteria, while peptidoglycan (murein) is found in both groups. LPS and peptidoglycan are unique to bacteria and recognized as danger signals by the immune system, which is activated on contact with these structures. Likewise, bacterial DNA also differs in structure from human DNA and acts as an additional danger signal to the immune system. Due to their different outer coatings, Gram-positive and negative bacteria have different effects on the immune system, with Gram-positives stimulating higher production of TNF-α, IFN-g and IL-12, suggesting that they induce a cytokine pattern than promotes Th1 effector functions. In contrast, lipopolysaccharide from Gram-negative bacteria stimulates higher production of IL-6, IL-8 and IL-10 (3, 4).

Characterization of the microbiota

From the beginning, people interested in studying bacteria have grown them in different kinds of media which provide food and other ingredients. This is known as culturing bacteria, and the original studies describing the gut microbiota in the 1970s were carried out using bacterial culture (5). However, standard culture methods are only able to detect 30-40% of the commensal bacteria in the colonic microbiota of an adult human being (6). More recently, the development of gene sequencing technologies has allowed detection of bacteria that are difficult to culture, particularly exquisitely anaerobic bacteria (6). Nonetheless, bacterial culture remains relevant to characterization of the microbiota, as outlined in more detail below. Both culture methods and 16S rRNA gene sequencing are used in this thesis. It should be noted that the anaerobes that colonize newborn infants can readily be cultured as they are relatively tolerant to ambient oxygen that is present in abundance in the intestine of the newborn infant.

Culture methods

Culture methods to identify microbes in a complex mixture such as the human microbiota rely on a range of different techniques, according to the following steps.

1) Firstly, microbiota samples, for example fecal samples in the case of studies of the gut microbiota, are cultured under various conditions suited to the particular growth requirements of individual microbial groups. As different

groups may have different nutritional needs, various types of growth media are used. Non-selective media such as Colombia blood agar support the growth of many different bacteria, while selective media are adapted to support the growth of individual microbial groups and to restrict the growth of others. For instance, Drigalski agar, which is used to culture facultative Gram-negative gut bacteria including *Enterobacteriaceae*, contains a mixture of suitable nutrients but also bile acid (deoxycholic acid) and other agents to inhibit Gram-positive bacteria. Cultures are incubated under either aerobic or anaerobic conditions. As facultative anaerobes grow best under aerobic conditions, they are cultured in air or in a CO_2-enriched atmosphere. Obligate anaerobes in the microbiota samples do not survive under aerobic conditions, although some aerotolerant anaerobes may persist. Thus, one can easily detect and enumerate facultatives with population sizes of 1000/gram of faeces, in the presence of 10^{11} anaerobic bacteria. Obligate anerobes are cultured under anaerobic conditions, in a nitrogen atmosphere, which in turn suppresses growth of facultatives. The facultatives also grow under anaerobic conditions and their counts must, hence, be subtracted from the counts on anaerobically incubated plates, to obtain the population counts of anaerobes.

2) In parallel with Step 1 above, the amount of microbes in a sample may be determined by a process called quantitative culture (**Figure 3**). In this process, a known amount of the microbiota sample is serially diluted and each dilution is cultured as described. Each microbial colony that forms originates from a single microbe, or from a small clump in the case of bacteria such as staphylococci that clump together. The originating single microbe or clump is known as the colony forming unit. Counting the number of colonies formed from a relevant dilution (which produces an amount of colonies which is not too few but also not so many that they can't be easily counted) and multiplying by the dilution factor gives the number of colony forming units of that particular bacterial group in the original microbiota sample.

3) Finally, the microbes which have been cultured are then identified. Some bacteria may be preliminary identified based on their growth in a given medium and atmosphere together with the colony morphology, and some bacterial groups may be identified based on Gram-stain and bacterial form (cocci, bacilli, endospores) as evaluated under a microscope. Biochemical tests based on the presence or absence of certain bacterial enzymes may aid in the identification of some groups. The enzyme catalase converts hydrogen

peroxide to oxygen, and presence of catalase is characteristic of e.g. staphylococci. Also, the enzyme coagulase converts fibrinogen to fibrin, which triggers clot formation in blood plasma, and this enzyme is found in *Staphylococcus aureus* but not other 'coagulase-negative' staphylococci such as *S. epidermidis* and *S. haemolyticus*. Biotyping is a method for species identification, which relies on different bacterial species having characteristic patterns of reaction to an array of biochemical tests. Biotyping kits are commercially available, for example API 20E for identification of *Enterobacteriaceae* species. Finally, matrix-assisted laser desorption ionization - time of flight mass spectrometry (MALDI-TOF) is a more recent, rapid, sensitive and inexpensive method for identification of bacteria and fungi, typically at the species level. Microbial colonies are bombarded by a laser beam, releasing microbial fragments which are detected by mass spectroscopy, and identification is carried out by comparison of the microbial fragment pattern against a reference library. Specific PCR methods may also be used to identify select groups of bacteria.

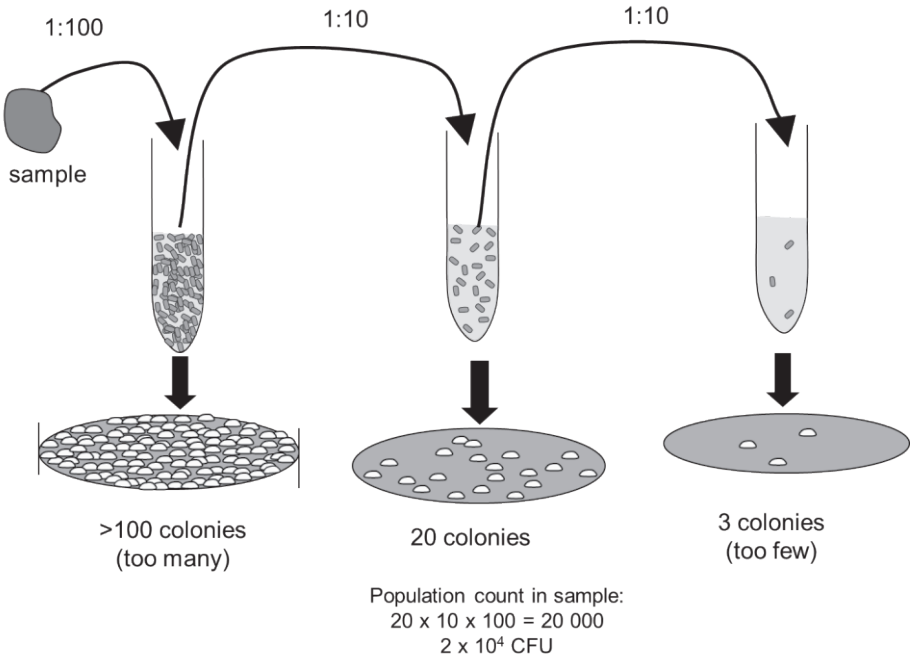

Figure 3: Quantitative culture. Microbial samples are serially diluted, the dilutions are cultured and the colonies counted. Multiplying the number of colonies by the dilution factor gives the microbial population count in the sample.

DNA-based methods

While various DNA-based methods exist for characterizing the microbiota, the main method currently used is 16S rRNA gene sequencing. This relies on the fact that the 16S subunit of the bacterial ribosome is highly conserved in bacteria, with only small variations in the 16S rRNA-gene within a genus or species, but with some variable gene regions that vary between bacterial genera and families. Sequencing of these variable regions of 16S rRNA genes in microbiota samples is carried out according to the following steps:

1) Firstly, DNA is extracted from the microbiota samples, and various commercial kits are available to do this. This entails rupture of the cell wall. As the Gram-positives have a much thicker and more sturdy cell wall than the Gram-negatives, it is much easier to extract DNA from the latter (7). Then, with the help of primers which bind to the desired variable region(s) of the 16S rRNA gene, copies of 16S rRNA gene fragments belonging to the many different types of bacteria in each sample are created using polymerase chain reaction (PCR).

2) Next, the DNA code of these copied gene fragments is sequenced (thousands of unique 16S rRNA gene fragments are sequenced simultaneously). One common sequencing method is called Illumina sequencing. In this method, the copied gene fragments generated above are separated and each attached to a specially-adapted surface in a reaction vessel called a flow cell and further copies of each fragment are made. Then in another PCR step, fluorescently-labelled nucleotides are added, which attach to the complementary nucleotides on the gene fragments to make new gene fragment copies containing only labelled nucleotides. Each nucleotide has a different colour fluorescent label, and laser excitation of these labels generates fluorescent light flashes showing which nucleotides have been added and in what order.

3) Bioinformatics processing of the information from these light signals generates the DNA sequences of the 16S rRNA gene fragments in the original microbiota samples. These sequences are then compared to reference libraries of known 16S rRNA gene sequences for various bacterial groups, generating a catalogue of the bacteria in the microbiota samples.

Comparison of culture and DNA-based methods

Despite the great advantage of DNA-based methods in allowing much of the bacterial diversity in the human microbiota to be detected, they also have some important limitations in comparison to culture methods, and conversely culture methods also have important limitations (**Table 1**). Firstly, while culture identifies only living microbes, 16S rRNA gene sequencing does not discriminate between living and dead cells. This may overestimate the proportions of certain bacterial groups in a sample, since the true population only corresponds to living cells. Further, when using culture methods, it is easy to include some media for the culture of yeasts, which are also part of the microbiota. 16S rRNA gene sequencing only detects bacteria, and additional methods, e.g. sequencing of the fungal 18S rRNA-gene, are needed to cover yeasts. Also, the taxonomic resolution available with 16S rRNA gene sequencing is limited by natural variation in the 16S rRNA gene, meaning that bacteria are, in the best case scenario, identified at the genus but not the species level. Often, the family level is used, and sometimes even higher taxonomic levels. In contrast, culture methods readily allow identification of certain clinically-relevant bacteria at the species level, e.g. *S. aureus*, which in 16SrRNA gene sequencing cannot be distinguished from *S. epidermidis* and other members of the *Staphylococcus* genus, collectively referred to as coagulase-negative staphylococci, which have very little in common with *S. aureus* in terms of pathogenicity and effects on the immune system. Another example is the highly pathogenic pneumococci that are interwoven within the viridans streptococci which they are related to and share similar 16S rRNA genes, but the pneumococci possess a capsule and several other virulence factors that makes it highly pathogenic. Along the same lines, *Escherichia* and *Shigella* are genetically identical, except for collections of virulence factors that makes *Shigella* a very dangerous pathogen.

With 16S rRNA gene sequencing, factors other than the detection of both live and dead bacteria may also introduce bias which lead to certain bacterial groups being either over- or under-represented in the resulting characterization of the microbiota. When extracting DNA from microbiota samples, more DNA is often obtained from Gram-negative than Gram-positive bacteria, as the latter have a thick cell wall which is less easily broken open, meaning that the proportions of various Gram-positive bacteria may be

Table 1: Comparison of bacterial culture and 16S rRNA gene sequencing.

	Culture	**16S rRNA gene sequencing**
Microbes detected	Facultative bacteria and a proportion of the anaerobes. Selective media may allow detection of less abundant bacteria. Also yeasts are detected.	Theoretically all bacteria. Detection of low abundance bacteria limited by sampling depth. Yeasts are not detected.
Taxonomic resolution	Species level. Several methods are available for identification to the species level.	Mostly genus level. Sometimes family level. (Species level possible with metagenome sequencing).
Estimation of bacterial proportions and population counts	Population counts of bacteria can be determined based on live microbes. Limited to culturable microbes. May be biased by culture and media limitations.	Population counts cannot be determined, quantification is proportion of 16S rRNA genes in population of unknown size. Live and dead bacteria detected. Biased by less efficient DNA extraction from Gram-positive bacteria and primers which do not bind equally well to all bacteria. Also biased by varying numbers of the 16S rRNA gene in different bacteria.
Quantification	Population counts based on living microbes	Proportion of 16S rRNA genes in population of unknown size, gene copy number varies between genera and species
Microbial characterization	Microbes may be isolated and physiology investigated	Bacteria not isolated, knowledge of bacterial function may be inferred from metagenomic studies

under-estimated. Also, although the primers which bind to the variable region(s) of the 16S rRNA gene are supposedly ‘universal’, meaning they are

effective for all kinds of bacteria, in fact they bind less well to some bacterial groups (eg. bifidobacteria) than others. Furthermore, the most abundant bacteria in the microbiota are most heavily sampled, and therefore depending on the sampling depth bacteria present at low abundance may be missed. One should recognize that the colonic microbiota contains microbes present at 10^5 CFU/g and others at 10^{11}, a million-fold difference. This means that after a million reads of species A, one read of species B will appear. As we discussed above, facultatives usually have population levels that are 1000 to 100,000 times lower than those of the anaerobes, and facultatives are the most relevant in terms of pathogenicity, one cannot ignore “low-abundance” populations in a complex ecosystem.

Another drawback with 16S rRNA gene sequencing is that it provides information on the proportion of 16S rRNA genes attributed to each bacterial group in microbiota samples but is not able to quantify the overall population size. As the microbial load of healthy adults can vary up to tenfold between individuals, this creates artefacts linked to the proportional nature of the data (8). Also, different bacterial genera, and different strains within the same genus, have widely varying numbers of the 16S rRNA gene, meaning the proportions of each genus may not be directly comparable within individuals and/or between groups. Furthermore, while many more different bacteria are detected by 16S rRNA sequencing or DNA-based methods than by culture, which is a great advantage of the former methods, knowledge on the function of these uncultured bacteria, and therefore their relevance to health and disease, is limited to what can be inferred from their genome. The number of different bacteria identified by culture may be greatly increased by the culturomics approach (9). However, this approach is limited to specialized laboratories with extended culturing facilities.

Microbiota on different body surfaces

The different parts of the body differ greatly in the conditions for enabling microbes to establish and replicate, and the number and type of microbes which flourish in these locations also varies profoundly (**Figure 4**). As the skin is dry and exposed to the external environment, is relatively inhospitable for microbes. On the other hand, the mucous membranes inside the body are protected, moist and rich in nutrients, often providing attractive locations for

microbes, and the mucosae of the gastrointestinal tract, upper respiratory tract and vagina all harbour their own unique microbiota. However, the mucosae of the lower respiratory tract and lungs, middle ear and sinuses are normally sterile, as are the mucosae of the uterus and most parts of the urinary tract.

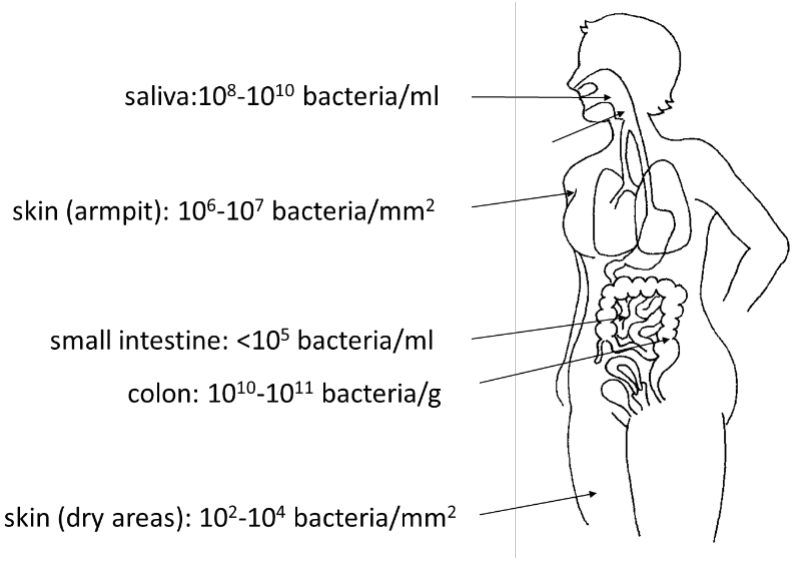

Figure 4: Bacterial density at different body sites

Alimentary tract

Most commensal bacteria are found in the alimentary tract, with density and composition varying from the mouth and along the gastro-intestinal tract.

Mouth

The mouth contains 10^8-10^{10} bacteria per milliliter of saliva and provides a variety of habitats to microbial communities including tooth surfaces, tooth pockets, gum, cheek and tongue. Saliva contains a mixture of bacteria from all these sources, and streptococci such as *S. mitis* dominate, with *Veillonella, Prevotella, Haemophilus, Gemella* and *Actinomyces* also being common colonizers (**Table 2**). Of these, *Haemophilus* is abundant in the cheek and *Actinomyces* and *Prevotella* on the gums, above and below the gum line

respectively (10, 11). As 5-10 L of saliva are swallowed every day comprising some 10^{11}-10^{15} bacteria in total, the bacterial in saliva seed the rest of the gastrointestinal tract. The stomach kills 99% of the swallowed bacteria, but this process is less efficient in people taking proton pump inhibitors or with an autoimmune elimination of the acid-producing cells (atrophic gastritis).

Stomach

The highly acidic stomach contains few microbes with only ~ 10 living cells per gram of content, reflecting the efficiency by which stomach acid kills microbes (bacteria, fungi and enveloped viruses), with only naked viruses being resistant. Although some species have a resident microbiota in the stomach (e.g. lactobacilli in the mouse), only a single microbe, *Helicobacter pylori* is known to be able to colonize the human stomach. However, the stomach may contain a substantial number of dead bacteria, from swallowed saliva. These will not be detected by bacterial culture, but may be detected by DNA-based methods which cannot distinguish between live and dead microbes, giving a false impression of organisms other than *H. pylori* colonizing the stomach (12).

Small intestine

The small intestine is also relatively sparsely populated, with around 10^3 bacteria per milliliter of content near the exit from the stomach rising to around 10^5-10^8 bacteria per milliliter near the ileocaecal valve (13). As the stomach contents enter the duodenum, few bacteria have survived and the rapid peristalsis does not permit bacteria to replicate fast enough to form large populations. Digestive enzymes and bile acids released from the pancreas and liver are added to the contents and antimicrobial peptides are released by the epithelia, which collectively may limit the chances for bacterial growth. Despite these challenges there are bacteria present, and biopsies show a mucosa-associated microbiota dominated by streptococci, *Neisseria* and *Sutterella*, with *Veillonella*, *Gemella*, *Actinomyces*, *Rothia* and *Haemophilus* also present, similar to the oral flora (14, 15). However, whether these bacteria are resident colonizers or represent transient oral bacteria remains to be determined. In the distal small intestine, obligately anaerobic bacterial groups abundant in the colon such as *Bacteroides* and clostridia are found near the ileocaecal valve, where the contents are more stagnant and

hence more anaerobic as oxygen is consumed by replicating facultative bacteria (13).

Table 2: Key bacterial taxa in the microbiota at different body sites

Phylum	Class	Genus	Facultative/ anaerobic	Gram
Gut (colon)				
Firmicutes	Bacilli	Enterococcus	Facultative	+
		Lactobacillus	Anaerobic	+
	Clostridia	Clostridium	Anaerobic	+
		Eubacterum	Anaerobic	+
		Ruminococcus	Anaerobic	+
		Faecalibacterium	Anaerobic	+
		Coprococcus	Anaerobic	+
	Negativicutes	Veillonella	Anaerobic	-
		Megasphaera	Anaerobic	-
Bacteroidetes	Bacteroidetes	Bacteroides	Anaerobic	-
		Prevotella	Anaerobic	-
Actinobacteria	Actinobacteria	Bifidobacterium	Anaerobic	+
	Coriobacteriia	Collinsella	Anaerobic	+
	Coriobacteriia	Eggerthella	Anaerobic	+
Proteobacteria	γ -Proteobacteria	Escherichia	Facultative	-
		Klebsiella	Facultative	-
		Enterobacter	Facultative	-
Verrucomicrobiota	Verrucomicrobiae	Akkermansia	Anaerobic	-
Mouth				
Firmicutes	Bacilli	Streptococcus	Facultative	+
		Gemella	Facultative	+
	Negativicutes	Veillonella	Anaerobic	-
Bacteroidetes	Bacteroidetes	Prevotella	Anaerobic	-
Actinobacteria	Actinobacteria	Actinomyces	Facultative	+
Proteobacteria	γ-Proteobacteria	Haemophilus	Facultative	-
Skin				
Firmicutes	Bacilli	Staphylococcus	Facultative	+
		Streptococcus	Facultative	+
Actinobacteria	Actinobacteria	Corynebacterium	Facultative	+
		Propionibacterium	Anaerobic	+
		Cutbacterium	Anaerobic	+
		Micrococcus	Facultative	+
		Brevibacterium	Facultative	+

Colon

The largest mass of commensal microbes are found in the colon, where bacteria make up over half of the fecal mass and are present at a density of 10^{11} bacteria per gram (13). The colonic contents are stagnant and permit replication, and there is abundant food supply. The adult gut (colonic) microbiota is completely dominated by obligate anaerobic bacteria (**Table 2**). Anaerobic Firmicutes such as *Faecalibacterium, Eubacterium* and *Ruminococcus* species make up around half the total abundance, and *Bacteroides* make up around a quarter. Other anaerobes such as *Bifidobacterium*, *Collinsella*, *Eggerthella*, *Lactobacillus* and *Akkermansia* are also abundant. Methane-forming archaea are the main group of archaea in the microbiota, of which *Methanobrevibacter smithii* are the most prevalent, making up between 1-10% of the gut microbiota (16, 17). Facultative anaerobes, mainly enterococci and members of the *Enterobacteriaceae* family, represent less than 1% of the total bacteria, and yeasts are detected at low abundance in around a third of people (13, 18, 19). Different habitats also exist within the colon and while many different species are found in the lumen, a more limited range reside in the mucus layer, eg. *B. fragilis*, *F. prausnitzii* and *A. mucinilphila* (20, 21). Also, gut microbiota composition varies with geography, and *Prevotella* is often observed in place of *Bacteroides* in adults from non-industrialized countries (22).

Skin

The skin is a much less hospitable environment to microbes than the mucosae, being dry, relatively cool, slightly acidic and salty. However, some skin environments are more favourable to microbes than others and microbial colonization of the skin varies depending on location, ranging from around 10^3 bacteria per square milimeter on drier areas to 10^6-10^7 bacteria per square milimeter in the armpit. The main bacterial colonizers on the skin (**Table 2**) are *Staphylococcus* (e.g. *S. epidermidis*, *S. capitis*, *S. hominis*, *S. haemolyticus* and *S. aureus*), *Streptococcus* (e.g. *S. mitis*, *S. oralis*), *Corynebacterium* (e.g. *C. tuberculostearicum*), and *Cutibacterium* (e.g. *C. acnes*), along with genera such as *Propionibacterium*, *Micrococcus* and *Brevibacterium*. Yeasts such as *Malassezia* are also often present (10, 23, 24) and *Candida* in low counts.

Development of the microbiota in infancy

The microbiota establishes on the mucous membranes and skin in early life (25). Typically, the first major encounter with the microbial world occurs at birth, when the fetal membranes are ruptured and the infant passes either through the birth canal, or through a surgical incision and into an operating theater in the case of caesarian delivery. The initial colonization involves the bacteria encountered during or immediately after birth, and is quite homogenous across body sites, i.e. the same bacterial species are detected on different mucosal sites and on the skin (26, 27). However, only some of these bacteria are able to survive, and within a few weeks, bacterial species compatible with the specific niches at different body sites have colonized the mucous membranes and skin (27). Over the next few years, bacterial colonization proceeds in a sequential manner until highly diversified site-specific microbiota are established. In the gut, the mature microbiota may contain hundreds of different bacterial species, somewhat fewer species make up the mature oral microbiota, and less species-rich bacterial communities reside in the upper airways and on the skin.

Infant microbiota at different body sites

Mouth

The newborn's mouth is often colonized within hours after birth, with one or a few bacterial species, and colonization proceeds quite rapidly (28) (**Figure 5**). Initially *Streptococcus* (*S. mitis* and *S. salivarius*) and *Gemella haemolysans* colonize the soft tissues of the oral mucosa and are detected in most neonates (28), and coagulase-negative staphylococci are also early colonizers. Within a few weeks anaerobic bacteria such as *Veillonella*, *Lactobacillus* and certain species of *Actinomyces* and *Prevotella* may be isolated from the infant's mouth, mostly from the crypts of the tongue. By 2-3 months of age, *Rothia mucilaginosa*, *Haemophilus parainfluenzae* and *Campylobacter concisus* also form part of the core oral microbiota (28). Once an infant's teeth come in, the mouth is colonized by bacteria able to live on their surface. These include facultatives such as *Streptococcus mutans*, *S. sanguinis* and *S. sobrinus*, and anaerobes such as *Prevotella nigrescens, P. pallens* and *P. intermedia*, certain *Actinomyces* species, *Leptotrichia* and *Peptostreptococcus*, among others. However, streptococci are characteristic

of the oral microbiota from infancy through to adulthood, due to their ability to metabolize sugars in the diet and also their production of surface adhesins which allow them to bind to other bacterial species and form adherent polymicrobial biofilms on teeth surfaces (27, 28).

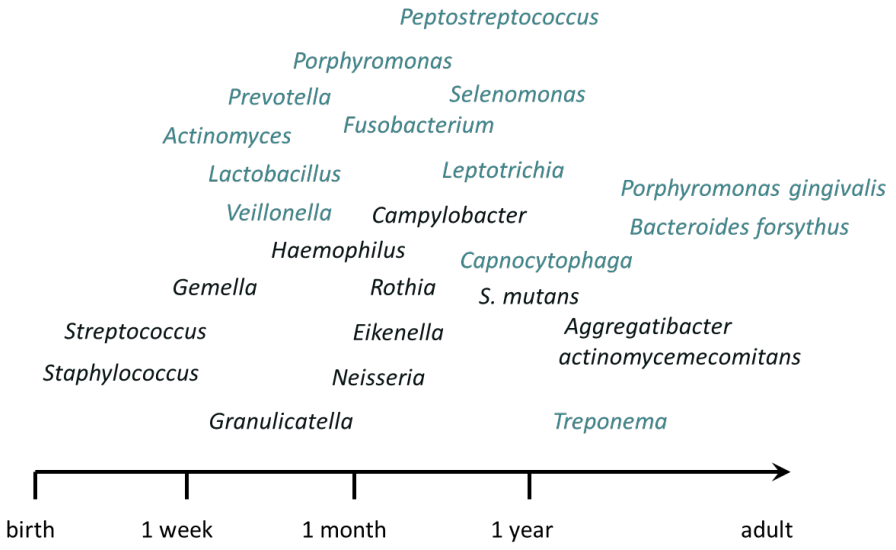

Figure 5: Development of the oral microbiota in early life. The first bacteria to colonize the infant's mouth are facultatively anaerobic bacteria (black) such streptococci and staphylococci. Then follow anaerobic bacteria (blue). Among the early anaerobes are various species of Bifidobacterium, Bacteroides and Clostridium. Anaerobes which colonize later include different groups of strictly anaerobic Firmicutes and also Akkermansia, and diversity increases up to around 3 years of age.

Gut

Normally, the infant is born sterile, unless the fetal membranes rupture prematurely allowing ascending bacteria from the perineum to colonize/infect the fetus/infant. *In-utero*, waste products from intestinal secretions, sloughed intestinal cells and ingested amniotic fluid collect in the gut and form meconium, which is excreted as the infant's first faeces within the first few days after birth. Meconium may contain bacteria when it is

excreted (27), which reflects exposure to bacteria during and immediately after delivery (29), and not exposure to bacteria *in-utero* as sometimes suggested (27).

Consistent with being a newly colonized habitat, the infant's gut microbiota is initially a much less complex ecosystem than that of an adult, with lower diversity and also greater differences between individuals (22, 30) (**Table 3**). At birth, the neonatal gut is an aerobic environment, and the first bacteria to establish are facultative anaerobes, including Gram-positives such as coagulase-negative staphylococci and enterococci, and Gram-negatives such as *Escherichia coli*, *Klebsiella* and other *Enterobacteriaceae* (13, 27, 30-32). *Staphylococcus aureus* is also a common early gut colonizer (33). Consumption of oxygen within the intestine by facultative anaerobes rapidly creates an increasingly anaerobic environment more suitable for anaerobic bacteria, which appear in succession, starting with the relatively aerotolerant ones such as bifidobacteria which are typically the dominant bacterial genus at around 3-6 months of age (34). *Bacteroides* and certain *Clostridium* species are also common in the early period, the latter because they are spore-formers and clostridial spores are found everywhere, even in hospital milieus (31).

Table 3: Differences between the infant and adult gut microbiota

	Infant	**Adult**
Diversity (richness)	Lower, increases over time	Higher
Interpersonal variation	Higher	Lower
Facultative:anaerobe ratio	Higher (1:1.5 in newborns), decreases over time	Lower (1:200)
Composition	Sequenced progression, from initial abundance of facultatives to an adult-like composition	Anaerobic Firmicutes and Bacteroidetes dominate stably over time

Over the next months levels of *Bacteroides* increase and populations of strictly anaerobic Firmicutes, especially those from the families *Lachnospiraceae* and *Ruminococcaceae*, begin to expand. During the toddler years, diversity continues to increase and levels of these groups continue to

rise. By 3 years of age they are dominant in the gut microbiota, and *Akkermansia* is also quite abundant, bringing the child's gut microbiota close to an adult composition (13, 31, 34-36) (**Figure 6**). It should be noted however that the majority of studies on gut microbiota development take place in industrialized countries and there is some variation with geography, for example with *Prevotella* often substituting for *Bacteroides* in children from non-industrialized countries (37, 38).

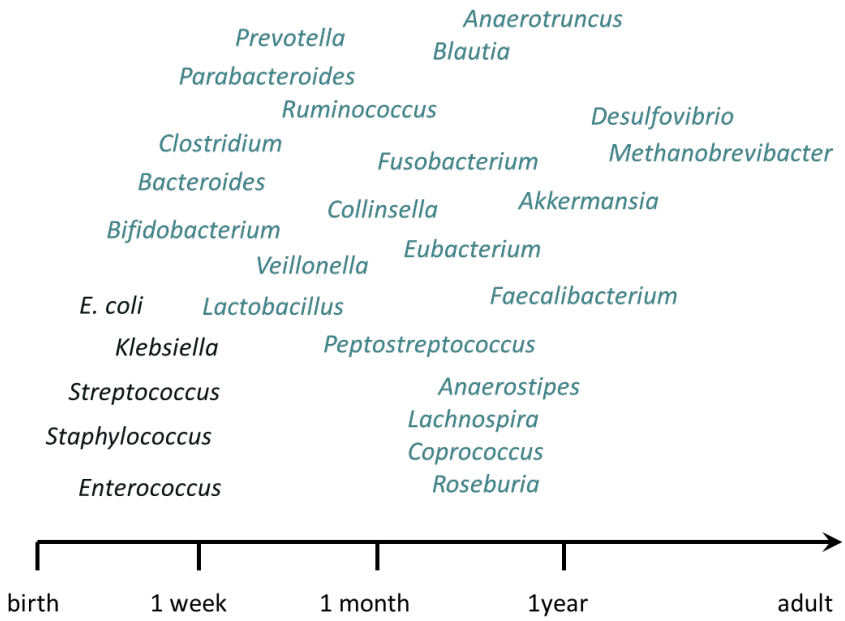

Figure 6: Development of the gut microbiota in early life. The first bacteria to colonize the infant's gut are facultatively anaerobic bacteria (shown in black) such as E. coli, enterococci and staphylococci. Then follow anaerobic bacteria (shown in blue). Among the early anaerobes are, for example, various species of Bifidobacterium, Bacteroides and Clostridium. Anaerobes which colonize later include different groups of strictly anaerobic Firmicutes and also Akkermansia, and diversity increases up to around 3 years of age.

As mentioned briefly above, the most well-characterized temporal effect in gut microbiota development is that the colonic microbiota becomes more and more anaerobic during infancy, as more and more strict anaerobes settle and change the environment in turn. The more the microbiota develops, the more

hostile it becomes to newcomers and also to facultatives, whose population sizes contract. Hence, the ratio of facultatives to anaerobes decreases from roughly 1:1.5 in newborn infants to 1:200 in adults (39), as evaluated by bacterial culture, and probably the ratio is considerably higher when also considering the many anaerobic species which cannot be cultured. Different facultatives differ in their ability to withstand the milieu created by the maturing colonic ecosystem. For example, *E. coli* is better at that than *Klebsiella* and therefore *E. coli* persists into adulthood, while *Klebsiella* mostly colonize small children (13). The gut population counts of staphylococci, which are poorly adapted to living in the colon, decrease much more and more rapidly than the population counts of *E. coli* (33).

The developmental trajectory of the gut microbiota during infancy and early childhood is often referred to as its maturation. The increasing anaerobic character, as measured by the ratio of anaerobes to facultatives, is one measure of microbiota maturity. Also, random forests machine learning algorithms of overall microbiota composition at different ages have been used to generate models of 'microbial age'. These can be applied to new gut microbiota samples to determine their 'estimated microbial age', which provides a measure of microbiota maturity in relation to a chosen reference population (40).

Skin

The skin represents an infant's largest interface with its new environment during and after delivery and harbors culturable levels of bacteria within minutes of birth (41). During the first 6 months there is considerable fluctuation in the skin microbiota, but a more stable composition establishes by 6-12 months of age (41). Up to around 6 months, staphylococci and streptococci are abundant and represent up to 40% of the skin microbiota (42). Among staphylococci, *S. epidermidis*, *S. haemolyticus* and *S. hominis* predominate, and *S. aureus* is also detected (43). During the first year, 25 different genera make up 80-85% of total bacterial abundance on average (42). *Cutibacterium acnes* is widespread (44) and micrococci such as *M. luteus* are also quite common (43). In one study, anaerobes such as *Finegoldia*, *Clostridium* and *Bacteroides* were found on the skin of the buttocks, as determined by 16S rRNA sequencing (42). However, this likely represented dead faecal bacteria. Similar to adults, most infants are colonized by yeasts,

with *Malassezia restricta* and *Candida albicans* the most common species (45, 46).

With time, there is some variation in the skin microbiota across different body locations. For example, as the skin on the infant's forehead becomes more sebaceous, a differentiated skin microbiota develops there (41). In one study, the skin microbiota of 6 weeks old infants was shown to have a quite similar composition to that of their mothers´ skin, indicating that close physical contacts contribute to skin microbiota development (27).

Sequenced microbiota development

The infant microbiota is an ecological community whose development is governed by general processes which also apply to other ecological contexts (47). From birth onwards, the infant is seeded by new taxa which arrive through various exposures in early life, and these taxa reproduce or die at different rates according to their fitness to the niche they are seeded into and their response to selection pressures such as host genetics and immune response (**Figure 7**).

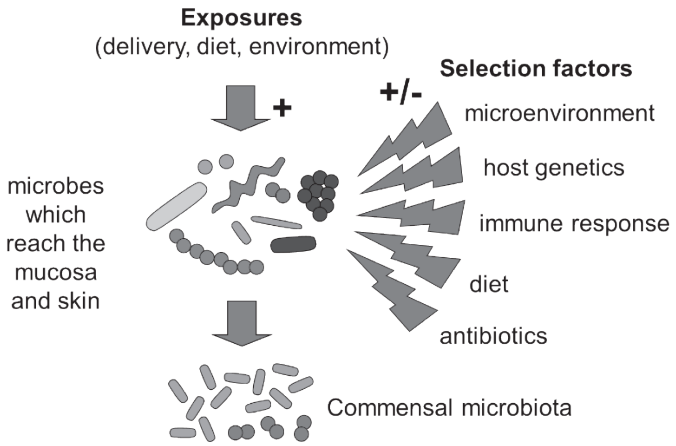

Figure 7: Microbiota development in relation to exposures and selection factors. Only some of the microbes to which the infant is exposed can establish in infant's microbiota. The local microenvironment varies considerably around the body with respect to levels of moisture, acid, salt, nutrients and so on, giving rise to different microbiota composition at different body sites. Other selection factors such as host genetics, immune response, diet and exposure to antibiotics also determine which microbes establish and thrive.

Especially in the gut, the infant microbiota develops according to an ordered sequence. This can be attributed to the fact that children themselves also undergo sequenced development, from being neonates to older infants, toddlers and young children, with different nutrient intakes and environmental exposures at each developmental stage. It also reflects the inherent nature of ecosystem development, in which early colonizers modify the environment and thereby create a habitat suitable for later colonizers. In the case of the gut microbiota, this occurs through factors such as the progressive decline in oxygen levels encouraging further growth of increasingly oxygen-sensitive obligate anaerobes, provision of nutrients for microbial cross-feeding relationships, and many others. In addition, community development is influenced by priority effects, where the order and timing with which different taxa arrive determine how they affect one another (47-49).

A detailed understanding of these phenomena may offer the potential to steer the developmental trajectory of the infant microbiota. At present, however, they remain incompletely understood. Current knowledge on the origins of microbes colonizing the infant and the selection factors determining which taxa thrive is discussed more fully below (**Figure 8**).

How does the infant get the microbiota?

Seeding of microbes at delivery

Vaginally delivered infants receive a body-wide inoculation of microbes at birth, during passage through the birth canal. Exposure may include not only vaginal, but also gut bacteria, since gut microbes may reside in the perianal area, and defecation during labour may also occur. Indeed, most maternally-transmitted species which persist in the infant's gut come from the mother's fecal flora, likely due to their capacity to survive in the gut environment. To a lesser extent, microbes from other maternal sources, such as *Veillonella parvula* resident in the mother's mouth, also persist in the infant's gut (26, 50).

Microbes acquired from the mother during a vaginal delivery and which have been shown to persist in the gut beyond the neonatal period include *E. coli*

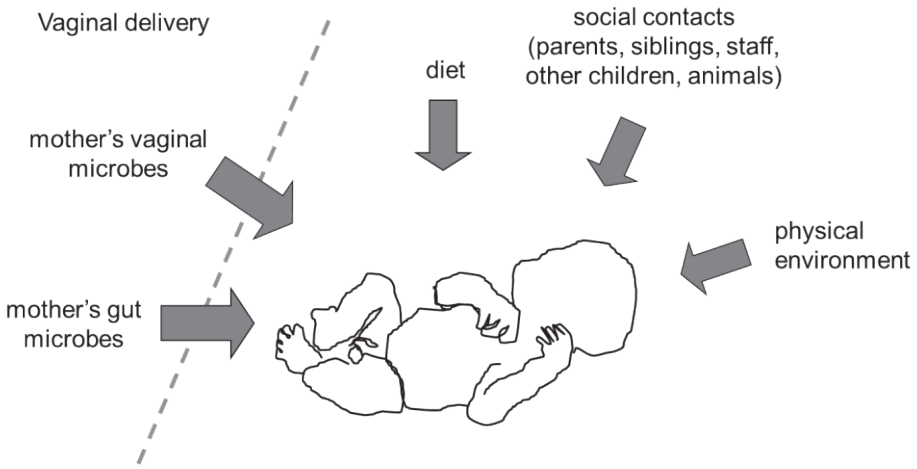

Figure 8: Factors affecting microbiota development in infants. Normally the in-utero environment is sterile, then during vaginal delivery the infant is exposed to the mother's vaginal and gut microbes. After delivery, the infant's diet, social contacts and physical environment affect their colonization by microbes.

(51) and *Bifidobacterium* species such as *B. breve* and *B. longum* subspecies *infantis*, which are able to metabolize human milk oligosaccharides. *Bacteroides* species such as *B. fragilis*, *B. dorei*, *B. thetaiotamicron*, *B. vulgatus* and *B. uniformis* are also transmitted and these are either able to metabolize human milk oligosaccharides, or are generalist feeders which are also abundant in the adult gut microbiota (50). *E. coli* strains of maternal origin are several times more likely to persist in the infant gut than other E. coli strains (52). Likewise, maternally transmitted strains of *Bifidobacterium* and *Bacteroides* show high stability in the infant gut, with 94% and 81% of strains still observable after one year respectively (53). However, bifidobacteria are likely quite easily acquired after delivery, and also from other individuals than the mother. In an early study, the bifidobacterial colonization pattern differed between different maternity wards, suggesting spread of these bacteria within the ward (54).

Infants born by caesarian section do not receive an initial inoculation of microbes from their mother's fecal and vaginal flora, and their acquisition of *Bacteroides*, *Bifidobacterium* and *E. coli* in the gut microbiota is delayed. In industrialized countries, it may take up to a year for *Bacteroides* and half a year for *E. coli* to attain the same abundance as in their vaginally-born

counterparts, while bifidobacterial colonization catches up earlier. In non-industrialized countries, infants delivered by caesarean section acquire the above bacteria almost as fast as vaginally delivered infants, reflecting higher levels of fecal bacteria circulating in the environment (13, 51-53). In the absence of initial colonization (and suppression of potential pathogens) by *Bacteroides* and *Bifidobacterium*, clostridia including *C. difficile* and *Enterobacteriaceae* such as *Klebsiella* and *Enterobacter* proliferate in the gut of caesarian-born infants (13) (**Table 4**).

Table 4: Bacterial genera which are more common in the gut microbiota of infants delivered either vaginally or by caesarian section.

Vaginal	Come from	Caesarian	Come from
E. coli	Mothers' gut microbiota, other people and animals	*Klebsiella, Enterobacter*	Other children, food, surroundings
Bacteroides	Mothers' gut microbiota, other people and animals	*Clostridium*	Surroundings, food, other people
Bifidobacterium	Mothers' gut microbiota, other people		

Diet

Consumption of breast milk or formula also influence the infant gut microbiota. Although lactobacilli in general are similarly abundant in the gut of breast- and formula-fed infants, one species, *L. rhamnosus*, seems to be favoured by breast-feeding, and staphylococci also tend to be more numerous in breastfed infants. Conversely, clostridia including *C. difficile*, *Bacteroides* (*B. fragilis*), enterococci and *Enterobacteriaceae*, especially *Klebsiella* and *Enterobacter*, tend to be more common in formula-fed infants (13, 55). Higher levels of streptococci and *Veillonella* are also observed in the latter group (30). While it is often believed that levels of bifidobacteria are higher in breast-fed infants, this is in fact not the case (13). Longer duration of breastfeeding is associated with lower gut microbiota diversity, and continued breastfeeding has a greater effect on the microbiota than introduction of solid food at around 6-12 months of age (34, 56). Probably, despite the influx of new bacteria with the new foods, the ability of the

lactoferrin and other defense factors consumed in breast milk to suppress growth of many bacteria exerts a more powerful effect (57). On cessation of breastfeeding, levels of many bacteria in the *Lachnospiraceae* family increase (56).

Social contacts with people and animals

Common environmental exposures are known to affect the microbiota within families (22, 58, 59). The skin microbiota of the parents is a source of *S. aureus* strains that persistently colonize the gut, possibly originating from the mother during breastfeeding when the infant may swallow skin bacteria from around the nipple, or from either of the parents during general care (33, 60). The role of fathers as sources for bacteria colonizing neonates has been poorly studied, but most likely, after delivery, they may be just as important sources as the mothers, if they contribute to a similar extent in the care of their infants. Members of the oral microbiota may also be transferred from parents to infants, e.g. by cleaning pacifiers with their mouths (61) and premastication of food (62) or kissing (58).

Other family members may also act as a source of colonizing microbes, and having elder siblings is associated with microbiota development in both the gut and the airways, especially during the first year. Infants with elder siblings have a more diverse gut microbiota containing higher abundance of bifidobacteria and lower abundance of *Clostridium* and *Enterobacteriaceae* other than *E. coli* in the first month, and a higher ratio of strict/facultative anaerobes, higher abundance of *Prevotella* and lower abundance of *Escherichia*, other *Enterobacteriaceae* and *Veillonella* at one year of age (31, 63). In the airways, infants with elder siblings have higher abundance of *Moraxella* and *Neisseria*, but lower abundance of staphylococci. In addition, the age gap to the closest elder sibling has a greater effect than the total number of elder siblings, possibly because of closer physical relationships and lower levels of hygiene during the early years (63).

Other people may also transfer microbes to the infant. In hospital wards, for example, enterococci, *E. coli*, *Klebsiella*, *Enterobacter* and *C. difficile* are transferred between infants by staff (13).

Day care

Once children begin day care, they become exposed to many unrelated children in a location outside of the family home for substantial periods of time. This expansion of the infant's microbial environment affects microbiota composition, with children attending day care having a more mature gut microbiota than children of the same age cared for at home, with higher diversity and species richness (64, 65).

Pets

Household pets bring their own microbes and also microbes from the surrounding environment into the family home, where they contribute to the indoor microbiota and colonize the skin of family members (59). Unsurprisingly, living with pets may also affect the gut microbiota of infants in the household (66, 67). This probably occurs through the infant's direct contact with pets, as well as through ingestion of pet-derived microbes from the surroundings and their own skin.

Physical environment

Microbes are abundant in the environment. Clostridial spores for instance are ubiquitous while *Enterobacteriaceae* other than *E. coli* and lactobacilli are commonly found in food and in natural environments. It has been shown that the microbes present in certain physical environments affect patterns of microbial colonization. For instance, higher circulating levels of enteric bacteria probably explain the fact that infants in Pakistan are colonized earlier by *E. coli* and have higher turnover of this organism in their gut than Swedish infants (52, 68). In general, infants living in less industrialized countries are also colonized earlier by other *Enterobacteriaceae*, enterococci, lactobacilli and eubacteria, although the specific source of these bacteria is not known, while coagulase-negative staphylococci, *S. aureus* and *C. difficile* are more common in infants in industrialized countries such as Sweden. The latter three bacteria proliferate poorly in a mature gut microbiota.

Farm living

Farms represent a physical environment of particular interest, as children growing up on traditional farms have low rates of allergic diseases (69, 70). Specific exposures known to protect against allergy include contact with farm

animals and consumption of unpasteurized milk (71, 72), while lower intake of margarine and vegetable oils and higher intake of full-fat dairy products in infants growing up on farms does not appear to affect allergy risk (73). Although the mechanisms by which farm-related exposures confer protection from allergies are not well understood, farms also represent rich microbial environments (72-74). Farm barns harbor high levels of microbial components such as endotoxin, as do farm houses compared to other rural or urban homes (75-77), and children living on farms are exposed to a wider range of microbes than other children (78). Consistent with the microbial exposure hypothesis, a more farm-like dust composition in non-farm homes is associated with lower risk of children developing asthma (79). Also, infants growing up on farms have a higher 'estimated microbial age' of their gut microbiota at one year of age, suggesting a higher microbiota maturity (40).

What does the microbiota do?

The human microbiota contributes to disease and also to some aspects of health (**Figure 9**). Among the negative effects is that it provides a reservoir for bacteria and fungi that may cause local or invasive infections, that it contributes to inflammation, and that it can harbour antibiotic resistant strains of bacteria and provide a niche for exchange of resistance genes (80). On the positive side, commensal bacteria may provide resistance to colonization and infection by more pathogenic bacteria, a phenomenon termed "colonization resistance". Further, the constant presence of bacteria and their products, and the turn-over of bacterial strains, provides a constant stimulus for the immune system, which may be of importance for its maturation during a critical window of development.

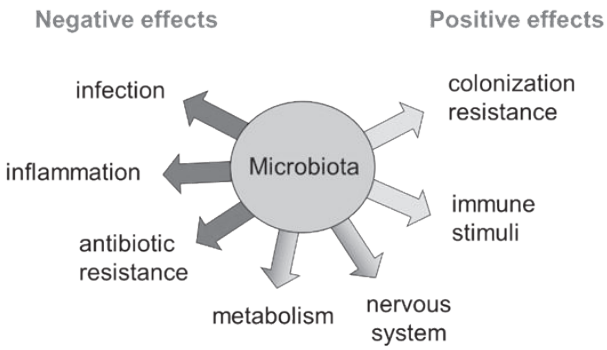

Figure 9: Effects of the microbiota on the body. The microbiota has both positive and negative effects. On the positive side, the microbiota provides resistance to colonization to pathogenic microbes and stimulation of the immune system, and on the negative side it contains microbes which may give rise to infection and inflammation, and also represents a reservoir of antibiotic-resistant strains and genes.

A reservoir for microbes causing infection

Commensal bacteria are often described as "good" and "harmless" inhabitants of the body. However, the great majority of all invasive infections are caused by bacteria from our own commensal microbiota, e.g. soft tissue infections, urinary tract infection, septicaemia, *etc...* Thus, most pathogens are also commensal bacteria that under certain conditions will cause disease. Examples are *E. coli*, *S. aureus* and pneumococci that can cause disease also in people with a well-functioning physical barriers and immune system. In addition, many commensal bacteria that are incapable of causing disease in otherwise healthy individuals, including coagulase-negative staphylococci, alpha-streptococci and bifidobacteria, can cause severe infections in people with a compromised or underdeveloped immune system, such as people on immunosuppressive drugs or premature infants who have not yet acquired full immune competence. The obligate anaerobes constituting the bulk of the gut microbiota do, however, very seldom cause infections, simply because they do not survive in the blood stream or tissues because of the too high oxygen content.

Colonization resistance against pathogens

A key function of the microbiota is helping prevent colonization by pathogenic microbes. In adults, the total surface area of the human body including the gastrointestinal tract, lungs and skin is over 100m^2, presenting a large area for potential pathogen invasion (81). The microbiota provides an additional layer of defence, which occurs by commensal bacteria (1) physically occupying available space, (2) competing for available food sources, and (3) producing toxic substances such as SCFAs and sulfur compounds to inhibit growth of pathogens. A striking illustration of this so-called colonization resistance comes from the use of fecal microbiota transplants for infection caused by overgrowth of *Clostridioides difficile* in the gut. In this procedure, administration of a healthy fecal microbiota into the patient's colon suppresses populations of *C. difficile* and cures the infection with high efficacy (82).

Regulation of the immune system

The microbiota is a potent regulator of immune system development and immune responses. This can be clearly seen in germ-free mice, who have underdeveloped lymphoid tissue with fewer Peyer's patches and lymphoid follicles in the gut wall, lower IgG and IgA concentrations in serum, fewer IgA-producing plasma cells and fewer CD4+, CD8+ and CD4+CD25+ T-cells in the gut wall (83, 84). Germfree animals have a more short-lived tolerance to ingested antigens (oral tolerance) (85-89) and their regulatory T cells have lower functional capacity than those of conventional animals (90).

Effect on metabolism

The colonic microbiota contains more than a kilogram of living cells producing different metabolites that can be of benefit or harm to the host (91). For example, the gut microbiota is the source of vitamin K (92) and germfree animals have to be supplemented with this vitamin. Also, gut bacteria break down complex polysaccharides such as dietary fiber which are indigestible to humans, forming various metabolites including short-chain fatty acids (SCFAs), which the body uses for energy and other purposes. Most polyphenols from fruits and vegetables are also relatively indigestible to humans and are likewise metabolized by gut microbes (92). The importance of the microbiota in extracting energy from our food can again be illustrated

in germ-free mice, who gain significant amounts of body fat after being colonized with a microbiota (93). It should also be noted that, as well as breaking down our undigested food, bacteria metabolize sloughed epithelial cells and mucus generated by the gastrointestinal tract. It has been estimated that 250 g of epithelial cells and about 10 L of mucus, of which 5-10% is solids, are shed daily into the lumen (72).

Besides those metabolites already mentioned, gut microbes also produce tryptophan and indole derivatives, choline metabolites, neurotransmitters, lipids and gases, and also modify the bile acids secreted by the liver via the gall bladder (91). As well as being produced directly from source substrates, metabolites may also be generated through secondary metabolic networks. For example, the archaea *Methanobrevi smithii* consumes hydrogen and CO_2 or formic acid generated by the bacterial metabolism of carbohydrates to produce methane (17).

Short chain fatty acids

What are SCFAs?

Short-chain fatty acids (SCFAs) are small molecules produced by bacteria in the gastrointestinal tract, some of which are also produced endogenously in the body. Structurally they are straight or branched-chain monocarboxylic acids with a chain length of up to six carbons, as shown in **Figure 10** below. Although not formally classified as SCFAs, succinic and lactic acid are similar small molecule bacterial metabolites. Succinic acid has two carboxy groups (see **Figure 10**), while lactic acid is similar to propionic acid except with a hydroxy group on the second carbon atom. The acid dissociation constant (pKa) of SCFAs ranges from 3.7 for formic acid to 4.9 for caproic acid. As this is significantly below physiological and intestinal pH (7.4 and 6.4-7.5 respectively (94)), SCFAs mostly exist in the dissociated form in the human body. Hence, they are often called by the name of their respective ions, e.g. formate, proprionate, succinate, valerate and so forth. SCFAs may be either straight-chain, e.g. acetic, proprionic and butyric acid, or branched, e.g. isobutyric and isovaleric acid, with the prefix ‘iso‘ denoting a branched structure.

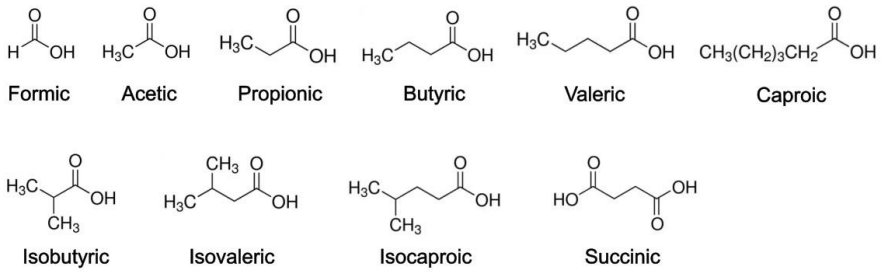

Figure 10: Chemical structure of short-chain fatty acids and succinic acid

Microbial production of SCFAs

Microbial production of SCFAs occurs in the colon through anaerobic fermentation of complex carbohydrates indigestible to humans, and also by the fermentation of intestinal mucous and sloughed epithelial cells, which contain both carbohydrate and proteins (95, 96). In infants, human milk oligosaccharides represent a source of complex carbohydrates and oligosaccharides are also added to some infant formulas (97, 98). At older ages, complex carbohydrates in the diet come from dietary fibre, with soluble fibre generating higher quantities of SCFAs than insoluble fibre (95, 99). In aerobic conditions, access to sufficient oxygen allows facultative bacteria to break down food sources to carbon dioxide and water, exactly like ourselves.

Short-chain fatty acids are produced *via* various biosynthesis pathways which are differentially represented in different bacterial groups**.** Therefore, production of the various SCFAs is dependent on both the composition of the diet and which bacteria are present in the gut (95, 100, 101). Many different bacterial groups produce acetate, while biosynthesis of the other SCFAs is more specific to certain bacteria (100, 101). Considering abundant bacterial groups in the gut (**Table 5**), bifidobacteria produce formic, acetic and also lactic acid from fermentation of complex carbohydrates (102). *Bacteroides* produce formic, acetic and propionic acid, with succinic acid generated as an intermediate in the biosynthesis of propionic acid (101, 102), and Firmicutes from the Negativicutes class such as *Veillonella* and *Dialister* also produce propionic acid (101). Clostridial bacteria such as *Faecalibacterium prausnizii*,

Eubacterium rectale, *Eubacterium hallii* and *Ruminococcus bromii* (101) are the main producers of butyric acid, and valeric and caproic acid (also known as pentanoic and hexanoic acid) are also produced by anaerobic Firmicutes (103-107). As mentioned above, other substrates besides complex carbohydrates are fermented to produce SCFAs. Intestinal mucous is metabolized by *Akkermansia mucinilphila* to produce acetic and propionic acid (108), amino acids are fermented by clostridia and peptostreptococci among various other groups to produce the branched SCFA isobutyric, isovaleric and isocaproic acid (96, 109) and fermentation of lysine by an *Intestinimonas* strain can produce butyric acid (110).

Table 5: Major SCFA fermentation products of common gut microbes.

Phylum	Genus	SCFA used	SCFA produced
Actinobacteria	*Bifidobacteria*		Formic, acetic, (lactic)
	Collinsella		Acetic, (lactic)
Bacteroidetes	*Bacteroides*		Formic, acetic, propionic, (succinic)
	Prevotella		Acetic, propionic, (succinic)
Firmicutes	*Eubacterium*	Acetic, (lactic)	Formic, butyric, (lactic)
	Roseburia	Acetic	Formic, butyric, (lactic)
	Anaerostipes	Acetic, (lactic)	Butyric
	Coprococcus	Acetic, (lactic)	Butyric
	Blautia		Acetic, (succinic)
	Faecalibacterium	Acetic	Formic, butyric, (lactic)
	Ruminococcus		Acetic, butyric
	Veillonella		Propionic
Verrucomicrobiota	*Akkermansia*		Acetic, propionic

“Cross-feeding”

Furthermore, production of some SCFAs involves cross-feeding pathways which consume SCFAs already generated by other bacteria in the gut. One biosynthesis route to produce butyric acid uses acetic acid as a substrate, while another uses lactic acid. These relationships help explain the successional colonization of the infant gut by acetic and lactic acid-generating bifidobacteria, followed by butyric acid-producing bacteria from clostridial

clusters IV and XIV (111-113). Valeric and caproic acid are generated by addition of a two-carbon unit from acetyl-CoA to propionic acid and butyric acid respectively (105, 106). Of note, isobutyric, valeric and/or isobutyric acid are required for growth of bacteria including *Bacteroides succinogenes*, *Ruminococcus flavefaciens* and *R. albus* (114), while acetic, propionic and butyric acid (but not formic acid) promote growth of *Bifidobacterium adolescentis* (115).

Dietary SCFAs

As well as being produced by the gut microbiota, SCFAs may be consumed in certain foods. For example, dairy products such cow's milk and butter contain SCFAs (116), while vinegar is dilute acetic acid and may be consumed directly or in various pickled and preserved foodstuffs. Ingested SCFAs are absorbed in the upper gastrointestinal tract. Also, alcohol is metabolized into acetaldehyde which is then converted into acetate.

SCFAs around the body

The highest SCFA levels in the body are found in the colon. Here, concentrations are ≥100mmol/kg in the caecum declining somewhat to reach 80 mmol/kg in the descending colon. SCFA concentration appears to be a main driver of colonic acidity, with pH correspondingly rising from 5.7 in the ascending colon to 6.6 in the descending colon. Not considering formic acid, which is not typically included in studies of SCFAs, acetic acid is the most abundant SCFA, followed by propionic and butyric acid in approximately equal amounts. Small quantities of longer and branched SCFA are also present, as are small quantities of succinic and lactic acid. SCFA concentrations in the rectal area are similar to those in the mid-colon and also those in feces, supporting the measurement of SCFA levels in feces as a proxy for colonic levels. In contrast, the SCFA concentration in the small intestine is much lower at <1 mmol/kg in the jejeunum, rising to ~10 mmol/kg in the ileum (117, 118).

SCFA are taken up from the colonic content primarily by active transport, via the monocarboxylate transporter 1 (MCT-1) or by a lesser extent via the sodium-coupled monocarboxylate transporter 1 (SMCT-1). They may also be taken up passively by epithelial cells (95).

After absorption from the gastrointestinal tract SCFAs pass rapidly through the body. In rodents, they reach the liver, kidneys, heart and brain within 5 minutes (119) and the skin within 45 minutes (120). However, the majority of absorbed SCFAs are used by the colonic epithelial cells for energy (especially butyric acid) and by the liver for production of glucose (propionic acid). Another smaller portion are taken up by peripheral tissues, as shown by a progressive drop in acetic acid concentration between the hepatic vein, forearm arterial blood and forearm venous blood (113, 117). In peripheral blood, SCFA concentrations are around 80 µmol/L, most of which is acetic acid (117, 121). (Note that the colonic concentrations are in millimole/L, which is approximately 1000 times higher than in the serum). In lactating mothers, short-chain fatty acids are also found in breast milk at concentrations of 190 µmol/L, with butyric acid the most abundant (122).

Endogenous SCFAs

Short-chain fatty acids are also endogenous to the human body, i.e. produced by ourselves. Formic acid derives from the endogenous metabolism of serine and other nutrients (123, 124), and succinic acid is a key intermediate in the tricarboxylic cycle (125), with blood levels increased by physical exercise (126). Acetyl coenzyme A (acetyl-CoA) is a central molecule in metabolism, and acetate is taken up from circulation and released through the actions of acetyl-CoA synthetase and hydrolase (127-130). Endogenous SCFAs therefore also contribute to circulating levels of SCFAs.

Functions of SCFAs

Short-chain fatty acids exert a multitude of effects on many body systems (131). As our interest is the potential of SCFAs to prevent allergic diseases, the effects of SCFAs on the immune system are discussed here. The immune system is made up of three lines of defence, consisting of physical barriers, an innate response and a specific adaptive response, and the role of SCFAs in each line of defence is outlined below.

Effect on barriers

With respect to physical barriers, in the gastrointestinal tract SCFAs promote epithelial tissue integrity and barrier function by providing energy for epithelial cells, regulating cell proliferation and differentiation, upregulating

various structural and tight junction proteins, and generating correct intestinal morphology. They also increase the production of mucus (132-136). For example, in a study where intestinal tissues of mice were damaged by radiation, oral administration of valeric acid was able to largely reverse this damage. Part of this reversal was due to upregulation by valeric acid of the structural protein keratin, which plays a significant role in epithelial integrity in the intestine (132). Also, in a mouse model of atopic dermatitis, orally administered butyric acid reached the skin within 45 minutes of ingestion (120) leading to reduced breach of the skin barrier, followed by less allergen ingress, sensitization and skin inflammation. This was likely due to the ability of SCFAs to upregulate differentiation of keratinocytes and production of several structural proteins (eg. filaggrin) and lipids (eg. ceramides), and increasing the thickness of the stratum corneum, the outer protective layer of the epidermis (120, 137, 138).

Effect on other bacteria

SCFA produced by anaerobic bacteria may contribute to the suppressed growth of facultative organisms such as *Staphylococcus aureus* and members of the *Enterobacteriaceae* family in the gut microbiota, while permitting the growth of commensal anaerobes such as bifidobacteria and lactobacilli (139, 140).

Effect on inflammation

Microbes that succeed in entering across the mucosal barrier are swiftly eliminated by macrophages and other cellular and soluble component of the innate immune system, collectively causing an inflammatory response. SCFAs regulate many cells of the innate immune system including neutrophils, macrophages, natural killer cells, eosinophils, basophils, and innate lymphocyte subsets (ILCs) (141, 142). G-protein-coupled receptors (GPR) are transmembrane receptors, and there are many types of these receptors expressed on many different cell types. In the case of innate immune cells (**Table 6**), GPR43 is expressed on neutrophil granulocytes, and its activation can trigger chemotactic recruitment to an inflammatory site. Also, GPR109a is expressed on colonic macrophages and dendritic cells, as well as colonic epithelial cells, and its activation promotes differentiation of regulatory T cells (Tregs) (143).

Table 6: Effects of SCFAs on innate immune cells via G-protein-coupled receptors

Cell type	Effect	Receptor	SCFA
Neutrophil granulocytes	Activate chemotactic recruitment to inflammatory site	GPR43	Propionic = butyric >> acetic
Colonic macrophages and dendritic cells (also colonic epithelial cells)	Induce production of Tregs	GPR109a	Butyric

Effects on the adaptive immune system

The third line of defence involves a specific response by T-cells and B-cells, and SCFAs also regulate these functions (141, 142). Notably, SCFAs act on macrophages and dendritic cells to promote maturation of naïve T cells into of FoxP3+ regulatory T-cells (Tregs) (141, 144-146). For example, oral administration of butyric (but not valeric) acid increases the number of mucosal Tregs in mice (147). Furthermore, valeric acid increases IL-10 production and reduces IL-17 expression in Th17 cells and inhibits generation of Th17 cells in the small intestine of mice (147). Some experiments have shown that SCFAs also regulate B-cell responses (142), for example by stimulating production of IL-10 by regulatory B-cells (141, 147), reducing serum IgE levels (148) and inhibiting IgE-mediated and also non-IgE-mediated mast cell degranulation.

Mechanistically, SCFAs act (i) *via* epigenetic pathways, as histone deacetylase inhibitors and substrates for histone acetylase transferases, (ii) through signaling as ligands for G protein-coupled receptors expressed on many body cells and tissues, and (iii) as modulators of cellular metabolism, providing acetyl-CoA units to fuel the Krebs cycle and for fatty acid synthesis, and boosting glycolysis (95, 111, 138, 147, 149). Overall, SCFAs act on multiple aspects of the immune system to promote an anti-inflammatory, tolerogenic phenotype.

Allergy

Allergy was once a very rare disease but now affects almost everyone, at least in countries with an affluent Western lifestyle and good sanitary conditions. One hypothesis to explain this is that changes in the normal microbiota may play a role in the "allergy epidemic". This thesis examines some aspects of the normal microbiota and its links to the risk of developing allergy or, conversely, staying healthy.

What is allergy?

Allergy means immune-mediated hypersensitivity, i.e. the allergic individual has an immune reaction to harmless substances in the environment that normally evoke no reaction, which results in symptoms in the individual.

Several different immune reactions may result in allergy. In atopic, or IgE-mediated allergy, the symptoms are caused by immune reactions mediated by IgE antibodies. The allergic individual produces IgE antibodies to common proteins in the environment, such as those found in pollens, foods, insect venoms or animal dander. Production of IgE antibodies to a harmless protein is termed *sensitization* and can be measured either by detecting specific IgE antibodies in the blood, or by a skin prick test, where the suspected allergen is injected into the skin. If the person has IgE antibodies to the allergen which are bound to mast cells in the dermis, the allergen will bind to these antibodies resulting in activation of the mast cell and release mediators that cause leakage of the blood vessels, resulting in a small blister. Sensitization is a prerequisite for the allergic reaction, but many individuals are sensitized without being allergic and a person may be sensitized to many different allergens, but only have allergic symptoms when encountering one or two of them. Naturally, a person may also have a negative test for sensitization but still be allergic, e.g. if antibodies are produced more locally in the gut and not being present on skin mast cells, etc…

IgE-mediated allergic symptoms can affect many organs (skin, airways, gastro-intestinal tract) with symptoms such as rash, urticaria, swelling of the tissues in the oral region and neck, a congested and runny nose ("hay fever") usually combined with running eyes, wheezing and difficulty to express the air from the airways (asthma), vomiting and diarrhea and, in the worst cases, a

systemic reaction termed anaphylaxis, characterized by low blood pressure and systemic vasodilatation that may be fatal.

The term atopy was coined in the 1920s long before IgE was discovered in the 1960s and was based on the observation that certain diseases – atopic diseases – tended to occur in the same individual or among members of the same family. Today, atopy is often used as a synonym to IgE-mediated allergies, including asthma and allergic rhino-conjunctivitis (hay fever). Atopic dermatitis (or atopic eczema) also counts among the atopic diseases, despite the fact that specific IgE antibodies to known allergens are usually not detected, although high levels of total IgE in the blood serum, without known specificity, is a common finding. Food allergies are immune mediated hypersensitivity to foods, that may be mediated by IgE or other less well-defined immune reactions (T cells and/or IgG). IgE-mediated food allergy usually produces symptoms of rapid onset, while non-IgE mediated food allergy often has a slower course with symptoms appearing over days. IgE-mediated food allergy is connected to the other atopic diseases, such that IgE-mediated food allergy to cow's milk may occur in the first years and then disappear, but then be replaced with another atopic disease, e.g. asthma or hay fever.

The phenomenon that allergic diseases may replace one another as the child ages, from atopic dermatitis and food allergy in infancy to asthma and allergic rhino-conjunctivitis later in childhood is termed "the atopic march" (150). Although the phenomenon is well known, its background is not known.

In this thesis, the terms atopy/atopic disease and IgE-mediated allergy will be used as synonyms. If not specified, "allergy" refers to IgE-mediated allergy.

The IgE-mediated allergic reaction

Allergic reactions are triggered when an allergen, for example a food or pollen protein, is absorbed intact across a mucosa and reacts with IgE antibodies specific for this protein on the surface of a mast cell situated in the mucosa. Binding of the allergen to at least two mast cell-bound IgE antibodies leads to release of the granules containing histamine, proteases and glycosaminoglycans such as heparin. Histamine causes vasodilatation and increased blood vessel permeability. Cross-linking of IgE by the allergen also induces production of mast cell leukotrienes that have the same function as

histamine but with prolonged action (151). Lastly, prostaglandin D2, unique to mast cells and cytokines from mast cells, macrophages and T cells, together recruit eosinophilic granulocytes and Th2 cells from the blood vessels into the tissue. In the airway and intestines, increased production and release of mucous occurs and the bronchial musculature contracts.

Sensitization

Sensitization means producing an immune response that may lead to allergy. In atopic allergy, this means production of IgE antibodies to one or several allergens, which are in almost all cases proteins. This takes place in lymph nodes to which the allergen is carried via the lymph and/or in antigen-presenting cells. Naïve T-cells and B cells with specificity for the allergen become activated and expand clonally. The Th2 cytokines IL-4 and IL-13 are of key importance to promote maturation of the naïve B-cell into an IgE-producing plasma cell. IgE reaches the mast cells *via* the blood and binds with high affinity to the mast cell surface via Fce receptors, where they are ready to react with their specific allergen if it enters into the mucosa. As mentioned previously, many people are sensitized to a particular antigen but do not experience allergic symptoms when they encounter the allergen and therefore do not have an allergic disease.

The epidemiology of allergy

Allergic diseases affect over a billion people worldwide (152) and as they typically manifest early in life, children are particularly affected. According to the latest International Study of Asthma and Allergies in Childhood (ISAAC) survey, 9% of children age 6-7 globally have eczema, 12% have asthma and 9% have rhino-conjunctivitis, while at 13-14 years of age these figures are 7%, 14% and 15% respectively (153). Patients with allergic diseases have a reduced quality of life, and asthma alone causes 250,000 deaths annually (152). In addition, there is a considerable economic cost from utilization of healthcare services as well as lost school and work days (152, 154).

The hygiene hypothesis

The dramatic rise in rates of allergic disease over a relatively short period of time cannot be explained by genetic factors alone and must therefore be related to factors in the environment which have changed over the same

period. The British epidemiologist David Strachan put forward the hygiene hypothesis in 1989, based on the observation that allergy was not only more common in children of affluent families, but also in only children or first-born children rather than children with many elder siblings. Strachan proposed that common to large and poor families was the constant spread of childhood infections. He also proposed that in some manner, these early childhood infections would educate the immune system in a manner that allergies would be avoided in the future (155, 156). The next important piece of evidence was produced by the Italian immunologist Paolo Matricardi who investigated serological evidence of childhood infections in Italian recruits and related this to whether they had hay fever or not. He found that antibodies against hepatitis A, a virus spread via the fecal-oral route and therefore a marker for lower hygienic standards, was associated with reduced rate of hay fever and asthma. Later on, he showed that also antibodies to *Toxoplasma gondii* and *Helicobacter pylori*, a parasite and a bacterium also spread under unsanitary conditions, were both linked to reduced risk of having hay fever, while antibodies to viruses transmitted through other routes were unrelated to protection against allergy (157-159). He concluded that exposure to microbes via the oral route was more protective than respiratory diseases in shaping the immune system in a manner that would prevent allergy.

Based on the findings that Swedish infants were much later colonized by common gut microbes than Pakistani infants (160) and also had a much slower turn-over of *E. coli* strains in the gut microbiota (52, 68), Wold proposed that a meager gut microbiota could be the cause of the allergy epidemic of Western countries (161). It is still not known which type of microbial stimulation that is needed to shape the infant's immune system in a manner preventing it to react excessively to harmless substances in the environment.

Oral tolerance

The reason why some people become sensitized and allergic to perfectly harmless proteins to which we are all exposed, most of us without becoming sensitized or having symptoms, is still not clear. The normal response to proteins that are ingested or inhaled is tolerance, termed oral (or mucosal tolerance) (**Figure 11**). This reaction was described already in 1911 by Wells

who found that guinea pigs could not be made hypersensitive and anaphylactic to a protein, if they had been fed on this protein as pups (162). Oral tolerance was demonstrated in humans in 1986 (163). Oral tolerance is upheld *via* antigen-specific regulatory T cells (T regs). They, in turn, are induced from naïve T cells by encounter with a special DC subset, the CD103+ dendritic cell. Strong immune activation by feeding superantigen, a bacterial T-cell-activating toxin, during early infancy leads to improved capacity to become orally tolerant to a new protein (164) and this is related to an enhanced functional activity of the CD103+ dendritic cells and an increased density of regulatory T cells in the gut lamina propria (165).

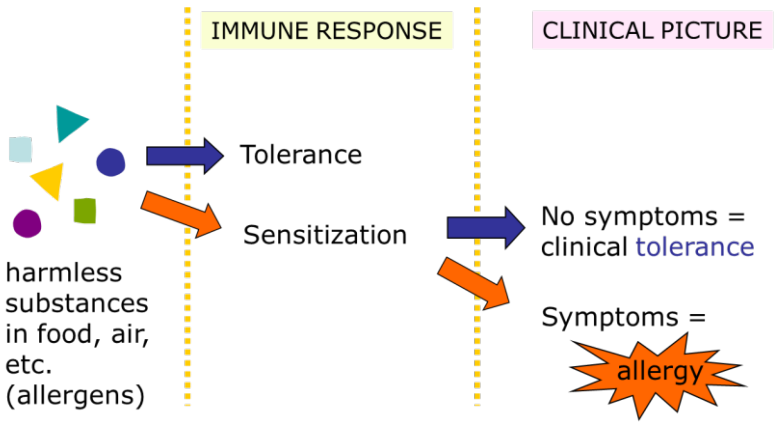

Figure 11: Tolerance, sensitization and allergy. Most people develop immune tolerance to the many harmless substances in our food and surroundings. In some people, however, exposure to these substances leads to sensitization. Among these people, some do not develop any symptoms and are described as clinically tolerant. Others show symptoms when encountering allergen(s) and these people are allergic.

One hypothesis, put forward by Gideon Lack, is that delayed introduction of foods into the baby's diet deprives the infant of the chance of developing oral tolerance and instead, the allergen may enter *via* the skin or airways, routes that are much more likely to lead to sensitization rather than allergy. Atopic eczema is strongly linked to high risk of a continued atopic march with sensitization to more and more allergens and symptoms from other organs.

This could be related to the skin barrier being breached by the eczematous reaction, enabling allergens to enter *via* the skin (120, 166).

Although allergy is by far the most common disease of young and middle-aged people in Western countries, we know very little regarding what causes more and more people to elicit immune reactions to things that our immune system should actively ignore. It is reasonably clear that a paucity of stimulation of the developing immune system by microbes plays a central role, but which microbes, where and when, is not known. In this thesis, we focus on a particular aspect of the microbial environment – the commensal gut microbiota and the short chain fatty acids that it produces.

Preterm infants and their microbiota

Preterm birth is defined as birth of an infant before 37 weeks of pregnancy, as opposed to full-term delivery at approximately 40 weeks. Infants are classified according to their degree of prematurity, as moderately preterm (32-36 weeks), very preterm (28-32 weeks) and extremely preterm (<28 weeks). With medical advances the limit of viability has become earlier over time, and currently the most preterm infant to survive was born at 21 weeks and one day of gestation, weighing just 420 grams. Here we are most concerned with extremely preterm infants, who represent the most vulnerable group.

Causes of preterm birth

The majority of preterm births follow either spontaneous preterm premature rupture of membranes or spontaneous preterm labour, with the remainder (approximately 30-35%) occurring when early delivery is required due to maternal or fetal indications (167). Intrauterine infection is a frequent mechanism leading to spontaneous preterm birth, while pre-eclampsia/eclampsia and intrauterine growth restriction are common reasons for indicated preterm delivery (167, 168). There are a multitude of risk factors for preterm birth, including pregnancy with twins or multiples, conception *via* assisted reproductive technologies, various maternal demographic characteristics such as ethnicity and socioeconomic status, nutrition, smoking, stress, periodontal disease and many others (167, 168).

Immaturity of the immune system and barriers

As mentioned above, immune defence against both pathogens and commensal microbes is composed of three layers, (i) anatomical, physical and chemical barriers to penetration by microbes, (ii) the innate immune system which kills microbes but also causes inflammation, and (iii) the acquired immune system. Newborn infants, especially those born prematurely, have a thin skin and low stomach acid. In the small intestine, Paneth cells, which produce large quantities of antimicrobial peptides to kill enteric pathogens, do not increase in numbers and become competent until 29 weeks of corrected gestational age. Regarding acquired immunity, the newborn infant is born without immunological memory. However, the fetus is supplied with maternal IgG through a transport system that starts in mid-gestation and a full-term baby actually has a higher IgG concentration in the blood than the mother. However, premature infants are born before having been filled up with IgG and the more premature, the lower their serum IgG level (**Figure 12**). This is considered the major reason for the extreme vulnerability to infection of the premature infant.

Human breast milk contains very little IgG, but massive amounts of secretory IgA (SIgA). SIgA is not taken up from the gut but protects the mucosa by hindering attachment of microbes and their toxins to the epithelium and thereby drastically reducing translocation. As it was shown that bottle-fed Swedish babies had substantially higher risk of septicaemia than breast-fed infants (169), all Swedish infants born prematurely receive breast milk, either their mother's own milk or banked donor milk.

Morbidity, mortality and diseases of prematurity

Complications from preterm birth are the leading cause of death in children under five years of age worldwide and account for 18% of the 5.3 million deaths in under-fives each year (170). Rates of morbidity and mortality are highest in the most preterm infants. In Sweden, where survival of preterm infants is higher than in many other industrialized countries, the survival rate of live-born infants at 22-26 weeks of gestation is 77% (based on data from 2014-16), and just 30% in infants born at 22 weeks (171, 172). Around half of surviving infants born at 22-26 weeks of gestation experience major neonatal morbidities such as necrotizing enterocolitis, severe bronchopulmonary

dysplasia or retinopathy of prematurity (172). Late-onset sepsis is another common cause of morbidity and mortality in preterm infants (173).

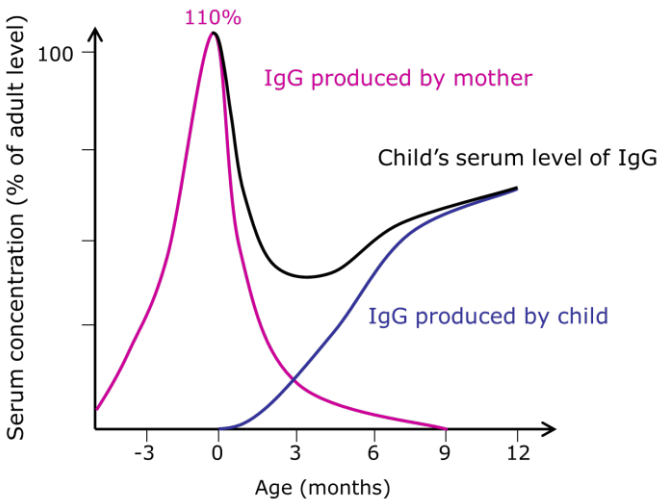

Figure 12: Blood IgG concentration in pregnancy and infancy. From around 20 weeks of gestation IgG is transferred from mother to child via the placenta, until at full-term birth the infant's IgG levels are higher than their mother's. From birth onwards, the infant's own production of IgG antibodies towards the antigens it encounters increases over time, while levels of IgG transferred from the mother decline.

Care of extremely preterm infants

Preterm infants undergo intensive medical intervention and extremely preterm infants spend at least some time in the neonatal intensive care unit. Pregnant mothers at risk of preterm birth before 34 weeks of gestation are given steroids where possible to help the infant's immature lungs develop more quickly. Even so, many preterm infants experience breathing difficulties and need respiratory support in the form of ventilation, continuous positive airway pressure (CPAP) and/or supplemental oxygen. As intrauterine infections are a common immediate cause of preterm birth, and also because preterm premature rupture of membranes can lead to infection, infants are often exposed *in-utero* to antibiotics administered to their mothers. After birth, it is also common practice to administer antibiotic prophylaxis to infants for the initial period to combat risk of sepsis, as well as further antibiotics in the case of infections or surgery, and also antifungals. The skin of preterm

infants is delicate and not fully developed, leading to trans-epidermal water loss, and the infants are unable to carry out effective thermoregulation, so are housed in incubators with controlled environmental conditions (174).

Initially, extremely preterm infants receive nutrition directly into their blood stream (parenteral nutrition) and then transition to receiving milk *via* a feeding tube (enteral nutrition). Depending on the country and hospital, enteral nutrition may consist of human breastmilk expressed by the infant's own mother and supplemented when required by milk from donor mothers, or infant formula. In Sweden, all preterm infants receive either their own mother's milk or banked milk from donors. Since the HIV epidemic, donor milk is pasteurized. Fortifiers, typically derived from cow's milk, are usually added. Once the infant is sufficiently developmentally advanced they may transition to direct breastfeeding, typically at 34 weeks of postmenstrual age. Again depending on the facility, infants may also receive kangaroo care, where they spend prolonged periods outside of their incubator swaddled on the chest of their parents, which has many benefits including improved physiological stability and relief of pain and stress (175).

The commensal microbiota in preterm infants

The gut microbiota of preterm infants shows low diversity and varies considerably between individuals. Facultatively anaerobic Firmicutes such as staphylococci and enterococci dominate initially, shifting to a dominance of proteobacteria and especially *Enterobacteriaceae* over time. Especially in the extremely preterm gut, obligate anaerobic bacteria such as bifidobacteria, *Bacteroides* and *Clostridia* which are abundant in the early gut microbiota of term infants are scarce (48, 176, 177). Key taxa often present in the microbiota of the preterm infant gut are shown in **Table 6**.

The oral microbiota of very preterm and extremely preterm infants is also characterized by facultative anaerobic Firmicutes, eg. *Staphylococcus epidermidis*, *Staphylococcs aureus* and streptococci, and proteobacteria eg. *Pseudomonas aerouginosa*, *Stenotrophomonas, Neisseria, Citrobacter koseri*, *E. coli* and *Halomonadaceae* (176, 178-181). Methicillin-resistant *S. aureus* (MRSA) can be common, and in one hospital colonized the mouth of around half of infants by 6 weeks postnatal age (181). However, MRSA is uncommon in the Nordic countries, due to energetic surveillance. Increased levels of

streptococci are a signature of maturation of the preterm oral microbiota, becoming predominant after 32 weeks of corrected gestational age (178).

Table 6: Key bacterial taxa in the preterm infant gut. Adapted from Clinics in Perinatology, 44 /2, Underwood MA and Sohn K, The Microbiota of the Extremely Preterm Infant, 407-27, Copyright (2017), with permission from Elsevier (170).

Phylum	Class	Genus	Facultative/ anaerobe	Gram
Firmicutes	Bacilli	Staphylococcus	Facultative	+
		Streptococcus	Facultative	+
		Enterococcus	Facultative	+
		Lactobacillus	Anaerobic	+
	Clostridia	Clostridium	Anaerobic	+
	Negativicutes	Veillonella	Anaerobic	-
	Mollicutes	Ureaplasma	Facultative	-
Proteobacteria	γ-Proteobacteria	Klebsiella	Facultative	-
		Escherichia	Facultative	-
		Proteus	Facultative	-
		Serratia	Facultative	-
		Enterobacter	Facultative	-
		Cronobacter	Facultative	-
		Pseudomonas	Facultative	-
		Acinetobacter	Facultative	-
Bacteroidetes	Bacteroidetes	Bacteroides	Anaerobic	-
Actinobacteria	Actinobacteria	Bifidobacterium	Anaerobic	+
		Propionibacterium	Anaerobic	+

The skin microbiota of preterm infants is dominated by coagulase-negative staphylococci, with species such as *S. epidermidis*, *S. capitis*, *S. hemolyticus* and *S. hominis* detected (49, 174, 182). After these Firmicutes, Actinobacteria (eg. *Corynebacterium*), Proteobacteria (eg. *E. coli*, *Enterobacter*) and Bacteroidetes (eg. *Prevotella*) are the most abundant phyla (174, 183). When compared to term infants, the skin microbiota of very preterm and extremely preterm infants in the neonatal intensive care unit (NICU) is less diverse and also appears to be relatively homogeneous across the body (183), which might be explained by the controlled temperature and humidity within incubators creating a consistent whole-body habitat for microbes. In preterm

as well as term infants, *Malassezia restricta* and *Candida albicans* are the most common fungi colonizing the skin (45).

Origin of the microbiota in preterm infants

Physiological immaturity in combination with a hospital environment, and medical practices including extensive treatment with antibiotics, seem to be key drivers of the altered microbiota in preterm infants. The hospital environment harbors many potential pathogens, and organisms such as staphylococci, enterococci, streptococci, *Escherichia*, *Enterobacter*, *Pseudomonas* and *Candida albicans* are found on NICU surfaces despite routine or even intensive cleaning regimes (176). Microbes found on NICU surfaces and on the hands of healthcare providers and parents resemble those found in the gut of premature infants (184), and the transfer of environmental microbes to infants is illustrated by bacteria on incubator surfaces, ventilation equipment and sink drains causing infectious outbreaks within NICUs (176). Indwelling feeding tubes colonized by *Enterobacteriaceae*, staphylococci, streptococci and other bacteria may also seed microbes to the preterm infant (185). As well as the gut, the preterm skin microbiota is also affected by the physical environment, as shown by the fact it can vary even between different NICU units on the same hospital campus (186).

Antibiotic treatment of term infants is associated with lower abundance of obligate anaerobes such as *Bacteroides* and bifidobacteria and higher abundance of proteobacteria in the gut microbiota, and frequent antibiotic treatments may contribute to the similar colonization patterns observed in the gut microbiota of preterm infants (177, 187). In addition, antibiotic treatment of preterm infants may be associated with decreased microbiota diversity and increased levels of streptococci and enterococci in the gut microbiota, which is also characteristic of the preterm infants, although the effects may vary considerably depending on the class of antibiotic and treatment regimen (187, 188). Effects of antibiotic treatment on the oral and skin microbiota of preterm infants has also been observed (183, 189, 190). Other medical practices besides antibiotic treatment may also affect the preterm microbiota. Antacid medications, which are commonly given for treatment of suspected gastroesophageal reflux, lower stomach acidity and can lead to bacterial overgrowth and increased colonization by *Candida* (191).

Respiratory support may induce an aerobic intestinal environment and contribute to the lack of gut colonization by obligate anaerobes (192). Also, kangaroo care is associated with a skin microbiota containing higher levels of typical commensal coagulase-negative staphylococci and lower levels of *Escherichia* (193), and also with an altered oral microbiota (178).

Delivery mode seems to have a less pronounced effect on the gut microbiota in preterm compared to term infants (188). This might be explained by the fact that caesarean delivery reduces exposure to the mother's microbiota and thus delays acquisition of anaerobic bacteria in the gut of term infants, whereas anaerobes are scarce in general in the most preterm infants, independent of delivery mode. In addition, no effect of delivery mode on the skin microbiota of preterm infants was noted in a study (183).

Diseases of prematurity and the microbiota

Extremely preterm infants are born before body systems such as the gastrointestinal tract, lungs and eyes and also immune system are fully developed. Even with the best hospital care after preterm birth, so far it is not possible to fully replicate the conditions the infant would experience *in-utero* during a normal term pregnancy, leaving the infant vulnerable to the major neonatal morbidities mentioned above (page 43). These diseases of prematurity may be related at least in part to the altered microbiota of preterm infants.

Necrotizing enterocolitis

Necrotizing enterocolitis (NEC) is a devastating inflammatory gastrointestinal condition which affects 5-10% of preterm infants born at <32 weeks of gestation and has a mortality rate of up to 30%. It is characterized by severe intestinal inflammation, gas within the intestinal wall and necrosis (194, 195). Although the exact pathogenesis of NEC remains unclear, it is believed to arise from the immature physiology of the preterm intestine and immune system combined with an inflammatogenic gut microbiota.

As well as the physiological and immune defects described above, toll-like receptor 4 (TLR-4), which recognizes lipopolysaccharide in the outer membrane of Gram-negative bacteria, is over-expressed on intestinal epithelial cells in the preterm gut (176, 177, 195). Studies of germ-free mice

have shown that a gut microbiota is required in order to develop NEC, while a combination of Paneth cell ablation and inoculation with *Klebsiella pneumoniae* in mouse models induces NEC-like injury (195, 196). In infants, NEC is often preceded by an abundance of proteobacteria in the gut, whereas high gut microbiota diversity and abundance of bifidobacteria are associated with protection from NEC, and certain probiotics are indicated for NEC prevention (176, 197). It may be that low levels of anaerobes in the preterm gut allow proliferation of facultatives including proteobacteria (198), which infiltrate the intestinal wall and trigger an intense TLR-4-mediated inflammatory response, causing NEC. Frequent administration of antacid medications to preterm infants for suspected gastroesophagal reflux may also contribute to the development of NEC (and sepsis), as these medications reduce the production of stomach acid which provides a natural defense against the proliferation of facultatives (191).

Late-onset sepsis

Neonatal sepsis, defined as a blood infection in the newborn, is a major cause of mortality and morbidty in preterm infants, including poor neurodevelopmental outcomes. Sepsis is divided into two groups, early- and late-onset sepsis. Early onset sepsis occurs within 72 hours of delivery, the causal organism being acquired from the mother before or during delivery, for example through exposure to chorioamnionitis or by transmission of group B streptococci from the vaginal microbiota. Late-onset sepsis occurs from 72 hours after delivery and the majority of episodes are caused by coagulase-negative staphylococci, although various other Gram-positive and Gram-negative bacteria such as enterococci, *S. aureus*, *E. coli* and *Klebsiella* may also cause late-onset sepsis. Transfer of coagulase-negative staphylococci or other organisms from the skin to the bloodstream *via* indwelling catheters was until recently believed to be a main mechanism underlying late-onset sepsis. Causative organisms may also be members of the infant gut (and oral) microbiota (199, 200). Especially in Gram-negative late-onset sepsis, the bacteria isolated in diagnostic blood culture was present in the gut microbiota beforehand (201). Therefore, it is now also believed that microbes in the gut can translocate into the bloodstream and cause septicaemia, facilitated by the intestinal and immune system immaturity described above in relation to NEC. Since translocation also occurs from the oral cavity (202), it is likely that oral bacteria may also reach the

blood stream by translocation and thereby cause septicaemia in preterm neonates.

Bronchopulmonary dysplasia and retinopathy of prematurity

Bronchopulmonary dysplasia and retinopathy of prematurity are two other diseases affecting preterm infants. Bronchopulmonary dysplasia is a lung disease causing long-term breathing difficulties, while retinopathy of prematurity involves abnormal growth of blood vessels in the retina which may cause blindness. The microbiota has also been implicated in these conditions, as decreased concentrations of microbial metabolites (acetic acid) in the gut have been associated with increased risk of bronchopulmonary dysplasia, and supplementation of acetic acid in a mouse model attenuates lung injury (203, 204). Also, retinopathy of prematurity has been associated with enrichment of *Enterobacteriaceae* in the preterm gut (205).

Aim

The overall aims were to:

1. Describe how the microbiota is established during early life, in different settings and in relation to various exposures
2. Investigate how the levels of individual fecal short chain fatty acids (SCFAs) changes with the age of the child and identify factors that determine the infant's plasma levels of SCFAs
3. Investigate if infant's microbiota and/or SCFA pattern relate to later allergy development

Specific aims:

- Identify how the gut microbiota establishes and develops over the first 6 months in infants in farming and non-farming families, in relation to various environmental exposures (Paper I)
- Describe how the faecal SCFA pattern changes with age and compare the pattern between infants from farming and non-farming families (Paper II)
- Determine whether the infant's gut microbiota (Paper I) and fecal SCFA pattern (Papers II and III) relate to subsequent allergy development
- Quantify the levels of a broad range of SCFAs in infant plasma at four months of age and relate these to those in corresponding samples of maternal plasma and breastmilk (Paper IV)
- Determine whether the SCFA pattern in infant's plasma at four months of age is related to subsequent sensitization and allergy development (Paper IV)
- Identify how the gut, oral and skin microbiota establishes and develops over the first 6 weeks in extremely preterm infants in a neonatal intensive care unit, in relation to various perinatal exposures (Paper V)

Patients and Methods

Birth cohort studies

Samples from children (and in some instances their mothers) from four different birth-cohorts were analyzed in the present thesis: The FARMFLORA, BAS, NICE and ÖFLORA birth-cohort studies. The essentials of these studies, as well as which samples that were used from them are summarized in **Table 1**. In three of these birth-cohort, allergy and/or sensitization was diagnosed. The criteria used for the different allergy diagnoses in the different studies are shown in **Table 2**. Below, a brief description of the cohorts is given including which samples that were analyzed for the present thesis.

Farmflora (Papers I & II)

The FARMFLORA birth cohort included 65 children, of whom 28 were raised on small family-owned dairy farms in Skaraborg County in South-West Sweden and 37 lived in the same rural area but not on farms (206-208). The purpose of the study was to describe and analyze, in depth, microbiota-associated or other factors that could contribute to the strongly protective effect against allergy of growing up on small dairy farms. Pregnant mothers were recruited in late pregnancy and healthy infants born at term were included. Exclusion criteria were coming from dairy farms with infrequent animal contact. For the present study, fecal samples collected at 1 week, 1 month and 6 months of age were analysed (by sequencing of bacterial 16S rDNA) for gut microbiota composition (Paper I) and fecal samples collected at 1 month and 1 and 3 years of age were analyzed for fecal SCFA levels (Paper II). Children were clinically examined for allergic diseases at 3 and 8 years of age and gut microbiota composition (Paper I) and faecal SCFA pattern (Paper II) were related to subsequent allergy development.

Barn-AllergiStudien (BAS) (Paper III)

The BAS (BarnAllergiStudien or Pediatric Allergy Study) birth cohort included all 1,228 surviving infants at Östersund hospital in northern Sweden during one year, from February 1996 to January 1997 (209-211). Children were assessed for allergic diseases and undertook skin prick tests against common allergens at 1, 4, and 13 years of age. Fecal samples were obtained from 139

children at one year of age and analysed for SCFA concentrations (209). For statistical purposes, this group was selected to include a higher proportion of atopic children than in the whole cohort, and 28 of these children (20%) had a positive skin prick test. Data on gestational age at birth, birth weight, breastfeeding, siblings, infections and use of antibiotics were collected by questionnaire, filled in by the parents. At 13 years of age, allergy was diagnosed using a questionnaire on symptoms, treatment and doctors' diagnoses of allergies which was filled in together by the adolescents and their parents (N = 834, i.e. 68% of the starting population). The questionnaire was modelled on that used the International Study of Asthma and Allergies in Childhood study (ISAAC) (212). Further data on parental allergies, siblings and pet ownership throughout childhood was also collected in the same questionnaire. In addition, adolescents undertook skin prick tests (N=794, i.e. 65% of the starting population). The population analyzed in this thesis included those children who provided a fecal sample in which SCFA levels were determined at 1 year of age, and answered the questionnaire and underwent a skin prick test at 13 years of age (N = 110).

NICE (Paper IV)

The NICE (Nutritional impact on Immunological maturation during Childhood in relation to the Environment) birth cohort included 651 children delivered at Sunderby Hospital in Northern Sweden. Pregnant women in the hospital catchment area were recruited in 2015-2018, and around 10% of all deliveries in this area were included in the NICE cohort (213). An analysis of the included women, compared to those who did not participate revealed that the included women were older, more educated and more often cohabitating with their partners than the non-included women (214). Families were followed with repeated collection of biological samples and questionnaires until the child reached six years of age (215). This included plasma samples from infants at 4 months of age, and plasma and breast milk samples were collected from mothers at the same time point (Paper IV). Infants were clinically examined for allergic diseases at one year of age. The study in Paper IV is based on a subgroup of 148 infants who either (i) provided plasma samples and also had corresponding mothers' plasma and breast milk samples available, or (ii) provided plasma samples and were diagnosed with allergic disease at one year of age.

Table 1: Birth cohorts from which samples and data were derived for the studies in the present thesis

	Birth cohort			
	FARMFLORA	**BAS**	**NICE**	**ÖFLORA**
General characteristics				
Site (Sweden)	Skaraborg County	Östersund county	Norrbotten county	Gothenburg
Inclusion period	Sept 2005 - May 2008	Feb 1996 – Jan 1997	Jun 2015 - Aug 2018	Mar 2013 - Jun 2015
No. in cohort	65	1228	651	89
Recruitment	Pregnancy, final inclusion at birth	Gestational wk 18, final inclusion at birth	Gestational wk 18-20, final inclusion at birth	Birth
Followed to	8 yr	13 yr (N=834)	6 yr	6 wks, or death or discharge from NICU
Inclusion criteria	Parents with family dairy farm (N=28), rural non-farm (N=37). Delivery at 36-42 wks of gestation. Good knowledge of Swedish		Planned delivery at Sunderby Hospital, Luleå. Good knowledge of Swedish	Extremely preterm (<28 wks of gestation). Admitted to NICU directly after birth
Exclusion criteria	From dairy farms with infrequent animal contact		Multiple births. Second pregnancies during the inclusion period	
Atopy diagnosis				
Sensitization		1, 4, yr [a] 13 yr [c]	12 mo SPT [a]	
Allergy diagnosis	18 mo [b], 3yr, 8yr	1, 4, 13 yr [c]	12 mo clinical examination	
References	(206-208)	(209-211)	(213-215)	

This thesis				
Samples used here				
Infants	Fecal samples 1 wk, 1 mo, 6 mo (16S rRNA gene sequencing); 1 mo, 1 yr, 3 yr (SCFAs)	Fecal samples 1 yr (examined previously)	Blood plasma 4 mo	Rectal swab/feces, oral swab, skin swab day 1, day 3-5, 1, 2, 3, 4 & 6 wks
Mothers			Infant blood plasma (N=148), mothers' blood plasma (N=142), breast milk (N=128) 4 mo post-partum	
N included here	65	110	148	42
Basis for selection	Available samples	Fecal sample at 1 yr + examined for sensitization and allergy at 13 yr	a) Full set of 4 mo samples: infant & mother plasma, mother's breast milk. b) Infant plasma available, diagnosed with allergy at 12 mo	Samples from oral, gut and skin microbiota cultured

[a] Skin prick test. Allergens: egg, milk, cat, dog, birch and timothy grass. [b] Not included in this thesis. [c] A total of 794 children were examined at 13 years of age with skin prick test for the following allergens: wheat, egg, milk, soy, fish, cat, dog, horse, birch and timothy grass. Responses to allergy-related questions in the 13-year follow-up questionnaire filled in by adolescents and their parents were assessed by pediatric allergologist Anna Sandin. [d] NICU = Neonatal intensive care unit.

ÖFLORA (Paper V)

The ÖFLORA birth cohort included 89 neonates born before 28 weeks of gestation and admitted directly after birth to the neonatal intensive care unit (NICU) at Queen Silvia Children's Hospital in Gothenburg, Sweden, between March 2013 and June 2015. (Of these, 71 were also participating in the Donna Mega randomized clinical trial to investigate the effect of parenteral fish oil supplementation on retinopathy of prematurity (216)). Infants were followed up to 6 weeks of postnatal age or until death or discharge from the NICU. Samples for microbial analysis were obtained at intervals from the mouth, feces and skin. For a sub-group of 42 infants, these samples were analysed for microbiota composition using quantitative culture methods. The subgroup was established due to the limited capacity of the research laboratory for microbial culturing and identification, and inclusion of an infant was based solely on the laboratory capacity at the infant´s date of birth.

Diagnosis of allergic (atopic) diseases

The different studies employed partly different criteria for diagnosis of different atopic diseases (atopic eczema/dermatitis, food allergy, asthma and allergic rhinoconjunctivitis). The different criteria are summarized in **Table 2**.

In the FARMFLORA study, all children were evaluated by study pediatricians at 18 months and at 3 and 8 years of age and all diagnoses were confirmed by a specialist pediatric allergologist. As a support for the diagnosis, specific IgE was in some instances measured in venous blood against common food (milk, egg, soy, fish, wheat and peanut) and inhalant allergens (birch, timothy grass, mugwort, cat, dog, horse and house dust mite). An allergen-specific IgE level ≥ 0.35 kU/L was considered positive. Skin prick tests were carried out using standard allergen extracts (birch, grass, mugwort, cat, dog, horse, rabbit, Dermatophagoides pteronyssinus, Dermatophagoides farinae and Cladosporium). However, sensitization was not mandatory for an allergy diagnosis (**Table 2**).

In the BAS study, adolescents (13 years old) were evaluated for atopic diseases based on their and/or their parents' responses to a questionnaire, supported by results from a skin prick test when relevant (**Table 2)**. In addition to the atopic conditions diagnosed in the other studies, respiratory allergy

was diagnosed, i.e. allergic asthma and/or rhinitis, the latter defined as runny nose and sneezing upon contact with an airborne allergen during the last 12 months, together with a positive skin prick test to relevant allergen(s).

In the NICE study, infants were clinically evaluated at 12 months of age by the study pediatrician to diagnose atopic eczema, food allergy, asthma and allergic rhino-conjunctivitis, based on the criteria listed in **Table 2**. Skin prick tests were carried out using standard allergen extracts.

Sampling

In the Farmflora study, parents collected fecal samples from their children at regular intervals and these were transported in gas-tight sachets in which an anaerobic atmosphere was generated after sealing (AnaeroGen Compact, Oxoid Ltd, Basingstoke, UK) to Gothenburg University, where they were stored at -80°C until analysis (Papers I & II).

In the BAS study, fecal samples were collected by the parents at home, placed immediately in a plastic tube and stored at −20◦C, until parents brought the samples to the hospital where they were stored frozen until analysis (Paper III).

In the NICE study, venous blood was collected in EDTA tubes from mothers and infants at the 4-month postpartum follow-up, and the plasma was separated and stored at -80^{o}C until analysis (215). Breast milk was collected by mothers at home, with mothers instructed to collect 30 mL of breast milk either by hand or with a pump just before feeding their infant in the morning and store the milk in their freezer. Mothers then brought the frozen milk to the clinic (Paper IV).

In the ÖFLORA study, the gut microbiota was sampled by either fecal or rectal swab samples. When available at the designated sampling time, freshly voided feces were collected from the infant's diaper and placed in a sterile tube, which was then placed in a plastic bag in which an anaerobic

Table 2: Diagnostic criteria for allergic diseases

	FARMFLORA study (3 yr, 8 yr)	NICE study (1 yr)	BAS study (13 years of age)
Atopic eczema			
	Williams' criteria (Williams 1996)	An itchy condition, usually manifested as scratching/rubbing, and ≥3 of: a) itchiness or visible eczema in typical areas such as cheeks, folds of the elbows or knees, or front of ankles or around the neck, b) general dryness of the skin, c) asthma, hay-fever, or another atopic disease, or d) a history of atopic disease in a first-degree relative	Questionnaire reported itching, chronic or chronically relapsing non-infectious dermatitis with typical features and localization, with symptoms in the last 12 months
Food allergy			
3 yr	An immediate or late onset reaction after intake of a specific food, followed by a clear and prompt improvement after elimination of the food, and any of: a) other signs of allergic disease, b) more than one organ system involved, c) supported by positive allergy tests, biopsies or food challenge tests	Typical symptoms appearing in relation to the intake of a triggering food(s) confirmed by ≥1 oral food challenge in which symptoms disappeared as the suspected offending food was eliminated from the infant's diet, and the same symptoms reappeared upon challenge	Questionnaire reported symptoms of food allergy
8 yr	Symptoms of food allergy supported by an open food challenge, planned or accidental		
Asthma			
3 yr	≥1 of the following criteria: a) ≥3 Wheezing episodes with the last episode after 2 years of age, and any of: i) symptoms outside of infections, or ii) other allergic manifestations, or b) Wheeze with onset after age 2, and any of: i) wheeze triggered by colds (in children with other allergic manifestation(s)), or ii) wheeze triggered by exercise	≥1 of the following criteria: a) Wheeze between infectious episodes, b) Persistent wheeze for at least 4 weeks, c) ≥3 Periods of wheeze in conjunction with infectious episodes, d) ≥1 Period of wheeze during an infectious episode in an infant having other atopic disease(s)	Doctor's-diagnosed asthma and/or use of asthma medication and/or wheezing symptoms during the last 12 months, together with a positive skin prick test to airborne allergen(s)

Asthma			
8 yr	Wheeze/heavy breathing, and any of: a) response to anti-inflammatory maintenance therapy, b) bronchial hyperresponsiveness with PD_{20} <0.6 mg methylcholine, or c) reversible forced expiratory volume in 1 second ≥12%		
Allergic rhinoconjunctivitis			
	Eye and/or nasal symptoms on exposure to pollen or animals, combined with a positive skin prick or IgE test against the suspected triggering allergen	Typical eye and/or nasal symptoms on ≥2 occasions when exposed to a specific allergen (usually animal dander or pollen) during an infection-free period. The reaction should disappear after removing the allergen (for pollen: out of season) or when taking medication	Reported itchy eyes and/or runny nose and sneezing on contact with an airborne allergen during the last 12 months, together with a positive skin prick test to airborne allergen(s).
Respiratory allergy			
			Allergic asthma and/or rhinitis, defined as runny nose and sneezing upon contact with an airborne allergen during the last 12 months, together with a positive skin prick test to relevant allergen(s)

atmosphere was generated after sealing (AnaeroGen Compact, Oxoid ltd, Basingstoke, UK). If a fecal sample was not available, a rectal swab sample was collected using the ESwab system (Copan Diagnostics Inc) by carefully inserting the swab into the anus and rolling it over the mucosa. Oral and skin swab samples were also collected using the ESwab system (Copan Diagnostics Inc, Murrieta, USA). Oral samples were collected by rolling the swab along the infant's oral mucosal surface of the mouth, inner cheeks and tongue, and skin samples were collected by rolling the swab over the skin close to the umbilical stump. All samples were transported to the laboratory and handled within 24 h after collection. For the 42 infants in our sub-cohort, fresh samples were cultured quantitatively as described below, and portions of samples from all infants were frozen at -80^{o}C for later analyses (Paper V).

Dietary assessment in the different cohorts

In the FARMFLORA study, dietary information was collected as described previously (73, 217). In brief, parents continuously recorded breastfeeding and formula practices in diaries until the infant reached 18 months of age. At 1 year of age (10–14 months), parents completed an unannounced 24-h dietary recall, followed by a 24-h food diary, of their child's diet. The collected dietary information was registered and nutritional content calculated using the software Dietist Net Pro (Kost och Näringsdata, version 15.11.02, Stockholm, Sweden) based on the food composition database of the Swedish National Food Agency (Papers I & II).

In the NICE study, mothers' dietary intake during the period 3-4 months postpartum was collected using a web-based, validated, semi-quantitative food frequency questionnaire (Meal-Q), which was sent out four months postpartum (218, 219). Daily intake in grams of a wide range of food items was estimated using images of portion sizes together with reported intake frequency (Paper IV).

Analytical methods

Microbiota characterization

Next-generation sequencing (Paper I)

After thawing of frozen fecal samples, DNA was extracted and the V3-4 region of the bacterial 16S rRNA gene was amplified as described previously (220). In brief, DNA was extracted using the QIAamp DNA stool mini kit (QIAGEN AB, Sollentuna Sweden) modified by adding a bead-beating step to improve DNA extraction from Gram-positive bacteria. Bacterial 16S rRNA genes were amplified by PCR using 341F and 785R primers (221). A second PCR was carried out to add the sequencing adaptors and multiplex identifiers, and all samples were pooled to create the library.

Sequencing of the 16S rRNA gene amplifications was carried out on an Illumina MiSeq platform (Illumina Corp. San Diego, CA, USA) at Science for Life Laboratory (Stockholm, Sweden) using 2 × 300 bp paired end reads. Raw sequencing data were converted from Bcl to FastQ format, demultiplexed and quality checked using FastQC, then further processed in QIIME2 (222).

Amplicon sequence variants (ASVs) were obtained using the DADA2 plugin and clustered into 97% operational taxonomic units (OTUs) using the vsearch plugin (223). The taxonomy of each OTU was assigned using the SILVA (v132) SSU rRNA database of reference sequence.

Quantitative bacterial culture (Paper V)

The oral, gut and skin microbiota were characterized using quantitative microbial culture (31, 51). In brief, fecal samples that had been transported in an anaerobic atmosphere were processed at the laboratory within 24 h after collection. A calibrated spoon of feces was serially diluted in sterile peptone water, plated on two non-selective media and eight selective media, and incubated under either aerobic or anaerobic conditions, as summarized in **Table 3** (224-235). From appropriate dilutions on each medium, free-lying colonies of different morphology were enumerated, Gram-stained, examined by microscope and sub-cultured for purity. Isolates were identified to the genus or species level using a range of methods, as summarized in **Table 4**.

Table 3: Media and culture conditions for culture-based analysis of oral, gut and skin commensal microbiota in premature infants.

	Medium (agar)	Days	Atmosphere	Ref	Sites cultured [a]
Facultatives					
Total	Colombia blood	2	Air	(224-5)	O, F, R, S
Gram-positive facultatives					
Staphylococcus	Staphylococcus	2	Air	(226)	O, F, R, S
Streptococcus, (*Enterococcus*[b]), *Gemella*	Streptococcus	2	5% CO_2[c]	(227)	O, F, R, S
Enterococcus	Enterococcel	2	Air	(228)	F, R
Micrococcus spp.	Colombia blood	2	Air		O, F, R, S
Corynebacterium spp.	Colombia blood	2	Air		O, F, R, S
Gram-negative facultatives					
Enterobacteriaceae, Pseudomonas, Acinetobacter etc [d]	Drigalski	2	Air	(225)	O, F, R, S
Hemophilus, Neisseria, Moraxella	HI-agar [e]	2	5% CO_2[c]	(229)	O
Yeasts					
Candida etc.	Saboraud	2d	Air	(230)	O, F, R, S
Malassezia	PO agar [f]			(231)	O, F, R, S
Anaerobes					
Total	Brucella blood	3d	Nitrogen	(232)	O, F, R, S
Gram-positive anaerobes					
Bifidobacterium	Beerens	3d	Nitrogen	(233)	F, R
Propionibacterium, *Actinomyces*, *Finegoldia*	Brucella blood	3d	Nitrogen	(232)	O, F, R, S
Lactobacillus (*Bifidobacteria* [g])	Rogosa	3d	Nitrogen	(234)	O, F, R

Anaerobes…					
Clostridium spp [h]	Brucella blood	3d	Nitrogen	(232)	R, F
C. difficile	CCFA [i]				
Gram-negative anaerobes					
Bacteroides fragilis group	BBE [j]	3d	Nitrogen	(235)	F, R
Veillonella, Parabacteroides, Prevotella, Fusobacterium	Brucella blood	3d	Nitrogen	(232)	O, F, R, S

Serially diluted swab and faecal samples were cultivated at 37°C on the media specified above for the isolation and enumeration of different bacterial groups and yeasts. For culturing of anaerobic bacteria agar plates were prereduced for 24 hours. [a] Sites cultured: O = oral, F = faeces, R = rectal and S = skin samples. [b] In oral and skin cultures, *Enterococcus* spp. were enumerated from *Streptococcus* agar. [c] Air atmosphere enriched with 5% CO_2. [d] Including other non-fermenting Gram-negative rods. [e] *Haemophilus influenzae* agar. [f] *Pityrosporum orbiculare* agar, introduced from May 2014. [g] In oral samples, *Bifidobacterium* spp. was enumerated from Rogosa agar. [h] For the isolation of spore forming *Clostridium* spp. undiluted rectal swab sample fluid or faeces diluted 1:10 was mixed 1:1 with 99% ethanol and incubated on a shaker at room temperature for 30 min to kill vegetative cells, where after the sample was diluted and plated. [i] Cycloserine Cefoxitin Fructose agar. [j] BBE= Bacteroides bile esculin.

Table 4: Identification of cultured microbes to the genus or species level

	Identification
Facultatives	
Enterobacteriaceae, *Pseudomonas, Acinetobacter*	API 20E biotyping [a]
Staphylococcus	Gram-stained appearance, catalase +, coagulase +/- [b]
Enterococcus	Gram-stained appearance, esculin hydrolysis, MALDI-TOF [a]
Streptococcus, Micrococcus, Corynebacterium, Actinomyces, Haemophilus, Neisseria	Gram-stained appearance, MALDI-TOF [a]
Anaerobes	
Bifidobacterium, Lactobacillus, Bacteroides, Clostridium, Clostridioides difficile, Cutibacterium, Propionibacterium, Finegoldia, Veillonella, Parabacteroides, Prevotella, Fusobacterium	Gram-stained appearance, MALDI-TOF [a]
Yeasts	
Yeasts [c]	Gram-stained appearance

[a] For identification to the species level. [b] Classified as *S. aureus* or "coagulase-negative staphylococci" based on a positive or negative coagulase test. [c] Yeasts were not identified to the genus or species level.

SCFA analysis (Papers II-IV)

SCFA in feces (Papers II & III)

Fecal concentrations of 8 SCFAs (acetic, propionic, iso-butyric, butyric, iso-valeric, valeric, iso-caproic and caproic acid) were measured by gas chromatography. In the FARMFLORA study (Paper II), analysis was carried out using a method modified from Zhao et al (236). In brief, samples were thawed and 200 mg of feces was mixed with 1 ml of water and homogenized, then dilute hydrochloric acid was added to achieve a pH of 2–3. Solids were separated out by centrifugation and 180 µL of supernatant was mixed with 20 µL of internal standard solution (7.9 mM 2-ethylbutyric acid in 12% v/v aqueous formic acid) to generate the solution for analysis. A fused silica capillary column with free fatty acid stationary phase was used for separation

and flame ionization was used for detection. In the BAS study (Paper III), a different method was used [21, 22]. In brief, samples were thawed and 0.5 g of feces was homogenized in 2 mL of distilled water containing 0.5 mmol/L of dilute sulfuric acid and 3 mmol/L of 2-ethylbutyric acid as an internal standard, then vacuum distilled and analysed for SCFAs on a gas chromatograph (Perkin Elemer Autosystem XL).

SCFA in plasma and breast milk

Plasma and breast milk concentrations of 9 SCFAs, of which 8 were "classic" SCFAs (formic, acetic, propionic, isobutyric, butyric, isovaleric, valeric, and caproic acid) and the 9th the di-carboxylic succinic acid, also produced by gut microbes, were measured by ultra-high performance liquid chromatography-mass spectroscopy (UPLC-MS) (237)(238). In brief, 10 µl plasma was incubated with 60 µl 75% methanol (LiChrosolv®), 10 µl 200 mM 3-NPH, and 10 µl 120 mM EDC-6% pyridine to derivatize the SCFAs, then the reaction was quenched with 10 µl of 200 mM quinic acid. Solids were separated out by centrifugation, 10% methanol in water was added to the supernatant up to 1 mL, and the solution was centrifuged again. Then 100 µl of the supernatant was mixed with 100 µl of internal standard (^{13}C6-3NPH). A Waters ACQUITY UPLC BEH C18 column was used for separation and a 6500+ QTRAP triple-quadrupole mass spectrometer (AB Sciex, 11432 Stockholm, Sweden) equipped with an APCI source and operated in the negative-ion MRM-mode was used for detection.

Statistics

For analysis of 16S rRNA gene sequencing data, estimation of alpha and beta diversity, and generation of random forests models to identify microbial taxa associated with various early life exposures, analyses were carried out using MicrobiomeAnalyst (https://www.microbiomeanalyst.ca) (239). Data was filtered to remove taxa containing all zero values or appearing in only one sample and normalized using total sum scaling. Alpha diversity was determined based on the number of observed OTUs and differences between groups were assessed using the Mann-Whitney U-test at each time point. To estimate beta diversity, rarefaction was carried out to the minimum library size and Bray-Curtis dissimilarity distances between groups were assessed using PERMANOVA. Random forests models based on 1,000 trees were used

to identify the top 15 discriminatory taxa between exposure groups. Differences in the proportion and prevalence (presence or absence) of these discriminatory taxa were tested by separate univariate analysis, and multiple linear and logistic regression models were used to identify associations independent of covarying exposures ($p<0.2$), with analysis carried out in SPSS. Estimated microbial age (EMA) was determined based on a method modified from Depner *et al* (40). Random forest models of gut microbiota composition in relation to the infants' exact age on collection of fecal samples were generated in a training set of 70% of samples using Scikit-Learn (240)/1.0.2 in python/3.9.9 using 2000 trees. The models were 5-fold cross-validated and the most important features were extracted using permutation importance in Scikit-Learn. The EMA of the gut microbiota composition in the remaining test set of samples was then determined, and bivariate analysis was used to identify associations between early life exposures and EMA.

For an overview of the relation between a multitude of microbial (X) variables with one or more exposure (Y) variables, we used orthogonal projection to latent structures (OPLS), a multivariate data analysis method based on linear regression, e.g. to identify the microbial groups associated with various perinatal factors at each body site in Paper V) (SIMCA P+ V17; Umetrics AB, Umeå, Sweden). The analyses were based on both the prevalence (presence or absence) of each bacterial group and their population count in colonized infants. For each exposure variable, associations with microbial variables identified in OPLS analysis were tested by separate univariate analysis and adjustment for false discovery rate was applied, with analysis carried out in SPSS (v26, IBM Corp., NY, USA) or GraphPad Prism (v.9.3.1, GraphPad Software, San Diego, USA).

In general, associations between categorical variables, a continuous and a categorical variable, or two continuous variables were assessed by Fisher's Exact test, Mann-Whitney U-test or Spearman rank correlation respectively. Statistical analyses were carried out using SPSS (v26, IBM Corp., NY, USA) or GraphPad Prism (v.9.3.1, GraphPad Software, San Diego, USA).

Results

Gut microbiota development in the FARMFLORA cohort (Paper I)

Infant gut microbiota composition up to 6 months of age

In Paper 1, the composition of the gut microbiota up to six months of age in infants living in rural Sweden (the FARMFLORA cohort) was characterized by next generation sequencing of bacterial 16S rRNA-genes in fecal samples. As shown in **Figure 1**, Bifidobacteriaceae was the most abundant bacterial family at 1 week of age (23% of total abundance), followed by Bacteroidaceae, both representing anaerobic bacterial groups. Facultative bacteria made up approximately half of the total bacterial abundance and were dominated by the Staphylococcaceae, Streptococcaceae and Enterobacteriaceae families. By 16S rRNA-sequencing, it is not possible to distinguish, e.g. *S. aureus* from coagulase-negative staphylococci, or to accurately distinguish *E. coli* from *Shigella*.

By one month of age, anaerobic bacteria outnumbered the facultatives. Bifidobacteriaceae were still the most common bacteria, but Bacteroidaceae had expanded and the obligate anaerobic Lachnospiraceae had also become more abundant. In contrast, the populations of facultatives such as Staphylococcaceae and Streptococcaceae had declined, while Enterobacteriaceae were less affected.

At six months of age, Bacteroidaceae had expanded further and outnumbered Bifidobacteriaceae. Other anaerobes belonging to the families Lachnospiraceae and Veillonellaceae, had also increased in abundance, making anaerobic bacteria completely dominant in the microbiota. At this stage, also Enterobacteriaceae had declined, and facultatives constituted less than a quarter of the total bacterial abundance.

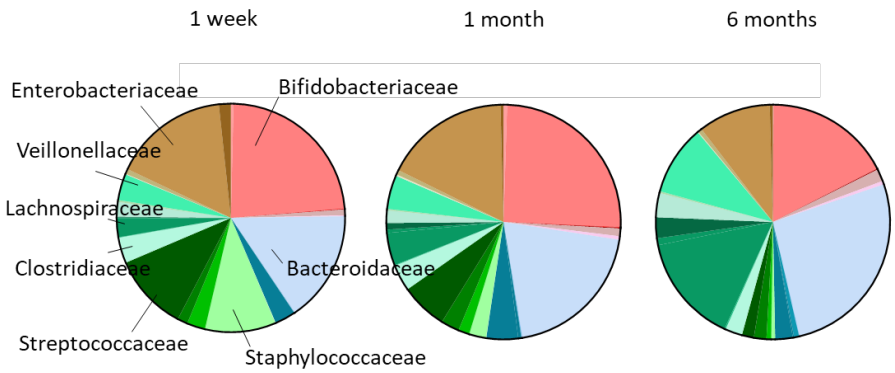

Figure 1: Relative abundance of bacterial taxa in the gut microbiota of the FARMFLORA infants at 1 week, 1 month and 6 months of age. Pie charts showing the mean relative abundance of bacterial taxa at the family level at 1 week, 1 month and 6 months of age. Figure reproduced from an article published by Taylor & Francis Group in Gut Microbes in 2020(12:1), available online: https://wwww.tandfonline.com/10.1080/19490976.2020.1847628." (241).

Effect of early life exposures on gut microbiota assembly in FARMFLORA

In Paper 1, the effect of various early life exposures on gut microbiota development was also investigated. Specifically, these were delivery mode, exposure to intrapartum antibiotics, formula feeding/weaning, elder siblings, growing up on a farm, and pets in the household. Firstly, the effect of each exposure on overall gut microbiota composition was determined, as shown in **Table 1**. Delivery mode, exposure to intrapartum antibiotics, presence of elder siblings and pets in the household all explained some of the observed microbial variance at 1 week of age, as evidenced by significant separation in Bray-Curtis dissimilarity distances. The effect of each early life exposure on the gut microbiota changed over time and by 6 months of age only the effect of exposure to antibiotics during delivery remained significant, although a significant effect of cessation of breastfeeding became apparent. Notably, farm living had no significant effect on microbial variance in this analysis.

Infants experience a number of these exposures during their early life, and a Microbial Exposure Index (MEI) was devised to investigate their cumulative

effect on the gut microbiota. In this index one point was assigned for each 'microbe-enriching' exposure, namely pets, farm living, elder siblings, formula feeding/weaning, vaginal delivery and lack of exposure to intrapartum antibiotics. To reflect varying patterns of infant nutrition over time, formula

Table 1: Effects of early life exposures on gut microbiota variance in the FARMFLORA infants

	r^2		
	1 week	1 month	6 months
Delivery mode	0.08 **	0.03 *	0.03
Intrapartum antibiotics	0.06 **	0.04 *	0.04 *
Formula/weaning [1]	0.02	0.03	0.04 *
Elder siblings	0.08 ***	0.04 *	0.01
Pets	0.04 *	0.02	0.02
Farm	0.03	0.01	0.02

[1] Defined as a mixture of breast and formula feeding as compared to exclusive breastfeeding at 1 week (N=7, 53) and at 1 month (N=12, 50) of age, and weaning (cessation of breastfeeding) as compared to continued partial or exclusive breastfeeding (N=15, 48) at 6 months of age.

feeding/weaning was defined at 1 week and 1 month of age as consumption of formula in addition to breastmilk, as compared to exclusive breastfeeding, and at 6 months of age as weaning (cessation of breastfeeding), as compared to continued consumption of at least some breastmilk. Two infants who only consumed formula at 1 week and/or 1 month of age were excluded. The resulting MEI score was related to (i) the anaerobic character of the microbiota, defined as the proportion of anaerobic bacteria, (ii) the proportion of major anaerobic bacterial groups in the infant gut, namely bifidobacteria and *Bacteroides*, and (iii) estimated microbial age (EMA) (40). EMA is a measure of microbiota maturity and was determined by relating gut microbiota composition to the exact age of infants at sampling in a training set of 70% of samples using a random forests model, and then applying this model to estimate the microbial age of the remaining 30% of samples. The higher the MEI score, ie. the greater the number of 'microbe-enriching' exposures (see **Figure 1d-g**, Paper I), the higher was the proportion of

anaerobes at 1 week and 1 month of age. MEI score was also strongly significantly associated with the proportion of *Bifidobacterium* at 1 week of age, although less so with the proportion of *Bacteroides*. Associations with MEI were most pronounced during the first month, but the relationship between higher MEI and higher proportions of anaerobes, *Bifidobacterium* and *Bacteroides* was still apparent at 6 months of age. In addition, higher MEI score tended to be linearly associated with higher EMA at 1 week of age, although not at later time points (see **Figure 3a-b**, Paper I).

Next, the effect of each early life exposure on gut microbiota composition was investigated individually, as shown in **Tables 2** and **3**, using a variety of measures including estimated microbial age, alpha diversity (number of OTUs) and anaerobic character. Also, the top 15 bacterial taxa which discriminated between exposure groups were identified using another random forests machine learning algorithm, and univariate associations with the prevalence and proportion of each of these taxa were determined separately. As many of the early life exposures were inter-related, the associations identified between elder siblings, pet ownership or farm residence and the prevalence and proportion of bacterial taxa and also the proportion of anaerobic bacteria were adjusted for covariates, where sample size allowed. However, the low number of infants delivered by caesarian section or exposed to intrapartum antibiotics and the high rates of breastfeeding in this cohort meant that sample size was not sufficient to adjust any findings related to these exposures for covariates.

As shown in **Table 2**, in infants delivered by caesarean section (N=10) tended to have a less 'mature' gut microbiota, as indicated by a lower microbial age, which also tended to be less anaerobic. However, various significant effects of caesarean delivery on microbial colonization were observed. Infants delivered by caesarean section were characterized by lower prevalence of *Bacteroides* and higher prevalence of the facultatives streptococci and enterococci in their microbiota during the first month after birth. However, they were more frequently colonized by *Veillonella* at 1 week, *Roseburia* at 1 month and *Clostridium* inocuum group at 6 months of age. Similarly, significant effects of exposure to intrapartum antibiotics were also observed. Infants exposed to intrapartum antibiotics (N=10, **Table 2**) had a less 'mature' gut microbiota in the first month than unexposed infants, which was less anaerobic, and less frequently contained *Bifidobacterium* and *Bacteroides* but

had higher prevalence of the facultatives *Enterococcus* and *Haemophilus*. These infants also had lower proportions of the anaerobic commensal *Eggerthella*, but higher prevalence of *Clostridium sensu stricto* 1, and also the facultative *Raoultella*, in their gut microbiota at 6 months of age.

Formula feeding and weaning were also clearly associated with gut microbiota composition (**Table 2**). Both alpha-diversity and estimated microbial age was significantly increased in pooled samples from formula fed or weaned infants. At 1 month of age, infants who consumed some formula in addition to breast milk, as opposed to being exclusively breastfed, had a significantly more diverse microbiota with a higher estimated microbial age characterized by higher prevalence and/or proportions of various anaerobic Firmicutes, especially clostridial bacteria, as well as higher proportions of enterococci and unclassified *Enterobacteriaceae*, but lower proportions of *Haemophilus*. Similarly, at 6 months of age, infants who had ceased breastfeeding as opposed to those who continued to receive at least some breast milk had a much more diverse microbiota with a tendency for higher estimated microbial age, which was again characterized by significantly higher levels of many clostridial bacteria and other anaerobic Firmicutes, but lower proportions of *Escherichia*.

In infants with elder siblings (**Table 3**), the gut microbiota showed no consistent difference in alpha diversity or estimated microbial age compared to first-born infants, although infants with elder siblings had a more diverse gut microbiota at 1 week of age. They were also significantly more frequently colonized by bifidobacteria and harbored higher proportions of *Atopobium*, *Corynebacterium* and *Dolosigranulum*, but had lower levels of enterococci during the first month. Also, while infants with elder siblings harbored significantly higher proportions of staphylococci initially, there was a tendency for fewer infants to be colonized by staphylococci at 6 months of age. Although it was not possible to adjust the association for covariates, infants with elder siblings were also significantly less frequently colonized by and harbored lower levels of *Clostridioides* at 1 month of age than first-born infants.

Table 2: Effect of delivery mode, exposure to intrapartum antibiotics and formula/weaning on gut microbiota composition in the FARMFLORA infants.

	All samples [a]	1 week samples	1 month samples	6 month samples
Caesarian				
α-div:	ns.	ns.	ns.	(↑)
EMA:	↓*	↓*	ns.	[b]
Bacteria ↑		*Streptococcus*, *Veillonella*	*Enterococcus, Roseburia*	*C. inocuum* group
Bacteria ↓		(Anaerobes), *Bacteroides*	(Anaerobes)	
Intrapartum antibiotics				
α-div:	ns.	↓*	ns.	ns.
EMA:	(↓)	↓*	ns.	ns.
Bacteria ↑		*Enterococcus, Haemophilus*		*C. sensu stricto* 1, *Raoultella*
Bacteria ↓		Anaerobes, *Bifidobacterium*, *Bacteroides*	Anaerobes	*Eggerthella*
Formula/weaning				
α-div:	↑***	ns.	↑*	↑***
EMA:	↑*	ns.	↑**	(↑)
Bacteria ↑		[c]	*Enterococcus*, unclassified Enterobacteriaceae, various Clostridia, *Veillonella*	various Clostridia, *Faecalitalea*
Bacteria ↓			*Haemophilus*	*Escherichia*

[a] In pooled samples from all time points. * p<0.05, ** p<0.01, *** p<0.001, Mann-Whitney U-test. Parentheses ( ) indicates a tendency for association, p>0.05 – p<0.10**.** ns. indicates no significant difference. Bacterial groups listed were significantly associated with the exposure unless indicated. [b] Not determined as there were no samples from infants delivered by caesarian section at 6 months of age in the test set of samples for determination of estimated microbial age. [c] Not determined as only 7 infants received formula in addition to breast milk at 1 week of age.

Table 3: Effect of elder siblings, pets and growing up on a farm on gut microbiota composition.

	All samples [a]	1 week samples	1 month samples	6 month samples
Siblings				
α-div:	ns.	↑ *	ns.	ns.
EMA:	ns.	ns.	ns.	ns.
Bacteria ↑		*Bifidobacterium, Staphylococcus, Corynebacterium*	*Bifidobacterium, Atopobium, Dolosigranulum, Corynebacterium*	
Bacteria ↓		*Enterococcus*	*Clostridioides*	*Staphylococcus*
Pets				
α-div:	↑ **	ns.	↑ *	(↑)
EMA:	ns.	(↑)	ns.	ns.
Bacteria ↑		Anaerobes	*Dialister*	
Bacteria ↓		*Corynebacterium*, *C. sensu stricto* 1	*Haemophilus*	
Farm				
α-div:	ns.	ns.	ns.	ns.
EMA:	(↑)	ns.	ns.	ns.
Bacteria ↑		*Lactobacillus*, *Sutterella*, *Ruminiclostridium*	*Butyricicoccus*, *Dialister*, *Megasphaera*	

[a] In pooled samples from all time points. * p<0.05, ** p<0.01, Mann-Whitney U-test. Parentheses ( ) indicates a tendency for association, p>0.05 – p<0.10. ns. indicates no significant difference. For bacterial groups, standard font indicates groups associated with early life exposures after adjustment for covariates, while an underline indicates groups with insufficient sample size to adjust for covariates but associated with early life exposures in univariate analysis.

Infants with pets (**Table 3**) had a more diverse microbiota than infants without pets, containing significantly higher proportion of anaerobes and lower levels of *Corynebacterium* and *Haemophilus* in the first month, and lower prevalence of *Clostridium sensu stricto* 1. Also, infants with pets had significantly higher levels of *Dialister* in their gut microbiota at 1 month of age, although it was not possible to adjust this finding for covariates.

In infants growing up on farms (**Table 3**), the gut microbiota showed no difference in alpha diversity compared to that of infants growing up in the same rural area but not on farms. However, farm infants tended to have a more 'mature' gut microbiota and harbored significantly increased proportions of lactobacilli at 1 week of age. Also, infants growing up on farms had significantly higher levels of *Sutterella*, *Butyricicoccus* and *Megasphaera*, and they more frequently harbored *Dialister* but less frequently *Ruminiclostridium* in their gut microbiota during the first month, although it was not possible to adjust these findings for covariates.

To summarize, pets and siblings as well as vaginal delivery were associated with measures of more rapid maturation of the microbiota with a higher proportion of anaerobes and/or bifidobacteria, while caesarean section and exposure to antibiotics tended to be associated with a less mature microbiota with increased levels of various facultatives.

SCFA patterns in early childhood

Fecal SCFAs up to 3 years of age in the FARMFLORA children

In Paper 2, fecal SCFA concentrations in the same cohort of children living in rural Sweden (FARMFLORA cohort) were measured at 4 weeks and 1 and 3 years of age by gas chromatography. Formic and succinic acid were not measured in this study. As shown in **Figure 2**, acetic acid was by far the most abundant SCFA at all time-points, although it should be noted that formic acid was not measured in this study. At 4 weeks of age, acetic acid made up >80% of the median total SCFA concentration and was present in all infants with a median concentration of 86 µmol/g. Propionic and butyric acid were detected in most infants although at much lower concentrations (6.1 and 1.9 µmol/g respectively), while the remaining SCFAs were detected in less than half of infants at several times lower median concentrations. With age, the median

total SCFA concentration increased, and by 3 years of age acetic, propionic and butyric acid were all present in all children, at median concentrations of

85, 32 and 28 µmol/g, respectively. Iso-butyric, iso-valeric and valeric acid were also present in almost all children, although at much lower median concentrations (2–3 µmol/g), and iso-caproic and caproic acid were present at even lower concentrations in a minority of children (data not shown).

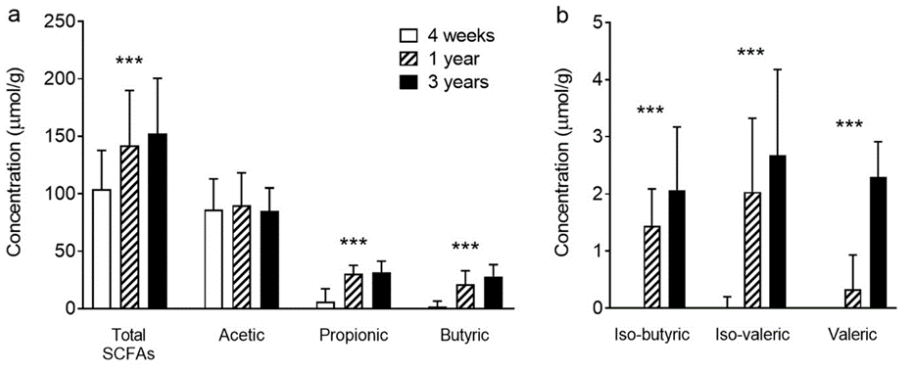

*Figure 2: Fecal SCFA at 4 weeks, 1 year and 3 years of age in the FARMFLORA infants Median (IQR) concentration of (a) total SCFAs and acetic, propionic and butyric acid, and (b) iso-butyric, iso-valeric and valeric acid. ** p<0.001, Friedman test. Reproduced with permission from Gio-Batta et al., Scientific Reports, 2020 (118).*

Approximately half of the children in the FARMFLORA cohort lived on small dairy farms. Farmers´ children and controls did not differ in total fecal SCFA concentrations or the concentration of any individual SCFAs at 4 weeks or 1 year of age, but at 3 years of age children growing up on farms had higher fecal concentrations of iso-butyric, iso-valeric and valeric acid (**Figure 3**).

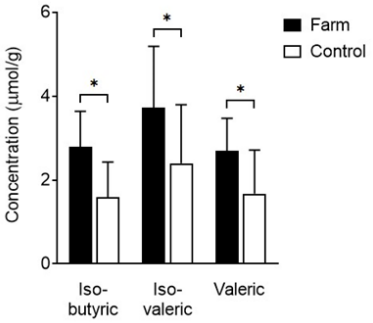

*Figure 3: Fecal SCFAs at 3 years of age in farm and control children, FARMFLORA. Median (IQR) concentration of iso-butyric, iso-valeric and valeric acid at 3 years of age in farmers´children and controls. * p<0.05 Mann-Whitney U-test. Reproduced with permission from Gio-Batta et al., Scientific Reports, 2020 (118).*

As shown in **Table 4**, multiple logistic regression models adjusted for covariates confirmed that children growing up on a farm had higher fecal concentrations of iso-butyric, iso-valeric and valeric acid at 3 years of age (aOR 17, 2.1–131, P = 0.008; aOR 6.7, 1.5–30, P = 0.01 and aOR 6.4, 1.4–31, P = 0.02 respectively). Part of the association between farm residence and fecal valeric acid concentration was explained by having pets in the household. Also, having elder siblings was associated with higher fecal valeric acid concentration at 3 years of age independent of farm residence (aOR 6.2; 95% CI 1.5–26; P = 0.01). Furthermore, protein intake was associated with higher levels of iso-butyric and iso-valeric acid (aOR 1.3, 1.1–1.5, P = 0.002 and aOR 1.1, 1.0-1.3, P = 0.01) while fiber intake was associated with lower levels (aOR 0.5, 0.3–0.8, P = 0.003 and aOR 0.7, 0.5-1.0, P = 0.04).

Plasma SCFAs at 4 months of age in infants in the NICE cohort, and in breast milk and plasma from their mothers

In Paper III, plasma concentrations of SCFAs were measured in infants in the NICE cohort at 4 months of age. In contrast to Paper II, formic acid was included in the analysis, as was succinic acid, which although not strictly a SCFA, is a small dicarboxylic acid microbial metabolite. Iso-caproic acid was not included in this study. The analysis was carried out by ultra-high performance gas chromatography-mass spectroscopy, which enabled detection of micromolar concentrations, as opposed to the millimolar levels measured in feces using regular gas chromatography in Paper II. As shown in **Table 5**, formic acid was the most abundant SCFA, making up 86% of total plasma SCFA abundance with a median concentration of 710 µmol/L. Acetic acid was the next most abundant SCFA, followed by propionic and succinic

acid, although these were found at substantially lower levels (median 97, 8 and 5 µmol/L respectively). The remaining SCFAs were found at median concentrations of <1 µmol/L.

Table 4: Logistic regression models exploring the effects of demographic, family and dietary characteristics on the associations between farm residence and high concentrations of fecal iso-butyric, iso-valeric and valeric acids at three years of age in the FARMFLORA cohort. Adapted with permission from Gio-Batta et al., Scientific Reports, 2020 (118).

SCFA [a]	Model [b]	Predictor variable(s)	OR	(95%CI)	*P*
Iso-butyric	Uni	Living on a farm	4.7	(1.3-16)	0.02*
	Multi [b]	Living on a farm	8.2	(1.6-43)	0.01*
		Protein [c]	1.1	(1.0-1.2)	0.005**
		Fruit and vegetables [c]	0.98	(0.97-1.00)	0.006**
Iso-valeric	Uni	Living on a farm	3.6	(1.0-12)	0.04*
	Multi [b]	Living on a farm	4.4	(1.1-18)	0.04*
		Protein [c]	1.1	(1.0-1.2)	0.02*
		Fruit and vegetables [c]	0.99	(0.98-1.00)	0.03*
Valeric	Uni	Living on a farm	6.9	(1.7-28)	0.008**
	Multi [b]	Living on a farm	6.4	(1.4-31)	0.02*
		Siblings	6.2	(1.5-26)	0.01*
		Cat or dog	3.3	(0.8-13)	0.10

* *P*<0.05, binary logistic regression, n=19 (farm) and 32 (control). [a] SCFA levels dichotomized as concentration of iso-butyric acid, > or ≤ 2.0 µmol/g; iso-valeric acid, > or ≤ 2.5 µmol/g; valeric acid, > or ≤ 2.0 µmol/g. [b] Final multivariate model. [b] Univariate or multivariate models. [c] Expressed as grams per day.

The average concentration of acetic acid in infant plasma in the NICE study, 97 µmol/L, can be roughly compared with the levels of acetic acid in feces in the FARMFLORA infants (**Figure 2**), which was approximately 80 µg/g of faeces, or 80 mg/kg feces. Thus, plasma levels of SCFA were approximately 1000 times lower than those in feces.

In Paper III, SCFAs (and succinic acid) concentrations were also measured in mother's breast milk and plasma collected when the infant was 4 months of age. As shown in **Table 5**, infants had a higher total SCFA concentration in

Table 5: Concentration of short chain fatty acids in infant plasma at 4 months of age and in maternal plasma and breast milk samples

Concentration (median, range) of SCFAs (µmol/L)									
	Infant plasma		Maternal plasma		Breast milk		p-value for comparison		
	N=148		N=142		N=128		IP-MP[a]	IP-BM[b]	MP-PM[c]
Formic	710	150-1300	460	100-1300	120	52-230	<0.001	<0.001	<0.001
Acetic	97	35-320	60	12-240	10	1.3-58	<0.001	<0.001	<0.001
Propionic	8.3	0.00-150	5.0	0.00-200	0.10	0.00-2.0	<0.001	<0.001	<0.001
Butyric	0.43	0.00-3.0	0.54	0.01-2.8	61	7.9-470	0.108	<0.001	<0.001
Isobutyric	0.18	0.00-2.4	0.20	0.00-1.0	0.88	0.18-7.7	0.043	<0.001	<0.001
Succinic	4.9	1.9-15	6.4	2.4-19	6.4	3.8-12	<0.001	<0.001	0.972
Valeric	0.05	0.00-0.63	0.05	0.00-0.34	0.15	0.00-1.8	0.690	<0.001	<0.001
Isovaleric	0.18	0.00-1.3	0.18	0.0-1.4	0.20	0.0-1.1	0.319	0.321	0.948
Caproic	0.58	0.09-1.5	0.58	0.09-1.3	50	7.1-350	0.310	<0.001	<0.001
Total SCFAs	850	210-1600	550	160-1300	250	97-1000	<0.001	<0.001	<0.001

Wilcoxon test was used to compare levels of [a] infant and maternal plasma, [b] infant plasma and breast milk, and [c] maternal plasma and breast milk.

their plasma compared to their mothers (median 850 vs. 550 µmol/L, $p<0.001$), which was attributable to higher levels of formic, acetic and propionic acid (all $p<0.001$). However, the relative concentration of the various SCFAs was similar between infants and mothers, i.e. formic > acetic > propionic, succinic > other SCFAs.

Compared to plasma of either infants or mothers, breast milk contained a lower total SCFA concentration (120 µmol/L) and the relative concentrations of the various SCFAs were quite different. Although formic acid was the most prevalent SCFA in both mother's milk and plasma, milk contained butyric and caproid acid as the next most abundant SCFAs. In fact, butyrate and caproate made up almost half of the total SCFA abundance in milk, but less than 1% of the SCFAs in mother's plasma. Thus, both butyric and caproic acid were enriched around 100 times in mother's milk compared to her blood plasma. The concentration of iso-butyric and valeric acid were also four times higher in breast milk than plasma.

Almost all infants were breast fed at 4 months of age, when the samples were collected. As breast milk was the sole, or a major food for the infant, we expected to see a correlation between the concentrations of individual SCFAs in the mother's milk and her infant's plasma. However, there were no significant positive correlations between concentrations of any individual SCFA in mothers' milk versus infants' plasma, and even a negative correlation for valeric acid (see **Figure 3** in Paper IV, middle row).

Quite surprisingly, concentrations of all individual SCFAs, except for butyric acid, were strongly correlated in infants' and mothers' plasma. The strongest correlations were observed for propionic, valeric and succinic acid. Caproic, formic, iso-valeric and iso-butyric acid all correlated moderately, and acetic acid levels correlated weakly (see **Figure 3** in Paper IV, bottom row). SCFA concentrations were not significantly correlated between mothers' plasma and mother's milk (see **Figure 3** in Paper IV, upper row).

Gut microbiota development and SCFAs in early childhood in relation to future allergy (Papers I, II, III & IV)

Gut microbiota in infancy and future allergy

In Paper 1, associations between gut microbiota development during infancy and allergy at 3 and 8 years of age were investigated. Despite few children in the FARMFLORA cohort developing allergy (N=11 and 10 at 3 and 8 years, respectively), limiting statistical power, there was an overall difference in gut microbiota composition at 1 week of age, but not at later time points, between children who were allergic and those who were not allergic at 8 years of age, as indicated by significant separation in Bray-Curtis dissimilarity index. As shown in **Table 6**, the gut microbiota was significantly less anaerobic and estimated microbial age tended to be lower at 1 week of age in infants who who later became allergic, while there was no difference in gut microbiota diversity in infants who were allergic or non-allergic at 3 or 8 years. Colonization by lactobacilli at 6 months was significantly negatively associated with allergy at 8 years of age. Conversely, higher proportions of *Rothia* at 1 week, *Clostridium sensu stricto* 1 at 1 month and *Lachnoclostridium* at 6 months, and colonization by *Rothia* at 1 week, *Ruminiclostridium* and unclassified *Peptostreptococcaceae* at 1 month and *Akkermansia* at 6 months were all significantly positively associated with allergy at either 3 or 8 years of age. Non-specific tendencies were seen for lower risk of allergy at 3 years of age with higher proportions of *Bacteroides* at 1 week and 1 month, and *Bifidobacterium* and *Escherichia* at 6 months.

Fecal and plasma SCFAs in early childhood and future allergy

The associations between the concentrations of SCFAs in faeces or plasma during infancy or early childhood and subsequent development of atopic disease was investigated in three cohorts: FARMFLORA, BAS and NICE. The results are summarized in **Table 7**.

Table 6: Gut microbiota composition up to 6 months of age in relation to future allergy development in the FARMFLORA cohort

	All samples [a]	1 week samples	1 month samples	6 month samples
Allergy 3y				
α-div:	ns.	ns.	↑*	ns.
EMA:		(↓)	ns.	ns.
Bacteria ↑		*Rothia*	Peptostreptococcaceae [b]	*Lachnoclostridium*, *Akkermansia*
Bacteria ↓		Anaerobes, (*Bacteroides*)	(*Bacteroides*)	(*Bifidobacterium*), (*Escherichia*)
Allergy 8y				
α-div:	ns.	ns.	ns.	ns.
EMA:		(↓)	ns.	ns.
Bacteria ↑			*C. sensu stricto* 1, *Ruminiclostridium*	
Bacteria ↓		Anaerobes		*Lactobacillus*

* p<0.05, Mann-Whitney U-test. Parentheses ( ) indicates a tendency for association, p>0.05 – p<0.10. ns. indicates no significant association. Bacterial groups listed were significantly associated with allergy unless indicated. [a] In pooled samples from all time points. [b] Unclassified.

FARMFLORA – fecal SCFA at 3 years versus allergy at 8 years of age

In Paper II, associations between fecal SCFA concentration at 3 years of age and allergy at 8 years of age were investigated in the FARMFLORA cohort. Median fecal valeric acid concentration at 3 years of age was lower in children who had atopic eczema at 8 years of age compared to children with no allergy/atopy diagnosis, i.e. atopic eczema, food allergy, asthma or allergic rhino-conjunctivitis (0.5 vs. 2.3 µmol/g, P = 0.007) (**Table 7** and **Figure 4**). No other significant associations were observed, but as the cohort is very small, this may not be surprising.

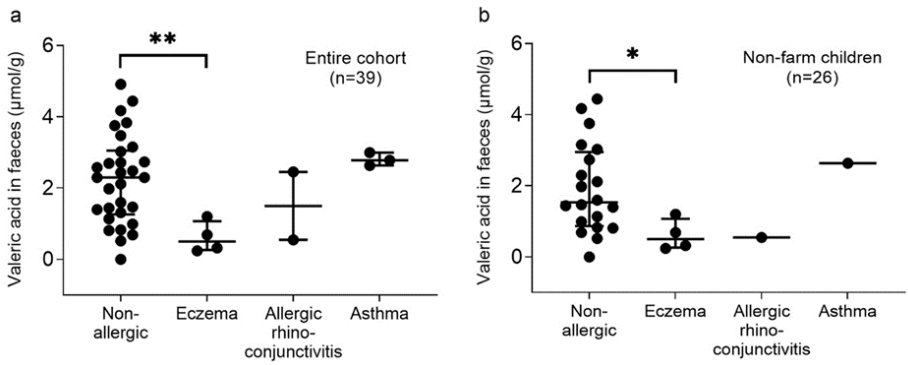

*Figure 4: Valeric acid in feces of FARMFLORA infants correlates with less future allergy development. Valeric acid concentration in faeces at 3 years of age in relation to allergic disease at 8 years of age, in (**a**) all children in the FARMFLORA cohort who were clinically evaluated for allergy at 8 years and also had fecal samples available for SCFA analysis (N=39) and (**b**) a sub-set including only control children not living on farms (N=26). Non-allergic children had no allergic/atopic disease at 8 years of age. *, p<0.05, **, p<0.01, Mann-Whitney U-test. Reproduced with permission from Gio-Batta et al., Scientific Reports, 2020 (118).*

Children growing up on farms have low rates of allergic diseases (69, 70). In the FARMFLORA cohort, farmers´ children had lower rate of allergy at 18 months and 3 years of age (206), but not at 8 years of age (208). As farmers´ children also had higher levels of fecal valeric acid at 3 years of age, a sub-group analysis was carried out in control children not living on farms to avoid possible confounding by farm residence. Again, median fecal valeric acid concentration at 3 years of age was lower in children with eczema at 8 years of age compared to non-allergic children (0.5 vs. 1.5 µmol/g, P = 0.03). No

significant associations were observed between other SCFAs and allergic/atopic diseases (no children in the FARMFLORA cohort had food allergy at 8 years of age).

Fecal SCFA at 1 year versus allergy at 13 years of age in the BAS cohort

In Paper III, associations between fecal SCFA concentrations at 1 year of age (determined in previous studies) (209, 242) and allergy at 13 years of age were investigated employing data from the BAS cohort, recruited in Jämtland in 1996-7 and evaluated for allergy at 13 years of age. In total, data from 56 children diagnosed as allergic at 13 years of age were analysed in the present study. Individuals with no reported allergic symptoms and a negative skin prick test formed the reference group (N=33). As also seen in the FARMFLORA cohort (above and **Table 7**), there was a correlation between low fecal levels of valeric acid and subsequent atopic eczema, with higher concentration of valeric acid at 1 year of age negatively associated with eczema at 13 years of age (OR 0.6, 95% CI: 0.4–1.0, p=0.049). In addition, valeric acid levels in feces tended to be negatively associated with food allergy at 13 years of age (OR 0.6, 95% CI: 0.4–1.0, p=0.057).

In sub-group analyses, fecal valeric acid concentration at 1 year of age was negatively associated with eczema and food allergy at 13 years in adolescents whose parent(s) had an allergic disease (OR 0.5, 95% CI: 0.3–1.0, p = 0.047 and OR 0.4, 95% CI: 0.2–0.9, p=0.03 respectively), but not in offspring of non-allergic parents. Valeric acid in feces at 1 year of age was also negatively associated with eczema at 13 years in girls (OR 0.4, 95% CI: 0.2–0.9, p = 0.03) but not boys.

Plasma SCFA at 4 months versus allergy at 1 year of age in the NICE cohort

In the NICE birth-cohort, associations between plasma SCFA concentrations at 4 months of age and allergy and sensitization at 1 year of age were investigated (Paper IV). As shown in **Table 7**, lower plasma levels of several SCFAs at 4 months were associated with increased prevalence of atopic eczema at 1 year of age. Acetic, succinic and isobutyric acid concentrations were significantly lower at 4 months of age in infants who had eczema at 1 year of age, compared to infants who were non-allergic and non-sensitized. Further, there were also significant associations between low plasma levels of SCFAs at 4 months of age and food allergy at 1 year of age, including acetic

Table 7: Summary of significant associations between SCFA levels and subsequent allergy development in three cohorts

	Birth cohort					
	FARMFLORA		**BAS**		**NICE**	
SCFA analysis	Feces 3 yr		Feces 1 yr		Plasma 4 mo	
Allergy evaluated	8yr		13 yr		1 yr	
Atopic eczema vs control		P [a]		P [a]		P [a]
Acetic					Lower	0.003
Succinic	NM		NM		Lower	0.03
Isobutyric					Lower	0.04
Valeric	Lower	0.03	Lower	<0.05		
Food allergy vs control						
Total					(Lower)	0.07
Acetic					Lower	0.04
Succinic	NM		NM		Lower	0.03
Valeric			(Lower)	0.06	(Lower)	0.08

[a] P-values. Only significant or near significant associations are shown, unfilled cells denote that no significant findings were found between allergic and non-allergic children. The following SCFAs were measured in all three cohorts: acetic, propionic, butyric, iso-butyric, valeric and caproic acid. In NICE, also formic and succinic acid were analyzed. In FARMFLORA and BAS, also iso-caproic acid was measured. NM= Not measured; succinic acid was only measured in the NICE cohort.

and succinic acid, and tendencies for valeric acid and total SCFAs. The associations between low levels of SCFA in infants' plasma and their tendency to develop allergy, remained when controlling for maternal allergy. No associations were observed between plasma SCFA concentration and asthma, and only two infants had allergic rhino-conjunctivitis at 1 year of age, precluding analysis.

In relation to sensitization, the plasma concentrations of formic, succinic, iso-butyric, valeric, caproic and total SCFAs at four months of age were significantly lower in infants who were sensitized at 1 year of age, compared to infants who were non-allergic and non-sensitized (see **Figure 7** in Paper IV).

As also the mothers' SCFA levels were measured in Paper IV, we could study whether allergic mothers had lower levels of SCFA in breast milk or serum,

compared to non-allergic mothers. Indeed, allergic mothers had lower concentrations of caproic acid in their breast milk than non-allergic mothers, and also tended to have lower concentrations of butyric acid (see **Table 3** in Paper IV). However, there were no associations between total or individual SCFA concentration in breast milk at 4 months postpartum and any allergic disease or sensitization in infants at 1 year of age. Allergic mothers also had lower plasma levels of several SCFA than non-allergic mothers, most consistently isobutyric and succinic acid, while isovaleric acid was generally higher in the plasma of allergic, as compared to non-allergic mothers (**Table 3** in Paper IV).

Oral, gut and skin microbiota establishment in extremely preterm infants in the ÖFLORA cohort (Paper V)

Microbiota composition up to 6 weeks of postnatal age

In Paper V, the establishment of the oral, gut and skin microbiota was studied in extremely preterm infants (the ÖFLORA cohort) who were followed up to six weeks of postnatal age during their stay at a neonatal intensive care unit. Microbial colonization was rapid (see **Figure 1** in Paper V), with detectable levels of live bacteria found on at least one body site in most infants on the day of birth. The mouth was the most common site of this early colonization. By 3-5 days after birth, almost all infants were colonized by bacteria at all three body sites, and around a third also harbored yeasts at one or more locations.

Considering major microbial groups (see **Figure 1** in Paper V), facultatively anaerobic bacteria were the first to establish and were the most prevalent group at all body sites throughout the first 6 postnatal weeks. Yeasts were also prevalent, being detected in the mouth and gut of over half of infants at 2-4 weeks of postnatal age, and up to 40% harbored yeasts on the skin. In contrast, obligate anaerobic bacteria were acquired late or not at all. Fewer than half of the infants had acquired any gut anaerobes by 4-6 weeks after birth, and anaerobes were scarce in oral and skin samples. Facultative

anaerobes were also the most abundant microbial group at all body sites throughout the study period. In the gut, mean population counts were around 10^{10} CFU/g or higher in colonized infants from a few days after birth through to 6 weeks postnatal age, while population counts of anaerobic bacteria were lower and only reached 10^{10} CFU/g of feces by 6 weeks of postnatal age. No absolute quantification of microbial population counts was possible in oral and skin swab samples, since the amount of swab sample material could not be carefully standardized. Still, bacterial counts per ml sample transport medium were calculated, but interpreted with caution.

As shown in **Table 3** in Paper V, coagulase-negative staphylococci (CoNS) were the most prevalent facultative anaerobes, colonizing the majority of infants over the first 4 postnatal weeks across all body sites. In the mouth, *S. aureus* and streptococci (mostly *S. mitis*) were the next most prevalent bacterial groups, and enterococci (mostly *E. faecalis*) were also quite prevalent. In the gut, enterococci were the second most prevalent bacterial group and became increasingly common over time, colonizing almost all infants by 6 weeks postnatal age, and *S. aureus* was also quite prevalent at the same time point. On the skin, CoNS remained dominant throughout the first 6 postnatal weeks, although *S. aureus*, *Enterococcus* and *Corynebacterium* each colonized up to around a third of infants. In contrast to Gram-positive facultative anaerobes, Gram-negative groups were less prevalent, especially at earlier time points. However, in the mouth *Haemophilus* and *Klebsiella* colonized over a third of infants by 6 weeks postnatal age, while in the gut *Klebsiella* was detected in two thirds of the infants and *Enterobacter* and *E. coli* were also somewhat prevalent at the same time point. In terms of abundance, staphylococci, streptococci and enterococci were each found at similar mean population counts in the mouth of colonized infants (see **Supplemental Table S3** in Paper V), while in the gut enterococci, *Klebsiella*, *Enterobacter* and *E. coli* all reached mean population counts of up to $10^{10\text{-}11}$ CFU/g in colonized infants (see **Table 4** in Paper V). With respect to obligate anaerobic bacteria, a quarter of the infants were colonized with *Clostridium* and *Bifidobacterium* in the gut by 6 weeks of postnatal age (see **Table 3** in Paper V), and genera such as *Prevotella*, *Propionibacterium*, *Finegoldia* and *Anaerococcus* were detected in oral and skin samples from a few infants at later time points.

Neonatal exposures and microbiota assembly

Associations between the following perinatal factors and the oral, gut and skin microbiota were also investigated: *in-utero* exposure to clinical chorioamnionitis and maternal antibiotics, latency period, gestational age at birth, delivery mode, APGAR score, and postnatal treatment with antibiotics, antifungals and steroids. Orthogonal projection to latent structures (OPLS) analysis was carried out including either the prevalence or population counts of each microbial group at each time point as the X-variables and the perinatal factor of interest as the Y-variable, and adjustment for false discovery rate (FDR) was applied to associations identified for each perinatal factor. Significant FDR-adjusted associations were identified in relation to postnatal treatment with antibiotics and gestational age at birth, although not in relation to other perinatal factors.

Antibiotics

Antibiotics may alter commensal colonization and antibiotics are prescribed very liberally to extremely preterm infants, both as prophylaxis, and as treatment on suspicion of sepsis or other infections. Investigating the effect of antibiotics in preterm infants is challenging due to their high levels of antibiotic exposure. In this study population, all infants received antibiotic prophylaxis (penicillin G and tobramycin, followed by cloxacillin) until at least 1 week of age, and broad-spectrum antibiotics such as vancomycin and/or meropenem were administered therapeutically as required in the case of infections. OPLS analysis was used to identify associations between antibiotic variables and microbial colonization, followed by univariate analyses for each association identified and subsequent correction for false discovery rate (FRD). Firstly, microbial colonization was compared in infants who received antibiotic prophylaxis only and infants who also received 'therapeutic' antibiotics in the first week after delivery (see **Figure 6 a**-**b** in Paper V). Infants in the latter group had significantly lower population counts of CoNS in the mouth at 1 week of age, but higher population counts of facultative anaerobes on the skin in the first week after delivery, possibly preceding treatment. Although not significant after FDR-adjustment, infants who received 'therapeutic' antibiotics were also more frequently colonized by yeast on the skin. As this colonization pattern was also observed in infants born at lower gestational age (as detailed below), we applied stratified analysis to investigate further. Skin colonization by yeast occurred almost

exclusively in infants born at <26 weeks of gestation, and within this group it was significantly more common in infants who received 'therapeutic' antibiotics in their first week compared to infants who received only prophylaxis (see **Figure 6 c-d** in Paper V).

Secondly, the effect of antibiotic compared to no antibiotic treatment on microbial colonization after the first postnatal week was investigated (see **Figure 3** in Paper V). Samples obtained from 11 infants after cessation of antibiotics were compared with those from infants who remained on antibiotics at all sampling time points until death or discharge. Samples from 8 infants who ceased antibiotics but were later treated again were excluded. Antibiotic treatment was associated with significantly lower prevalence of *S. aureus* and *S. mitis* in the mouth; *S. aureus*, *E. coli*, anaerobic bacteria and *Bifidobacterium* in the gut; and *S. aureus*, *E. faecalis* and anaerobic bacteria on the skin. Antibiotic treatment was also significantly associated with lower population counts of facultative anaerobes in the mouth and gut, and in particular with lower counts of enterococci in the gut and CoNS on the skin.

Prematurity

Infants in this study were born between 22^{+6} and 27^{+5} weeks of gestational age, and the effect of the degree of prematurity on microbial colonization was investigated (see **Figure 4** in Paper V). The most preterm infants, born at <26 weeks of gestation, were less frequently colonized by obligate anaerobic bacteria, including *C. perfringens*, in the gut by 4 weeks of age as compared to infants born at >26 weeks of gestation, and they were more often colonized by yeasts on the skin. In addition, infants born at <26 weeks of gestation had lower population counts of enterococci in the gut (if colonized) at 4 weeks of postnatal age, but higher population counts of facultative anaerobes including CoNS on the skin at 1 week. Although not significant after FDR-adjustment, infants born at <26 weeks of gestation also showed delayed colonization by *S. mitis* in the mouth and bifidobacteria in the gut, and were more frequently colonized by *S. aureus* on the skin during the first week after birth (see **Figure 4 e-g** in Paper V).

As colonization of the gut by anaerobes was reduced in infants born at lower gestational age and also in those treated with antibiotics in the week preceding sampling, we again applied stratified analysis to investigate further. Colonization by anaerobes was generally uncommon in antibiotic-treated

infants regardless of the degree of prematurity, while in infants not treated with antibiotics colonization by bifidobacteria occurred almost exclusively in those born at ≥26 weeks of gestation (see **Figure 5** in Paper V).

Discussion

Allergy has increased dramatically in prevalence in industrialized regions, particularly between 1970 and 1990, when the incidence of different allergic diseases tripled in many Western countries, including Sweden (243), whereas lower rates of allergy are seen in non-industrialized regions and also in traditional communities within industrialized countries such as the Amish (244, 245), Old Order Mennonites (246) and followers of an anthroposophic lifestyle (247). One theory which has been suggested to account for this increase is the hygiene hypothesis, which proposes that progressive loss of microorganism diversity and turnover is responsible for poor control of unwanted immune responses and increased rates of allergic diseases (248), but also inflammatory bowel disease and organ-specific autoimmunity (249). Exposure to farming environment and to pets, larger family sizes and poor sanitation are the factors that have been most convincingly linked to reduced risk of allergy development and are also likely to be associated with greater exposure to a wide variety of microbes. Conversely, modern lifestyle factors such as delivery by caesarian section (250) and exposure to antibiotics (251) are associated with increased risk of allergy.

To investigate the effect of several of these exposures on the bacterial colonization of the infant gut, the establishment of the gut microbiota up to six months of age was characterized by 16S rRNA gene sequencing, in a cohort of infants living in rural Sweden. Infants from families living in the same rural regions, but not on farms served as controls. In the cohort as a whole, *Bifidobacterium* and *Bacteroides* were abundant across the whole time period, with bifidobacteria the most abundant genus initially, and *Bacteroides* the most abundant genus by 6 months of age. *Bacteroides* is a major member of both infants' and adults' colonic microbiota and chiefly derives from the maternal fecal flora which is transferred during a vaginal delivery, as shown by the fact that infants delivered by caesarean section may show a paucity of *Bacteriodes* still at six or twelve months of age (13, 31). Bifidobacteria are a dominant member of the gut microbiota in infants and also found in adults, but usually in lower population levels than *Bacteroides* spp and obligately anaerobic members of the Firmicutes phylum. Bifidobacteria are also more easily acquired from non-maternal sources than *Bacteroides* and sectio-delivered infants rapidly catch up with respect to bifidobacterial colonization

(13, 51). As well as a consistent abundance of *Bifidobacterium* and *Bacteroides* throughout the first six months there was also a clear progression in gut microbiota composition over time, characterized by a decline in the early abundance of facultative anaerobes and an expansion of numerous genera of strict anaerobes, mostly clostridial bacteria and other anaerobic Firmicutes but also other groups such as *Akkermansia*, with a corresponding increase in bacterial richness. These observations correspond with findings from previous studies showing a similar development trajectory of the gut microbiota (13, 30, 252). However, it should be noted that the proportion of anaerobes in the infant microbiota is likely greater than was shown in our analysis. Fecal samples from the FARMFLORA cohort were also cultured (Ljung *et al*, manuscript) and by this method the proportion of anaerobes was found to be >90% at 1 week of age, compared to around 50% by 16S rRNA analysis, even though bacterial culture does not allow detection of many groups of strict anaerobes. Probably this can be explained by the fact that Gram-positives such as bifidobacteria are less well detected by 16S rRNA gene sequencing (7).

As infants experience numerous overlapping exposures in early life, we considered the cumulative effect of 'microbe-enriching' exposures (vaginal delivery, farm living, pets, elder siblings, formula/weaning, and lack of intrapartum exposure to antibiotics) on the infant gut microbiota. The greater the number of 'microbe-enriching' exposures an infant experienced, the higher the proportion of obligate anaerobic bacteria and especially bifidobacteria in their gut microbiota, particularly in the first month after birth. Our findings correspond with previous studies which found that bifidobacteria are more abundant in the gut microbiota of infants in non-industrialized compared to industrialized countries (253-256), and are also enriched in the gut microbiota of Old Order Mennonite infants (257). We also considered the effect of these 'microbe-enriching' exposures on estimated microbial age (40), a measure of the maturation of the gut microbiota from a simple one early in life to a complex anaerobe-dominated microbiota more like the one found in adults, and found that most of the above exposures were also associated or showed a tendency to be associated with higher estimated microbial age. In turn, higher estimated microbial age was associated with lower populations of facultative anaerobes such as streptococci and staphylococci and also *Clostridium sensu stricto* 1 (which includes *C. disporicum*, *C. perfringens* and *C. botulinum*), especially during the first

month. Analysis by 16S rRNA gene sequencing does not make it possible to determine the species, but our previous studies have shown *C. disporicum* to be the most common clostridial species in Swedish infants' microbiota. Higher estimated microbial age was also positively associated with increased populations of various strictly anaerobic clostridial bacteria, especially at 1-6 months of age.

In addition, 'microbe-enriching' exposures were individually associated with enrichment of various core anaerobic groups in the infant gut, among other effects. Independent of covariates, elder siblings were associated with higher levels of bifidobacteria but lower levels of enterococci and also lower levels of staphylococci at 6 months of age, and growing up on a farm was associated with accelerated colonization by lactobacilli. Furthermore, growing up with pets was independently associated with enrichment of anaerobic bacteria but lower levels of *Clostridium sensu stricto* 1, while weaning (defined as the cessation of breastfeeding) heralded the expansion of diverse clostridial bacteria. These findings correspond with previous studies showing enrichment of bifidobacteria, particularly *B. infantis* and *B. breve*, in children with elder siblings (63, 257, 258), and a more pronounced effect of farm living on overall gut microbiota composition from around one year of age onwards, including a more mature gut microbiota at one year of age in children growing up on farms (34, 40). In contrast, delivery by caesarean section and exposure to intrapartum antibiotics were associated with depletion of bifidobacteria and/or *Bacteroides* and enrichment of facultatives and *Clostridium sensu stricto* 1, among other effects, consistent with previous findings (13, 259). Therefore, together with others, we show that traditional lifestyle exposures overall confer a more anaerobic gut microbiota composition, enriched in bifidobacteria and other core anaerobic taxa of the early infant gut microbiota, and showing an accelerated trajectory towards a mature adult-type microbiota characterized by strictly anaerobic Firmicutes.

At 1 week of age there was an overall separation in gut microbiota composition between infants who did and did not have allergy at 8 years, characterized by a higher proportion of obligate anaerobes (OR 0.6) in the latter group, and increased colonization by lactobacilli at 6 months of age was negatively associated with allergy in later childhood. The "estimated microbial age" index tended to be higher at 1 week of age in infants who did not develop allergy in their first 8 years of life (OR 0.5-0.6), and reduced odds

of allergy were also observed for higher proportions of other typical infant commensals such as *Bifidobacterium*, *Bacteroides* and *Escherichia*, although not reaching statistical significance. Thus, despite the low statistical power in this study due to few children in the cohort developing allergy (N≤11), there was a clear relationship between early life gut microbiota composition and later allergy development. Our findings with respect to lactobacilli (260-263) and bifidobacteria (264-271) are consistent with numerous previous studies and also the conclusion of a recent systematic review (272), although some studies have shown enrichment (273, 274), no effect (275) or mixed effect of different species (276, 277) of bifidobacteria in association with allergy. Also, higher relative abundance of *Bacteroides* and *Parabacteroides* at 2 months of age were also associated with protection from subsequent asthma development in another rural cohort including many farm infants (40). Therefore, our findings combine with others to indicate an allergy-protective effect of an early gut microbiota rich in typical anaerobic commensals. Consistent with the hygiene hypothesis, this microbiota composition appears to be promoted by various traditional lifestyle exposures that increase microbial exposure. As well as identifying features of the early gut microbiota associated with protection from allergy, we also observed that various bacterial groups were positively associated with allergy development, and most of these - namely higher levels of *Clostridium sensu stricto* 1 and unclassified Peptostreptoccaceae at 1 month and *Lachnoclostridium* at 6 months of age - were also associated with less exclusive or prolonged breastfeeding. Since longer duration of exclusive breastfeeding was associated with lower rates of allergy in this cohort (217), these bacterial groups may be markers of this relationship.

Ever since the hygiene hypothesis was formulated in 1989, there has been a continuous search to pinpoint which microbial exposures that are beneficial regarding maturation of the immune system and which mechanisms that operate. In principle, a few selected microbes could have this propensity, or it is the overall complexity of the microbiota, or a general high turn-over of the microbiota, reflecting constant exposure to new strains that could activate the immune system, that is important. One candidate for the link between a complex microbiota dominated by anaerobes and a beneficial immune maturation during the critical period of early infancy would be the SCFAs produced in the infant's colon. SCFAs are highly bioactive metabolites with a range of immunoregulatory effects which are produced by the gut

microbiota mostly through anaerobic fermentation of complex carbohydrates and proteins, and therefore act as intermediaries between the gut microbiota and immune system (141, 142). Fecal SCFA production changes with age in parallel with the maturation of the colonic microbiota. In analyses not including formic or lactic acid, acetic acid (2C) is the predominant SCFA initially, followed by propionic (3C) and butyric acid (4C), with the longer and branched SCFAs, isobutyric (i4C), valeric (5C), isovaleric (i5C) and caproic (6C) acid, increasingly produced from the end of the first year although only in small amounts (209, 278). This corresponds with a progression from a more simple gut microbiota in early months to an increasingly anaerobic gut environment containing an expanding diversity of obligate anaerobes able to produce these latter SCFAs, some of them through cross-feeding pathways which rely on existing production of shorter SCFAs by earlier colonizers (102, 105, 279, 280). We could replicate this pattern and in addition we could demonstrate that two exposures were associated with a more mature microbiota SCFA pattern, namely growing up on a farm and having elder siblings. Farm children had higher levels of iso-butyric, iso-valeric and valeric acid at 3 years of age than rural controls, which, to the best of our knowledge is a new finding. Also, children with elder siblings had higher levels of valeric acid at 3 years of age, which was independent of the farm effect and also shown previously in the BAS cohort (209). Further, higher concentration of isobutyric, isovaleric and valeric acid at 3 years of age correlated with a higher proportion of anaerobes in the gut microbiota at 1 week, but not 1 month or 6 months, and therefore higher levels of these SCFAs reflect a distinct development trajectory of the microbiota beginning at or shortly after birth.

The branched SCFAs (isobutyric and isovaleric acid) are formed from fermentation of amino acids and, indeed, we saw a positive association between intake of protein and the concentration of these SCFAs in the feces, while intake of fruit and vegetables at one year of age was negatively associated with the fecal concentrations of these SCFAs in feces at age 3 years. Probably this is because a diet rich in fruit and vegetables likely includes less protein. Unexpectedly, there was also no effect of dietary fibre intake on plasma levels of any individual SCFA in the NICE study, and instead consumption of refined, carbohydrate-rich foods was associated with higher plasma levels of some SCFA, with reported intake of biscuits, rusks and cookies correlating positively with plasma butyric acid, for reasons which are unclear. Although it is popular to link high-fiber diets to production of SCFA,

a study actually showed very little effect on the SCFA pattern of feeding a diet enriched in fiber (281). Actually, there is plenty of substrate in the colon for fermentation (mucus, sloughed epithelial cells) and the contribution from non-fermented fibers and resistant starch may not be so significant that it alters the SCFA production. It is estimated that some 250 g of epithelial cells and about 10 L of mucus, of which 5-10% is solids (72), are shed daily into the intestinal lumen (of an adult) which is, of course, far more than the amounts of non-digested carbohydrate.

The SCFAs are produced in the colon because the fermentation process requires an anaerobic atmosphere. In order to influence the infant's immune system, SCFAs need to be taken up into the circulation and distributed between the cells and tissues of the body. Several transporters ensure that the SCFAs are taken up into colonic epithelial cells. Butyric acid in particular is used as an energy source of these cells (117), but SCFAs are also transported via transporters on the basolateral side along the concentration gradient into the blood. Via the portal circulation, they reach the liver where propionic acid is consumed as a substrate for glycolysis (117). The remaining portion of SCFAs enter circulation where they interact with peripheral organs, tissues and cells (117, 119, 120, 282, 283). The half-life of SCFA in the circulation is only a couple of hours (284) showing that they are consumed rapidly after having been absorbed.

SCFA in plasma have been much less studied than SCFA in the colonic contents, because the concentrations are about a thousand-fold lower, in the microgram per liter range, rather than the milligram per liter range in the colon. Using highly sensitive ultra-high performance liquid chromatography-mass spectroscopy which allowed detection of micromolar plasma SCFA concentrations, we measured plasma concentrations of nine SCFAs in >100 pairs of 4-month-old infants and their mothers in the NICE cohort and, in addition, in the mothers' breast milk. To the best of our knowledge, this is the first study examining samples from infants and their mothers simultaneously, and also to measure SCFA both in blood (plasma) and breast milk of the same women. In addition to the SCFAs analyzed in the BAS and FARMFLORA cohorts, we included two additional SCFAs, namely formic acid/formate, the smallest SCFA with only one hydrogen atom in addition to the carboxylic group, and succinic acid. Succinic acid has two carboxyl groups and is, hence, formally not a SCFA, but it is similar small molecule produced as a metabolite

of the colonic microbiota. It is also a molecule in the citric acid (Krebs) cycle and, thus, a molecule produced also by ourselves. Formic acid (not commonly being included in studies of SCFAs) was the most abundant SCFA in the plasma of 4-month-old infants, followed by acetic and propionic acid, while butyric acid and the longer and branched SCFAs were found in very low concentrations. Comparison of concentrations of acetic acid in feces (Paper II) and plasma (Paper IV) confirmed that levels were approximately a thousand times lower in plasma, consistent with previous comparisons of fecal and plasma SCFA levels in adults (117, 121).

The colonic microbiota is comparatively simple during the first weeks and months and the SCFAs produced are mainly formic and acetic acid, providing SCFAs to the infant, SCFAs may be also provided by particular foods, including cow's milk and other dairy products (116), as well as mother's milk (122, 285-287). Ingested SCFAs are absorbed rapidly in the upper gastrointestinal tract, and we therefore considered whether levels of SCFAs consumed in breast milk were related to the infant's plasma SCFA levels. However, no significant positive correlation was observed between concentrations of any individual SCFA in breast milk and infant plasma among the whole cohort, and only weak correlations when considering exclusively breastfed infants only. The amount of SCFAs consumed in breast milk may be too low, as compared to the amounts produced by the infant's own colonic microbiota, to have any influence on the plasma levels. Another explanation, not mutually exclusive, is that most of the SCFAs supplemented via maternal milk is used by the epithelia and do not end up in the blood plasma. Nevertheless, we assume that milk SCFA may provide an important nutrient for the infant's mucosal epithelia in the oral cavity and upper gastro-intestinal tract and, perhaps, may even aid in shaping the immune function of the infant, e.g. by promoting maturation of naïve T cells into induced Treg, a function carried out in the mucosa-associated lymphoid tissue (288).

A striking finding was that, when SCFA concentrations were compared between mothers' breast milk and plasma, butyric and caproic acid were found at 100-fold higher concentrations in breast milk, and valeric and isobutyric acid were also enriched although to a considerably lesser extent, while levels of the other SCFAs and also the total SCFA concentration were up to 10 times lower. These findings are in accordance with previous studies which reported concentrations of formic, acetic, and butyric acid in breast

milk (122, 285-287) with formic acid being the most abundant milk SCFA, followed by butyric and caproic acid, then acetic and succinic acid, and then the other SCFAs. Thus, the SCFA distribution was very different to that found in either plasma or feces, and higher levels of butyric acid in breast milk as compared to serum have been previously noted when comparing findings from two separate studies (289). However, to the best of our knowledge, our study is the first to directly compare the SCFA in both breast milk and serum in the same woman and clearly demonstrating that butyrate and caproate must be either locally produced in the mammary epithelium, or selectively transported into the milk.

As breast milk has been the main source of nutrition for infants during evolution, it seems plausible that at least some SCFAs are deliberately provided as nutrients for the developing infant, for example to promote epithelial tissue integrity and barrier function. It is also possible that the immune system surrounding the oral cavity and the inductive sites for mucosal responses in the small intestine could be influenced (95). Although no active transport protein is so far known to bring SCFAs into the mammary gland, one possible candidate may be the breast cancer resistance protein (BCRP), which is known to transport butyric acid into colonic epithelial cells (290) and is expressed in alveolar epithelial cells during late pregnancy and lactation where it pumps a variety of small compounds against a concentration gradient (291). Supporting this hypothesis, mutations in BCRP/ABCG2 have major effects on yield and fat composition of cow's milk (292). Furthermore, BCRP is expressed more in the duodenum than in the colon (293), which also suggests that breast milk SCFAs may nourish the upper gastrointestinal epithelia, which would otherwise be less exposed to SCFAs than the colonic epithelia where most bacterial SCFA production occurs. This notion is supported by the fact that butyrate and caproate in cow's milk are esterified almost entirely to the *sn*-3 position of the triacylglycerol, which is then preferentially hydrolysed by lingual lipases in the mouth and lingual and gastric lipases in the stomach, resulting in targeted release of these SCFAs in the upper gastrointestinal tract (294), although no analogous esterification pattern has been shown to date in humans. Finally, butyric and caproic acid are produced at low levels by the early infant microbiota, as shown above, further supporting a hypothesis of their deliberate enrichment in breast milk.

Another striking finding was the consistent and often strong correlations between infant and maternal levels of all SCFAs, with the sole exception of butyric acid. This correlation cannot be explained by similar colonic production of SCFAs, as the gut microbiota composition of young infants is substantially different to that of adults, with infants lacking the diversity of strict anaerobes characteristic of the adult microbiota and consequently generates a different distribution of SCFAs, as described above. Furthermore, the diet of young infants and their mothers is quite different, especially since almost all the infants in this cohort were breastfed (78% exclusively and 17% partially). Likewise, transport of SCFAs from mother to infant during pregnancy (295) cannot explain a correlation between infant and maternal plasma SCFA levels at four months postpartum, as the half-life of SCFAs in plasma is just a couple of hours (284). As well as being produced by the microbiota and consumed in the diet, endogenous production of SCFAs also occurs in the body. Formic acid is a source molecule for the one-carbon metabolism, and a whole-body loop exists for its endogenous production and regeneration (123, 124). Also, acetic acid in the form of acetyl coenzyme-A is a fundamental molecule in cellular metabolism (296), and succinic acid is an intermediate in the Krebs cycle (125). Therefore, it appears that tightly regulated uptake and/or endogenous production might keep plasma SCFA within certain limits, which may also be under genetic control. This notion is supported by unpublished data from a pilot study of a few children in the Farmflora cohort (not included in this thesis), which showed relatively constant concentrations of individual SCFAs in the plasma of children with increasing age (at birth (cord blood), four months and three years of age).

The SCFA valeric acid, one of the longer SCFAs not produced in significant quantities until the microbiota has reached a certain level of complexity and anaerobiosis, was negatively related to allergy in all our three examined cohorts. Fecal valeric acid levels at 3 years of age were negatively associated with eczema at 8 years of age in FARMFLORA (Paper II) and fecal valeric acid at 12 months was negatively associated with eczema and possibly also food allergy at age 13 in the BAS cohort. Lastly, plasma valeric acid levels (along with other SCFAs) at 4 months of age were negatively correlated with sensitization and possibly also with food allergy at 12 months of age in the NICE cohort. Interestingly, median fecal valeric acid concentration was higher in children living on farms than in healthy non-farm children, and also higher in children with siblings, another factor that is associated with reduced risk of

allergy development (156). Rather scanty information is available regarding the effects of valeric acid, as most focus has been on butyric acid. As for other SCFAs, it is antimicrobial and therefore able to suppress populations of facultative anaerobes (140) and supports epithelial barrier function including upregulation of tight junction proteins (132, 297). Also, it upregulates keratin, which is a component of the intermediary filaments that are necessary to keep the shape and strength of epithelial cells (132). This is interesting with respect to valeric acid being particularly associated with the skin condition eczema, not only in the very young infants where this is the most common atopic condition, but also among the 13-years old adolescents in the BAS study. Although no *in-vivo* studies have been carried out to investigate an effect of valeric acid on tolerance development or allergy, it has anti-inflammatory effects as administration of valeric acid or its esters to animals protects against colitis (132) and necrotic enteritis (135). Further, transfer of valerate-treated regulatory B cells together with naive T cells into Rag1-deficient mice protects against colitis and experimental autoimmune encephalitis (147). Although a first wave of research into SCFAs focused primarily on acetic, propionic and butyric acid, as these are the most abundant SCFAs in the gut, interest in valeric acid has bloomed and valeric acid has now also been shown in rodents to heal radiation injuries (132), combat liver cancer (298), reduce liver cholesterol (299) and lower blood pressure (119), illustrating its systemic effects and further reinforcing its importance in the body.

Apart from valeric acid, we observed that plasma concentrations of formic, succinic, iso-butyric, valeric and caproic acid in 4-month-old infants in the NICE cohort were negatively associated with sensitization at 1 year of age; acetic, succinic and iso-butyric acid were negatively associated with atopic eczema; and acetic and succinic acid were negatively associated with food allergy, with valeric acid and total SCFA concentration also tending to be negatively associated with food allergy. It cannot be excluded that there is a degree of reverse causation involved in these associations as allergic mothers had lower plasma concentrations of succinic, isobutyric, and to a lesser extent formic, propionic and total SCFAs compared to non-allergic mothers. In a family with allergic persons, the diet and lifestyle might be affected. Furthermore, an infant that has displayed certain signs of potential allergy may be put on hypoallergenic formula. An interesting finding was that the allergic mothers had lower levels of only two SCFAs in the breast milk, namely

butyric and caproic acid, the two SCFAs that were hundredfold enriched in milk compared to blood plasma. However, we observed no beneficial effect on allergy in infants whose mothers had more of any particular SCFAs in breast milk.

Extremely preterm infants, born before 28 weeks of gestational age, live in an environment that lacks many of the types of exposures that promotes a rich microbiota. Not surprisingly, they had a markedly different gut microbiota as compared to infants born at term in the FARMFLORA cohort where obligate anaerobes made up around half of the total bacterial abundance (or >90% as determined by bacterial culture (Ljung *et al*, manuscript)) and became more dominant over time. In contrast, the gut microbiota of extremely preterm infants was almost entirely composed of facultative anaerobes until at least six weeks of age, as has also been observed previously (176). Coagulase-negative staphylococci (CoNS) colonized almost all extremely preterm infants in the gut within a few days after birth and were highly prevalent over the following weeks. Colonization by *Enterococcus*, *S. aureus* and *Klebsiella* increased over time, as also did colonization by *Enterobacter* and *E. coli* to a lesser extent, although *Enterococcus* and *E. coli* were slower to establish than in term infants (when compared to the microbiota of term infants in the ALLERGYFLORA (31) and FARMFLORA cohorts (Ljung *et al.* in manuscript), studied using similar culture protocols), possibly reflecting more limited exposure to these bacteria in the NICU. From around 1 week after birth the infant gut microbiota was often dominated by one of these groups, although the composition was unstable and the dominant microbe(s) often changed between sampling time points a few days or a week apart. This instability might be due to antibiotic treatment (300), which was the factor that had the broadest impact on bacterial colonization pattern. Although total population counts of facultatives were similar in extremely preterm and term infants, enterococci were around 100 times more abundant in the extremely preterm gut, as also were *Klebsiella* and *S. aureus* by around 1 month of age although not before. Yeasts, namely *Candida* and *Malassezia*, were much more common in extremely preterm than in term infants, and were in fact more common than obligate anerobes, which colonized fewer than half of infants by one month of age. Among obligate anaerobes, spore formers from the genus *Clostridium*, most commonly *C. perfringens*, were more prevalent than bifidobacteria. Clostridial spores are ubiquitous in all environments, including in hospital and other very hygienic environments. In the absence of

colonization by bacteria transferred via feces, clostridia will be the dominant anaerobes (13).

Our study also characterized the oral and skin microbiota of extremely preterm infants. As in the gut, CoNS were by far the most prevalent initial colonizers, indicating the transfer of skin bacteria from the parents, other family and staff. CoNS remained the most common microbial group on the skin up to six weeks of age, consistent with their high abundance in the healthy adult skin microbiota (10, 23, 24). In the mouth, although CoNS remained common, colonization by *S. aureus* and enterococci increased over time, as also did colonization by streptococci and in particular *S. mitis*, a typical inhabitant of the oral cavity (10, 11). Colonization by yeasts was common in both locations.

The unique microbiota of extremely preterm infants could be the result of many factors, including developmental immaturity, medical treatments, higher rates of caesarean delivery in preterm infants, exposure to hygiene-resistant bacteria in the hospital environment and lack of normal contact with family members. We found that antibiotic treatment, that was extremely prevalent in these infants, was the most influential factor with respect to the "NICU colonization pattern". Thus, infants who were still on antibiotics when the sample was taken had a microbiota that differed in many aspects from infants who no longer received antibiotics. Apart from this, and independent of antibiotic treatment, was an effect likely due to prematurity *per se*, giving a more "immature" microbiota with less colonization by bifidobacteria. This was not due to an effect of caesarean section, because, quite to the contrary, the most extremely premature infants were almost always delivered vaginally, while those who were a little less premature were more often delivered by caesarean section, in line with the clinical guidelines in place in the NICU at Sahlgrenska University Hospital. Other factors, including *in-utero* exposure to clinical chorioamnionitis, maternal antibiotics, latency period, APGAR score and postnatal treatment with antifungals or steroids did not substantially influence gut microbiota composition, but a larger study might be needed to reveal an effect of these factors.

Necrotizing enterocolitis (NEC) is a devastating disease of prematurity and is believed to arise from the immature physiology of the preterm intestine and immune system combined with a particular microbiota. NEC is often

preceded by an abundance of Proteobacteria in the gut (176, 177, 195) and the paucity of anaerobes in the microbiota of extremely preterm infants enable the facultatives to attain high population numbers. This was evident in our cohort, where *Klebsiella* was common in the gut microbiota and was also associated with subsequent NEC development (data not in thesis). Also, in order to increase colonization by anaerobes, certain probiotics including bifidobacteria are currently given to all preterm infants in Sweden born at >32 weeks of gestation as a means to reduce the risk of development of NEC (176, 197). Late-onset sepsis is also linked to microbiota composition. CoNS is a dominant cause of late-onset sepsis and may derive from the skin microbiota via indwelling catheters but translocation of CoNS from the gut microbiota has also been suggested (199-201). CoNS were the major cause of late-onset sepsis also in our cohort. An advantage of our study was that we studied commensal colonization not only in the gut, or on the skin, but simultaneously in the oral cavity, the gut and on the skin. CoNS were dominant across all body sites, and almost one in four infants was diagnosed with CoNS septicaemia, and a majority of the CoNS septicaemia strains could be traced back to the commensal microbiota of the skin, oral cavity and/or gut of the infected infant (Nowrouzian *et al*, in manuscript). The absence of a complex anaerobic microbiota facilitates the ability of dominant facultatives to translocate from the gut to the mesenteric lymph nodes and from there to the blood stream. Translocation over the oral mucosa has been less studied. However, in an animal model, translocation of oral bacteria to cervical lymph nodes was at least as frequent as translocation of gut bacteria to mesenteric lymph nodes (202). Thus, translocation of oral bacteria could possibly also result in septicaemia in preterm infants. Interestingly, we found that oral colonization by a certain group of CoNS was increased in extremely preterm infants who developed CoNS septicaemia (Nowrouzian *et al*, in manuscript).

Conclusion and future perspectives

In summary, our studies have shown effects of traditional lifestyle exposures on the development of the gut microbiota and fecal SCFAs during early life in healthy infants born at term, some of which were linked in turn with lower risk of allergy in later childhood. These findings help bring us closer to establishing a molecular basis for the hygiene hypothesis. Also, our findings of enrichment of certain SCFAs in the mothers' breast milk, although not directly linked to protection from allergy in our cohort, suggest a potential mechanism for health promotion and induction of immune tolerance in the infant. Interesting next steps would be to carry out *in-vivo* studies to evaluate the effect of oral supplementation of butyric, valeric and caproic acid on eczema, typically the first allergic manifestation in childhood.

Our studies have also provided important baseline information on the development of the microbiota in extremely preterm infants, who are at high risk of morbidity and mortality from microbiota-related diseases of prematurity. We identified a pronounced effect of antibiotics on microbial colonization, supporting current efforts towards antibiotic stewardship. We also noted that colonization often began in the mouth (and in a parallel study in this cohort showed that oral colonization by certain microbes was associated with sepsis), suggesting that future interventions to modulate the microbiota of extremely preterm infants consider targeting the oral cavity. Finally, and consistent with previous knowledge, we observed that facultative anaerobes dominated the extremely preterm gut, likely contributing to pathogenesis of NEC and sepsis. Interesting next steps would be to carry out *in-vivo* studies to evaluate the effect of measures to suppress populations of facultative bacteria on these diseases, for example by enteral supplementation with antimicrobial SCFAs.

Acknowledgement

I am grateful to many people who have made this work possible, and especially I would like to thank:

The infants, children and their families who took part in the studies included in this thesis, and the study teams involved.

Agnes, Fei and Ia for your exceptional kindness and support. Thanks for the opportunity to work on these fascinating studies, and for your guidance, enthusiasm, scientific mentorship and always-open doors. I feel very lucky to have you as my supervisors and am deeply grateful.

Hardis for fun times working together as my immunological partner in crime, your off-site meetings are the best. Forough, for very nice discussions about the Öflora study data and many other things besides.

All my collaborators, including Ann-Sofie Sandberg, Malin Barman, Anders Elfvin, Thordur Thordarson and Anna Sandin, for such nice collaborative experiences. Thanks also to Assi and Malin for the warm welcome with fika for project meetings.

Rikard Fristedt for excellent analysis of short chain fatty acids in plasma and breast milk.

Mia Stråvik for cheerful and prompt assistance with nutritional data.

Sanna Abrahamsson at Gothenburg University's Bioinformatics Core Facility for the analysis of estimated microbial age and patience with my many follow-up emails.

Henrik Imberg at Statistiska Konsultgruppen for expert statistical advice, and Jens Michelsen at Statistiska Konsultgruppen for producing an aesthetically pleasing network analysis figure.

Sanuja Ahammu for expertise in arrangement of 16S rRNA gene sequencing data for analysis using MicrobiomeAnalyst, and bringing in delicious celebratory meals to share.

Anneli for valuable help with administration as well as our nice conversations. Gunilla for help with booking meeting rooms and for brightening up the stairwell with flowers. Mike for help locating samples in freezers and many nice chats with a fellow Brit.

My fellow students who have kept me company for the duration. Leo, for your kind nature and knowing all the best online tools. Linda, for bringing your natural elegance and also Viggo, who is such a joy. Emil for taking the time to familiarize me with plasmids. Sofie, Xisco, Dani and Piyush, your company has brightened my day.

Bea for lifting my spirits with stories about cats, brunches in town and trips to the spa. Roger for being a friendly face around the building.

And everyone at the Department of Infectious Diseases and on the 6th floor of Guldhedsgatan 10A for creating a pleasant working environment that feels like home.

These studies were supported by the Swedish Research Council for Environmental, Agricultural Sciences and Spatial Planning (Formas); Swedish Research Council for Health, Working Life and Welfare (FORTE); the Ekhaga Foundation; the Innovation Unit at Region Norrbotten; the Västra Götaland Region's Food and Health Concept Centre; the Swedish Research Council (Vetenskapsrådet); and the Swedish federal government under the LUA/ALF agreement.

References

1. Wells CL, Maddaus MA, Reynolds CM, Jechorek RP, Simmons RL. Role of anaerobic flora in the translocation of aerobic and facultatively anaerobic intestinal bacteria. *Infect Immun*. 1987;**55**(11):2689-94.
2. Steffen EK, Berg RD, Deitch EA. Comparison of translocation rates of various indigenous bacteria from the gastrointestinal tract to the mesenteric lymph node. *J Infect Dis*. 1988;**157**(5):1032-8.
3. Hessle CC, Andersson B, Wold AE. Gram-positive and Gram-negative bacteria elicit different patterns of pro-inflammatory cytokines in human monocytes. *Cytokine*. 2005;**30**(6):311-8.
4. Hessle C, Andersson B, Wold AE. Gram-positive bacteria are potent inducers of monocytic interleukin-12 (IL-12) while gram-negative bacteria preferentially stimulate IL-10 production. *Infect Immun*. 2000;**68**(6):3581-6.
5. Finegold SM, Sutter VL, Mathisen GE. CHAPTER 1 Normal Indigenous Intestinal Flora. Human Intestinal Microflora in Health and Disease1983. p. 3-31.
6. Hayashi H, Sakamoto M, Benno Y. Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods. *Microbiol Immunol*. 2002;**46**(8):535-48.
7. Sjoberg F, et al. Are all faecal bacteria detected with equal efficiency? A study using next-generation sequencing and quantitative culture of infants' faecal samples. *J Microbiol Methods*. 2020;**177**:106018.
8. Vandeputte D, et al. Quantitative microbiome profiling links gut community variation to microbial load. *Nature*. 2017;**551**(7681):507-11.
9. Lagier JC, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect*. 2012;**18**(12):1185-93.
10. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012;**486**(7402):207-14.
11. Bach LL, Ram A, Ijaz UZ, Evans TJ, Lindström J. A Longitudinal Study of the Human Oropharynx Microbiota Over Time Reveals a Common Core and Significant Variations With Self-Reported Disease. 2021;**11**.
12. Ohno H, Satoh-Takayama N. Stomach microbiota, Helicobacter pylori, and group 2 innate lymphoid cells. *Exp Mol Med*. 2020;**52**(9):1377-82.
13. Adlerberth I, Wold AE. Establishment of the gut microbiota in Western infants. *Acta Paediatr*. 2009;**98**(2):229-38.

14. Ou G, et al. Proximal small intestinal microbiota and identification of rod-shaped bacteria associated with childhood celiac disease. *Am J Gastroenterol*. 2009;**104**(12):3058-67.
15. Cheng J, et al. Duodenal microbiota composition and mucosal homeostasis in pediatric celiac disease. *BMC Gastroenterol*. 2013;**13**:113.
16. Mafra D, et al. Archaea from the gut microbiota of humans: Could be linked to chronic diseases? *Anaerobe*. 2022;**77**:102629.
17. Catlett JL, et al. Metabolic Synergy between Human Symbionts Bacteroides and Methanobrevibacter. *Microbiology spectrum*. 2022;**10**(3):e0106722.
18. Qin J, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*. 2010;**464**(7285):59-65.
19. Duncan SH, Louis P, Flint HJ. Cultivable bacterial diversity from the human colon. *Lett Appl Microbiol*. 2007;**44**(4):343-50.
20. Juge N. Relationship between mucosa-associated gut microbiota and human diseases. *Biochem Soc Trans*. 2022;**50**(5):1225-36.
21. Huang JY, Lee SM, Mazmanian SK. The human commensal Bacteroides fragilis binds intestinal mucin. *Anaerobe*. 2011;**17**(4):137-41.
22. Yatsunenko T, et al. Human gut microbiome viewed across age and geography. *Nature*. 2012;**486**(7402):222-7.
23. Byrd AL, Belkaid Y, Segre JA. The human skin microbiome. *Nat Rev Microbiol*. 2018;**16**(3):143-55.
24. Costello EK, et al. Bacterial community variation in human body habitats across space and time. *Science*. 2009;**326**(5960):1694-7.
25. Gilbert JA, et al. Current understanding of the human microbiome. *Nat Med*. 2018;**24**(4):392-400.
26. Dominguez-Bello MG, et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A*. 2010;**107**(26):11971-5.
27. Chu DM, et al. Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat Med*. 2017;**23**(3):314-26.
28. Blum J, et al. Temporal development of the infant oral microbiome. *Crit Rev Microbiol*. 2022;**48**(6):730-42.
29. Kennedy KM, et al. Fetal meconium does not have a detectable microbiota before birth. *Nature Microbiology*. 2021;**6**(7):865-73.
30. Arrieta M-C, Stiemsma LT, Amenyogbe N, Brown EM, Finlay B. The Intestinal Microbiome in Early Life: Health and Disease. 2014;**5**.

31. Adlerberth I, et al. Gut microbiota and development of atopic eczema in 3 European birth cohorts. *J Allergy Clin Immunol*. 2007;**120**(2):343-50.
32. Selma-Royo M, et al. Perinatal environment shapes microbiota colonization and infant growth: impact on host response and intestinal function. *Microbiome*. 2020;**8**(1):167.
33. Lindberg E, Nowrouzian F, Adlerberth I, Wold AE. Long-time persistence of superantigen-producing Staphylococcus aureus strains in the intestinal microflora of healthy infants. *Pediatr Res*. 2000;**48**(6):741-7.
34. Stewart CJ, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature*. 2018;**562**(7728):583-8.
35. Arrieta MC, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med*. 2015;**7**(307):307ra152.
36. Lozupone CA, et al. Meta-analyses of studies of the human microbiota. *Genome Res*. 2013;**23**(10):1704-14.
37. de Goffau MC, et al. Gut microbiomes from Gambian infants reveal the development of a non-industrialized Prevotella-based trophic network. *Nat Microbiol*. 2022;**7**(1):132-44.
38. Xiao L, Wang J, Zheng J, Li X, Zhao F. Deterministic transition of enterotypes shapes the infant gut microbiome at an early age. *Genome Biol*. 2021;**22**(1):243.
39. Ellis-Pegler RB, Crabtree C, Lambert HP. The faecal flora of children in the United Kingdom. *J Hyg (Lond)*. 1975;**75**(1):135-42.
40. Depner M, et al. Maturation of the gut microbiome during the first year of life contributes to the protective farm effect on childhood asthma. *Nat Med*. 2020;**26**(11):1766-75.
41. Dwyer LR, Scharschmidt TC. Early life host-microbe interactions in skin. *Cell Host Microbe*. 2022;**30**(5):684-95.
42. Capone KA, Dowd SE, Stamatas GN, Nikolovski J. Diversity of the human skin microbiome early in life. *J Invest Dermatol*. 2011;**131**(10):2026-32.
43. Carr DL, Kloos WE. Temporal study of the staphylococci and micrococci of normal infant skin. *Appl Environ Microbiol*. 1977;**34**(6):673-80.
44. Leyden JJ, McGinley KJ, Mills OH, Kligman AM. Age-related changes in the resident bacterial flora of the human face. *J Invest Dermatol*. 1975;**65**(4):379-81.
45. Paul AA, et al. Fungal cutaneous microbiome and host determinants in preterm and term neonates. *Pediatr Res*. 2020;**88**(2):225-33.
46. Nagata R, et al. Transmission of the major skin microbiota, Malassezia, from mother to neonate. *Pediatr Int*. 2012;**54**(3):350-5.

47. Sprockett D, Fukami T, Relman DA. Role of priority effects in the early-life assembly of the gut microbiota. *Nature Reviews Gastroenterology & Hepatology*. 2018;**15**(4):197-205.
48. Rao C, et al. Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature*. 2021;**591**(7851):633-8.
49. Hellmann KT, Tuura CE, Fish J, Patel JM, Robinson DA. Viability-Resolved Metagenomics Reveals Antagonistic Colonization Dynamics of Staphylococcus epidermidis Strains on Preterm Infant Skin. *mSphere*. 2021;**6**(5):e0053821.
50. Browne HP, Shao Y, Lawley TD. Mother-infant transmission of human microbiota. *Curr Opin Microbiol*. 2022;**69**:102173.
51. Adlerberth I, et al. Reduced enterobacterial and increased staphylococcal colonization of the infantile bowel: an effect of hygienic lifestyle? *Pediatr Res*. 2006;**59**(1):96-101.
52. Adlerberth I, et al. High turnover rate of Escherichia coli strains in the intestinal flora of infants in Pakistan. *Epidemiol Infect*. 1998;**121**(3):587-98.
53. Korpela K, et al. Selective maternal seeding and environment shape the human gut microbiome. *Genome Res*. 2018;**28**(4):561-8.
54. Lundequist B, Nord CE, Winberg J. The composition of the faecal microflora in breastfed and bottle fed infants from birth to eight weeks. *Acta Paediatr Scand*. 1985;**74**(1):45-51.
55. Penders J, et al. Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics*. 2006;**118**(2):511-21.
56. Galazzo G, et al. Development of the Microbiota and Associations With Birth Mode, Diet, and Atopic Disorders in a Longitudinal Analysis of Stool Samples, Collected From Infancy Through Early Childhood. *Gastroenterology*. 2020;**158**(6):1584-96.
57. Wold AE, Hanson LÅ. Defense factors in human milk. 1994;**10**(6):652-8.
58. Kort R, et al. Shaping the oral microbiota through intimate kissing. *Microbiome*. 2014;**2**(1):41.
59. Song SJ, et al. Cohabiting family members share microbiota with one another and with their dogs. *eLife*. 2013;**2**:e00458.
60. Lindberg E, et al. High rate of transfer of Staphylococcus aureus from parental skin to infant gut flora. *J Clin Microbiol*. 2004;**42**(2):530-4.
61. Hesselmar B, et al. Pacifier cleaning practices and risk of allergy development. *Pediatrics*. 2013;**131**(6):e1829-37.
62. Han CS, et al. Salivary microbiomes of indigenous Tsimane mothers and infants are distinct despite frequent premastication. *PeerJ*. 2016;**4**:e2660.

63. Christensen ED, et al. The developing airway and gut microbiota in early life is influenced by age of older siblings. *Microbiome*. 2022;**10**(1):106.
64. Amir A, et al. Gut microbiome development in early childhood is affected by day care attendance. *npj Biofilms and Microbiomes*. 2022;**8**(1):2.
65. Thompson AL, Monteagudo-Mera A, Cadenas MB, Lampl ML, Azcarate-Peril MA. Milk- and solid-feeding practices and daycare attendance are associated with differences in bacterial diversity, predominant communities, and metabolic and immune function of the infant gut microbiome. *Frontiers in cellular and infection microbiology*. 2015;**5**:3.
66. Tun HM, et al. Exposure to household furry pets influences the gut microbiota of infants at 3–4 months following various birth scenarios. *Microbiome*. 2017;**5**(1):40.
67. Nermes M, Endo A, Aarnio J, Salminen S, Isolauri E. Furry pets modulate gut microbiota composition in infants at risk for allergic disease. *J Allergy Clin Immunol*. 2015;**136**(6):1688-90.e1.
68. Nowrouzian F, et al. Escherichia coli in infants' intestinal microflora: colonization rate, strain turnover, and virulence gene carriage. *Pediatr Res*. 2003;**54**(1):8-14.
69. Alfvén T, et al. Allergic diseases and atopic sensitization in children related to farming and anthroposophic lifestyle--the PARSIFAL study. *Allergy*. 2006;**61**(4):414-21.
70. Genuneit J. Exposure to farming environments in childhood and asthma and wheeze in rural populations: a systematic review with meta-analysis. *Pediatr Allergy Immunol*. 2012;**23**(6):509-18.
71. Brick T, Hettinga K, Kirchner B, Pfaffl MW, Ege MJ. The Beneficial Effect of Farm Milk Consumption on Asthma, Allergies, and Infections: From Meta-Analysis of Evidence to Clinical Trial. *The journal of allergy and clinical immunology In practice*. 2020;**8**(3):878-89.e3.
72. von Mutius E, Radon K. Living on a farm: impact on asthma induction and clinical course. *Immunol Allergy Clin North Am*. 2008;**28**(3):631-47, ix-x.
73. Jonsson K, et al. Diet in 1-year-old farm and control children and allergy development: results from the FARMFLORA birth cohort. *Food Nutr Res*. 2016;**60**:32721.
74. Jonsson K, et al. Fat intake and breast milk fatty acid composition in farming and nonfarming women and allergy development in the offspring. *Pediatr Res*. 2016;**79**(1-1):114-23.
75. Roy SR, Schiltz AM, Marotta A, Shen Y, Liu AH. Bacterial DNA in house and farm barn dust. *J Allergy Clin Immunol*. 2003;**112**(3):571-8.

76. Karvonen AM, et al. Exposure to microbial agents in house dust and wheezing, atopic dermatitis and atopic sensitization in early childhood: a birth cohort study in rural areas. *Clin Exp Allergy*. 2012;**42**(8):1246-56.
77. Schram D, et al. Bacterial and fungal components in house dust of farm children, Rudolf Steiner school children and reference children--the PARSIFAL Study. *Allergy*. 2005;**60**(5):611-8.
78. Ege MJ, et al. Exposure to environmental microorganisms and childhood asthma. *N Engl J Med*. 2011;**364**(8):701-9.
79. Kirjavainen PV, et al. Farm-like indoor microbiota in non-farm homes protects children from asthma development. *Nat Med*. 2019;**25**(7):1089-95.
80. Penders J, Stobberingh EE, Savelkoul PH, Wolffs PF. The human microbiome as a reservoir of antimicrobial resistance. *Front Microbiol*. 2013;**4**:87.
81. Gallo RL. Human Skin Is the Largest Epithelial Surface for Interaction with Microbes. *J Invest Dermatol*. 2017;**137**(6):1213-4.
82. Minkoff NZ, et al. Fecal microbiota transplantation for the treatment of recurrent Clostridioides difficile (Clostridium difficile). *Cochrane Database Syst Rev*. 2023(4).
83. Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. *Nature Reviews Immunology*. 2009;**9**(5):313-23.
84. Johansson Malin EV, et al. Normalization of Host Intestinal Mucus Layers Requires Long-Term Microbial Colonization. *Cell Host Microbe*. 2015;**18**(5):582-92.
85. Sudo N, et al. The requirement of intestinal bacterial flora for the development of an IgE production system fully susceptible to oral tolerance induction. *J Immunol*. 1997;**159**(4):1739-45.
86. Moreau MC, Gaboriau-Routhiau V. The absence of gut flora, the doses of antigen ingested and aging affect the long-term peripheral tolerance induced by ovalbumin feeding in mice. *Res Immunol*. 1996;**147**(1):49-59.
87. Moreau MC, Coste M, Gaboriau V, Dubuquoy C. Oral tolerance to ovalbumin in mice: effect of some parameters on the induction and persistence of the suppression of systemic IgE and IgG antibody responses. *Adv Exp Med Biol*. 1995;**371b**:1229-34.
88. Moreau MC, Corthier G. Effect of the gastrointestinal microflora on induction and maintenance of oral tolerance to ovalbumin in C3H/HeJ mice. *Infect Immun*. 1988;**56**(10):2766-8.

89. Rask C, Evertsson S, Telemo E, Wold AE. A full flora, but not monocolonization by Escherichia coli or lactobacilli, supports tolerogenic processing of a fed antigen. *Scand J Immunol*. 2005;**61**(6):529-35.
90. Ostman S, Rask C, Wold AE, Hultkrantz S, Telemo E. Impaired regulatory T cell function in germ-free mice. *Eur J Immunol*. 2006;**36**(9):2336-46.
91. Liu J, et al. Functions of Gut Microbiota Metabolites, Current Status and Future Perspectives. *Aging Dis*. 2022;**13**(4):1106-26.
92. Rowland I, et al. Gut microbiota functions: metabolism of nutrients and other food components. *Eur J Nutr*. 2018;**57**(1):1-24.
93. Turnbaugh PJ, et al. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*. 2006;**444**(7122):1027-31.
94. Evans DF, et al. Measurement of gastrointestinal pH profiles in normal ambulant human subjects. *Gut*. 1988;**29**(8):1035-41.
95. van der Hee B, Wells JM. Microbial Regulation of Host Physiology by Short-chain Fatty Acids. *Trends Microbiol*. 2021;**29**(8):700-12.
96. Smith EA, Macfarlane GT. Enumeration of amino acid fermenting bacteria in the human large intestine: effects of pH and starch on peptide metabolism and dissimilation of amino acids. *FEMS Microbiol Ecol*. 1998;**25**(4):355-68.
97. Wiciński M, Sawicka E, Gębalski J, Kubiak K, Malinowski B. Human Milk Oligosaccharides: Health Benefits, Potential Applications in Infant Formulas, and Pharmacology. *Nutrients*. 2020;**12**(1).
98. Bührer C, et al. Infant formulas with synthetic oligosaccharides and respective marketing practices: Position Statement of the German Society for Child and Adolescent Medicine e.V. (DGKJ), Commission for Nutrition. *Molecular and Cellular Pediatrics*. 2022;**9**(1):14.
99. den Besten G, et al. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res*. 2013;**54**(9):2325-40.
100. Flint HJ, Duncan SH, Scott KP, Louis P. Links between diet, gut microbiota composition and gut metabolism. *Proc Nutr Soc*. 2015;**74**(1):13-22.
101. Nogal A, Valdes AM, Menni C. The role of short-chain fatty acids in the interplay between gut microbiota and diet in cardio-metabolic health. *Gut microbes*. 2021;**13**(1):1-24.
102. Rios-Covian D, et al. Different metabolic features of Bacteroides fragilis growing in the presence of glucose and exopolysaccharides of bifidobacteria. *Front Microbiol*. 2015;**6**:825.

103. Zhu X, et al. Production of high-concentration n-caproic acid from lactate through fermentation using a newly isolated Ruminococcaceae bacterium CPB6. *Biotechnology for Biofuels*. 2017;**10**(1):102.
104. Jeon BS, Kim S, Sang BI. Megasphaera hexanoica sp. nov., a medium-chain carboxylic acid-producing bacterium isolated from a cow rumen. *Int J Syst Evol Microbiol*. 2017;**67**(7):2114-20.
105. Kim H, Jeon BS, Sang B-I. An Efficient New Process for the Selective Production of Odd-Chain Carboxylic Acids by Simple Carbon Elongation Using Megasphaera hexanoica. *Sci Rep*. 2019;**9**(1):11999.
106. Jeon BS, Choi O, Um Y, Sang BI. Production of medium-chain carboxylic acids by Megasphaera sp. MH with supplemental electron acceptors. *Biotechnol Biofuels*. 2016;**9**:129.
107. Iino T, Mori K, Tanaka K, Suzuki KI, Harayama S. Oscillibacter valericigenes gen. nov., sp. nov., a valerate-producing anaerobic bacterium isolated from the alimentary canal of a Japanese corbicula clam. *Int J Syst Evol Microbiol*. 2007;**57**(Pt 8):1840-5.
108. Kim S, et al. Mucin degrader Akkermansia muciniphila accelerates intestinal stem cell-mediated epithelial development. *Gut microbes*. 2021;**13**(1):1892441.
109. Guo CJ, et al. Depletion of microbiome-derived molecules in the host using Clostridium genetics. *Science*. 2019;**366**(6471).
110. O'Keefe SJ. Diet, microorganisms and their metabolites, and colon cancer. *Nat Rev Gastroenterol Hepatol*. 2016;**13**(12):691-706.
111. Parada Venegas D, et al. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. *Front Immunol*. 2019;**10**:277.
112. Rivière A, Selak M, Lantin D, Leroy F, De Vuyst L. Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Front Microbiol*. 2016;**7**:979.
113. Boets E, et al. Systemic availability and metabolism of colonic-derived short-chain fatty acids in healthy subjects: a stable isotope study. *J Physiol*. 2017;**595**(2):541-55.
114. Dehority BA, Scott HW, Kowaluk P. Volatile fatty acid requirements of cellulolytic rumen bacteria. *J Bacteriol*. 1967;**94**(3):537-43.
115. Kaneko T, Mori H, Iwata M, Meguro S. Growth stimulator for bifidobacteria produced by Propionibacterium freudenreichii and several intestinal bacteria. *J Dairy Sci*. 1994;**77**(2):393-404.
116. Parodi PW. Cows' Milk Fat Components as Potential Anticarcinogenic Agents. *The Journal of Nutrition*. 1997;**127**(6):1055-60.

117. Cummings JH, Pomare EW, Branch WJ, Naylor CP, Macfarlane GT. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut*. 1987;**28**(10):1221-7.
118. Gio-Batta M, et al. Fecal short chain fatty acids in children living on farms and a link between valeric acid and protection from eczema. *Sci Rep*. 2020;**10**(1):22449.
119. Onyszkiewicz M, et al. Valeric acid lowers arterial blood pressure in rats. *Eur J Pharmacol*. 2020;**877**:173086.
120. Trompette A, et al. Gut-derived short-chain fatty acids modulate skin barrier integrity by promoting keratinocyte metabolism and differentiation. *Mucosal Immunol*. 2022;**15**(5):908-26.
121. Müller M, et al. Circulating but not faecal short-chain fatty acids are related to insulin sensitivity, lipolysis and GLP-1 concentrations in humans. *Sci Rep*. 2019;**9**(1):12515.
122. Stinson LF, et al. Human Milk From Atopic Mothers Has Lower Levels of Short Chain Fatty Acids. *Front Immunol*. 2020;**11**:1427.
123. Meiser J, et al. Serine one-carbon catabolism with formate overflow. *Science advances*. 2016;**2**(10):e1601273.
124. Pietzke M, Meiser J, Vazquez A. Formate metabolism in health and disease. *Molecular metabolism*. 2020;**33**:23-37.
125. Grimolizzi F, Arranz L. Multiple faces of succinate beyond metabolism in blood. *Haematologica*. 2018;**103**(10):1586-92.
126. Hochachka PW, Dressendorfer RH. Succinate accumulation in man during exercise. *Eur J Appl Physiol Occup Physiol*. 1976;**35**(4):235-42.
127. Knowles SE, Jarrett IG, Filsell OH, Ballard FJ. Production and utilization of acetate in mammals. *Biochem J*. 1974;**142**(2):401-11.
128. Skutches CL, Holroyde CP, Myers RN, Paul P, Reichard GA. Plasma acetate turnover and oxidation. *J Clin Invest*. 1979;**64**(3):708-13.
129. Pouteau E, et al. Kinetic aspects of acetate metabolism in healthy humans using [1-13C] acetate. *Am J Physiol*. 1996;**271**(1 Pt 1):E58-64.
130. van Hall G, Sacchetti M, Rådegran G. Whole body and leg acetate kinetics at rest, during exercise and recovery in humans. *J Physiol*. 2002;**542**(Pt 1):263-72.
131. Xiong RG, et al. Health Benefits and Side Effects of Short-Chain Fatty Acids. *Foods (Basel, Switzerland)*. 2022;**11**(18).
132. Li Y, et al. Gut commensal derived-valeric acid protects against radiation injuries. *Gut microbes*. 2020;**11**(4):789-806.
133. Burger-van Paassen N, et al. The regulation of intestinal mucin MUC2 expression by short-chain fatty acids: implications for epithelial protection. *Biochem J*. 2009;**420**(2):211-9.

134. Kleuskens MTA, et al. Butyrate and propionate restore interleukin 13-compromised esophageal epithelial barrier function. *Allergy*. 2022;**77**(5):1510-21.
135. Onrust L, et al. Valeric acid glyceride esters in feed promote broiler performance and reduce the incidence of necrotic enteritis. *Poult Sci*. 2018;**97**(7):2303-11.
136. Bartholome AL, Albin DM, Baker DH, Holst JJ, Tappenden KA. Supplementation of total parenteral nutrition with butyrate acutely increases structural aspects of intestinal adaptation after an 80% jejunoileal resection in neonatal piglets. *JPEN J Parenter Enteral Nutr*. 2004;**28**(4):210-22; discussion 22-3.
137. Leon Carrion S, Sutter CH, Sutter TR. Combined treatment with sodium butyrate and PD153035 enhances keratinocyte differentiation. *Exp Dermatol*. 2014;**23**(3):211-4.
138. Xiao X, et al. The role of short-chain fatty acids in inflammatory skin diseases. *Front Microbiol*. 2022;**13**:1083432.
139. van Limpt C, Crienen A, Vriesema A, Knol J. 134 Effect of Colonic Short Chain Fatty Acids, Lactate and PH on The Growth of Common Gut Pathogens. *Pediatr Res*. 2004;**56**(3):487-.
140. Kovanda L, et al. In Vitro Antimicrobial Activities of Organic Acids and Their Derivatives on Several Species of Gram-Negative and Gram-Positive Bacteria. *Molecules*. 2019;**24**(20).
141. Liu XF, et al. Regulation of short-chain fatty acids in the immune system. *Front Immunol*. 2023;**14**:1186892.
142. Kim CH. Control of lymphocyte functions by gut microbiota-derived short-chain fatty acids. *Cell Mol Immunol*. 2021;**18**(5):1161-71.
143. Singh N, et al. Activation of Gpr109a, receptor for niacin and the commensal metabolite butyrate, suppresses colonic inflammation and carcinogenesis. *Immunity*. 2014;**40**(1):128-39.
144. Thorburn AN, et al. Evidence that asthma is a developmental origin disease influenced by maternal diet and bacterial metabolites. *Nature communications*. 2015;**6**:7320.
145. Bellinghausen I, Khatri R, Saloga J. Current Strategies to Modulate Regulatory T Cell Activity in Allergic Inflammation. 2022;**13**.
146. Martin-Gallausiaux C, et al. Butyrate Produced by Commensal Bacteria Down-Regulates Indolamine 2,3-Dioxygenase 1 (IDO-1) Expression via a Dual Mechanism in Human Intestinal Epithelial Cells. *Front Immunol*. 2018;**9**:2838.

147. Luu M, et al. The short-chain fatty acid pentanoate suppresses autoimmunity by modulating the metabolic-epigenetic crosstalk in lymphocytes. *Nature communications*. 2019;**10**(1):760.
148. Tan J, et al. Dietary Fiber and Bacterial SCFA Enhance Oral Tolerance and Protect against Food Allergy through Diverse Cellular Pathways. *Cell Rep*. 2016;**15**(12):2809-24.
149. Kim M, Qie Y, Park J, Kim CH. Gut Microbial Metabolites Fuel Host Antibody Responses. *Cell Host Microbe*. 2016;**20**(2):202-14.
150. Spergel JM, Paller AS. Atopic dermatitis and the atopic march. *J Allergy Clin Immunol*. 2003;**112**(6, Supplement):S118-S27.
151. Allen JE, Sutherland TE. Host protective roles of type 2 immunity: parasite killing and tissue repair, flip sides of the same coin. *Semin Immunol*. 2014;**26**(4):329-40.
152. Pawankar R. Allergic diseases and asthma: a global public health concern and a call to action. *World Allergy Organ J*. 2014;**7**(1):12.
153. Mallol J, et al. The International Study of Asthma and Allergies in Childhood (ISAAC) Phase Three: a global synthesis. *Allergol Immunopathol (Madr)*. 2013;**41**(2):73-85.
154. Doll RJ, et al. Epidemiology of Allergic Diseases. Allergy and Asthma: The Basics to Best Practices. Cham: Springer International Publishing; 2019. p. 31-51.
155. Perkin MR, Strachan DP. The hygiene hypothesis for allergy - conception and evolution. *Frontiers in allergy*. 2022;**3**:1051368.
156. Strachan DP. Hay fever, hygiene, and household size. *BMJ*. 1989;**299**(6710):1259-60.
157. Matricardi PM, et al. Cross sectional retrospective study of prevalence of atopy among Italian military students with antibodies against hepatitis A virus. *BMJ*. 1997;**314**(7086):999-1003.
158. Matricardi PM, et al. Exposure to foodborne and orofecal microbes versus airborne viruses in relation to atopy and allergic asthma: epidemiological study. *BMJ*. 2000;**320**(7232):412-7.
159. Matricardi PM. The role of early infections, hygiene and intestinal microflora. *Pediatr Pulmonol Suppl*. 2004;**26**:211-2.
160. Adlerberth I, et al. Intestinal colonization with Enterobacteriaceae in Pakistani and Swedish hospital-delivered infants. 1991.
161. Wold AE. The hygiene hypothesis revised: is the rising frequency of allergy due to changes in the intestinal flora? *Allergy*. 1998;**53**(46 Suppl):20-5.

162. Wells HG, Osborne TB. The Biological Reactions of the Vegetable Proteins I. Anaphylaxis. *The Journal of Infectious Diseases*. 1911;**8**(1):66-124.
163. Husby S, Mestecky J, Moldoveanu Z, Holland S, Elson CO. Oral tolerance in humans. T cell but not B cell tolerance after antigen feeding. *J Immunol*. 1994;**152**(9):4663-70.
164. Lönnqvist A, et al. Neonatal exposure to staphylococcal superantigen improves induction of oral tolerance in a mouse model of airway allergy. *Eur J Immunol*. 2009;**39**(2):447-56.
165. Stern A, Wold AE, Östman S. Neonatal mucosal immune stimulation by microbial superantigen improves the tolerogenic capacity of CD103(+) dendritic cells. *PLoS One*. 2013;**8**(9):e75594.
166. Scadding GK, Scadding GW. Innate and adaptive immunity in allergic airway disease. *Curr Opin Allergy Clin Immunol*. 2022;**22**(1):10-5.
167. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. *Lancet*. 2008;**371**(9606):75-84.
168. Muglia LJ, Katz M. The enigma of spontaneous preterm birth. *N Engl J Med*. 2010;**362**(6):529-35.
169. Winberg J, Wessner G. Does breast milk protect against septicaemia in the newborn? *Lancet*. 1971;**1**(7709):1091-4.
170. Perin J, et al. Global, regional, and national causes of under-5 mortality in 2000-19: an updated systematic analysis with implications for the Sustainable Development Goals. *The Lancet Child & adolescent health*. 2022;**6**(2):106-15.
171. Morgan AS, et al. Birth outcomes between 22 and 26 weeks' gestation in national population-based cohorts from Sweden, England and France. *Acta Paediatr*. 2022;**111**(1):59-75.
172. Norman M, et al. Association Between Year of Birth and 1-Year Survival Among Extremely Preterm Infants in Sweden During 2004-2007 and 2014-2016. *JAMA*. 2019;**321**(12):1188-99.
173. Flannery DD, Edwards EM, Coggins SA, Horbar JD, Puopolo KM. Late-Onset Sepsis Among Very Preterm Infants. *Pediatrics*. 2022;**150**(6).
174. Marissen J, et al. The Delicate Skin of Preterm Infants: Barrier Function, Immune-Microbiome Interaction, and Clinical Implications. *Neonatology*. 2023;**120**(3):295-307.
175. Campbell-Yeo ML, Disher TC, Benoit BL, Johnston CC. Understanding kangaroo care and its benefits to preterm infants. *Pediatric Health, Medicine and Therapeutics*. 2015;**6**:15-32.
176. Underwood MA, Sohn K. The Microbiota of the Extremely Preterm Infant. *Clin Perinatol*. 2017;**44**(2):407-27.

177. Healy DB, Ryan CA, Ross RP, Stanton C, Dempsey EM. Clinical implications of preterm infant gut microbiome development. *Nature Microbiology*. 2022;**7**(1):22-33.
178. Hendricks-Muñoz KD, et al. Skin-to-Skin Care and the Development of the Preterm Infant Oral Microbiome. *Am J Perinatol*. 2015;**32**(13):1205-16.
179. Olm MR, et al. Identical bacterial populations colonize premature infant gut, skin, and oral microbiomes and exhibit different in situ growth rates. *Genome Res*. 2017;**27**(4):601-12.
180. Li H, et al. Impacts of delivery mode on very low birth weight infants' oral microbiome. *Pediatr Neonatol*. 2020;**61**(2):201-9.
181. Shimizu A, Shimizu K, Nakamura T. Non-pathogenic bacterial flora may inhibit colonization by methicillin-resistant Staphylococcus aureus in extremely low birth weight infants. *Neonatology*. 2008;**93**(3):158-61.
182. Govindarajan V, Devadas S, Shah PA, Diggikar S. Impact of Kangaroo Mother Care on Skin Microbiome of Very Preterm Infants - A Pilot Study. *Indian J Pediatr*. 2023.
183. Pammi M, et al. Development of the cutaneous microbiome in the preterm infant: A prospective longitudinal study. *PLoS One*. 2017;**12**(4):e0176669.
184. Brooks B, et al. Microbes in the neonatal intensive care unit resemble those found in the gut of premature infants. *Microbiome*. 2014;**2**(1):1.
185. Vongbhavit K, et al. A comparison of bacterial colonization between nasogastric and orogastric enteral feeding tubes in infants in the neonatal intensive care unit. *J Perinatol*. 2022;**42**(11):1446-52.
186. Hourigan SK, et al. Comparison of Infant Gut and Skin Microbiota, Resistome and Virulome Between Neonatal Intensive Care Unit (NICU) Environments. 2018;**9**.
187. Morreale C, et al. Effects of Perinatal Antibiotic Exposure and Neonatal Gut Microbiota. *Antibiotics (Basel, Switzerland)*. 2023;**12**(2).
188. Aguilar-Lopez M, Dinsmoor AM, Ho TTB, Donovan SM. A systematic review of the factors influencing microbial colonization of the preterm infant gut. *Gut microbes*. 2021;**13**(1):1-33.
189. Makhoul IR, et al. Factors influencing oral colonization in premature infants. *Isr Med Assoc J*. 2002;**4**(2):98-102.
190. Orchanian SB, et al. Multiomic Analyses of Nascent Preterm Infant Microbiomes Differentiation Suggest Opportunities for Targeted Intervention. *Advanced biology*. 2022;**6**(8):e2101313.

191. Levy EI, Hoang DM, Vandenplas Y. The effects of proton pump inhibitors on the microbiome in young children. *Acta Paediatr*. 2020;**109**(8):1531-8.
192. Zwittink RD, et al. Metaproteomics reveals functional differences in intestinal microbiota development of preterm infants. *Mol Cell Proteomics*. 2017;**16**(9):1610-20.
193. Govindarajan V, Devadas S, Shah PA, Diggikar S. Impact of Kangaroo Mother Care on Skin Microbiome of Very Preterm Infants - A Pilot Study. *Indian J Pediatr*. 2023.
194. Pammi M, et al. Intestinal dysbiosis in preterm infants preceding necrotizing enterocolitis: a systematic review and meta-analysis. *Microbiome*. 2017;**5**(1):31.
195. Masi AC, Stewart CJ. The role of the preterm intestinal microbiome in sepsis and necrotising enterocolitis. *Early Hum Dev*. 2019;**138**:104854.
196. Zhang C, et al. Paneth cell ablation in the presence of Klebsiella pneumoniae induces necrotizing enterocolitis (NEC)-like injury in the small intestine of immature mice. *Dis Model Mech*. 2012;**5**(4):522-32.
197. van den Akker CHP, et al. Probiotics and Preterm Infants: A Position Paper by the European Society for Paediatric Gastroenterology Hepatology and Nutrition Committee on Nutrition and the European Society for Paediatric Gastroenterology Hepatology and Nutrition Working Group for Probiotics and Prebiotics. 2020;**70**(5):664-80.
198. Berg RD. Promotion of the translocation of enteric bacteria from the gastrointestinal tracts of mice by oral treatment with penicillin, clindamycin, or metronidazole. *Infect Immun*. 1981;**33**(3):854-61.
199. Stewart CJ, et al. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. *Microbiome*. 2017;**5**(1):75.
200. Shaw AG, et al. Late-Onset Bloodstream Infection and Perturbed Maturation of the Gastrointestinal Microbiota in Premature Infants. *PLoS One*. 2015;**10**(7):e0132923.
201. El Manouni El Hassani S, et al. Profound Pathogen-Specific Alterations in Intestinal Microbiota Composition Precede Late-Onset Sepsis in Preterm Infants: A Longitudinal, Multicenter, Case-Control Study. *Clin Infect Dis*. 2021;**73**(1):e224-e32.
202. Von Bültzingslöwen I, Adlerberth I, Wold AE, Dahlén G, Jontell M. Oral and intestinal microflora in 5-fluorouracil treated rats, translocation to cervical and mesenteric lymph nodes and effects of probiotic bacteria. *Oral Microbiol Immunol*. 2003;**18**(5):278-84.

203. Frazer LC, Yakah W, Martin CR. Decreased Acetic Acid in the Stool of Preterm Infants Is Associated with an Increased Risk of Bronchopulmonary Dysplasia. *Nutrients*. 2022;**14**(12).
204. Zhang Q, Ran X, He Y, Ai Q, Shi Y. Acetate Downregulates the Activation of NLRP3 Inflammasomes and Attenuates Lung Injury in Neonatal Mice With Bronchopulmonary Dysplasia. *Frontiers in pediatrics*. 2020;**8**:595157.
205. Skondra D, et al. The early gut microbiome could protect against severe retinopathy of prematurity. *J AAPOS*. 2020;**24**(4):236-8.
206. Lundell AC, et al. High proportion of CD5+ B cells in infants predicts development of allergic disease. *J Immunol*. 2014;**193**(2):510-8.
207. Lundell AC, et al. Higher B-cell activating factor levels at birth are positively associated with maternal dairy farm exposure and negatively related to allergy development. *J Allergy Clin Immunol*. 2015;**136**(4):1074-82.e3.
208. Strombeck A, et al. Allergic disease in 8-year-old children is preceded by delayed B cell maturation. *Clin Exp Allergy*. 2017;**47**(7):918-28.
209. Sandin A, Bråbäck L, Norin E, Björkstén B. Faecal short chain fatty acid pattern and allergy in early childhood. *Acta Paediatr*. 2009;**98**(5):823-7.
210. Sandin A, Björkstén B, Bråbäck L. Development of atopy and wheezing symptoms in relation to heredity and early pet keeping in a Swedish birth cohort. *Pediatr Allergy Immunol*. 2004;**15**(4):316-22.
211. Barman M, et al. High levels of both n-3 and n-6 long-chain polyunsaturated fatty acids in cord serum phospholipids predict allergy development. *PLoS One*. 2013;**8**(7):e67920.
212. Asher MI, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J*. 1995;**8**(3):483-91.
213. Ögge LE, et al. Maternal characteristics and pregnancy outcomes in the NICE birth cohort: an assessment of self-selection bias. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet*. 2022:1-9.
214. Ögge LE, et al. Maternal characteristics and pregnancy outcomes in the NICE birth cohort: an assessment of self-selection bias. *J Matern Fetal Neonatal Med*. 2022;**35**(25):9014-22.

215. Barman M, et al. Nutritional impact on Immunological maturation during Childhood in relation to the Environment (NICE): a prospective birth cohort in northern Sweden. *BMJ Open*. 2018;**8**(10):e022013.
216. Najm S, et al. Effects of a lipid emulsion containing fish oil on polyunsaturated fatty acid profiles, growth and morbidities in extremely premature infants: A randomized controlled trial. *Clinical nutrition ESPEN*. 2017;**20**:17-23.
217. Jonsson K, et al. Late introduction of fish and eggs is associated with increased risk of allergy development - results from the FARMFLORA birth cohort. *Food Nutr Res*. 2017;**61**(1):1393306.
218. Stråvik M, et al. Maternal Intake of Cow's Milk during Lactation Is Associated with Lower Prevalence of Food Allergy in Offspring. *Nutrients*. 2020;**12**(12).
219. Stråvik M, et al. Food and Nutrient Intake during Pregnancy in Relation to Maternal Characteristics: Results from the NICE Birth Cohort in Northern Sweden. *Nutrients*. 2019;**11**(7).
220. Rabe H, et al. Neonatal gut colonization by Bifidobacterium is associated with higher childhood cytokine responses. *Gut microbes*. 2020;**12**(1):1-14.
221. Klindworth A, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res*. 2013;**41**(1):e1.
222. Bolyen E, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 2019;**37**(8):852-7.
223. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*. 2016;**4**:e2584.
224. Ellner PD, Stoessel CJ, Drakeford E, Vasi F. A new culture medium for medical bacteriology. *Am J Clin Pathol*. 1966;**45**(4):502-4.
225. Chapman GH. Comparison of Ludlam's medium with Staphylococcus medium number 110 for the isolation of staphylococci that clot blood. *J Bacteriol*. 1949;**58**(6):823.
226. Petts DN. Colistin-oxolinic acid-blood agar: a new selective medium for streptococci. *J Clin Microbiol*. 1984;**19**(1):4-7.
227. Isenberg HD, Goldberg D, Sampson J. Laboratory studies with a selective Enterococcus medium. *Appl Microbiol*. 1970;**20**(3):433-6.
228. Kauffmann F. The bacteriology of enterobacteriaceae : collected studies of the author and his co-workers. Copenhagen: Copenhagen; 1966.

229. Chapin KC, Doern GV. Selective media for recovery of Haemophilus influenzae from specimens contaminated with upper respiratory tract microbial flora. *J Clin Microbiol*. 1983;**17**(6):1163-5.
230. Sandven P, Lassen J. Importance of selective media for recovery of yeasts from clinical specimens. *J Clin Microbiol*. 1999;**37**(11):3731-2.
231. Leeming JP, Notman FH. Improved methods for isolation and enumeration of Malassezia furfur from human skin. *J Clin Microbiol*. 1987;**25**(10):2017-9.
232. Slots J. Bacterial specificity in adult periodontitis. A summary of recent work. *J Clin Periodontol*. 1986;**13**(10):912-7.
233. Silvi S, Rumney CJ, Rowland IR. An assessment of three selective media for bifidobacteria in faeces. *J Appl Bacteriol*. 1996;**81**(5):561-4.
234. Rogosa M, Mitchell JA, Wiseman RF. A selective medium for the isolation and enumeration of oral and fecal lactobacilli. *J Bacteriol*. 1951;**62**(1):132-3.
235. Livingston SJ, Kominos SD, Yee RB. New medium for selection and presumptive identification of the Bacteroides fragilis group. *J Clin Microbiol*. 1978;**7**(5):448-53.
236. Zhao G, Nyman M, Jonsson JA. Rapid determination of short-chain fatty acids in colonic contents and faeces of humans and rats by acidified water-extraction and direct-injection gas chromatography. *Biomed Chromatogr*. 2006;**20**(8):674-82.
237. Han J, Lin K, Sequeira C, Borchers CH. An isotope-labeled chemical derivatization method for the quantitation of short-chain fatty acids in human feces by liquid chromatography-tandem mass spectrometry. *Anal Chim Acta*. 2015;**854**:86-94.
238. Iversen KN, et al. The Effects of High Fiber Rye, Compared to Refined Wheat, on Gut Microbiota Composition, Plasma Short Chain Fatty Acids, and Implications for Weight Loss and Metabolic Risk Factors (the RyeWeight Study). *Nutrients*. 2022;**14**(8).
239. Dhariwal A, et al. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res*. 2017;**45**(W1):W180-w8.
240. Barupal DK, Fiehn O. Generating the Blood Exposome Database Using a Comprehensive Text Mining and Database Fusion Approach. *Environ Health Perspect*. 2019;**127**(9):97008.
241. Rabe H, et al. Neonatal gut colonization by Bifidobacterium is associated with higher childhood cytokine responses. *Gut microbes*. 2020;**12**(1):1-14.

242. Böttcher MF, Nordin EK, Sandin A, Midtvedt T, Björkstén B. Microflora-associated characteristics in faeces from allergic and nonallergic infants. *Clin Exp Allergy*. 2000;**30**(11):1590-6.
243. Bråbäck L, Hjern A, Rasmussen F. Trends in asthma, allergic rhinitis and eczema among Swedish conscripts from farming and non-farming environments. A nationwide study over three decades. *Clin Exp Allergy*. 2004;**34**(1):38-43.
244. Holbreich M, et al. Amish children living in northern Indiana have a very low prevalence of allergic sensitization. *J Allergy Clin Immunol*. 2012;**129**(6):1671-3.
245. Stein MM, et al. Innate Immunity and Asthma Risk in Amish and Hutterite Farm Children. *N Engl J Med*. 2016;**375**(5):411-21.
246. Martina C, Looney RJ, Marcus C, Allen M, Stahlhut R. Prevalence of allergic disease in Old Order Mennonites in New York. *Ann Allergy Asthma Immunol*. 2016;**117**(5):562-3.e1.
247. Fagerstedt S, et al. Anthroposophic lifestyle is associated with a lower incidence of food allergen sensitization in early childhood. *J Allergy Clin Immunol*. 2016;**137**(4):1253-6.e3.
248. Rook GA, Lowry CA, Raison CL. Microbial 'Old Friends', immunoregulation and stress resilience. *Evolution, medicine, and public health*. 2013;**2013**(1):46-64.
249. Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. *N Engl J Med*. 2002;**347**(12):911-20.
250. Xiong Z, et al. Prevalence of eczema between cesarean-born and vaginal-born infants within 1 year of age: a systematic review and meta-analysis. *Eur J Pediatr*. 2022;**181**(6):2237-47.
251. Baron R, et al. The relationship of prenatal antibiotic exposure and infant antibiotic administration with childhood allergies: a systematic review. *BMC Pediatr*. 2020;**20**(1):312.
252. Stewart CJ, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature*. 2018;**562**(7728):583-8.
253. Grześkowiak Ł, et al. Distinct gut microbiota in southeastern African and northern European infants. *J Pediatr Gastroenterol Nutr*. 2012;**54**(6):812-6.
254. Young SL, et al. Bifidobacterial species differentially affect expression of cell surface markers and cytokines of dendritic cells harvested from cord blood. *Clin Diagn Lab Immunol*. 2004;**11**(4):686-90.
255. Huda MN, et al. Bifidobacterium Abundance in Early Infancy and Vaccine Response at 2 Years of Age. *Pediatrics*. 2019;**143**(2).

256. Balamurugan R, et al. Faecal bifidobacteria in Indian neonates & the effect of asymptomatic rotavirus infection during the first month of life. *Indian J Med Res*. 2010;**132**(6):721-7.
257. Seppo AE, et al. Infant gut microbiome is enriched with Bifidobacterium longum ssp. infantis in Old Order Mennonites with traditional farming lifestyle. *Allergy*. 2021;**76**(11):3489-503.
258. Ismail IH, et al. Early gut colonization by Bifidobacterium breve and B. catenulatum differentially modulates eczema risk in children at high risk of developing allergic disease. *Pediatr Allergy Immunol*. 2016;**27**(8):838-46.
259. Hoang DM, Levy EI, Vandenplas Y. The impact of Caesarean section on the infant gut microbiome. *Acta Paediatr*. 2021;**110**(1):60-7.
260. Johansson MA, Sjögren YM, Persson JO, Nilsson C, Sverremark-Ekström E. Early colonization with a group of Lactobacilli decreases the risk for allergy at five years of age despite allergic heredity. *PLoS One*. 2011;**6**(8):e23031.
261. Penders J, et al. Intestinal lactobacilli and the DC-SIGN gene for their recognition by dendritic cells play a role in the aetiology of allergic manifestations. *Microbiology (Reading, England)*. 2010;**156**(Pt 11):3298-305.
262. Sjögren YM, Jenmalm MC, Böttcher MF, Björkstén B, Sverremark-Ekström E. Altered early infant gut microbiota in children developing allergy up to 5 years of age. *Clin Exp Allergy*. 2009;**39**(4):518-26.
263. Björkstén B, Naaber P, Sepp E, Mikelsaar M. The intestinal microflora in allergic Estonian and Swedish 2-year-old children. *Clin Exp Allergy*. 1999;**29**(3):342-6.
264. Arrieta MC, et al. Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting. *J Allergy Clin Immunol*. 2018;**142**(2):424-34.e10.
265. Fujimura KE, et al. Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med*. 2016;**22**(10):1187-91.
266. Hong PY, et al. Comparative analysis of fecal microbiota in infants with and without eczema. *PLoS One*. 2010;**5**(4):e9964.
267. Mah KW, et al. Distinct pattern of commensal gut microbiota in toddlers with eczema. *Int Arch Allergy Immunol*. 2006;**140**(2):157-63.
268. Murray CS, et al. Fecal microbiota in sensitized wheezy and non-sensitized non-wheezy children: a nested case-control study. *Clin Exp Allergy*. 2005;**35**(6):741-5.

269. Sepp E, Julge K, Mikelsaar M, Bjorksten B. Intestinal microbiota and immunoglobulin E responses in 5-year-old Estonian children. *Clin Exp Allergy*. 2005;**35**(9):1141-6.
270. Watanabe S, et al. Differences in fecal microflora between patients with atopic dermatitis and healthy control subjects. *J Allergy Clin Immunol*. 2003;**111**(3):587-91.
271. Björkstén B, Sepp E, Julge K, Voor T, Mikelsaar M. Allergy development and the intestinal microflora during the first year of life. *J Allergy Clin Immunol*. 2001;**108**(4):516-20.
272. Zimmermann P, Messina N, Mohn WW, Finlay BB, Curtis N. Association between the intestinal microbiota and allergic sensitization, eczema, and asthma: A systematic review. *J Allergy Clin Immunol*. 2019;**143**(2):467-85.
273. Storrø O, et al. Temporal variations in early gut microbial colonization are associated with allergen-specific immunoglobulin E but not atopic eczema at 2 years of age. *Clin Exp Allergy*. 2011;**41**(11):1545-54.
274. Suzuki S, Shimojo N, Tajiri Y, Kumemura M, Kohno Y. Differences in the composition of intestinal Bifidobacterium species and the development of allergic diseases in infants in rural Japan. *Clin Exp Allergy*. 2007;**37**(4):506-11.
275. Penders J, et al. Molecular fingerprinting of the intestinal microbiota of infants in whom atopic eczema was or was not developing. *Clin Exp Allergy*. 2006;**36**(12):1602-8.
276. Stsepetova J, et al. Molecularly assessed shifts of Bifidobacterium ssp. and less diverse microbial communities are characteristic of 5-year-old allergic children. *FEMS Immunol Med Microbiol*. 2007;**51**(2):260-9.
277. Ouwehand AC, et al. Differences in Bifidobacterium flora composition in allergic and healthy infants. *J Allergy Clin Immunol*. 2001;**108**(1):144-5.
278. Midtvedt AC, Midtvedt T. Production of short chain fatty acids by the intestinal microflora during the first 2 years of human life. *J Pediatr Gastroenterol Nutr*. 1992;**15**(4):395-403.
279. Elsden SR, Hilton MG. Volatile acid production from threonine, valine, leucine and isoleucine by clostridia. *Arch Microbiol*. 1978;**117**(2):165-72.
280. Lanjekar VB, Marathe NP, Ramana VV, Shouche YS, Ranade DR. Megasphaera indica sp. nov., an obligate anaerobic bacteria isolated from human faeces. *Int J Syst Evol Microbiol*. 2014;**64**(Pt 7):2250-6.

281. Oliver A, et al. High-Fiber, Whole-Food Dietary Intervention Alters the Human Gut Microbiome but Not Fecal Short-Chain Fatty Acids. *mSystems*. 2021;**6**(2).
282. Lucas S, et al. Short-chain fatty acids regulate systemic bone mass and protect from pathological bone loss. *Nature communications*. 2018;**9**(1):55.
283. Trompette A, et al. Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat Med*. 2014;**20**(2):159-66.
284. Boets E, et al. Systemic availability and metabolism of colonic-derived short-chain fatty acids in healthy subjects: a stable isotope study. *The Journal of physiology*. 2017;**595**(2):541-55.
285. Prentice PM, et al. Human Milk Short-Chain Fatty Acid Composition is Associated with Adiposity Outcomes in Infants. *J Nutr*. 2019;**149**(5):716-22.
286. Wu J, et al. NMR-based metabolite profiling of human milk: A pilot study of methods for investigating compositional changes during lactation. *Biochem Biophys Res Commun*. 2016;**469**(3):626-32.
287. Paparo L, et al. Butyrate as a bioactive human milk protective component against food allergy. *Allergy*. 2021;**76**(5):1398-415.
288. Arpaia N, et al. Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature*. 2013;**504**(7480):451-5.
289. Stinson LF, George AD. Human Milk Lipids and Small Metabolites: Maternal and Microbial Origins. *Metabolites*. 2023;**13**(3).
290. Gonçalves P, Gregório I, Martel F. The short-chain fatty acid butyrate is a substrate of breast cancer resistance protein. *Am J Physiol Cell Physiol*. 2011;**301**(5):C984-94.
291. Jonker JW, et al. The breast cancer resistance protein BCRP (ABCG2) concentrates drugs and carcinogenic xenotoxins into milk. *Nat Med*. 2005;**11**(2):127-9.
292. Cohen-Zinder M, et al. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res*. 2005;**15**(7):936-44.
293. Gutmann H, Hruz P, Zimmermann C, Beglinger C, Drewe J. Distribution of breast cancer resistance protein (BCRP/ABCG2) mRNA expression along the human GI tract. *Biochem Pharmacol*. 2005;**70**(5):695-9.
294. Månsson HL. Fatty acids in bovine milk fat. *Food Nutr Res*. 2008;**52**.

295. Hu M, et al. Decreased maternal serum acetate and impaired fetal thymic and regulatory T cell development in preeclampsia. *Nature communications*. 2019;**10**(1):3031.
296. Moffett JR, Puthillathu N, Vengilote R, Jaworski DM, Namboodiri AM. Acetate Revisited: A Key Biomolecule at the Nexus of Metabolism, Epigenetics and Oncogenesis-Part 1: Acetyl-CoA, Acetogenesis and Acyl-CoA Short-Chain Synthetases. *Front Physiol*. 2020;**11**:580167.
297. Nguyen TD, Hållenius FF, Lin X, Nyman M, Prykhodko O. Monobutyrin and Monovalerin Affect Brain Short-Chain Fatty Acid Profiles and Tight-Junction Protein Expression in ApoE-Knockout Rats Fed High-Fat Diets. *Nutrients*. 2020;**12**(4).
298. Han R, Nusbaum O, Chen X, Zhu Y. Valeric Acid Suppresses Liver Cancer Development by Acting as a Novel HDAC Inhibitor. *Molecular therapy oncolytics*. 2020;**19**:8-18.
299. Nguyen TD, Prykhodko O, Hållenius FF, Nyman M. Monobutyrin Reduces Liver Cholesterol and Improves Intestinal Barrier Function in Rats Fed High-Fat Diets. *Nutrients*. 2019;**11**(2).
300. Yassour M, et al. Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. *Sci Transl Med*. 2016;**8**(343):343ra81.