

# The nature of gut microbiota in early life

## Citation for published version (APA):

van Best, N. (2021). *The nature of gut microbiota in early life: origin and impact of pioneer species*. Gildeprint Drukkerijen. <https://doi.org/10.26481/dis.20210430nb>

## Document status and date:

Published: 01/01/2021

## DOI:

[10.26481/dis.20210430nb](https://doi.org/10.26481/dis.20210430nb)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# **THE NATURE OF GUT MICROBIOTA IN EARLY LIFE**

Origin and impact of pioneer species

**Niels van Best**



**Copyright © 2021 Niels van Best**

All rights reserved. For articles published or accepted for publication, the copyright has been transferred to the respective publisher. No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without prior permission of the author, or where appropriate, the publisher of the manuscript.

**ISBN:** 978-94-6419-162-2

**Cover:** Elif Swinkels & Maya van Best i.s.m. Bregje Jaspers | Proefschriftontwerp.nl

**Layout:** Wendy Bour-van Telgen

**Printing:** Gildeprint

The research presented in this thesis was conducted within RWTH Aachen and NUTRIM School of Nutrition and Translational Research in Metabolism of Maastricht University.

Printing of this thesis was financially supported by: Nederlandse Vereniging voor Gastroenterologie (NVGE), the Netherlands Society of Medical Microbiology (NVMM) and the Royal Netherlands Society for Microbiology (KNVM).

# **THE NATURE OF GUT MICROBIOTA IN EARLY LIFE**

Origin and impact of pioneer species

DISSERTATION

to obtain the degree of Doctor at Maastricht University and RWTH Aachen University,  
on the authority of the Rector Magnificus,  
Prof.dr. Rianne M. Letschert  
in accordance with the decision of the Board of Deans,  
to be defended in public  
on Friday April 30<sup>th</sup> 2021, at 14:00 hours

by

**Niels van Best**

**Supervisors:**

Prof. Dr. P.H.M. Savelkoul

Prof. Dr. M.W. Hornef (RWTH University Hospital Aachen, Germany)

**Co-supervisor:**

Dr. J. Penders

**Assessment Committee**

Prof. Dr. D.M.A.E Jonkers (chair)

Prof. Dr. T. Clavel (RWTH University Hospital Aachen, Germany)

Prof. dr. L.J. Hall (Technical University of Munich, Germany)

Prof. Dr. M.C. Jenmalm (Linköping University, Sweden)

Dr. A.C.E. Vreugdenhil

**Voor mama en papa**

*The disturbance of ecosystems, whether small or big, could have dramatic implications  
and therefore we need to cherish them*



## Table of contents

<b>Chapter 1</b>	General introduction	9
<b>Chapter 2</b>	On the origin of species: factors shaping the establishment of infant's gut microbiota	21
<b>Chapter 3</b>	Bile acids drive the newborn's gut microbiota maturation	45
<b>Chapter 4</b>	Gut colonization by methanogenic archaea is associated with organic dairy consumption in children	85
<b>Chapter 5</b>	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	113
<b>Chapter 6</b>	Gut microbiota in wheezing preschool children and the association with childhood asthma	159
<b>Chapter 7</b>	Probiotic-mediated changes in the developing neonatal microbiota	177
<b>Chapter 8</b>	General discussion and summary	211
<b>Addendum</b>	Samenvatting	236
	Zusammenfassung	241
	Impact	247
	Acknowledgements	250
	About the author	254
	List of Publications	255





# **Chapter 1**

**General introduction**

**The (in)visible microbial world**

More than 300 years ago, the Dutch Antonie van Leeuwenhoek stared through his self-made microscope and discovered the first microbes colonizing human body surfaces. Since van Leeuwenhoek's pioneering work on microbial life, microbiota research has revolutionized the area of biology, and at this moment microbiology is one of the fastest growing fields in science. The role of microorganisms in non-communicable and infectious diseases has been emphasized ever since. These discoveries all challenged our fundamental view on the microbiota and its complexity, and enormously increased the awareness for our resident microbes as part of ourselves. Here, I will outline the size and complexity of the microbiota in our intestinal tract, its importance in health and disease and its developmental origin in early life. In addition, I highlight the methodologies used to study the microbiota and define the aims of this thesis.

We harbor thousands of different microbial species and trillions of cells in our gut, including archaea, bacteria and eukaryotic microorganisms [1]. Collectively, the microbial cells even exceed the number of human cells. Recent calculations suggest that we are composed of approximately 44% human cells [2] [3], albeit the small size of these microbial cells only add a few kilograms to the adult body weight. Although our usual thought considers the human genome to characterize us as a *Homo sapiens*, our own genes are outnumbered at least 100 times by the 2 to 20 million microbial genes and thus comprise only a minor portion of the total gene pool within our body. In contrast to the human genome, this metagenome only starts to settle from birth onwards [1, 4, 5]. As such, we actually are holobionts: hosts that are accompanied with a living microbiome.

The different genes of the gut microbiota encode for molecules involved in pathways and functions that affect a wide range of regulatory circuits such as immunity, nutrition and metabolism [6-8]. No other microbiome in the human body is as numerous and diverse as the gut microbiome. The complex interactions of the gut microbiome with the continuous flow of environmental microorganisms, the wide variety of dietary compounds and host factors is a research field of high interest. The gut microbiota is able to provide a competitive barrier and thereby protect against infections. The gut microbiota is also of importance in food fermentation and involved in the maturation of the immune system. Additionally, increasing evidence suggests that the gut microbiota profoundly influences the etiology and progression of immune-mediated disorders like allergies, neurological disorders like Alzheimer's disease and autism, and metabolic disorders such as diabetes and obesity [9]. In 1989, Dr. David Strachan formulated the so-called 'hygiene hypothesis' [10] to explain the

rising incidence of many non-communicable diseases over the past 50 years in industrialized countries. His initial hypothesis attributed the lower allergy risk among children with older siblings to their increased exposure to childhood infections. In the following decades this hypothesis has been revisited multiple times and led to more rational explanations for the link between microbial exposure and inflammatory diseases, such as the old friends hypothesis introduced by Graham Rook [10]. The lack of exposure to all kinds of microbes, as a result of e.g. antibiotic (over)use, rising cesarean sections rates and western (dietary) lifestyle, has been suggested to make young people nowadays more susceptible to non-communicable diseases, particularly allergies. Moreover, the observation, that certain conditions are associated with specific changes in the intestinal microbiota composition makes it an intriguing target for new treatments. Whereas a lot of research has focused on adults, there is more and more interest in the development of the early-life microbiota. In contrast to the microbiota of healthy adults, which is quite stable and resilient over time (e.g. due to colonization resistance), this early age-window is characterized by profound changes in indigenous microbial composition. Moreover, the microbiota in early life seems to be particularly crucial for immune development and maturation and perturbations have been associated with adverse outcome. As such, the neonatal window may provide the best opportunity for targeted manipulation of the microbiota, potentially even with long-lasting positive health consequences.

### **Studying the microbiota: methodologies and study designs**

Until the 1990s, knowledge of the gut microbiota was limited because traditional culturing of microorganisms represented the main approach to study its composition [11]. Since then, the scientific community has developed innovative, mainly molecular, methods to investigate the microbial world. The human gut mainly harbors strictly anaerobic microbial species that are difficult to culture and to date only a minor part of the microbiota has been cultured and thus functionally investigated. Instead, several culture-independent techniques have been introduced in recent decades to analyze composition and complexity of the intestinal microbiota. These techniques include (quantitative) polymerase chain reaction PCR [(q)PCR], PCR followed by denaturing gradient gel electrophoresis (DGGE), fluorescent in situ hybridization (FISH), and DNA microarrays that hybridize ribosomal RNA (rRNA) gene sequences with oligonucleotide probes. Although these methods are useful for microbial profiling, they usually do not provide detailed taxonomic data. For that reason, the most important techniques nowadays are high-throughput next-generation

sequencing followed by in-depth bioinformatic analysis. These sequence-based methods are commonly based on the analysis of the 16S rRNA gene, present in all bacteria and archaea, consisting of nine unique hypervariable regions (V1–V9) flanked by highly conserved regions [12]. Despite the fact that other genes have been proposed, sequencing of 16S rRNA gene regions remains the gold standard to analyze the microbiota in view of the completeness of reference databases, low costs, and advanced bioinformatics software available. Different next-generation sequence equipment can be used to analyze 16S rRNA gene amplicons. However, to date Illumina sequencing (Illumina Inc., San Diego, CA) is the most commonly used platform. This widely popular approach has high coverage and a low cost per sample [13]. It should be noted that the selection of primers, the number of sequences per sample, costs, and aim of the experiment are strongly interdependent. Therefore, it is essential to balance these factors to achieve the optimum amount of information [14].

As opposed to sequencing of marker genes such as the 16S rRNA gene, sequencing of the entire genomic content of microorganisms (the ‘microbiome’) within the microbiota provides more specific information on their potential functional roles. These metagenomic or metatranscriptomic (when sequencing mRNA rather than DNA) analyses, referred to as shotgun sequencing, are especially appropriate for linking microbial communities with functional potential and activity. In addition, metabolomics—the quantitative assessment of metabolic responses of organisms to genetic or pathophysiological changes—is a powerful approach and also key to unraveling specific host-microbe interactions. Metagenomic approaches relate to the microbiome, and metabolomic approaches to the metabolic profile with disease phenotypes [15]. The latter analyses are principally based on proton nuclear magnetic resonance ( $^1\text{H-NMR}$ ) spectroscopy and mass spectrometry methods. Ideally, these analytical techniques should be complemented for the most detailed characterization of microbial metabolites obtained from feces, blood, urine, or intestinal tissues.

As for the design of studies, experimental studies and large-scale well-defined prospective cohorts of patients that have been carefully phenotyped are essential. In these studies, factors known to influence gut microbiota composition should be carefully taken into account. In particular, dietary aspects and use of antibiotics or probiotics should be well-documented and should be considered when setting up microbiota studies. Cross-sectional studies are neither able to unravel if exposure to determinants precede changes in the microbiome nor if changes in the microbiome occur prior to disease manifestation. Longitudinal studies are therefore of high importance to increase our understanding, because they are able to examine both

the (in)stability of the microbiome within individuals and differences between subjects and allow us to investigate how the microbiota changes over time. These study designs are able to determine the development of the microbiome and to unravel associated determinants as well as diseases. In contrast to the majority of other microbiome publications, my thesis is solely based on studies with a longitudinal design, involving both animal models and human studies and provides more insight on the development of the microbiome in early life and its role for the onset of inflammatory diseases.

### **Aims and outline of the thesis**

A better understanding of the interplay between the gut microbiota, us as a host and the factors that contribute to changes in the microbial composition may provide novel interventional strategies to support human health. Early childhood appears to represent a critical age-window, since major alterations occur during this time and the microbiota establishes to form a life-long stable ecological system. To this end, this thesis aimed to unravel environmental and host-related determinants that drive development and inter-individual variations in the infant microbiota. Identifying these determinants is pivotal in order to better understand the role of the microbiota in health and disease.

To achieve this aim, the following objectives have been addressed:

- (1) Identify (metabolic and nutritional) factors that influence the establishment of the microbiota in the neonate host;
- (2) examine how changes in the microbial composition during early life relate to the onset of allergic diseases, and;
- (3) explore if manipulation of the neonatal microbiota can be achieved in order to improve health.

**In chapter 2**, the literature on the current knowledge of pioneer bacteria and the evolutionary factors that influence the development of the gut microbiota in infants is reviewed. In addition, future perspectives including techniques to manipulate and promote the succession of the microbiota for prevention and treatment of disorders during infancy will be discussed.

The postnatal development of both the intestine and liver is accompanied by increased enzymatic, absorptive and metabolic capacity [16, 17]. The related changes may contribute to luminal substrate availability and could thereby influence the bacterial succession and overall composition [18]. Although it is known that metabolites impact the microbial environment in the gut, data on how and when metabolites affect the colonization of the microbiota are lacking. In **chapter 3**, we present a study on the impact of the developing host metabolism, in particular the hepatic synthesis of bile acids, on the site-specific gut microbiota composition. First, we determined the gut microbiota development in the small intestine and colon from birth, early infancy and weaning until adulthood in mice. Subsequently, we screened the metabolites followed by correlation analysis to identify bile acids as strong drivers of the microbiota development. Finally,

oral administration of bile acids was performed to confirm their causal role herein.

Although a variety of microbes have been studied within the human infant gut, important members that are often ignored are the methanogenic archaea despite their potential role in the promotion of a number of gastrointestinal and metabolic diseases [19]. Archaea have been shown to be mostly absent during infancy while omnipresent in school-aged children [20], suggesting that colonization may result from environmental exposure during childhood. Therefore, in **chapter 4**, we explored potential determinants associated with the acquisition of methanogenic archaea in children by analyzing fecal samples combined with extensive data on environmental factors within the context of the KOALA Birth Cohort Study.

Asthma and allergic diseases are the most common pediatric chronic disorders affecting more than 300 million children globally. Recently, the influence of microbial changes in the gut on the onset of asthma gained widespread attention. However, it is of particular importance taking both early samples and potential confounders into account with a longitudinal design. For instance, it remains unknown when and which microbes could potentially influence the onset of allergies, asthma or atopic dermatitis. In **chapter 5** we therefore investigated the dynamics of the infant's microbiome by monitoring the development of the gastrointestinal microbiota during infancy at several timepoints in the first year of life and once at school-age and examined a possible association with subsequent onset of asthma and allergies.

Continuing on allergic diseases, little is known on the association between the gut microbiome and the progression to asthma in wheezing preschool children. As reliable biomarkers to predict which preschool children with wheezing symptoms will ultimately develop asthma are currently lacking, the identification of potential gut microbial biomarkers is of particular relevance. In **chapter 6**, we examined the gut microbiome in wheezing children and healthy controls aged 2-4 years and the link with immune modulators, sensitization and asthma at the age of 6 years within the Asthma DEtection and Monitoring (ADEM) study.

Oral administration of specific beneficial bacteria, so-called probiotics, might represent a way for targeted manipulation of the gut microbiota. Following preterm birth, the neonate host is more susceptible to infections and overwhelming inflammatory diseases such as necrotizing enterocolitis (NEC), an inflammatory disease of the intestinal tissue associated with high morbidity and mortality [21-23]. Although supplementation of probiotics has been reported to decrease the incidence of NEC [24], the precise influence on the microbiota during and following the intervention has remained incompletely understood. In particular, the question whether and to what extent orally administered probiotics persistently colonize in preterm



neonates and how they impact the microbial ecosystem and the development of this disease has not been fully addressed. In **chapter 7**, we examined these questions by comparing the microbiome of preterm neonates in the absence of as well as before, during and after oral supplementation of two different probiotic regimens within the context of a natural experiment.

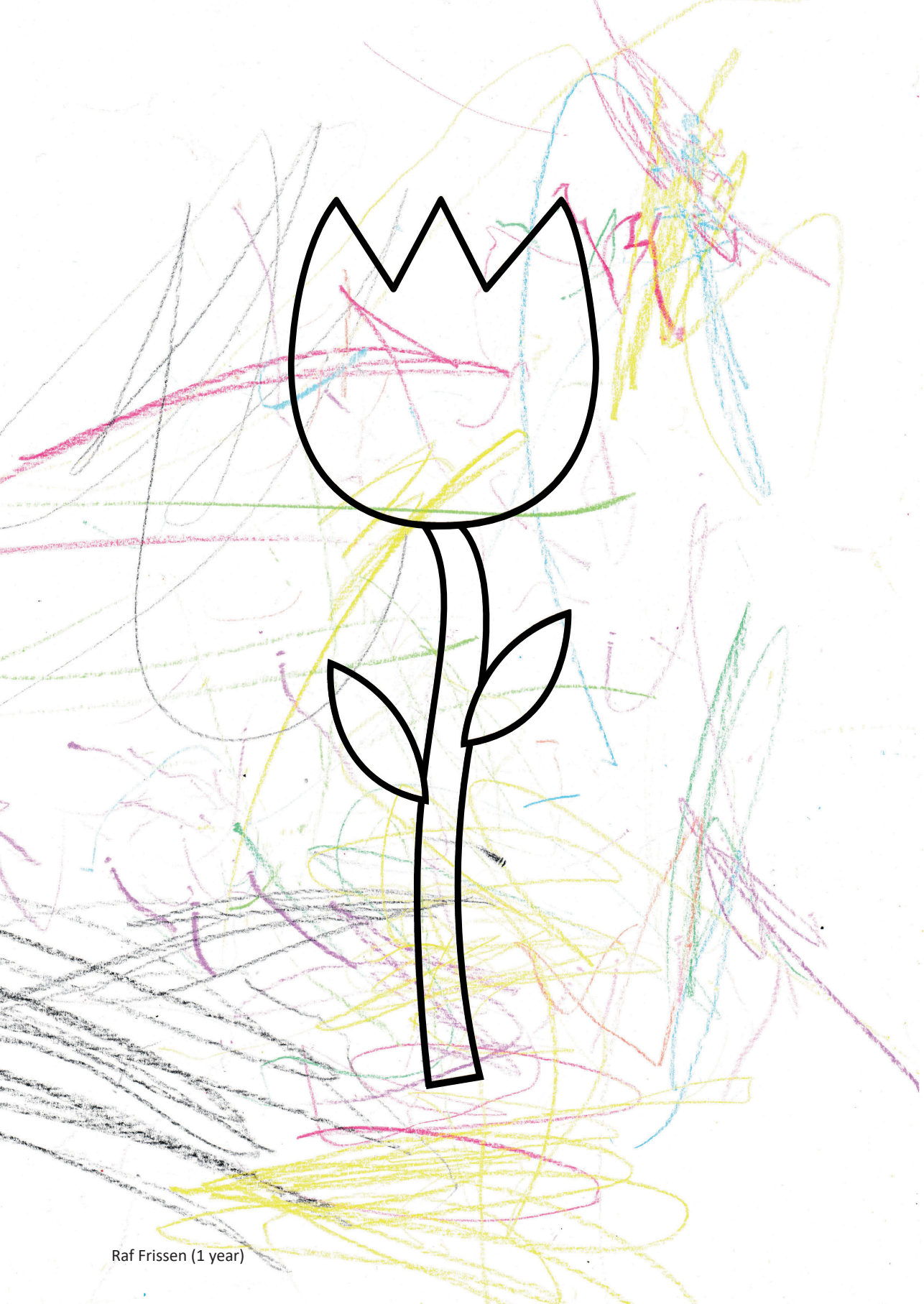
Finally, in **chapter 8**, all findings and results of the studies are summarized and discussed. In addition, the contribution of this research to our knowledge of the basic mechanisms of the gut microbiome establishment in early life and maintenance of the life-long host microbial interactions is evaluated. Moreover, future perspectives and novel strategies for medical interventions in this unique period of early infancy are discussed.

## References

1. Qin, J., et al., *A human gut microbial gene catalogue established by metagenomic sequencing*. Nature, 2010. 464(7285): p. 59-65.
2. Sender, R., S. Fuchs, and R. Milo, *Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans*. Cell, 2016. 164(3): p. 337-40.
3. Sender, R., S. Fuchs, and R. Milo, *Revised Estimates for the Number of Human and Bacteria Cells in the Body*. PLoS Biol, 2016. 14(8): p. e1002533.
4. Li, J., et al., *An integrated catalog of reference genes in the human gut microbiome*. Nat Biotechnol, 2014. 32(8): p. 834-41.
5. Turnbaugh, P.J., et al., *The human microbiome project*. Nature, 2007. 449(7164): p. 804-10.
6. Cho, I. and M.J. Blaser, *The human microbiome: at the interface of health and disease*. Nat Rev Genet, 2012. 13(4): p. 260-70.
7. Clemente, J.C., et al., *The impact of the gut microbiota on human health: an integrative view*. Cell, 2012. 148(6): p. 1258-70.
8. Kau, A.L., et al., *Human nutrition, the gut microbiome and the immune system*. Nature, 2011. 474(7351): p. 327-36.
9. Schroeder, B.O. and F. Backhed, *Signals from the gut microbiota to distant organs in physiology and disease*. Nat Med, 2016. 22(10): p. 1079-1089.
10. Strachan, D.P., *Family size, infection and atopy: the first decade of the "hygiene hypothesis"*. Thorax, 2000. 55 Suppl 1: p. S2-10.
11. Schaedler, R.W., R. Dubos, and R. Costello, *The Development of the Bacterial Flora in the Gastrointestinal Tract of Mice*. J Exp Med, 1965. 122: p. 59-66.
12. Woese, C.R., *Bacterial evolution*. Microbiol Rev, 1987. 51(2): p. 221-71.
13. Caporaso, J.G., et al., *Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample*. Proc Natl Acad Sci U S A, 2011. 108 Suppl 1: p. 4516-22.
14. van Best, N., P.L. Jansen, and S.S. Rensen, *The gut microbiota of nonalcoholic fatty liver disease: current methods and their interpretation*. Hepatol Int, 2015. 9(3): p. 406-15.
15. Morgan, X.C. and C. Huttenhower, *Meta'omic analytic techniques for studying the intestinal microbiome*. Gastroenterology, 2014. 146(6): p. 1437-1448 e1.
16. Henning, S.J., *Ontogeny of enzymes in the small intestine*. Annu Rev Physiol, 1985. 47: p. 231-45.
17. Grijalva, J. and K. Vakili, *Neonatal liver physiology*. Semin Pediatr Surg, 2013. 22(4): p. 185-9.
18. Backhed, F., et al., *Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life*. Cell Host Microbe, 2015. 17(5): p. 690-703.
19. Horz, H.P., *Archaeal Lineages within the Human Microbiome: Absent, Rare or Elusive?* Life (Basel), 2015. 5(2): p. 1333-45.
20. Dridi, B., et al., *Age-related prevalence of Methanomassiliococcus luminyensis in the human gut microbiome*. APMIS, 2012. 120(10): p. 773-7.
21. Stewart, C.J., et al., *Temporal bacterial and metabolic development of the preterm gut reveals specific signatures in health and disease*. Microbiome, 2016. 4(1): p. 67.

22. Hill, C.J., et al., *Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort*. *Microbiome*, 2017. 5(1): p. 4.
23. Gopalakrishna, K.P., et al., *Maternal IgA protects against the development of necrotizing enterocolitis in preterm infants*. *Nat Med*, 2019. 25(7): p. 1110-1115.
24. Dermyshe, E., et al., *The "Golden Age" of Probiotics: A Systematic Review and Meta-Analysis of Randomized and Observational Studies in Preterm Infants*. *Neonatology*, 2017. 112(1): p. 9-23.





# Chapter 2

## **On the origin of species: factors shaping the establishment of infant's gut microbiota**

**Niels van Best**, Mathias Hornef, Paul H.M. Savelkoul, John Penders

*Birth Defects Res C Embryo Today. 2015 Dec;105(4):240-51*

## Abstract

The human gut microbiota is a complex and dynamic ecosystem, which naturally lives in a symbiotic relationship with the host. Perturbations of the microbial composition (dysbiosis) and reduced diversity may promote disease susceptibility and recurrence. In contrast to the mature intestinal microbiota of healthy adults, which appears relatively stable over time, the infant's microbiome only establishes and matures during the first years of life. In this respect, early childhood seems to represent a crucial age-window in disease prevention since microbial diversification and maturation of the microbiome primarily occurs during this period of life. A better understanding of ecological processes and pioneer consortia in microbial development is crucial in order to support the development of a beneficial microbiota. Various deterministic and stochastic aspects seem to shape the microbiome in early life, including maternal, environmental and host factors. Here, we review the current understanding of the origin of pioneer bacteria and the evolutionary factors that influence the development of the gut microbiota in infants. In addition, future perspectives including manipulating and promoting the succession of initial bacteria during infancy will be highlighted.

## Introduction

The human gut microbiota is a complex and dynamic ecosystem consisting of hundreds to thousands of distinct bacterial species, as well as microbial eukaryotes, viruses, phages and a limited number of Archaea, which naturally live in a symbiotic relationship with the host [1]. In the past decades, the knowledge on the role of the gut microbiota and its development as well as host-microbial interactions in human health and disease has rapidly increased due to the advancement of modern molecular technologies. On one hand, the microbiota has a profound impact on its host by providing a competitive barrier against invading pathogens, utilizing undigested food components and producing essential metabolites, modulating immune responses and stimulating intestinal maturation [2, 3]. On the other hand, the host's immune system needs to tolerate the luminal microbiota, but avoid their overgrowth and translocation of bacteria into the subepithelial tissue. As such, a perturbation of the microbial communities (dysbiosis) has the ability to promote disease susceptibility and progression. Indeed, alterations in microbiota composition and function have been associated with many diseases, including metabolic diseases (i.e. obesity, non-alcoholic fatty liver disease) [4, 5] and immune-related diseases (i.e. inflammatory bowel disease, allergies) [6-8]. Moreover, dysbiosis has also been linked to neoplastic diseases [9], mental disorders and autism-spectrum disorder [10], although it has to be noted that for many of these diseases causality has still to be proven. The prevalence of many of these inflammatory non-communicable diseases (NCD's) has increased rapidly over the past few decades, especially in western countries. As such, microbiota perturbations as a consequence of modern lifestyle have often held (in part) accountable for the rise in prevalence of NCD's [11].

Our knowledge on the intestinal microbial ecophysiology and the processes contributing to microbial dysbiosis has a direct impact on our ability to manage and maintain human health. In this respect, early childhood appears a crucial age-window since diversification and maturation of the microbiota primarily occurs during this period of life. In contrast to the mature intestinal microbiota of healthy adults, which seems to be relatively stable over time, the infant microbiota waxes and wanes as diet changes and the host develops. Longitudinal studies monitoring the dynamics of the newborn microbiota [12-14] showed that periods of relative stability were interrupted by abrupt changes in microbial community structure and diversity. In some, but not all cases, these changes can be linked to preceding life events (e.g. weaning, fever, medication), indicating that both deterministic and stochastic aspects



shape the microbiota in early life.

Here, we review the origin of pioneer bacteria and evolutionary factors that influence the development of the gut microbiota in infants. In addition, the potential of manipulating the microbiota for prevention and treatment of disorders in early life will be discussed.

## Factors influencing the intestinal microbiota development

### Maternal exposure

#### *In utero transmission*

Until recently, the initial microbial colonizers of the neonatal intestine were supposed to originate from the maternal perineal, vaginal and fecal microbiota. However, the development of the microbiota composition might already start before birth by microbial transfer through the placental barrier. Recent sequencing studies revealed the presence of DNA derived from a wide variety of microbial taxa in the human placenta, umbilical cord blood, amniotic fluid and meconium [15-18]. Although the presence of bacteria such as *Mycoplasma* and *Ureaplasma* have long been implicated in invasion of the amniotic cavity and detrimental health effects, such as preterm birth and necrotizing enterocolitis [18-20], these recent findings suggest that normal colonization may already start before rupture of membranes and birth. The placental microbiota has been suggested to resemble the oral microbiota from the Firmicutes, Tenericutes, Proteobacteria, Bacteroidetes and Fusobacteria phyla, indicating a possible mode of bacterial transmission to the fetus [15]. This evidence is, however, indirect as the investigators did not compare the placental and oral microbiota of the same subjects but rather examined the oral microbiota in non-pregnant individuals. An experimental animal study using labeled *Enterococcus faecium* further evidenced that maternal bacteria can be transferred to fetuses *in utero* via the gastrointestinal tract [16]. The underlying mechanism for this bacterial translocation remains unclear, but several possibilities have been proposed, i.e. via penetration of dendritic cells through the gut epithelium [17]. Interestingly, bacterial translocation of the intestinal mucosa seems to be highly increased in pregnant mice in comparison to non-pregnant animals [21].

Moreover, a number of studies detected a low abundant microbiota in the meconium of preterm and term infants using culture as well as molecular methods [16, 22], implying that the neonatal gut may not be sterile at birth. Culture techniques revealed that *Staphylococcus* and *Enterococcus* predominated in the meconium [16]. DNA-based techniques confirmed the predominance of these genera together with *Enterobacteriaceae* [22]. The source of meconial bacteria remains unidentified, but they might have an intrauterine origin and be derived from swallowed amniotic fluid. The observation that the major genera observed in meconium are similar to those detected in amniotic fluid [18] and less related to the bacteria prevailing in

the maternal oral or vaginal microbiome, is supportive for an intrauterine origin [22].

Although these studies indicate that the neonatal gut might already become colonized with bacteria before birth, the majority of studies relied on culture-independent techniques and therefore culture-based methods are ultimately required to establish a *prepartum* microbiota. The detected bacterial DNA sequences may originate from dead bacteria or free bacterial DNA in amniotic fluid, placenta and meconium specimens. Moreover, not birth, i.e. transit through the birth canal, but rather rupture of the membranes exposes the fetus to the environment and thus needs to be considered as critical endpoint. Finally, the actual origin of these bacteria or bacterial remnants needs to be considered. Sampling of material from the same pregnant women at multiple sites may help to exclude local infection or sample contamination. Clearly, further research is necessary to clarify whether the potential *in utero* bacterial transmission of a low abundant inoculum truly contributes to the development of a healthy neonatal microbiome, immune maturation and metabolism or merely reflects the vulnerability that pregnancy inflicts to the maternal host.

### ***Mode of delivery***

The microbial colonization of the human intestine has been traditionally assumed to begin at birth and further develop during early life to a diverse and mature adult microbiota. Within a couple of days after birth, the neonatal intestine becomes rapidly colonized by maternal and environmental bacteria, strongly affected by the delivery mode [23, 24]. Infants born vaginally (i.e. passing the birth canal in intimate contact with the vaginal mucosa) display a microbiota similar to that of the maternal vagina while infants delivered by cesarean (C-) section demonstrate a microbiota mainly consisting of commensal skin bacteria [23]. Accordingly, the initial microbiota in vaginally delivered newborns is determined by the vaginal milieu of the mother, which is predominated by *Lactobacillus* as well as by a high abundance of *Prevotella* and *Sneathia spp.* [23, 25, 26]. Infants born by C-section seem to have in general delayed and decreased colonization of *Bifidobacterium spp.* and *Bacteroides spp.* and increased rates of *Clostridium difficile* and typical skin bacteria such as *Staphylococcus*, *Corynebacterium*, and *Propionibacterium spp.* compared to vaginally delivered infants [23, 24, 27-30]. Increased rates of *C. difficile* may result from environmental exposure since *C. difficile* spores can be found widespread in the hospital setting. Consistently, *C. difficile* has been shown to be more prevalent in neonates delivered in the hospital as compared to those born at home. Moreover, the risk of colonization by this bacterium increased with each

additional day of hospitalization [24]. Lacking the inoculum of the maternal microbiota, the hospital environment and staff may represent an important source of bacteria colonizing the intestinal tract of C-section delivered infants. Longitudinal studies showed that the delayed and altered colonization pattern in C-section delivered infants persisted at least to the age of one year [28, 30], and even after seven years of age minor differences could be detected [31]. However, a potential functional impact on infant's development and wellbeing has not been examined. In addition, Salminen et al. visualized specific bacteria by *fluorescent in situ hybridization* (FISH) and therefore comprehensive data on a fully established microbiota in later childhood in relation to the mode of delivery are still lacking.

Given the profound impact of the mode of delivery on the establishment of the neonatal microbiome and its potential consequences, the continuously increasing C-section rates in many part of the world are worrisome. Recent findings based on the ecological association of cesarean delivery rates and perinatal outcomes from 19 developed countries [32], corroborated previous suggestions of the WHO expert panel that population-level C-section delivery rates above 10–15 percent are hardly justified by medical reasoning. Yet, during the past decades C-section rates have rapidly increased far above 10-15% in many countries mainly attributable to non-evidence-based indications, professional convenience, maternal request, and over-medication of child-birth [33]. In many European countries, the United States, the larger cities in China and Australia rates have risen up to above 30%, while rates as high as 45% have been reported for Brazil [34].

The increased risk of obesity [35] and childhood asthma [36] in children born by C-section has been attributed to an altered intestinal colonization pattern in these children. Indeed, within the context of two birth cohort studies, mode of delivery was shown to strongly impact the infant microbiota and affect the risk of allergies and asthma [37]. Moreover, the decreased abundance of *Bifidobacterium* and higher numbers of *S. aureus* in the intestine during early life has been associated with subsequent overweight and obesity in childhood [38], emphasizing the critical role of the first exposure.

### ***Infant feeding***

Breastfeeding plays both directly, by containing living bacteria, and indirectly, by providing prebiotic nutrients and bioactive components, an important role in the establishment of the neonatal microbiota. The first abundant bacteria of the newborn's gut are generally facultative anaerobes such as *Escherichia*, *Streptococcus* and *Enterococcus* probably due

to the oxygen-rich environment in the neonatal gut lumen [30, 39]. Although some studies observed a gradual shift in microbial composition and increase in diversity [40], other studies demonstrated chaotic shifts over time including a transient drop in diversity within the first week of life [39, 41]. The lack of extensive bacterial competition results in high population levels in neonates, exceeding the numbers found in adults [42]. Facultative anaerobes utilize oxygen and thereby create a reduced environment, making it more favorable for subsequent colonization of strict anaerobe bacteria. Therefore, the microbiota development continues in the first weeks with proliferation of strictly anaerobic bacteria that are able to convert lactic acid derived from to the milk-rich diet. Infants that are exclusively breast-fed in general harbor a microbiota dominated by bifidobacteria, while exclusively formula-fed infants harbor a more diverse microbiota with increased abundance of *Escherichia coli*, *Clostridia* and *Bacteroides* [24, 27, 29, 30, 43]. A recent study revealed by deep sequencing that formula-fed infants displayed enhanced populations of *Ganulicatella*, *Citrobacter*, *Enterobacter* and *Bilophila* whereas breast-fed infants showed increased *Lactobacillus* species apart from bifidobacteria [30]. The beneficial effects of human milk oligosaccharides (HMO's), a structurally diverse family of unconjugated glycans that are highly abundant in human milk and absent in formula nutrition, might cause the differences in the microbiota composition between formula-fed and breast-fed infants. The maternal HMO's have a prebiotic role and promote the growth of *Bifidobacterium*, increase microbial diversity and are able to protect the infant against a number of enteropathogenic microorganisms [44]. The quantity and composition of these oligosaccharides varies over the course of the lactation period, but also between individual women, mainly driven by their genotype [45, 46].

Additionally, breastmilk contains other immunologic compounds, i.e. IgA, lactoferrin, lysozyme [47]. All these factors are able to interact with pathogens, reduce their viability and prevent their attachment to the intestinal mucosal surface, assisting the expansion of commensal bacteria. Therefore, factors such as the IgA repertoire might shape a distinct neonatal microbiota [48]. Vitamin D, which is part of infant formulas and recommended supplement in breastfed infants, is another compound that has been suggested to influence the developing microbiome [49], given its role in the development of regulatory T-cells and dendritic cells [50]. Direct evidence, however, is limited to a single study correlating a lower vitamin D intake to a different microbiota composition in adult African Americans as compared to Caucasian Americans [51]. The results need to be confirmed in larger longitudinal studies.

Interestingly, recent studies also revealed the presence of a microbiota in maternal

breastmilk. With concentrations of  $10^2$ - $10^4$  viable bacterial per mL breast milk, an exclusively breastfed infant will consume the significant amount of about  $10^5$ - $10^7$  commensal bacteria per day [52]. A study on variations in the breastmilk microbiome over the course of lactation indicated that staphylococci, streptococci, and lactobacilli derived from the milk ducts and skin predominated in colostrum, whereas after six months a higher abundance of oral taxa, including *Veillonella* and *Prevotella*, were found in breastmilk [45]. Other studies hypothesize an entero-mammary pathway, postulating that maternal gut bacteria could translocate through the intestinal epithelial barrier, migrate to the mammary gland and subsequently become secreted with the milk and colonize the breastfed neonate [53]. Using a combination of culture-dependent and -independent methods, Jost et al. examined whether viable strains of gut-associated obligate anaerobes were shared between the maternal and neonatal gut microbiota and transferred via breastfeeding. Next to facultative anaerobes, obligate anaerobic genera, including *Bifidobacterium*, *Bacteroides* and *Clostridia* were detected in maternal faeces, breast milk and neonatal feces. Moreover, pulse-field gel electrophoreses (PFGE) of *Bifidobacterium breve* isolates from maternal feces, breast milk and neonatal feces of a single mother-child pair revealed identical restriction profiles, suggesting transfer of these strains from the maternal to the neonatal gut via breast milk [54]. Alternatively, the maternal gut might represent a common source for both the mammary skin/gland duct bacteria as well as the neonatal intestine.

The third stage in the development of the microbiota is defined by alterations of the microbial composition and function due to the introduction of solid food 4-6 months post-partum (weaning). Upon weaning, the microbiota functionally matures by a decrease in the relative abundance of genes involved in the degradation of sugars from breast milk and towards enrichment of genes involved in the degradation of complex sugars and starch. The major pectin-degrading enzyme, pectinesterase, has been shown to be enriched in infants by the age of 12 months, likely in response to the increased intake of foods rich in pectin by this age [30]. The particular effects of solid foods, however, strongly depend on the geography linked to major differences in dietary habits around the world. In western countries, the introduction of a more complex diet rapidly increases the levels of *Bacteroides* and *Clostridium*, alters the lactobacilli communities and decreases the *Bifidobacterium* levels [30, 39, 41, 55]. The intestinal microbiota closely resembles the diverse adult-like composition at the age of three years [41, 55], even though older data indicated that microbial stability is already achieved at the age of one year [12, 40]. Interestingly, a longitudinal study of 98 Swedish mothers and

their infants also compared the bacterial composition between the introduction of solid food with and without continuation of breastfeeding [30]. This comparison suggested that the cessation of breastfeeding rather than the introduction of solid food represents the major factor driving the microbiota composition towards a more adult-like ecology. Altogether, these data emphasize the role of breastmilk and diet in the succession and shaping of the intestinal microbiota in early life.

### **Environmental Exposure**

Besides the influence of the mother and the associated exposures during delivery, the child's living environment, including the indoor home environment, pets and older siblings as sources for microbial exposure, has been shown to affect the intestinal microbial development. A recent longitudinal study by the Home Microbiome Consortium, showed substantial interaction among the human, home and pet microbiota [56]. Moreover, strikingly similar dynamics in the development of the enteric microbiota of a fraternal twin pair, including the coincidental transient appearance of specific organisms in both twins, suggest that stochastic environmental influences can also play an important role during the assembly of microbial communities [12].

### **Siblings**

Singletons seem to have a distinct colonization pattern in comparison to infants that grow up together with older siblings, although data are scarce and not always consistent. The Canadian Healthy Infant Longitudinal Development (CHILD) cohort showed a lower abundance of *C. difficile* and its family *Peptostreptococcaceae* [27], whereas a German study reported decreased colonization rates of Clostridium cluster I in newborns with older siblings as compared to singletons [57]. The higher levels of *Clostridium* species in singletons are in accordance with another study comprising three European birth cohorts [58]. A decrease in bifidobacteria in infants without older siblings has also been reported [24]. Moreover, a higher microbial richness and diversity has recently been observed in infants with older siblings [59]. Altogether, these data suggests that firstborn infants harbor a less mature microbiota than infants with older siblings at a certain age, which may impact the risk on allergies [37]. Whether this occurs due to direct or indirect bacterial transfer between siblings or due to other factors, such as altered vaginal microbiota and breastmilk composition in multiparous as compared to primiparous women, remains to be elucidated.

### ***Pets & house dust***

There is limited data available on the role of in-house microbial exposure in humans, but an association between the bacteria found in house dust and in children has recently been reported [60]. Moreover, Azad et al. observed that the presence of pets might affect infant colonization [61], in contrast to studies which did not find a relation between the presence of household pets and the human gut microbiota [24, 60]. Pets might also cause the occurrence of certain bacteria in house dust and transfer of specific communities to humans. Infants living with pets showed mainly increased *Peptostreptococcaceae* (*C. difficile*) and decreased abundance of bifidobacteria at 4 months of age. The discrepancy between studies might indicate the minor role of pets and more research is necessary to confirm the association and identify transfer pathways.

### ***Geography***

The microbial ecology also differs between infants from various geographical habitats. Children born in developing countries are characterized by enhanced levels of *Prevotella* and decreased abundance of *Bacteroides* in early life as compared to infants living in western countries [41, 62-64]. The geographical variation in colonization patterns might mainly result from differences in diet, as segregation of the microbiota composition into distinct geographic clusters becomes apparent only after weaning, i.e. the start of consumption of solid food [62]. The food in developing countries is low in fat and protein, and rich in fibers compared to the typical western diet. In addition, the advanced sanitary conditions in western countries compared to a less advanced infrastructure and crowded settings in developmental countries could also influence differences in indigenous microbial community structures. Furthermore, intra-continental differences between western countries have also been reported with lower proportions of *Bifidobacterium* and higher proportions of *Akkermansia*, *Clostridium* and *Bacteroides* in German infants than Finish infants [65]. The latter is in line with another European study that showed overrepresentation of *Bifidobacterium* in Northern European infants whereas increased diversity was detected in Southern European infants [29]. However, whether these differences are driven by environmental exposure, genetic make-up, or dietary factors is currently unclear.

### ***Host genetics***

Host genotype is another factor with a major impact on the microbial composition, as shown



by studies comparing the microbiota of individuals with varying degrees of relatedness, in particular monozygotic as compared to dizygotic twins. Several studies using traditional microbial fingerprinting techniques in both children and adults reported host genetic effects on the microbiota composition [66-68]. A recent study analyzing the microbiota development by next-generation sequencing in three dichorionic triplet sets showed that at the age of one month the monozygotic twin pairs shared a more similar microbiota as compared to their fraternal siblings. By month 12 however, the profile became more uniform between the three infants, suggesting that host genetics initially play a significant role, while environmental determinants dominate by the age of 12 months [69].

By far the most comprehensive study to date examined the influence of host genetics on the microbiota using fecal samples of 416 pairs of adult twins from the TwinsUK cohort [70]. The microbiota were more similar between twin pairs as compared to unrelated individuals and were also more similar between monozygotic twins than dizygotic twins. By estimating the heritability of individual microbial taxa, it was shown that the community structure within the phylum Bacteroidetes was mainly shaped by environmental factors. Heritability appeared strongest for families within the phylum Firmicutes, including *Ruminococcaceae* and *Lachnospiraceae*, with *Christensenellaceae* being the most heritable family. *Christensenellaceae* formed a co-occurrence network with other microbes such as methanogenic Archaea and was associated with lean BMI. The functional relevance was illustrated by subsequent transfer of *Christensenellaceae* species to germ-free mice resulting in reduced body weight. Interestingly, the authors were able to validate their findings of heritability of the above-mentioned bacterial families by applying their statistical models to the data of two studies that had previously failed to reveal significant genotype effects on the microbiota [41, 71].

Furthermore, studies using a candidate gene approach examined the involvement of individual genes. For instance, individuals that have a functional fucosyltransferase 2 (*FUT2*) gene, referred as secretors, displayed different microbial communities compared to non-secretors [72]. The secretor-status regulates the existence of mucosal ABH and Lewis blood group antigens, but also the glycosylation of the intestinal mucus and breastmilk oligosaccharides. In addition, fucosyltransferase 3 (*FUT3*) produces the Lewis A antigens in non-secretors and B antigens in secretors. The profiles of the human milk oligosaccharides has been shown to differ in accordance to maternal secretor status as well as Lewis blood group antigen due to genetic variations of *FUT2* and *FUT3* respectively [73]. In accordance, a recent study observed earlier colonization and higher levels of bifidobacteria in breastfed infants

of maternal secretors [74]. Many other candidate genes have been suggested to influence microbiota composition. Among them in particular genes encoding immune mediators such as *NOD2* and *MEFV* – reviewed in detail elsewhere [75]. The nucleotide-binding oligomerization domain 2 (*NOD2*) receptor recognizes bacterial cell wall-derived muramyl dipeptide fragments and regulates the release of  $\alpha$ -defensins by Paneth cells. Consistent with an influence of enteric antimicrobial peptides on microbiota composition [76], expression of *NOD2* has been shown to affect the microbial ecology in mice [77]. In humans, *NOD2* mutations are a risk factor for the development of inflammatory bowel disease and are associated with microbial dysbiosis [78]. Mutations in the human gene *MEFV* that encodes pyrin results in familial Mediterranean fever, an autoimmune disease with uncontrolled interleukin (IL)-1 release. Mutations in *MEFV* are associated with changes in the gut microbiota including lower bacterial diversity [79]. These findings highlight that the genotype of the host influences the intestinal microbiota composition, which in turn, may influence host metabolism.

### Antibiotics

Although the majority of common pediatric infections, e.g. upper respiratory tract infections, are self-limiting and caused by viruses, large volumes of broad-spectrum antibiotics are still being prescribed for these conditions in primary care settings [80] with little evidence for clinical benefit [81]. Next to being the main factor driving the development and dissemination of antimicrobial resistance, (over)use of antibiotics has profound and direct effects on the indigenous microbiota. The exact effects on the microbiota composition are dependent on the specific type of antibiotic, but also differ between individuals as was demonstrated in a study on the effects of ciprofloxacin treatment in healthy volunteers [82]. This latter study showed that the microbiota is relatively resilient to perturbations induced by a single course of antibiotics. A subsequent study, however, indicated that the microbiota composition may permanently shift to an alternative state upon multiple courses of ciprofloxacin [83]. In newborns exposed to antibiotics, obligate anaerobic genera such as *Bacteroides spp.* and *Bifidobacterium spp.* and *Lactobacillus spp.* are depleted, whereas levels of Proteobacteria increase [24, 84, 85]. Not surprisingly, also maternal antibiotic use during pregnancy or breastfeeding has been shown to affect the infant's microbiota composition with decreased levels of *Bacteroides spp.* and *Atopbium spp.* [29]. Interestingly, the maternal vaginal microbiota is affected by antibiotic use during pregnancy, which might contribute to a distinct neonatal microbiome [86].

To date, few studies have examined the effect of intrapartum antibiotic prophylaxis (IAP), a practice routinely used in pregnant women positive for group B streptococci (GBS) on the vaginal mucosa, on the establishment of the neonatal microbiome. GBS represents the most important causative agent of neonatal sepsis with often fatal outcome. In the absence of an available vaccine, prophylactic antibiotic treatment represents the recommended measure to prevent infection. Two recent studies, quantifying a selected group of bacterial genera, revealed a decrease in bifidobacterial abundance as well as a shift in bifidobacterial species composition in newborns whose mothers did receive IAP as compared to newborns without maternal IAP exposure [87, 88]. In addition, newborns are (indirectly) exposed to antibiotics during labor as prophylaxes prior to C-section. The effects on the neonatal microbiome have, however, not been characterized.

Although there is little direct evidence from human studies, it is highly likely that the developing infant microbiota is far less resilient to antibiotic-induced perturbations than the microbiota of adults. This is supported by a recent study mimicking pediatric antibiotic use in a mouse-model by providing early-life administration of antibiotics at therapeutic doses (beta-lactam and macrolides). Progressive and long-term alterations in the gut microbiota composition and metagenomic content were observed, varying with the number of courses and class of antibiotic treatment [89]. Moreover, murine antibiotic exposure in early life showed an altered host metabolism regarding energy harvesting and fat storage [89, 90]. Therefore, antibiotic use during infancy has been suggested to increase subsequent weight gain and enhance the risk of childhood overweight [91-93]. The latter studies indicate that perturbation of the microbiota in early life results in different metabolic effects even in the presence of a similar diet. The lack of acquisition or the extinction of specific species during the neonatal period might therefore represent a critical factor, emphasizing the key role of pioneer bacteria.

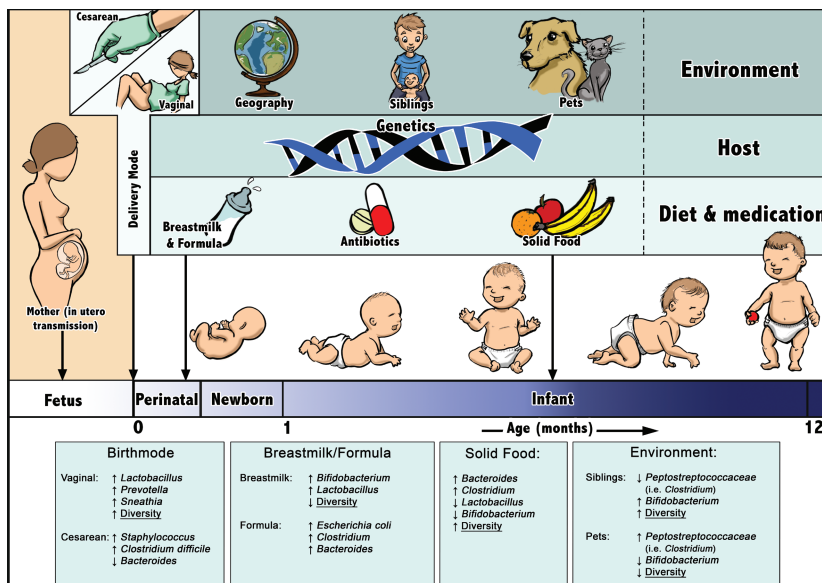
Data from human infants are commonly based on stool samples which deviate considerably from colonic mucosal microbiota [94]. Recent high-throughput microbial profiling also showed that the mucosal microbiota even varies along the complete length of the gut [95]. Therefore, data from animals of small intestinal and colonic microbiota provides additional insight on the development as well as specific host-microbe interactions. Although mice and rats have an increased exposure to environmental and fecal microbiota, the essence of the initial succession of gut bacteria in rodents reflects that in humans. In mice, the neonatal gut is mainly dominated by Lactobacillaceae, Streptococcus and Enterobacteriaceae [96, 97]. The colon in mice is first established by bacteria that resemble the vaginal microbiota, followed by a decrease in microbial diversity and shift in dominance from Streptococcus to Lactobacillus [97]. The microbial abundance seems to be related to the anatomical complexity of the intestine, resulting in very low microbial diversity in a simple and small gut after birth [96]. After the introduction of solid food, the murine microbiota increases in strictly anaerobes such as Bacteroides, Lachnospiraceae and Clostridium. Altogether, animal studies further indicate that the microbiota development is multiple-phasic and are pivotal in unraveling host-microbiota relationships.

## Window of opportunity for microbiota manipulation

The therapeutic manipulation of the gut microbiota is promising since the microbiota plays a crucial role in human health. For instance, fecal transplantation is a procedure that has recently raised a lot of scientific awareness and many intervention studies are currently ongoing to treat a variety of diseases. However, this method has so far only been clinically accepted as an effective method to eradicate *C. difficile* infection [98]. Moreover, numerous randomized placebo-controlled trials have been conducted on the efficacy of pre- and probiotic supplementation in pregnant women and newborns in the primary prevention of diseases such as necrotizing enterocolitis (NEC), atopic diseases, and (extra)intestinal infections. For most clinical outcomes the efficacy of pre- and probiotic supplementation is inconclusive, although there is evidence that probiotics might reduce the risk of atopic eczema in infants [99] and NEC in preterm neonates [100]. Yet, there is great heterogeneity between studies in terms of the probiotic strains used, dosing, duration and timing of probiotic therapy. The understanding and selection of the appropriate probiotics is crucial and should be determined by fundamental host-microbiota interactions. Importantly, a better knowledge

on the bacteria that are essential for specific health effects might also stimulate the use of selected diets or drugs as a way to modify the microbiota. Further research is necessary to identify not only beneficial bacteria, but also other microbes such as Archaea and viruses and their interactions. This may include approaches that use phage therapy as a tool against specific bacteria or pathogens [101].

In addition, a better understanding of ecological processes and pioneer consortia in microbial development will be crucial to design efficacious therapies and standardized manipulations. The microbial maturation relies on the succession of initial bacteria during infancy, making early life a particular suitable period to modify the microbiota in favor of health-promoting bacteria. The identification of specific neonatal populations that will benefit most from microbiota manipulation could also be relevant in terms of personalized strategies e.g. following C-section delivery, intra-partum antibiotic prophylaxis or neonatal infections. In this respect, trials are under way in which newborns delivered via C-section are being exposed to the maternal vaginal microbiota directly after birth. However, further studies are needed to unravel the critical shaping factors of intestinal bacteria in early life in order to support future targeted microbial therapies to reduce the risk of disease.



**Figure 1 - Main maternal, environmental and host determinants that impact the establishment of the intestinal microbiota during the first year of life.** A non-exclusive listing of bacterial taxa affected by these determinants is depicted at the bottom of the figure. Note: the timeline is colored ranging from white to blue corresponding to a change in microbial diversity from low to high, respectively. Perinatal starts at rupture of the membranes.

## Conclusion and future perspectives

Our knowledge on the maturation dynamics of the gut microbiota during infancy has advanced considerably during the past decades. The microbiota develops from the initial inoculum to a complex and diverse ecosystem during the first years of life and is influenced by a large number of perinatal environmental and maternal factors. The colonization of essential bacteria in the neonatal gut is important, since their lack might result in increased risk of metabolic or immune-associated disorders such as obesity or allergy, respectively. The establishment of the microbiota and its function during early life is strongly affected by the mode of delivery, breastmilk *versus* formula, and diet. Studies also emphasize the destructive role of antibiotic exposure in the succession of initial bacteria. In addition, the child's living location, including the indoor home environment, the presence of pets and older siblings as sources for microbial exposure, have been suggested to impact the infant microbiota. Further studies highlighted the critical role of the host's genotype in the intestinal microbiota composition, although these results were mostly derived from animal studies. Importantly, besides the determinants that select or shape the microbiota, the order of acquiring specific bacteria – the historical contingency - should also be taken into account. The historical contingency might be crucial because it defines how bacteria interact with each other and with the host, and may in turn influence the entire community [102]. However, the order of microbial arrivals is mainly stochastic and unknown, resulting in highly unpredictable microbial interactions. Therefore, experimental studies providing insight in the extent to which historical contingency contributes to microbial assembly are pivotal. Large-scale birth cohort studies generally assess the growth or stability of the microbiome by sampling at intervals ranging from months or years, but whether these dynamics sufficiently reflect the variability over a shorter period is unclear. As such, more regular sampling moments (weekly or even daily) within the context of longitudinal birth cohort studies is warranted. Moreover, further research is needed to unravel the functional and metabolic development of the neonatal microbiome. The integration of metatranscriptomics and metabolomics in future longitudinal human studies will greatly improve our understanding of the role of the infant's microbiome. It will also be crucial to determine the impact of early microbial changes on the incidence and severity of certain diseases in infancy as well as in later life. Consequently, new longitudinal studies should define in more detail the host-microbe interactions in regard to infant's health and provide sufficient follow-up to monitor the onset of non-communicable diseases. For instance, C-section may

have a long-term impact on child's health, regarding immune-mediated disorders [37, 103] as well as increased BMI in offspring and obesity in adulthood [104]. The latter emphasizes the crucial role of the first inoculum and suggests that the vaginal microbiota has the ability to induce metabolic and immune maturation in the newborn over a prolonged time period. Insight into the multifaceted interactions between the microbiota and ecological as well as host factors are still in its infancy and need to be expanded to facilitate therapeutic or prophylactic microbiota manipulations. Altogether, the overall challenge will be to promote certain bacterial populations in early life, prevent their eradication by antibiotics, and in this way prevent the development of non-communicable diseases.

## **Acknowledgements**

We are thankful to Mayk Lucchesi for figure designing and formatting.

## References

1. Hooper, L.V. and J.I. Gordon, *Commensal host-bacterial relationships in the gut*. Science, 2001. 292(5519): p. 1115-8.
2. Macpherson, A.J. and N.L. Harris, *Interactions between commensal intestinal bacteria and the immune system*. Nat Rev Immunol, 2004. 4(6): p. 478-85.
3. Turnbaugh, P.J., et al., *The human microbiome project*. Nature, 2007. 449(7164): p. 804-10.
4. Turnbaugh, P.J., et al., *An obesity-associated gut microbiome with increased capacity for energy harvest*. Nature, 2006. 444(7122): p. 1027-31.
5. van Best, N., P.L. Jansen, and S.S. Rensen, *The gut microbiota of nonalcoholic fatty liver disease: current methods and their interpretation*. Hepatol Int, 2015. 9(3): p. 406-15.
6. Penders, J., et al., *Gut microbiota composition and development of atopic manifestations in infancy: the KOALA Birth Cohort Study*. Gut, 2007. 56(5): p. 661-7.
7. Erickson, A.R., et al., *Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease*. PLoS One, 2012. 7(11): p. e49138.
8. Rajilic-Stojanovic, M., et al., *Intestinal microbiota and diet in IBS: causes, consequences, or epiphenomena?* Am J Gastroenterol, 2015. 110(2): p. 278-87.
9. Petra, L., L.H. Georgina, and J.F. Harry, *The gut microbiota, bacterial metabolites and colorectal cancer*. Nature Reviews Microbiology, 2014. 12(10): p. 661-672.
10. Mayer, E.A., K. Tillisch, and A. Gupta, *Gut/brain axis and the microbiota*. J Clin Invest, 2015. 125(3): p. 926-38.
11. West, C.E., et al., *The gut microbiota and inflammatory noncommunicable diseases: associations and potentials for gut microbiota therapies*. The Journal of allergy and clinical immunology, 2015. 135(1): p. 3.
12. Palmer, C., et al., *Development of the human infant intestinal microbiota*. PLoS Biol, 2007. 5(7): p. e177.
13. Favier, C.F., et al., *Molecular Monitoring of Succession of Bacterial Communities in Human Neonates*. Applied and Environmental Microbiology, 2002. 68(1): p. 219-226.
14. Jeremy, E.K., et al., *Succession of microbial consortia in the developing infant gut microbiome*. Proceedings of the National Academy of Sciences, 2011. 108 Suppl 1: p. 4578-4585.
15. Aagaard, K., et al., *The placenta harbors a unique microbiome*. Sci Transl Med, 2014. 6(237): p. 237ra65.
16. Jimenez, E., et al., *Is meconium from healthy newborns actually sterile?* Res Microbiol, 2008. 159(3): p. 187-93.
17. Funkhouser, L.J. and S.R. Bordenstein, *Mom knows best: the universality of maternal microbial transmission*. PLoS Biol, 2013. 11(8): p. e1001631.
18. DiGiulio, D.B., *Diversity of microbes in amniotic fluid*. Seminars in fetal & neonatal medicine, 2012. 17(1): p. 2-11.
19. Goldenberg, R.L. and J.F. Culhane, *Infection as a cause of preterm birth*. Clinics in perinatology, 2003. 30(4): p. 677-700.
20. Okogbule-Wonodi, A.C., et al., *Necrotizing enterocolitis is associated with ureaplasma colonization in preterm infants*. Pediatric research, 2011. 69(5 Pt 1): p. 442-447.
21. Perez, P.F., et al., *Bacterial imprinting of the neonatal immune system: lessons from maternal cells?* Pediatrics, 2007. 119(3): p. e724-32.
22. Ardissonne, A.N., et al., *Meconium microbiome analysis identifies bacteria correlated with premature birth*.



- PLoS One, 2014. 9(3): p. e90784.
23. Dominguez-Bello, M.G., 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): p. 11971-5.
  24. Penders, J., et al., *Factors influencing the composition of the intestinal microbiota in early infancy*. Pediatrics, 2006. 118(2): p. 511-21.
  25. Romero, R., et al., *The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women*. Microbiome, 2014. 2(1): p. 4.
  26. MacIntyre, D.A., et al., *The vaginal microbiome during pregnancy and the postpartum period in a European population*. Sci Rep, 2015. 5: p. 8988.
  27. Azad, M.B., et al., *Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months*. CMAJ, 2013. 185(5): p. 385-94.
  28. Jakobsson, H.E., et al., *Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section*. Gut, 2014. 63(4): p. 559-66.
  29. Fallani, M., et al., *Intestinal microbiota of 6-week-old infants across Europe: geographic influence beyond delivery mode, breast-feeding, and antibiotics*. J Pediatr Gastroenterol Nutr, 2010. 51(1): p. 77-84.
  30. Backhed, F., et al., *Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life*. Cell Host Microbe, 2015. 17(5): p. 690-703.
  31. Salminen, S., et al., *Influence of mode of delivery on gut microbiota composition in seven year old children*. Gut, 2004. 53(9): p. 1388-9.
  32. Ye, J., et al., *Searching for the optimal rate of medically necessary cesarean delivery*. Birth, 2014. 41(3): p. 237-44.
  33. Robson, M., L. Hartigan, and M. Murphy, *Methods of achieving and maintaining an appropriate caesarean section rate*. Best practice & research. Clinical obstetrics & gynaecology, 2013. 27(2): p. 297-308.
  34. Gibbons, L., et al., *The global numbers and costs of additionally needed and unnecessary caesarean sections performed per year: overuse as a barrier to universal coverage*. World health report, 2010. 30: p. 1-31.
  35. Kuhle, S., O.S. Tong, and C.G. Woolcott, *Association between caesarean section and childhood obesity: a systematic review and meta-analysis*. Obes Rev, 2015. 16(4): p. 295-303.
  36. Huang, L., et al., *Is elective cesarean section associated with a higher risk of asthma? A meta-analysis*. J Asthma, 2014. 52(1): p. 16-25.
  37. Penders, J., et al., *New insights into the hygiene hypothesis in allergic diseases: mediation of sibling and birth mode effects by the gut microbiota*. Gut Microbes, 2014. 5(2): p. 239-44.
  38. Kalliomäki, M., et al., *Early differences in fecal microbiota composition in children may predict overweight*. The American journal of clinical nutrition, 2008. 87(3): p. 534-538.
  39. Avershina, E., et al., *Major faecal microbiota shifts in composition and diversity with age in a geographically restricted cohort of mothers and their children*. FEMS Microbiol Ecol, 2014. 87(1): p. 280-90.
  40. Koenig, J.E., et al., *Succession of microbial consortia in the developing infant gut microbiome*. Proc Natl Acad Sci U S A, 2011. 108 Suppl 1: p. 4578-85.
  41. Yatsunenko, T., et al., *Human gut microbiome viewed across age and geography*. Nature, 2012. 486(7402): p. 222-7.
  42. Stark, P.L. and A. Lee, *The microbial ecology of the large bowel of breast-fed and formula-fed infants during the first year of life*. J Med Microbiol, 1982. 15(2): p. 189-203.
  43. Bezirtzoglou, E., A. Tsiotsias, and G.W. Welling, *Microbiota profile in feces of breast- and formula-fed newborns*

- by using fluorescence in situ hybridization (FISH). *Anaerobe*, 2011. 17(6): p. 478-82.
44. Zivkovic, A.M., et al., *Human milk glyco biome and its impact on the infant gastrointestinal microbiota*. *Proc Natl Acad Sci U S A*, 2011. 108 Suppl 1(Supplement\_1): p. 4653-8.
  45. Cabrera-Rubio, R., et al., *The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery*. *Am J Clin Nutr*, 2012. 96(3): p. 544-51.
  46. Bode, L., *Human milk oligosaccharides: every baby needs a sugar mama*. *Glycobiology*, 2012. 22(9): p. 1147-1162.
  47. Pacheco, A.R., et al., *The impact of the milk glyco biome on the neonate gut microbiota*. *Annu Rev Anim Biosci*, 2015. 3: p. 419-45.
  48. Rogier, E.W., et al., *Secretory antibodies in breast milk promote long-term intestinal homeostasis by regulating the gut microbiota and host gene expression*. *Proc Natl Acad Sci U S A*, 2014. 111(8): p. 3074-9.
  49. Ly, N.P., et al., *Gut microbiota, probiotics, and vitamin D: interrelated exposures influencing allergy, asthma, and obesity?* *J Allergy Clin Immunol*, 2011. 127(5): p. 1087-94.
  50. Griffin, M.D., N. Xing, and R. Kumar, *Vitamin D and its analogs as regulators of immune activation and antigen presentation*. *Annu Rev Nutr*, 2003. 23: p. 117-45.
  51. Volker, M., et al., *Associations between dietary habits and body mass index with gut microbiota composition and fecal water genotoxicity: an observational study in African American and Caucasian American volunteers*. *Nutrition Journal*, 2009. 8(1): p. 49.
  52. Heikkila, M.P. and P.E. Saris, *Inhibition of Staphylococcus aureus by the commensal bacteria of human milk*. *J Appl Microbiol*, 2003. 95(3): p. 471-8.
  53. Fernandez, L., et al., *The human milk microbiota: origin and potential roles in health and disease*. *Pharmacol Res*, 2013. 69(1): p. 1-10.
  54. Jost, T., et al., *Vertical mother-neonate transfer of maternal gut bacteria via breastfeeding*. *Environ Microbiol*, 2014. 16(9): p. 2891-904.
  55. Bergstrom, A., et al., *Establishment of intestinal microbiota during early life: a longitudinal, explorative study of a large cohort of Danish infants*. *Appl Environ Microbiol*, 2014. 80(9): p. 2889-900.
  56. Lax, S., et al., *Longitudinal analysis of microbial interaction between humans and the indoor environment*. *Science*, 2014. 345(6200): p. 1048-1052.
  57. Penders, J., et al., *Establishment of the intestinal microbiota and its role for atopic dermatitis in early childhood*. *J Allergy Clin Immunol*, 2013. 132(3): p. 601-607 e8.
  58. Adlerberth, I., et al., *Gut microbiota and development of atopic eczema in 3 European birth cohorts*. *J Allergy Clin Immunol*, 2007. 120(2): p. 343-50.
  59. Laursen, M.F., et al., *Having older siblings is associated with gut microbiota development during early childhood*. *BMC Microbiol*, 2015. 15(1): p. 154.
  60. Konya, T., et al., *Associations between bacterial communities of house dust and infant gut*. *Environ Res*, 2014. 131: p. 25-30.
  61. Azad, M.B., et al., *Infant gut microbiota and the hygiene hypothesis of allergic disease: impact of household pets and siblings on microbiota composition and diversity*. *Allergy Asthma Clin Immunol*, 2013. 9(1): p. 15.
  62. De Filippo, C., et al., *Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa*. *Proc Natl Acad Sci U S A*, 2010. 107(33): p. 14691-6.
  63. Grzeskowiak, L., et al., *Distinct gut microbiota in southeastern African and northern European infants*. *J Pediatr Gastroenterol Nutr*, 2012. 54(6): p. 812-6.

64. Clemente, J.C., et al., *The microbiome of uncontacted Amerindians*. *Sci Adv*, 2015. 1(3).
65. Grześkowiak, L., et al., *The impact of perinatal probiotic intervention on gut microbiota: double-blind placebo-controlled trials in Finland and Germany*. *Anaerobe*, 2012. 18(1): p. 7-13.
66. Van de Merwe, J.P., J.H. Stegeman, and M.P. Hazenberg, *The resident faecal flora is determined by genetic characteristics of the host. Implications for Crohn's disease?* *Antonie van Leeuwenhoek*, 1983. 49(2): p. 119-124.
67. Stewart, J.A., V.S. Chadwick, and A. Murray, *Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children*. *J Med Microbiol*, 2005. 54(Pt 12): p. 1239-42.
68. Zoetendal, E.G., et al., *The Host Genotype Affects the Bacterial Community in the Human Gastrointestinal Tract*. *Microbial Ecology in Health & Disease*, 2001. 13(3).
69. Murphy, K., et al., *The gut microbiota composition in dichorionic triplet sets suggests a role for host genetic factors*. *PLoS One*, 2015. 10(4): p. e0122561.
70. Goodrich, J.K., et al., *Human genetics shape the gut microbiome*. *Cell*, 2014. 159(4): p. 789-99.
71. Turnbaugh, P.J., et al., *A core gut microbiome in obese and lean twins*. *Nature*, 2009. 457(7228): p. 480-4.
72. Rausch, P., et al., *Colonic mucosa-associated microbiota is influenced by an interaction of Crohn disease and FUT2 (Secretor) genotype*. *Proc Natl Acad Sci U S A*, 2011. 108(47): p. 19030-5.
73. Totten, S.M., et al., *Comprehensive profiles of human milk oligosaccharides yield highly sensitive and specific markers for determining secretor status in lactating mothers*. *J Proteome Res*, 2012. 11(12): p. 6124-33.
74. Lewis, Z.T., et al., *Maternal fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants*. *Microbiome*, 2015. 3(1): p. 13.
75. Marietta, E., A. Rishi, and V. Taneja, *Immunogenetic control of the intestinal microbiota*. *Immunology*, 2015. 145(3): p. 313-22.
76. Salzman, N.H., et al., *Enteric defensins are essential regulators of intestinal microbial ecology*. *Nat Immunol*, 2010. 11(1): p. 76-83.
77. Petnicki-Ocwieja, T., et al., *Nod2 is required for the regulation of commensal microbiota in the intestine*. *Proc Natl Acad Sci U S A*, 2009. 106(37): p. 15813-8.
78. Frank, D.N., et al., *Disease phenotype and genotype are associated with shifts in intestinal-associated microbiota in inflammatory bowel diseases*. *Inflamm Bowel Dis*, 2011. 17(1): p. 179-84.
79. Khachatryan, Z.A., et al., *Predominant role of host genetics in controlling the composition of gut microbiota*. *PLoS One*, 2008. 3(8): p. e3064.
80. Hersh, A.L., et al., *Antibiotic prescribing in ambulatory pediatrics in the United States*. *Pediatrics*, 2011. 128(6): p. 1053-61.
81. Smith, S.M., J. Smucny, and T. Fahey, *Antibiotics for acute bronchitis*. *JAMA*, 2014. 312(24): p. 2678-9.
82. Dethlefsen, L., et al., *The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing*. *PLoS Biol*, 2008. 6(11): p. e280.
83. Dethlefsen, L. and D.A. Relman, *Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation*. *Proc Natl Acad Sci U S A*, 2011. 108 Suppl 1: p. 4554-61.
84. Savino, F., et al., *Faecal microbiota in breast-fed infants after antibiotic therapy*. *Acta Paediatr*, 2011. 100(1): p. 75-8.
85. Fouhy, F., et al., *High-throughput sequencing reveals the incomplete, short-term recovery of infant gut microbiota following parenteral antibiotic treatment with ampicillin and gentamicin*. *Antimicrob Agents Chemother*, 2012. 56(11): p. 5811-20.

86. Stokholm, J., et al., *Antibiotic use during pregnancy alters the commensal vaginal microbiota*. Clin Microbiol Infect, 2014. 20(7): p. 629-35.
87. Corvaglia, L., et al., *Influence of Intrapartum Antibiotic Prophylaxis for Group B Streptococcus on Gut Microbiota in the First Month of Life*. J Pediatr Gastroenterol Nutr, 2015. Publish Ahead of Print.
88. Aloisio, I., et al., *Influence of intrapartum antibiotic prophylaxis against group B Streptococcus on the early newborn gut composition and evaluation of the anti-Streptococcus activity of Bifidobacterium strains*. Appl Microbiol Biotechnol, 2014. 98(13): p. 6051-60.
89. Nobel, Y.R., et al., *Metabolic and metagenomic outcomes from early-life pulsed antibiotic treatment*. Nat Commun, 2015. 6: p. 7486.
90. Cox, L.M., et al., *Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences*. Cell, 2014. 158(4): p. 705-21.
91. Ajslev, T.A., et al., *Childhood overweight after establishment of the gut microbiota: the role of delivery mode, pre-pregnancy weight and early administration of antibiotics*. Int J Obes (Lond), 2011. 35(4): p. 522-9.
92. Trasande, L., et al., *Infant antibiotic exposures and early-life body mass*. Int J Obes (Lond), 2013. 37(1): p. 16-23.
93. Cox, L.M. and M.J. Blaser, *Antibiotics in early life and obesity*. Nat Rev Endocrinol, 2015. 11(3): p. 182-90.
94. Zoetendal, E.G., et al., *Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces*. Appl Environ Microbiol, 2002. 68(7): p. 3401-7.
95. Zhang, Z., et al., *Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota*. ISME J, 2014. 8(4): p. 881-93.
96. Del Chierico, F., et al., *A metaproteomic pipeline to identify newborn mouse gut phylotypes*. J Proteomics, 2014. 97: p. 17-26.
97. Pantoja-Feliciano, I.G., et al., *Biphasic assembly of the murine intestinal microbiota during early development*. ISME J, 2013. 7(6): p. 1112-5.
98. Drekonja, D., et al., *Fecal Microbiota Transplantation for Clostridium difficile Infection: A Systematic Review*. Ann Intern Med, 2015. 162(9): p. 630-8.
99. Cuello-Garcia, C.A., et al., *Probiotics for the prevention of allergy: A systematic review and meta-analysis of randomized controlled trials*. J Allergy Clin Immunol, 2015. In Press.
100. AlFaleh, K. and J. Anabrees, *Probiotics for prevention of necrotizing enterocolitis in preterm infants*. Evid Based Child Health, 2014. 9(3): p. 584-671.
101. Viertel, T.M., K. Ritter, and H.-P.P. Horz, *Viruses versus bacteria-novel approaches to phage therapy as a tool against multidrug-resistant pathogens*. The Journal of antimicrobial chemotherapy, 2014. 69(9): p. 2326-2336.
102. Dethlefsen, L., et al., *Assembly of the human intestinal microbiota*. Trends Ecol Evol, 2006. 21(9): p. 517-23.
103. Decker, E., et al., *Cesarean delivery is associated with celiac disease but not inflammatory bowel disease in children*. Pediatrics, 2010. 125(6): p. e1433-40.
104. Darmasseelane, K., et al., *Mode of delivery and offspring body mass index, overweight and obesity in adult life: a systematic review and meta-analysis*. PLoS One, 2014. 9(2): p. e87896.



Elif Swinkels (2 years)

# Chapter 3

## **Bile acids drive the newborn's gut microbiota maturation**

**Niels van Best**, Ulrike Rolle-Kampczyk, Frank G. Schaap, Marijana Basic,  
Steven Olde Damink, André Bleich, Paul H.M. Savelkoul, Martin von Bergen,  
John Penders‡, Mathias W. Horne‡

‡equal contribution

*Nat Communications. 2020 Jul 23;11(1):3692*

## Abstract

Following birth, the neonatal intestine is exposed to maternal and environmental bacteria that successively form a dense and highly dynamic intestinal microbiota. Whereas the effect of exogenous factors has been extensively investigated, endogenous, host-mediated mechanisms have remained largely unexplored. Concomitantly with microbial colonization, the liver undergoes functional transition from a hematopoietic organ to a central organ of metabolic regulation and immune surveillance. The aim of the present study was to analyze the influence of the developing hepatic function and liver metabolism on the early intestinal microbiota. Using metabolomic and microbial profiling in combination with multivariate analysis we characterized the colonization dynamics and liver metabolism in the murine gastrointestinal tract (n=6-10 per age group). We observed major age-dependent microbial and metabolic changes and identified bile acids as potent drivers of the early intestinal microbiota maturation. Consistently, oral administration of tauro-cholic acid or  $\beta$ -tauro-murocholic acid to newborn mice (n= 7-14 per group) accelerated postnatal microbiota maturation.

## Introduction

All mammalian host body surfaces and in particular the intestinal tract are colonized by a variety of microorganisms. The intestinal microbiota of the healthy host represents a dense and competitive microbial ecosystem and provides a core set of metabolic activities and immunostimulatory molecules, which together significantly contribute to the host's metabolism, tissue development and immune maturation and protect from infection with enteropathogenic microorganisms [1]. Perturbations of the human intestinal microbiota composition have been associated with highly prevalent metabolic and immune-mediated diseases. Therefore, a detailed understanding of the mechanisms that establish and maintain a beneficial microbiota composition in the healthy host is fundamental. In particular, this relates to the immediate postnatal and early infant period during which initial bacterial colonization accompanies dynamic fluctuations in the microbiota composition [2-5]. Notably, the early postnatal colonization process is of critical importance for the long-term microbiota composition [6], and immune maturation and thereby the susceptibility to highly prevalent non-communicable diseases [7]. Microbiota alterations during this time window have been associated with adverse consequences and an enhanced susceptibility to disease [8, 9].

The fetal mucosal surfaces are sterile [10]. With birth, exposure to the maternal vaginal, fecal and skin microbiota establishes an early set of bacteria. With time, additional bacteria are incorporated generating an increasingly diverse bacterial ecosystem with significant resilience to exogenous perturbations [11]. The process of intestinal bacterial colonization after birth has been intensively studied during the last years and exogenous factors which significantly impact the early microbiota composition have been identified [3, 5, 12-15]. However, exogenous factors can explain only a limited amount of the inter-individual variation in microbiota composition [5, 16].

Few studies have attempted to characterize host-mediated endogenous mechanisms such as for example genetic factors that shape the microbial ecosystem early after birth [17-19]. The identification and characterization of such endogenous mechanisms is expected to provide insight in the functional relevance of particular aspects of the microbiota since they result from a long process of host-microbial coevolution [20]. In the present study, we aimed to unravel the impact of the developing host hepatic metabolism on the murine intestinal microbiota composition. To characterize the intestinal microbiota maturation, we first profiled the microbial composition of the small and large intestine obtained from animals within the same litter during the immediate postnatal period, early infancy and weaning until



adulthood. Subsequently, we performed a global metabolic screen in the liver tissue of the same animals to provide information on the site-specific metabolites. Multi-omics analysis identified specific bile acids associated with maturation of the early intestinal microbiota and subsequent interventional studies functionally confirmed and specified the role of specific bile acids as important drivers of microbiota development.

## Material and Methods

### Ethic statement

All animal experiments were performed in compliance with the German animal protection law (TierSchG) and approved by the local animal welfare committees, the Landesamt für Natur, Umwelt und Verbraucherschutz, North Rhine Westfalia (84-02.04.2016.A207 and 84-02.04.2015.A293). All C57BL6J wildtype mice were bred locally and held under specific pathogen-free conditions at the Institute of Laboratory Animal Science at RWTH Aachen University Hospital. The day of birth was considered day 0, i.e. animals screened at day 1 were approximately 24 h old and verified to have ingested breast milk (abdominal milk spot). Mice were weaned at PND21.

### In vivo study design

To monitor microbiota and host metabolic development throughout the neonatal period into adulthood, intestinal and hepatic tissues were obtained from C57BL/6J mice in two different approaches. First, total small intestinal, colon and liver tissues were obtained from C57BL/6J mice aged 1, 7, 14, 21, 28, and 56 days (n=6 per timepoint). In a separate set of experiments, similar tissues were obtained from mice aged 0, 6, 12, 18 and 24 hours (n=10 per timepoint). To rule out potential litter and cage effects, we obtained tissues from a single animal of one litter for a given age group and repeated this by selecting a new animal from the same litter for every other age group (Fig. 1a). Every animal was only examined once at the indicated age (PND). For oral bile acid administration, PND7 animals received the indicated bile acid (Sigma-Aldrich, Biomol) at a concentration of 70 µg/g body weight or PBS daily for 3 days by oral gavage (average 5µl) with 100% succession (Lorenzo-Zuniga et al, 2003). Body weight and food/water consumption was monitored daily. To determine the effect of bile acid administration, the intestine was aseptically cut into 10-20 parts and alternately assigned to two collections for microbial profiling and assessment of metabolites. Samples from an additional group of adult 8-12 week old animals (n=5) were processed in parallel to provide an adult microbiota control. Tissues were transferred into sterile micro-centrifuge tubes and stored at -80°C before analysis. Liver tissues of 7-, 14-, 21- and 56-day-old germ-free mice were obtained from the Institute for Laboratory Animal Science at Hannover Medical School. Tissues were collected and stored in sterile tubes stored at -80°C until further analysis.

### **DNA isolation and generation of sequence data**

Total metagenomic DNA was isolated from snap frozen small intestinal and colonic tissues by repeated-bead-beating combined with chemical lysis plus a column-based purification method as described in detail elsewhere [21]. For each DNA isolation batch, additional isolation was performed on PCR-grade water as a negative control. Generation of amplicon libraries and sequencing was performed according to a previously published protocol [22]. Briefly, the V4 region of the 16S rRNA gene was PCR amplified from each DNA sample in duplicate. Pooled amplicons from the duplicate reactions were purified using AMPure XP purification (Agencourt, Massachusetts, USA) according to the manufacturer's instructions and subsequently quantified by Quant-iT PicoGreen dsDNA reagent kit (Invitrogen, New York, USA). Amplicons were mixed in equimolar concentrations, to ensure equal representation of each sample, and sequenced on an Illumina MiSeq instrument using the V3 reagent kit (2x 250 cycles). All V4 16S rDNA bacterial sequences generated in this study have been submitted to the Qiita and ENA databases under accession No. 10719 and ERP116798 respectively.

### **Sequence analysis**

Data demultiplexing, length and quality filtering, pairing of reads and clustering of reads into Operational Taxonomic Units (OTUs) at 97% sequence identity was performed using the online Integrated Microbial Next Generation Sequencing (IMNGS, [www.imngs.org](http://www.imngs.org)) platform using default settings [23]. Removal of primers and technical reads resulted in fragments of approximately 250 bases. Sequencing was performed from both the 3' and 5' side resulting in sufficient resolution. IMNGS is a UPARSE based analysis pipeline [24]. Pairing, quality filtering and OTU clustering (97% identity) was done by USEARCH 8.0 [25]. The analysis was based on OTUs rather than amplicon sequence variants (ASVs) since we aimed at aggregating taxa at a higher level and wanted to avoid overestimation of prokaryotic diversity due to intragenomic heterogeneity of 16S rRNA genes [26]. Chimera filtering was performed by UCHIME (with RDP set 15 as a reference database [27]). Taxonomic classification was done by RDP classifier version 2.11 training set 15.8. Sequence alignment was performed by MUSCLE and treeing by FastTree [28, 29]. A total of 21,372,397 V4 reads were generated over two runs. After trimming, quality filtering, removal of potential chimeric reads, de-multiplexing and removal of low abundant operational taxonomic units (OTUs), 15,692,587 sequences belonging to 478 OTUs were retained for downstream analysis. Negative controls were evaluated based on their number of sequences and composition compared to other samples. We used for each

batch, sampling blank controls, DNA blank extraction controls and no-template amplification controls, and monitored the lack of contaminant bacterial DNA load herein by a gel-based principle. Moreover, we compared the acquired OTUs composition from the negative controls to our low abundant microbial samples to ensure that our findings were not driven by potential contaminant taxa. Subsequently, samples of the negative controls and with low sequencing depth (less than 6,749 sequences/sample) were excluded from subsequent analysis. For the remaining samples, the number of sequences per sample ranged from 6,749 to 239,395 (median 87,227).

### **Richness, diversity, taxonomy and enterotype analyses**

Data normalization, diversity, taxonomical binning and group comparisons were performed using the Rhea package [30]. In order to not discard informative data, normalization in Rhea was performed by dividing OTU counts per sample for their total count (sample depth) followed by multiplying all of the obtained relative abundance for the lowest sample depth (6,749 reads/sample). Alpha- (observed species and Shannon index) and the generalized Unifrac beta-diversity index were calculated using the Rhea package [31]. Additional beta-diversity indices (weighted Unifrac, unweighted Unifrac, and Bray-Curtis distance) were calculated using the R package Phyloseq [32]. Ordination of samples according to their microbial composition expressed as Hellinger transformed genus abundance data or beta-diversity indices was visualized using Principal Component Analysis (PCA) and Principal Coordinate Analysis (PCoA), respectively. All ordinations were constructed using the R package Phyloseq and included 95% confidence ellipses. Dirichlet multinomial mixtures models (DMM) were used to calculate genus-level enterotypes [33]. When including samples from all time-points, Laplace approximation revealed an optimal number of three clusters.

### **Downstream microbial analyses and presentation**

Smoothing of the kinetic for dominant taxa (Fig. 1e and f as well as Fig. S1d and e) was generated using the `geom_smooth` function of the `ggplot` package with default settings. The lines reflect the mean values of the relative abundance. The appearance and disappearance of OTUs of the dominant phyla (Actinobacteria, Proteobacteria, Firmicutes, Bacteroidetes) with age was visualized in Sankey-plots (SankeyMATIC.com). For readability of Sankey-plots, only OTUs present in >10% of all samples per timepoint and with a prevalence of >20% in the entire dataset, were included. Ecosystem specific functional metagenome predictions

were created by the novel PICRUSt-iMGMC workflow with the *de novo* picked OTUs and using mouse metagenome-assembled genomes linked to 16S rRNA genes [34]. The derived KEGG orthologs were mapped into multiple pathways or modules. Differentially changed KEGG modules were identified using the pathways enrichment analyses from MicrobiomeAnalyst with default settings [35]. The identification of *lactobacillus*-related OTUs were performed using EZbiocloud and used to construct a phylogenetic tree of lactobacilli by MEGA7 (MUSCLE) for alignment and iTOL v4 for the final annotations (itol.embl.de) with default settings.

### **Liver metabolomics and bile acid analyses**

The metabolome analyses were carried out with the AbsoluteIDQ® p180 Kit (Biocrates Life Science AG, Austria). The kit allowed identification and quantification of 188 metabolites from 5 compound classes (acyl carnitines, amino acids, glycerophospho- and sphingolipids, biogenic amines and hexoses). The kit used flow injection tandem mass spectrometry (FIA) for the non-polar metabolites and LC-MS/MS for the more polar compounds. The integrated MetIDQ Software streamlined the data analysis by automated calculation of metabolite concentrations. Quantification of analytes utilized stable isotope-labelled or chemically homologous internal standards (IS). Controls were included for 3 different concentration levels. For calibration, the kit contained a calibrator mix at 7 different concentrations. The measurements were carried out with an ABI Sciex API5500 Q-TRAP mass spectrometer via Electrospray ionization (ESI) by Multi Reaction Monitoring (MRM) mode for high specificity and sensitivity. 158 MRM pairs were measured in positive ion mode (13 IS) and 2 MRM pairs were measured in negative mode (1 IS). The following additional chemicals for LC-MS were used: water, Millipore; PITC, Fluka; pyridine, Fluka (p.a.); methanol, Merck; Lichrosolv for LC/MS; acetonitrile, Merck; formic acid, Sigma Aldrich. Metabolites were extracted from liver samples by adding H<sub>2</sub>O/acetonitrile (1:1,v:v) per mg sample followed by homogenization with a tissue disruptor (10 min, 30Hz, 4 steel balls). The samples were centrifuged (1400g, 2 min) and the supernatant analysed. The targeted analysis was performed by adding 10 µL of extracted liver sample to the AbsoluteIDQ® p180 Kit (Biocrates Life Science AG, Innsbruck, Austria), following the vendor's instructions [36]. For bile acid measurements the MS-based Bile Acids Kit (Biocrates Life Sciences AG, Innsbruck, Austria), a 96-well plate format assay, was used as described in the manufacturer's instructions with normalization based on mass [37].

The contribution of each metabolite to metabolomic variation was derived from age-constrained redundancy analysis (RDA) based on all metabolites stratified in group levels with

additional scaling by normalization based on z-scores (Phyloseq package). Moreover, PCA was used to illustrate changes in the composition of the different metabolic groups during the postnatal period for both PC1 and PC2, and PC2 and PC3. For the bile acids the 2nd and 3rd component were chosen for plotting, since the first component was mainly driven by the strong separation observed between the first and subsequent time points (PND1 vs. PND7-56).

### Multi-omics analyses

Regularized canonical correlation analyses (rCCA) were performed (Mixomics package) [38] to unravel specific correlations between bile acids and OTUs with a minimal presence of 20% in all samples. Samples were excluded from the microbiota-data if they were not measured in the bile acid analyses (i.e. PND1 of litter 1 & 2). Prior to rCCA a hyperbolic sine transformation was used on OTU-counts and a log-transformation for the bile acids. For the estimation of regularization (penalization) parameters  $\lambda_1$  and  $\lambda_2$ , the cross-validation procedure [39] method was used. We used a  $\lambda_1 = 0.0001$ ,  $\lambda_2 = 1$  with a CV-score = 0.4779644 and 2 components. OTUs with a correlation between -0.3 and 0.3 on the first 2 components were filtered out to optimize the rCCA.

The Spearman's rank correlation coefficient was calculated between bile acids (weight corrected) and bacterial genera (relative abundances) with a minimal presence of 20% in all samples. Benjamini & Hochberg FDR correction was performed to correct for multiple testing ( $p > 0.05$ ). For the heatmap only significant correlations with adjusted p-value of  $< 0.05$  are shown and sorted according to their PCA loading scores.

### Hepatic gene expression analyses

Total RNA was isolated from liver tissue using an RNeasy Plus Universal Kit (Qiagen). RNA concentration and purity were assessed by UV absorption measurement, and integrity was checked using agarose gel electrophoresis. 750 ng of total RNA was converted into single-stranded cDNA using a SensiFAST™ cDNA Synthesis Kit (Bioline). cDNA equivalent to ca. 16 ng of total RNA was used as input for real-time PCR analysis, employing Sybr Green chemistry (SensiMix™ SYBR® Hi-ROX Kit, Bioline) and a LightCycler® 480 System [10]. All primer pairs (Sigma-Aldrich) were designed using PrimerBlast, and correct amplicon size was verified by agarose gel electrophoresis. LinRegPCR was used for analysis of real-time qPCR data. Gene expression data was normalized to the geometric mean of two reference genes, viz. *36b4* and *Hprt* [40].

### **Absolute abundances (16S rDNA gene qPCR)**

The 16S rDNA gene copy number was determined to quantify the amount of bacterial genomes present in the samples and convert the relative OTU abundances into absolute bacterial numbers. The PCR primers 5'-CCTACGGGNGGCWGCAG-3' (16S\_341\_F) and 5'-GACTACHVGGGTATCTAATCC-3' (16S\_805\_R) were used as described earlier [41]. In short, the real-time PCR was performed on a MyiQ™ System (BioRad, USA) in a reaction-volume of 25 µl with 12.5 µl iQ™ SYBR Green (Biorad), 2 µl template DNA and 0.75 µl primers (10 µM). The cycling conditions were 95°C for 3 min followed by 35 cycles of 95°C for 15 s; 55°C for 20 s and 72°C for 30 s. Total 16S gene copy numbers were calculated by comparing the CT value to a standard curve with known concentrations of a plasmid encoding the target 16S gene sequences. Absolute OTU numbers were calculated by multiplying the relative abundances with the 16S rDNA gene copy numbers per sample.

### **In vitro culture assays**

*Lactobacillus johnsonii* DSM 100219 similar to BA-OTU 4, *Lactobacillus reuteri* DSM 32035 similar to BA-OTUs 7 & 364, *Lactobacillus murinus* DSM 100193 similar to the BA-OTUs 2 & 798 and *Escherichia coli* DSM 28618 previously isolated from the intestine of mice [42] were employed for *in vitro* growth assays. *Lactobacillus* and *E. coli* strains were cultured overnight in MRS (Merck, Germany) and BHI (Sigma-Aldrich) medium, respectively. Cultures were diluted 1:100 and incubated in a 96-well plate (at 100 µl/well) in the absence or presence of 1% (w/v) UDCA, TCA, GCA or βTMCA (Sigma-Aldrich, Taufenkirchen and Biomol, Hamburg) and incubated at 37° C for 24h under anaerobic conditions. Bacterial growth was determined by optical density (OD) measurements at 570 nm (OD<sub>570</sub>) and expressed as OD<sub>570</sub> in medium with 1% bile acid/OD<sub>570</sub> in medium with no bile acid supplement.

### **Statistical analysis**

A permutational multivariate analysis of variance (PERMANOVA) using distance matrices (vegan::adonis) was performed in each case to determine if the separation of groups was significant. The Kruskal-Wallis or Mann-Whitney U test were used to test for differences with respect to the relative abundance of bacterial phyla, families and genera, as well as for differences in microbial richness and diversity. For these analyses, apart from the default settings we used no cutoff for relative abundance, no cutoff for median abundance, and zero values were included in statistics. In addition, significance of a trend was evaluated to test for

differences in alpha-diversity with age, by including it as a continuous variable in a regression model. All statistical tests were performed by two-sided approaches. Data analyses were performed with R and GraphPad Prism software. p values are indicated as follows: \*\*\*\*,  $p < 0.0001$ ; \*\*\*,  $p < 0.001$ ; \*\*,  $p < 0.01$ ; and \*,  $p < 0.05$ .

## Results

### Postnatal kinetic of the intestinal microbiota composition

To investigate the postnatal gut microbial development, we obtained the complete small intestinal and colonic tissue from mice aged 0, 6, 12, 18 and 24 h ( $n=10$ /age group) as well as 1, 7, 14, 28 and 56 days ( $n=6$ /age group) after birth to cover the early developmental stages until adulthood. To avoid litter effects, tissues for all time points in a given kinetic were taken consecutively from animals from one litter (Fig. 1a and Fig. S1a – schematic overview of study design). Amplicon sequencing of the 16S rRNA V4 gene region of these samples generated in total 15,692,587 high-quality sequences with a median of 87,227 [IQR 77,784] reads per sample. All sequences were subsequently clustered based on de novo assembly into 478 operational taxonomic units (OTUs) that were taxonomically classified to the genus level.

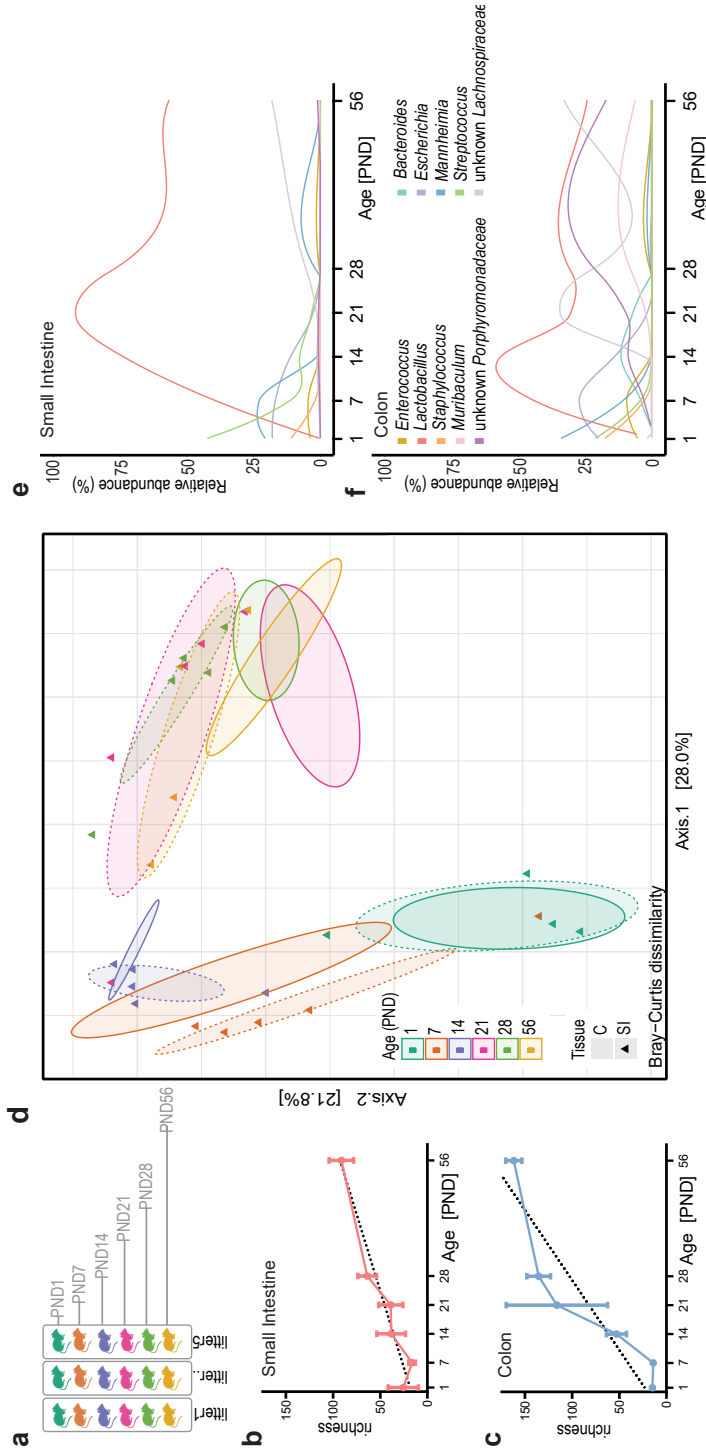
No substantial changes in the general community structure were observed during the first 24 hours (Figure S1b). The bacterial richness and diversity decreased significantly in both small intestine and colon during the first 24 h after birth (Fig. S1c and d, Fig. S2a). In line with the reduced diversity, the relative abundance of almost all genera decreased during this early time window in both the small intestine and colon. Only the abundance of the genera *Corynebacterium* and *Mannheimia* increased temporarily with a peak at 6-18h and 24h or PND1, respectively. (Figure S1e and f). The initial decrease in richness most likely indicates the transient passage of bacterial species ingested during birth and shortly thereafter which were unable to permanently colonize the intestinal tract (e.g. *Corynebacterium*). The developmental phase thus starts as indicated by the diversification from 24 h onwards (Fig. 1b-f) and we therefore decided to focus all subsequent analyses on these later time points.

Microbial richness increased from PND7 onwards in both small intestine and colon (Fig. 1b and c). While the bacterial richness was similar in both organs at birth (PND1), a significantly higher colonic richness was detected after weaning (i.e. PND28,  $FDR_{\text{adjusted}} = 1.7 \times 10^{-5}$ , Mann-Whitney U test). Similar findings were observed for microbial diversity using the Shannon index (Fig. S2b). The presence, origin and loss of OTUs in small intestinal



and colon tissue were visualized in Sankey plots for all postnatal time points as well as for the maternal source (Fig. S3a and b). This analysis illustrated the contribution of OTUs from all four major phyla to the microbial ecosystem at the different time points after birth. Notably, it revealed that a significant proportion of Proteobacteria originating from an environmental (non-maternal) source (78.2% of colonic and 73.2% of small intestinal OTUs) contributed to the microbiota early after birth (PND1). In contrast, the majority of OTUs from the Actinobacteria, Bacteroidetes and Firmicutes phylum persisted throughout the examined time window and were shared with the mother (53.3-94.2% and 45.0-86.5% for the colon and small intestine, respectively). The early (PND1) Proteobacteria contribution was most evident in the colon, whereas the small intestine was most strongly influenced by Proteobacteria and Actinobacteria.

Clear separation of the small intestinal and colonic microbial composition according to age was detected by principal coordinate analysis (PCoA) (Fig. 1d). The general microbiota community structure of the small intestine and colon was indistinguishable at the early time points (PND 1-21) but diverged significantly after weaning leading to a site-specific intestinal microbiota. This result was supported by the Dirichlet multinomial mixtures (DMM) modelling in which small intestinal and colonic samples from PND 1 and 7 clustered together, whereas two distinct clusters dominated the small intestinal and colonic microbiota from PND 21 onwards (Fig. S2c). The taxonomic changes confirmed the late emergence of an organ-specific microbiota. Whereas the abundance of the most prevalent genera *Mannheimia*, *Streptococcus*, *Escherichia*, *Staphylococcus* and *Enterococcus* gradually decreased in small intestine and colon during the postnatal period, a relative enrichment in *Lactobacillus* abundance was observed in both organs (Fig. 1e and f). After weaning, the relative increase in lactobacilli continued in the small intestine whereas a more complex composition, with expansion of *Bacteroides* and members of the Porphyromonadaceae (a.o. *Muribaculum*) and Lachnospiraceae families, developed in the colon.



**Figure 1 - Postnatal dynamics of the intestinal microbiota composition.** (A) Image illustrating the study outline; one mouse from each one litter was analyzed for each individual time point (PND). (B and C) Richness (observed species) of the small intestinal (B) and colonic (C) microbiota exhibiting a gradual increase with age ( $n=5$  per PND for all subsequent analyses, mean and SD,  $p<0.0001$ , linear regression  $R^2=0.7702$  and  $p<0.0001$ , linear regression  $R^2=0.7281$ , respectively). (D) Principal Coordinate Analysis (PCoA) based on Bray-Curtis dissimilarity indicating a gradual shift in microbial community structure along PC2 during the neonatal period PND1-PND14 and distinct structures for the small intestine and colon along PC2 at PND21-56 ( $p<0.001$ , PerMANOVA for age and  $p<0.01$  PerMANOVA for tissue). Squares and solid line: colon (C); triangles and dashed line: small intestine (SI). (E and F) Relative abundances of the 10 most abundant genera listed in (F) between PND1-56 for the small intestine (E) and colon (F).

### **Metabolic factors with potential influence on the microbiota**

The enzymatic and absorptive capacity of the intestine and the metabolic capacity of the liver mature significantly during the postnatal period [43-45]. The associated changes contribute to luminal substrate availability and could thereby influence the bacterial succession and overall composition [3]. To allow analysis of the same mice used for the long-term microbiota analysis presented above we investigated a panel of hepatic metabolites as a proxy for the small intestinal metabolome [46]. Major age-dependent differences were noted in particular between PND7 and PND21, 28 and 56 (Fig. 2a-d). Amino acids, biogenic amines, acylcarnitines and some glycerophospholipids showed a moderate but significant decrease between these timepoints, most likely due to the switch in diet during weaning (Fig. 2b-d), which is normally completed after day PND21. In contrast, a number of bile acids increased significantly after weaning in particular at PND7 vs. PND28 as well as PND7 vs. PND56 (Fig. 2c and d).

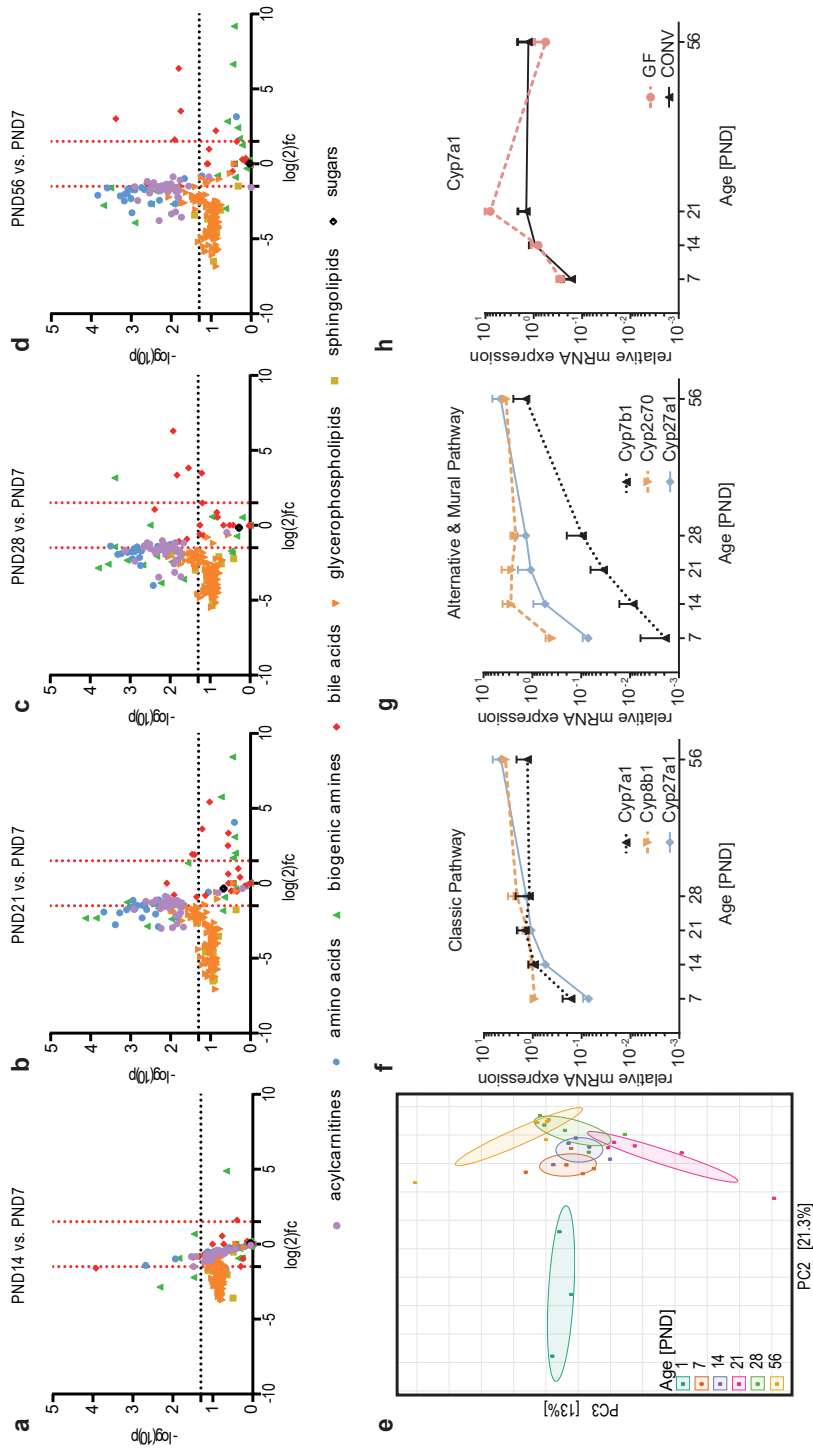
Constraining the ordination of all metabolites at the group level for age using Redundancy Analysis (RDA) indicated that variation in acylcarnitines, amino acids, glycerophospholipids, and sphingolipids was mainly the result of the marked differences between PND1 and all other age-groups (Fig. S4a-c, Fig. S5). In contrast, enterohepatic bile acids explained a major part of the age-dependent metabolic variation thereafter as depicted by RDA2 and RDA3 (Fig. S4d and e). These age-dependent metabolic shifts were further supported using unconstrained ordination of all metabolic groups separately, illustrating the compositional changes (Fig. S5a-i).

### **Postnatal maturation of bile acid synthesis, modification and reabsorption**

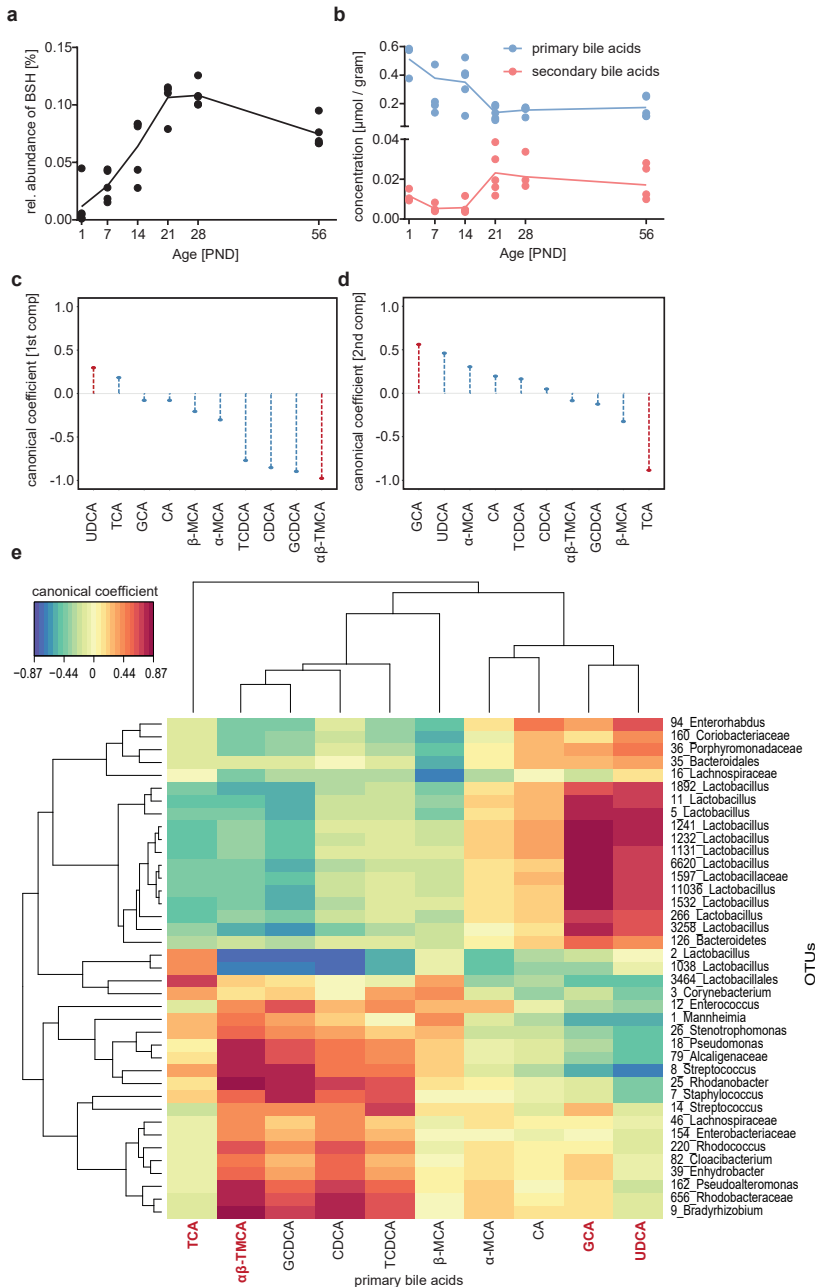
Consistently, the bile acid composition revealed a clear separation according to age along the second and third component with strong differences between birth and adulthood based on PCA (Fig. 2e). The global compositional shifts were supported by transcriptional changes of genes involved in bile acid synthesis in hepatic tissue. Transcripts for key enzymes of the classic pathway of bile acid synthesis involved in cholic acid and chenodeoxycholic acid formation such as *Cyp7a1* and *Cyp27a1* increased early (PND7-14) followed by increased *Cyp8b1* mRNA (Fig. 2f). Moreover, the key enzyme of the alternative pathway leading to chenodeoxycholic acid formation, *Cyp7b1*, was barely expressed until PND28 and unlikely to significantly contribute to synthesis of this bile acid within this time period. In contrast, the early elevation of *Cyp2c70* mRNA may in particular drive increased conversion of chenodeoxycholic acid to muricholic acids, along with cholic acid major bile acid species in mice (Fig. 2g). Notably, postnatal

upregulation of the key enzyme of the classical bile acid synthesis pathway *Cyp7a1* was also observed in germ-free animals. Germ-free animals are bred and raised in the absence of viable bacteria and thus this postnatal upregulation occurred in a developmental, microbiota-independent fashion (Fig. 2h).

Age-dependent changes in the concentration of bile acids were noted within the time window of major alterations in the microbiota composition. Since the metabolome analysis was performed on liver tissue to reflect the situation in the small intestine, these changes throughout the postnatal time period primarily reflected the abundance of primary bile acids. Nevertheless, prediction of the metabolic capacity of the developing microbiota by means of Phylogenetic Investigation of Communities by Reconstruction (PICRUSt) using the novel integrated Mouse Gut Metagenomic Catalog (iMGMC) revealed an age-dependent enrichment of bacteria encoding for genes involved in bile acid metabolism in adults (PND56) compared to neonates (Fig. S6). More specifically, the predicted abundance of genes for the key enzyme involved in bacterial bile acid metabolism, the bile salt hydrolase (BSH, KO1442), increased steadily with age reaching the highest small intestinal levels approximately at PND 21-28 (Fig. 3a). Notably, much higher abundances of BSH activity harboring bacterial taxa are observed in the large intestine. In addition, enhanced expression of genes encoding proteins involved in the hepatobiliary transport of bile acids such as *Ntcp (Slc10A1)*, *Abcb11*, and *Abcb4* as well as proteins involved in the Fxr-Fgf15-Fgfr4 negative feedback loop regulating *de novo* hepatic bile acid synthesis was observed after birth in total liver tissue as measured by RT-PCR (Fig. S7). Consistent with enhanced BSH activity and maturation of the hepatobiliary transport system, the hepatic concentration of bacterially modified secondary bile acids increased with age, whereas the concentration of primary bile acids decreased (Fig. 3b). Thus, the maturation of the host's bile acid biosynthetic capacity and hepatobiliary transport system in combination with an enhanced bacterial bile acid metabolism led to a marked change of the total bile acid pool as well its composition after weaning. The small intestine during the immediate postnatal period, however, was characterized by the presence of mainly conjugated, primary bile acids.



**Figure 2 - Metabolic changes during the postnatal period.** (A-D) Volcano plots depicting acylcarnitines, glycerophospholipids, sphingolipids, sugars, bile acids, amino acids, and biogenic amines in liver tissue at PND14 vs. PND7 (A), PND21 vs. PND7 (B), PND28 vs. PND7 (C), and PND56 vs. PND7 (D) measured by mass spectrometry (n=5 per group, two-tailed independent student's t-test, fold-changes were calculated based on mean values by dividing the PNDx to PNDy concentrations and log2 the values). PND7 was selected as baseline for comparisons due to the high interindividual variation at PND1. (E) Principal Component Analysis (PCA) illustrating changes in the composition of the bile acid pool during the postnatal period. (F-G) Relative hepatic expression of the key-enzymes of (F) the classical (Cyp7a1, Cyp8b1 and Cyp27a1) and (G) the alternative and murine pathway (Cyp7b1, Cyp2c70 and Cyp27a1) (mean and SD). (H) Relative expression of Cyp7a1 in total liver tissue of germ-free (GF) and conventionally (CONV) raised mice).



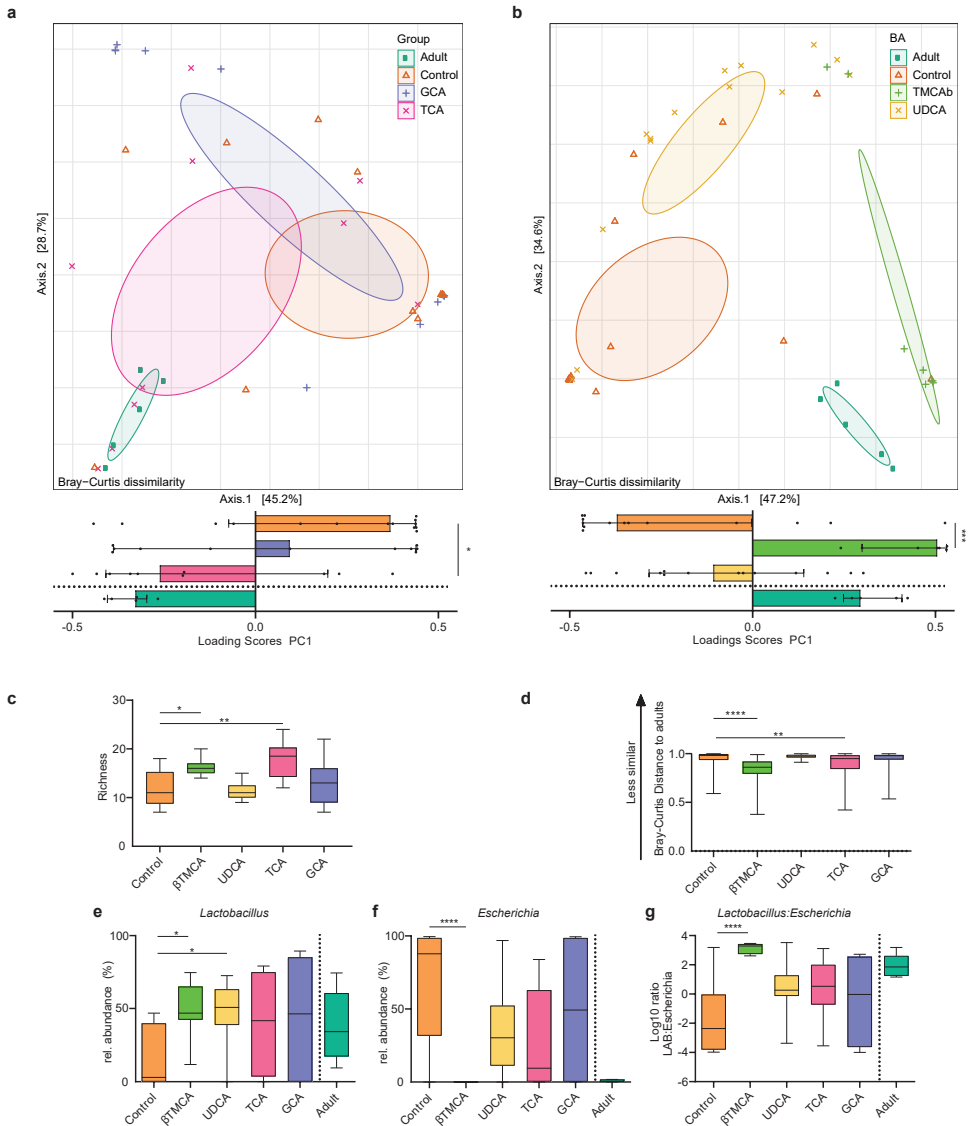
**Figure 3 - Enterohepatic cooperation in the maturation of the microbiota. (A)** Relative abundance of the predicted bacterial bile salt hydrolyses (KO1442) in the small intestinal microbiota (n=5 per group, means connected and replicates). **(B)** Total concentration [ $\mu\text{mol} / \text{gram}$ ] of primary (blue) and secondary (red) bile acids (n=3-5 per group, means connected and replicates). **(C-D)** Coefficients of the regularized canonical correlation analyses (rCCA) indicating the relationship between primary bile acids and small intestinal OTUs on the first component **(C)** and second component **(D)** (selected bile acids are highlighted in red). **(E)** Correlation heatmap based on the coefficients of the rCCA between hepatic primary bile acid concentrations and relative abundances of small intestinal OTUs indicated a strong positive effect (red) of GCA and UDCA on many Lactobacillus OTUs.

### **Integrative analysis of metabolomic and microbiota data**

To investigate the potential influence of bile acid exposure on the composition of the small intestinal microbiota, we next characterized the relationship between individual bile acids and microbial OTUs using regularized canonical correlation analysis (rCCA). rCCA aims at maximizing the correlation of two data sets obtained from the same animals. The first two components explained most of the variance with canonical correlations  $>0.3$  (27.1% of variation in OTUs and 54.1% of variation in bile acids). Ursodeoxycholic acid (UDCA) and glycine-conjugated cholic acid (GCA) obtained the highest canonical coefficient on the first and second component, respectively (Fig. 3c and d). Taurine-conjugated  $\alpha/\beta$ -muricholic acid (TMCA) and taurine-conjugated cholic acid [5] exhibited the strongest negative coefficient. Consistently, UDCA and GCA negatively correlated with early colonizers such as *Mannheimia* and *Streptococcus* and positively correlated with bacteria capable to metabolize bile acids such as *Enterorhabdus* and *Lactobacillus* (Fig. 3e and Fig. S8). In contrast, TMCA showed the opposite effect exhibiting a negative correlation with members of the Coriobacteriaceae family and a positive correlation with *Mannheimia* and *Streptococcus*. Notably, these patterns are in line with the observed changes in the concentration of specific bile acids between pre- and post-weaning but also early after birth and after weaning (Fig. S9). Whereas UDCA and GCA increased in molarity towards adulthood, the concentration of TMCA and TCA decreased over time. Thus, UDCA, GCA, TMCA and TCA are significantly associated with the compositional changes of the microbiota with age and represent good candidates of host-derived factors that drive postnatal establishment of the enteric microbiota.

### **Bile salt-mediated maturation of the neonatal gut microbiota**

Next, we therefore administrated the four bile acids UDCA, GCA,  $\beta$ TMCA and TCA orally to neonate mice and evaluated the impact on the developing small intestinal microbiota. Neonate mice received the indicated bile acid or PBS as control by oral gavage for three consecutive days between PND7-PND9. At PND9, the site-specific small intestinal microbiota composition was determined. PCoA revealed that both TCA and  $\beta$ TMCA shifted the global microbiota composition towards a more adult-like microbiota composition (Fig. 4a-b). Consistently, administration of TCA and  $\beta$ TMCA significantly increased the richness of the small intestinal microbiota (Fig. 4c) and decreased the distance to the adult microbiota composition based on the Bray-Curtis dissimilarity analysis (Fig. 4d).



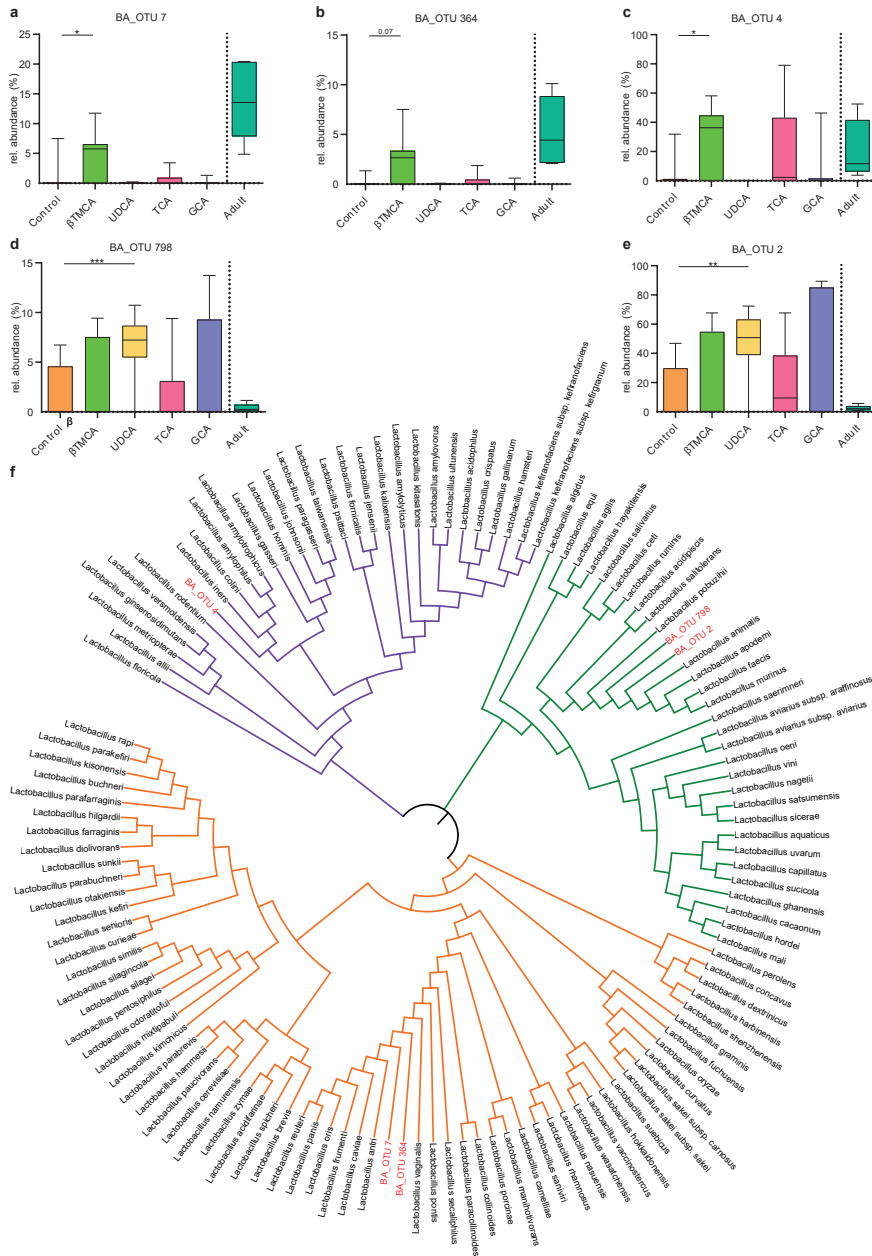
**Figure 4 - Oral administration of bile acids to neonate mice accelerates microbial maturation. (A-B)** PCoA and scores of the 1st axis based on Bray-Curtis dissimilarity indicating shifts in microbial community structure in the small intestine of adult mice (Adult), untreated neonate mice (Control) and neonate mice after oral administration of **(A)** GCA or TCA and **(B)** betaTMCA or UDCA (Kruskal-Wallis test to controls with Dunn's post-test and correction for multiple comparisons, median and interquartile range; \*, p<0.05; \*\*\*, p<0.001. n=7-14, and for subsequent analyses). **(C)** Richness (observed species) of the intestinal microbiota of untreated mice (Control) and mice after oral administration of GCA, TCA, betaTMCA or UDCA (Kruskal-Wallis test to controls with Dunn's post-test and correction for multiple comparisons, mean and SD; \*, p<0.05; \*\*, p<0.01). **(D)** Bray-Curtis distance of untreated neonate mice (Control) and neonate mice after oral administration of GCA, TCA, betaTMCA or UDCA compared to adult animals (Kruskal-Wallis test to controls with Dunn's post-test and correction for multiple comparisons, mean and SD; \*\*, p<0.01; \*\*\*\*, p<0.0001). **(E-G)** Relative abundance of the two most abundant genera *Lactobacillus* and *Escherichia* **(E, F)** and ratio of the abundance of *Lactobacillus* OTUs to *Escherichia* OTUs **(G)** in the microbiota of untreated neonate mice (Control) as compared to neonate mice after oral administration of GCA, TCA, betaTMCA or UDCA (Kruskal-Wallis test to controls with Dunn's post-test and correction for multiple comparisons, mean and SD; \*, p<0.05; \*\*\*\*, p<0.0001). The abundance of these genera **(E, F)** and the ratio of the abundance of *Lactobacillus* OTUs to *Escherichia* OTUs **(G)** in adult mice (Adult) was added to allow visual comparison but was not included in the statistical evaluation (n=5).



In contrast, GCA and UDCA had no impact on bacterial richness (Fig. 4c) or microbiota maturity i.e. the similarity to the adult microbiota (Fig. 4d).

A bile acid-induced shift of the small intestinal microbiota composition towards a more mature phenotype was further supported by the relative abundances of the two most abundant genera in the small intestine, *Lactobacillus* and *Escherichia* (Fig. 4e-f). Here, inoculation of  $\beta$ TMCA and UDCA enhanced the abundance of *Lactobacillus* whereas  $\beta$ TMCA significantly decreased the abundance of *Escherichia*. Using the *Lactobacillus:Escherichia* ratio as a proxy for microbiota maturity,  $\beta$ TMCA administration led to the highest ratio with a value similar to adult animals (Fig. 4g).

Interestingly, members (OTUs) of the *Lactobacillus* genus responded differently to individual bile acids. Whereas  $\beta$ TMCA enhanced OTUs that were found abundantly in adults (Fig. 5a-c), UDCA influenced only OTUs with a minor contribution of the adult microbiota (Fig. 5d, e). Moreover, the annotation of these OTUs into a phylogenetic tree with their closely related *Lactobacillus* ancestors clearly indicated a genetic distinction between the 'UDCA-responders' and ' $\beta$ TMCA-responders' by falling into two separated key branches (Fig. 5f). Transformation of the OTU abundance in absolute bacterial numbers confirmed that  $\beta$ TMCA increased the colonization density of *Lactobacillus* OTUs found abundantly in the adult microbiota (Fig. S10a-c) and decreased the colonization of *Escherichia* (Fig. S10e). Although UDCA also decreased the absolute abundance of *Escherichia* (Fig. S10e), it increased the colonization density of *Lactobacillus* OTUs not typically found abundantly in the adult microbiota (Fig. S10a-c). This finding may explain why UDCA does not enhance the similarity of the neonatal microbial composition to the adult microbiota (Fig. 4d). Notably, *in vitro* analysis of the effect of UDCA, GCA,  $\beta$ TMCA and TCA on murine *Lactobacillus* isolates revealed a direct growth promoting effect of  $\beta$ TMCA on the ' $\beta$ TMCA-responders' *L. johnsonii* (BA\_OTU 4) and *L. reuteri* (BA\_OTU 7 & 364) but not on the 'UDCA-responder' *L. murinus* (BA\_OTU 796 & 2) consistent with our *in vivo* results (Fig. S11). UDCA exerted an inhibitory effect on the murine *Escherichia* isolate both *in vitro* and *in vivo* (Fig. S10e, Fig. S11). In contrast, an inhibitory effect of UDCA on *Lactobacillus* isolates was observed *in vitro* but not *in vivo* (Fig. S10 and Fig. S11).



**Figure 5. Differences in the impact of bile acids on the phylogenetic lactobacillus clusters. (A-E)** Relative abundance of the Lactobacillus OTUs **(A)** BA\_OTU 7, **(B)** BA\_OTU 364, **(C)** BA\_OTU 4, **(D)** BA\_OTU 798, and **(E)** BA\_OTU 2 (Kruskal-Wallis test to controls with Dunn's post-test and correction for multiple comparisons, mean and SD; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ) in the small intestinal microbiota of adult mice (Adult), untreated neonate mice (Control) and neonate mice after oral administration of GCA, TCA,  $\beta$ TMCA or UDCA. **(F)** Phylogenetic tree of lactobacilli (Ezbiocloud) based on the 16S rDNA gene. The three colored branches indicate the main clusters; identified OTUs are assigned to the clusters and highlighted in red (purple, BA\_OTU 4; green, BA\_OTU 798 and BA\_OTU 2; orange, BA\_OTU 7 and BA\_OTU 364).

## Discussion

Together, our results identify bile acids as host factors driving the postnatal small intestinal microbiota composition. They are in accordance with the results of a recent large human twin study that associated metabolic genes with the intestinal bacterial colonization [18]. The postnatal increase in hepatic bile acid synthesis and maturation of the enterohepatic cycle might therefore significantly contribute to shape the early intestinal microbiota. Alterations in the intestinal microbiota composition have previously been observed when supplementing bile acids to adult mice [47], following bile duct ligation in mice [48] as well as in human adults with cholestasis [49]. Yet, these changes have not been linked to any microbiota-driven functional benefit for the host. In contrast, our results in mice identify specific bile acids that promote characteristics of a mature and beneficial microbiota during the particularly important immediate postnatal period.

Careful microbiota profiling throughout the postnatal period revealed major age-dependent compositional changes. Consistent with previous reports, monitoring the immediate postnatal time window revealed an initial reduction of microbial richness within the first 24 hours after birth [2, 50, 51]. The reduction in microbial diversity immediately after birth most likely illustrates the inability of many bacterial species ingested during or shortly after birth to find their niche in the neonatal intestine and thrive. Alternatively, it may reflect the perinatal ingestion and intestinal passage of bacterial DNA remnants. Although not studied in detail so far, it is consistent with the results from previous studies in animals and humans that included samples from this early time window after birth [2, 52, 53]. Few bacterial genera such as *Mannheimia* and *Corynebacterium* exhibited a transient increase in abundance early after birth. The underlying mechanisms are unclear but might involve changes in exposure or altering environmental conditions in the gut lumen such as oxygen concentration, changes in substrate availability and bacterial competition as well as the induction of mucosal host responses. Global bacterial diversification and increasing richness started only thereafter and continued until after weaning in both small intestine and colon.

We analyzed the hepatic metabolome of the identical animals with the goal to reveal the site-specific spectrum of metabolites. This strategy has previously been shown to reliably reflect the metabolic situation in the small intestine [46, 54]. The complexity and inter-individual variation of both the metabolome and the gut microbiota prompted us to perform an integrative analytical approach by rCCR analysis. This analysis identified bile acids as potential drivers of

the microbiota composition. Concomitant with the transition from a hematopoietic tissue to a central metabolic organ, the capacity of the human liver for bile acid biosynthesis increases [45]. Our metabolomic and transcriptional analyses are consistent with the previous notion that the human neonatal liver is able to produce and secrete a significant amount of bile acids into the small intestine [55]. Conjugated primary bile acids dominate during the immediate postnatal period and also after weaning are present at high concentrations in the small intestine. They may well exert a significant influence particularly on the small intestine where a critical influence of the microbiota composition on the maturation of the mucosal immune system in the neonate has been demonstrated [39, 56]. Continuous maturation of the microbiota composition subsequently initiates the well-described bacteria-mediated modification of bile acids as illustrated by the rise in bile salt hydrolyse (BSH) encoding genes during weaning.

Bile acids due to their amphipathic character compromise the bacterial cell wall integrity and thereby exert a potential direct antibacterial effect, consistent with our results for UDCA [57]. *In vivo*, they might additionally synergize with other antibacterial molecules such as bacteriocins or host-derived antimicrobial peptides present in the mammalian intestinal lumen that also target the bacterial surface [58, 59]. In accordance with our results, a great species-specific variability in the bile acid tolerance has been reported, although some general phylogenetic differences exist [60] [57]. Additionally, specific bile acids in mice and men act as agonists or antagonists of mucosal receptor molecules that regulate the secretion of soluble mediators and influence the mucosal immune system and the global host metabolism. Induced changes may in turn indirectly influence the postnatal intestinal microbiota [61]. Specific direct or indirect effects might explain why the overall effect size of a given bile acid might not primarily depend on the concentration. Also, the described activities are expected to differentially affect bacterial taxa and thus, bile acids via a number of mechanisms could influence the composition of the intestinal microbiota.

Oral administration of two major conjugated bile acids,  $\beta$ TMCA and TCA confirmed their functional activity. Both taurine-conjugated bile acids enhanced the richness of the neonatal small intestinal microbiota and rendered it more similar to an adult composition.  $\beta$ TMCA and TCA may suppress bacteria of the neonatal microbiota thereby generating new niches for incoming bacteria and foster the abundance of members of the adult microbiota. Consistently,  $\beta$ TMCA lowered the abundance of *Escherichia* and enhanced the abundance of lactobacilli inverting the *Lactobacillus:Escherichia* ratio in an adult-like manner. Notably, bile acid administration did not influence all members of the *Lactobacillus* genus in a uniform

manner.  $\beta$ TMCA administration *in vivo* enhanced the abundance *Lactobacillus* OTUs that according to available databases contain BSHs and were closely related to dominant murine *Lactobacillus* species, including *L. reuteri* (BA\_OTU 7 and 364) and *L. johnsonii* (BA\_OTU 4) [62]. In contrast, OTUs related to species not containing BSHs (*L. fermentum* and *L. pontis*) remained unaffected [63]. BSHs have different catalytic efficiencies and substrate preferences [64]. Some species are known to contain multiple BSHs with different substrate specificity whereas others only carry a single BSH (e.g. *L. reuteri*) [63]. The carriage of multiple BSHs by the majority of *L. johnsonii* strains might explain why TCA also enhanced the abundance of BA\_OTU 4 *in vivo* while it did not affect the OTUs related to *L. reuteri*. Indeed, it has recently been shown that BSHs from *L. johnsonii* significantly vary in substrate specificity [65]. Consistently,  $\beta$ TMCA exposure *in vitro* strongly promoted the growth of *L. johnsonii* and, to a lesser extent, *L. reuteri*. In contrast, the growth of the BSH-negative bacteria *L. murinus* and *E. coli* remained unaffected.

Notably, the administered bile acid may exert its effect not only directly but also following bacterial modification to CA or DCA or via the induction of metabolic host changes as discussed above. In contrast to  $\beta$ TMCA and TCA, UDCA and GCA had no influence on richness and microbiota maturation. UDCA is highly soluble and generally considered a non-toxic bile acid with little activity on the host's bile acid receptors. GCA is a representative of the glycine-conjugated bile acids, which are common in humans but rare in mice. Substrate specificity and low glycine-deconjugating enzymatic capacity in the murine host's microbiota may therefore explain the lack of a significant effect of GCA on microbiota richness and maturation [66].

However, associations between microbial and bile acid species as observed in our multi-omics rCCR analysis did not completely match with the microbial changes induced upon oral bile acid administration. These differences may reflect the time course of bile acid exposure, the circadian luminal concentration and the dosing. In addition, the metabolic transformation of administered bile acids e.g. by deconjugation or transformation of GCA and TCA to CA and DCA may contribute. Although human neonates other than human adults were shown to produce a more taurine-dominated bile acid conjugation (similar to what is seen in mice), differences exist between the dominant bile acid species in mice and men and our findings may not completely reflect the human setting. For instance, the classic pathway is the predominant pathway for synthesis of bile acids in human liver, whereas all pathways including the mural pathway contribute about equally to bile acid synthesis in rodents. However, ethical concerns

impede the analysis of healthy neonatal liver and small intestinal tissue samples in human neonates. Human fecal samples cannot be used since they mainly contain deconjugated and secondary bile acids [55]. Likewise, the fecal bacterial composition does not reflect the small intestinal microbiota [55]. Future studies need to find ways to analyse human neonatal liver and small intestinal tissue samples and extend the analysis to rare bile acid species. Moreover, further insight in the underlying mechanisms through which bile acids influence the microbiota maturation at the various anatomical sites of the intestinal tract is needed.

Insight into the ecological factors that shape the microbiome during early life is particularly important. These factors act during the non-redundant time-window after birth that critically primes the host's immune system with life-long consequences on the susceptibility to inflammatory and immune-mediated diseases [7, 13]. From the perspective of ecological theory, bile acid concentrations and species profiles could potentially promote processes such as environmental filtering or niche-based interactions. The extent to which bile acids contribute to the inter-individual variation in microbiota composition in humans should be the focus of future research.

In conclusion, we identify bile acids as host factors that shape the postnatal intestinal microbiota composition. The transition of the liver from a hematopoietic tissue to a central organ of metabolic regulation after birth might therefore significantly contribute to the postnatal establishment of host-microbial homeostasis. Our results might help to identify strategies to compensate for the adverse effects of necessary medical interventions such as early-life antibiotic treatment and to promote a microbial composition that is consistent with a beneficial host-microbial homeostasis.

## Acknowledgment

We would like to thank Xinwei Chang and Martina Bernecker for technical support.

## References

1. Integrative, H.M.P.R.N.C., *The Integrative Human Microbiome Project*. Nature, 2019. 569(7758): p. 641-648.
2. Palmer, C., et al., *Development of the human infant intestinal microbiota*. PLoS Biol, 2007. 5(7): p. e177.
3. Backhed, F., et al., *Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life*. Cell Host Microbe, 2015. 17(5): p. 690-703.
4. Hill, C.J., et al., *Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort*. Microbiome, 2017. 5(1): p. 4.
5. Stewart, C.J., et al., *Temporal development of the gut microbiome in early childhood from the TEDDY study*. Nature, 2018. 562(7728): p. 583-588.
6. Martinez, I., et al., *Experimental evaluation of the importance of colonization history in early-life gut microbiota assembly*. Elife, 2018. 7.
7. Torow, N. and M.W. Hornef, *The Neonatal Window of Opportunity: Setting the Stage for Life-Long Host-Microbial Interaction and Immune Homeostasis*. J Immunol, 2017. 198(2): p. 557-563.
8. Ahmadizar, F., et al., *Early-life antibiotic exposure increases the risk of developing allergic symptoms later in life: A meta-analysis*. Allergy, 2018. 73(5): p. 971-986.
9. Cox, L.M., et al., *Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences*. Cell, 2014. 158(4): p. 705-721.
10. Leiby, J.S., et al., *Lack of detection of a human placenta microbiome in samples from preterm and term deliveries*. Microbiome, 2018. 6(1): p. 196.
11. Lozupone, C.A., et al., *Diversity, stability and resilience of the human gut microbiota*. Nature, 2012. 489(7415): p. 220-30.
12. van Best, N., et al., *On the origin of species: Factors shaping the establishment of infant's gut microbiota*. Birth Defects Res C Embryo Today, 2015. 105(4): p. 240-51.
13. Tamburini, S., et al., *The microbiome in early life: implications for health outcomes*. Nat Med, 2016. 22(7): p. 713-22.
14. Chu, D.M., 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): p. 314-326.
15. Vatanen, T., et al., *Genomic variation and strain-specific functional adaptation in the human gut microbiome during early life*. Nat Microbiol, 2019. 4(3): p. 470-479.
16. Falony, G., et al., *Population-level analysis of gut microbiome variation*. Science, 2016. 352(6285): p. 560-4.
17. Benson, A.K., et al., *Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors*. Proc Natl Acad Sci U S A, 2010. 107(44): p. 18933-8.
18. Goodrich, J.K., et al., *Genetic Determinants of the Gut Microbiome in UK Twins*. Cell Host Microbe, 2016. 19(5): p. 731-43.
19. Fulde, M., et al., *Neonatal selection by Toll-like receptor 5 influences long-term gut microbiota composition*. Nature, 2018. 560(7719): p. 489-493.
20. Dominguez-Bello, M.G., et al., *Role of the microbiome in human development*. Gut, 2019. 68(6): p. 1108-1114.
21. Salonen, A., et al., *Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis*. J Microbiol Methods, 2010.

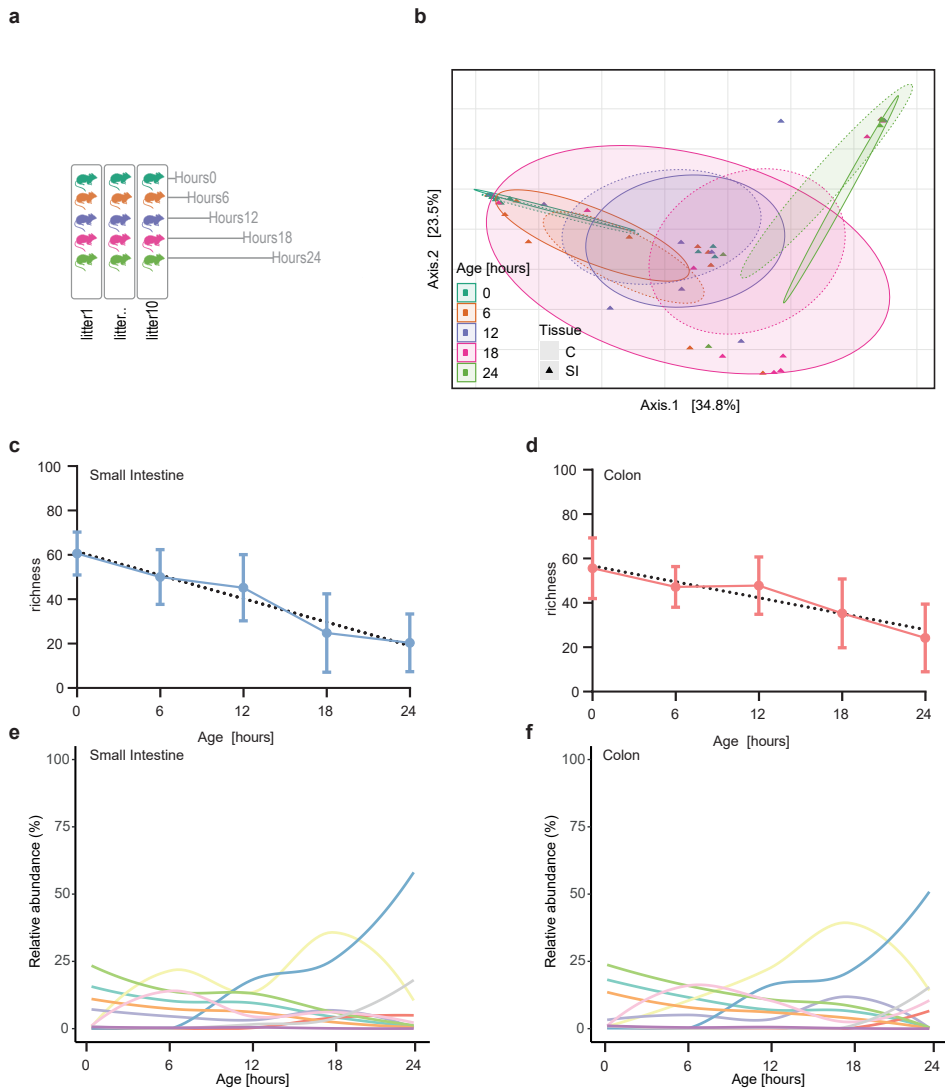
- 81(2): p. 127-34.
22. Caporaso, J.G., et al., *Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms*. ISME J, 2012. 6(8): p. 1621-4.
  23. Lagkouvardos, I., et al., *IMGs: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies*. Sci Rep, 2016. 6: p. 33721.
  24. Edgar, R.C., *UPARSE: highly accurate OTU sequences from microbial amplicon reads*. Nat Methods, 2013. 10(10): p. 996-8.
  25. Edgar, R.C., *Search and clustering orders of magnitude faster than BLAST*. Bioinformatics, 2010. 26(19): p. 2460-1.
  26. Sun, D.L., et al., *Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity*. Appl Environ Microbiol, 2013. 79(19): p. 5962-9.
  27. Edgar, R.C., et al., *UCHIME improves sensitivity and speed of chimera detection*. Bioinformatics, 2011. 27(16): p. 2194-200.
  28. Edgar, R.C., *MUSCLE: multiple sequence alignment with high accuracy and high throughput*. Nucleic Acids Res, 2004. 32(5): p. 1792-7.
  29. Price, M.N., P.S. Dehal, and A.P. Arkin, *FastTree 2--approximately maximum-likelihood trees for large alignments*. PLoS One, 2010. 5(3): p. e9490.
  30. Lagkouvardos, I., et al., *Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons*. PeerJ, 2017. 5: p. e2836.
  31. Chen, J., et al., *Associating microbiome composition with environmental covariates using generalized UniFrac distances*. Bioinformatics, 2012. 28(16): p. 2106-13.
  32. McMurdie, P.J. and S. Holmes, *Phyloseq: a bioconductor package for handling and analysis of high-throughput phylogenetic sequence data*. Pac Symp Biocomput, 2012: p. 235-46.
  33. Holmes, I., K. Harris, and C. Quince, *Dirichlet multinomial mixtures: generative models for microbial metagenomics*. PLoS One, 2012. 7(2): p. e30126.
  34. Lesker, T.R., et al., *An Integrated Metagenome Catalog Reveals New Insights into the Murine Gut Microbiome*. Cell Rep, 2020. 30(9): p. 2909-2922 e6.
  35. Chong, J., et al., *Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data*. Nat Protoc, 2020. 15(3): p. 799-821.
  36. Kenéz, Á., et al., *A metabolomics approach to characterize phenotypes of metabolic transition from late pregnancy to early lactation in dairy cows*. Metabolomics, 2016. 12(11): p. 165.
  37. Pham, H.T., et al., *Inter-Laboratory Robustness of Next-Generation Bile Acid Study in Mice and Humans: International Ring Trial Involving 12 Laboratories*. The Journal of Applied Laboratory Medicine, 2019. 1(2): p. 129-142.
  38. Rohart, F., et al., *mixOmics: An R package for 'omics feature selection and multiple data integration*. PLoS Comput Biol, 2017. 13(11): p. e1005752.
  39. Constantinides, M.G., et al., *MAIT cells are imprinted by the microbiota in early life and promote tissue repair*. Science, 2019. 366(6464).
  40. Ruijter, J.M., et al., *Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data*. Nucleic Acids Res, 2009. 37(6): p. e45.
  41. 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): p. e1.



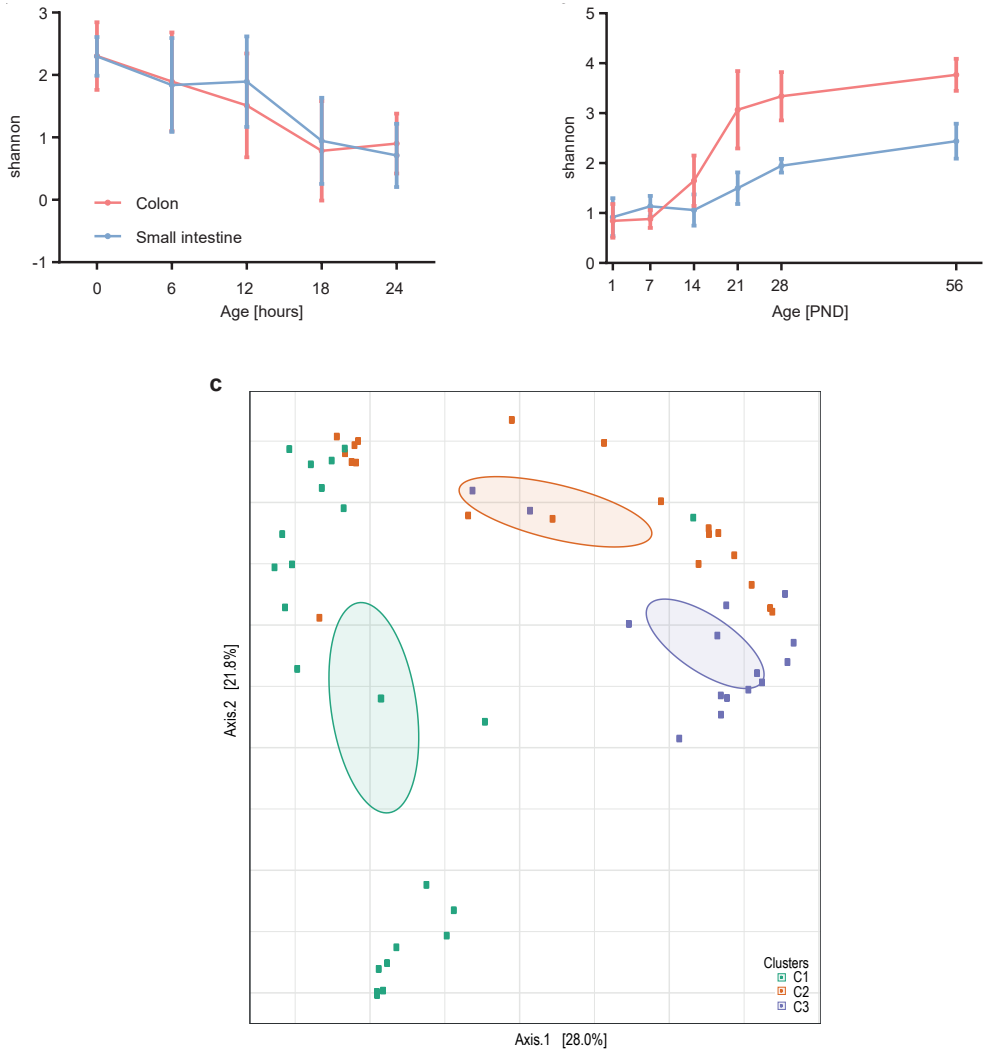
42. Lagkouvardos, I., et al., *The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota*. *Nat Microbiol*, 2016. 1(10): p. 16131.
43. Muncan, V., et al., *Blimp1 regulates the transition of neonatal to adult intestinal epithelium*. *Nat Commun*, 2011. 2: p. 452.
44. Henning, S.J., *Ontogeny of enzymes in the small intestine*. *Annu Rev Physiol*, 1985. 47: p. 231-45.
45. Grijalva, J. and K. Vakili, *Neonatal liver physiology*. *Semin Pediatr Surg*, 2013. 22(4): p. 185-9.
46. Quinn, R.A., et al., *Global chemical effects of the microbiome include new bile-acid conjugations*. *Nature*, 2020. 579(7797): p. 123-129.
47. Van den Bossche, L., et al., *Ursodeoxycholic Acid and Its Taurine- or Glycine-Conjugated Species Reduce Colitogenic Dysbiosis and Equally Suppress Experimental Colitis in Mice*. *Appl Environ Microbiol*, 2017. 83(7).
48. Cabrera-Rubio, R., et al., *Cholestasis induced by bile duct ligation promotes changes in the intestinal microbiome in mice*. *Sci Rep*, 2019. 9(1): p. 12324.
49. Li, Y., et al., *Bile acids and intestinal microbiota in autoimmune cholestatic liver diseases*. *Autoimmun Rev*, 2017. 16(9): p. 885-896.
50. Mueller, N.T., et al., *Delivery Mode and the Transition of Pioneering Gut-Microbiota Structure, Composition and Predicted Metabolic Function*. *Genes (Basel)*, 2017. 8(12).
51. Wampach, L., et al., *Colonization and Succession within the Human Gut Microbiome by Archaea, Bacteria, and Microeukaryotes during the First Year of Life*. *Front Microbiol*, 2017. 8: p. 738.
52. Pantoja-Feliciano, I.G., et al., *Biphasic assembly of the murine intestinal microbiota during early development*. *ISME J*, 2013. 7(6): p. 1112-5.
53. Costa, M.C., et al., *Development of the faecal microbiota in foals*. *Equine Vet J*, 2016. 48(6): p. 681-688.
54. Sayin, S.I., et al., *Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist*. *Cell Metab*, 2013. 17(2): p. 225-35.
55. Tanaka, M., et al., *The association between gut microbiota development and maturation of intestinal bile acid metabolism in the first 3 y of healthy Japanese infants*. *Gut Microbes*, 2020. 11(2): p. 205-216.
56. Cahenzli, J., et al., *Intestinal microbial diversity during early-life colonization shapes long-term IgE levels*. *Cell Host Microbe*, 2013. 14(5): p. 559-70.
57. Watanabe, M., S. Fukiya, and A. Yokota, *Comprehensive evaluation of the bactericidal activities of free bile acids in the large intestine of humans and rodents*. *J Lipid Res*, 2017. 58(6): p. 1143-1152.
58. Tremblay, S., et al., *Bile Acid Administration Elicits an Intestinal Antimicrobial Program and Reduces the Bacterial Burden in Two Mouse Models of Enteric Infection*. *Infect Immun*, 2017. 85(6).
59. Aymeric, L., et al., *Colorectal cancer specific conditions promote *Streptococcus gallolyticus* gut colonization*. *Proc Natl Acad Sci U S A*, 2018. 115(2): p. E283-E291.
60. Ridlon, J.M., et al., *Consequences of bile salt biotransformations by intestinal bacteria*. *Gut Microbes*, 2016. 7(1): p. 22-39.
61. Wahlstrom, A., et al., *Intestinal Crosstalk between Bile Acids and Microbiota and Its Impact on Host Metabolism*. *Cell Metab*, 2016. 24(1): p. 41-50.
62. Schwab, C., et al., *Gene expression of lactobacilli in murine forestomach biofilms*. *Microb Biotechnol*, 2014. 7(4): p. 347-59.
63. O'Flaherty, S., et al., *The Lactobacillus Bile Salt Hydrolase Repertoire Reveals Niche-Specific Adaptation*. *mSphere*, 2018. 3(3).
64. Foley, M.H., et al., *Bile salt hydrolases: Gatekeepers of bile acid metabolism and host-microbiome crosstalk in*

- the gastrointestinal tract*. PLoS Pathog, 2019. 15(3): p. e1007581.
65. DiMarzio, M., et al., *Identification of a mouse Lactobacillus johnsonii strain with deconjugase activity against the FXR antagonist T-beta-MCA*. PLoS One, 2017. 12(9): p. e0183564.
  66. Dong, Z. and B.H. Lee, *Bile salt hydrolases: Structure and function, substrate preference, and inhibitor development*. Protein Sci, 2018. 27(10): p. 1742-1754.

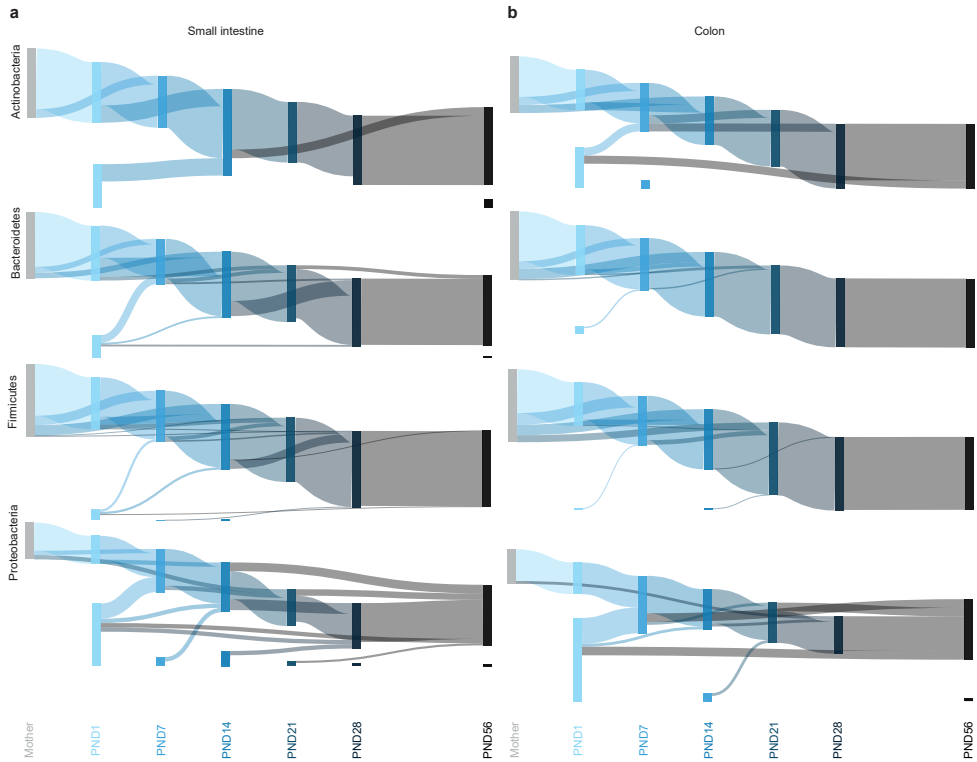
## Supplementary data



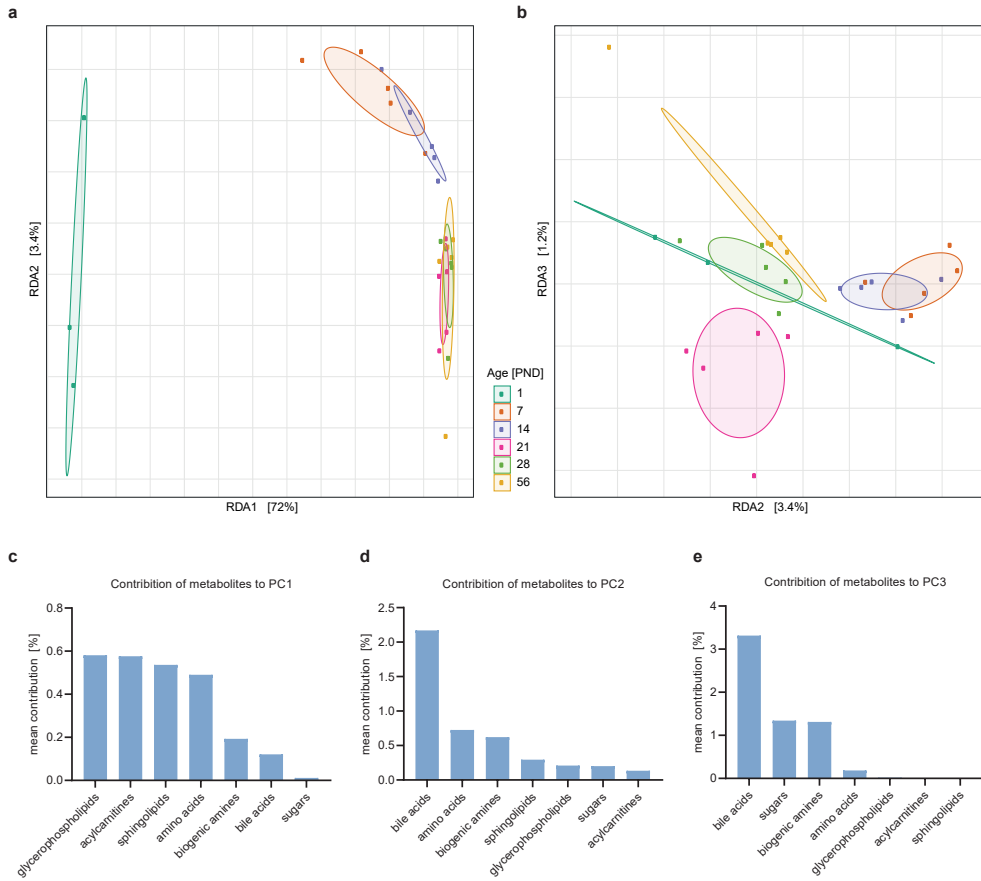
**Figure S1 - Dynamics of the intestinal microbiota composition during the first 24 h after birth. (A)** Image illustrating the study outline; one mouse from each one litter was analyzed for each individual time point (Hours). **(B)** PCoA based on Bray-Curtis dissimilarity illustrating the small intestinal (SI) and colonic **(C)** microbial community structure along PC1 and PC2 during the first 24 hours ( $n=6-10$  per group, also for subsequent analyses). **(C and D)** Richness (observed species) of the **(C)** small intestinal ( $p<0.0001$ , linear regression  $R^2=0.3548$ ) and **(D)** colonic ( $p<0.0001$ , linear regression  $R^2=0.5360$ ) microbiota during the first 24 hours after birth. **(E-F)** Changes in the relative abundance of the 10 most abundant genera for both the small intestine **(E)** and colon **(F)** during the first 24 h after birth.



**Figure S2: Postnatal dynamics in microbial diversity.** (A-B) Shannon index for the small intestinal (blue) and colonic (red) microbiota (A) between 0-24 h after birth (n=6-10 per group) and (B) between 1-56 days (PND1-56) after birth (n=5 per group). (C) Three DMM-clustering profiles confirming the distinct microbial community structures of mainly small intestine and colon PND1-14 (C1, green symbols), small intestine PND21-56 (C2, orange symbols) and colon PND21-56 (C3, blue symbols) ( $p < 0.001$ , Permanova). The ellipses depict the 95% confidence interval.

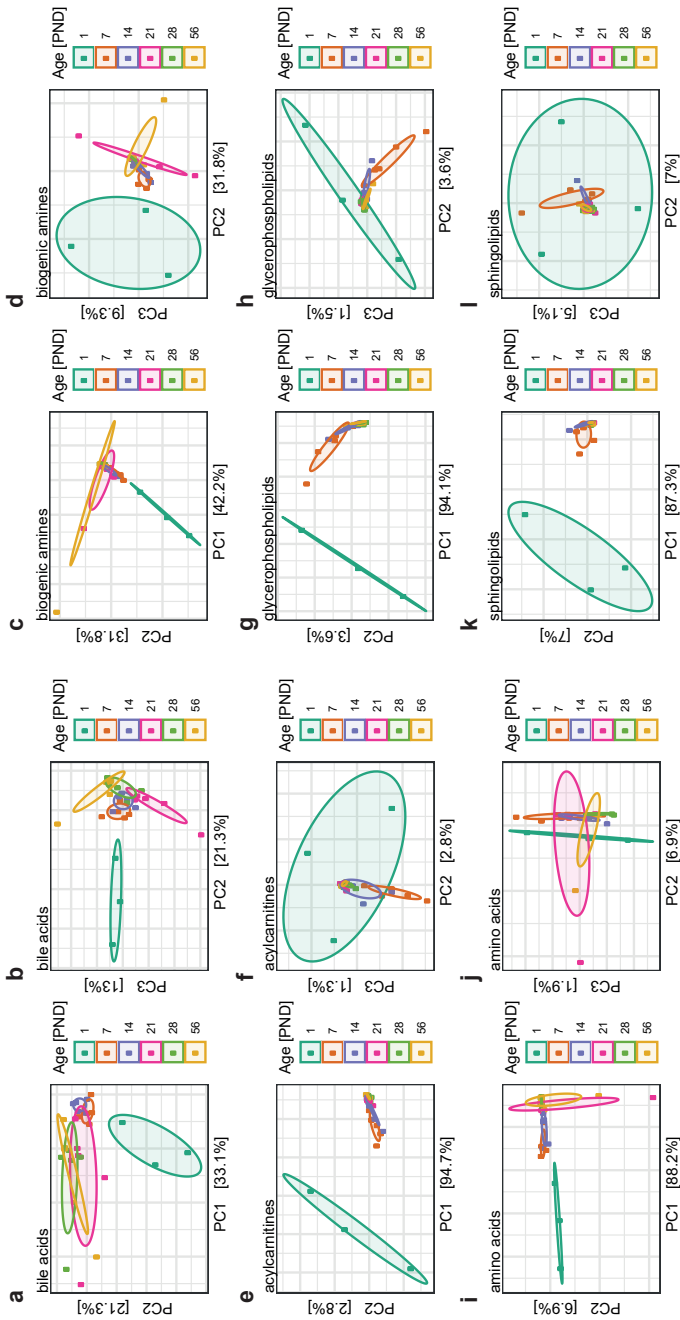


**Figure S3 - Sankey plot tracking OTUs between PND1-PND56. (A and B)** Sankey plot illustrating the origin and transmission of OTUs between mother and offsprings at the indicated time points after birth (postnatal day, PND). All OTUs shared by at least by 20 percent of the population in small intestinal (A) and colonic (B) microbiota were tracked and are shown individually for the four major phyla. The rectangle height indicates the relative number of OTUs (number of observed OTUs at a given time-point divided by the total number of OTUs within a phylum across all time-points) and the rectangle color reflects the age of the animals. Light grey rectangles represent the fecal microbiota of maternal animals. Lines represent the transfer of OTUs between different time points; the thickness of the line is indicative of the number of OTUs transferred.

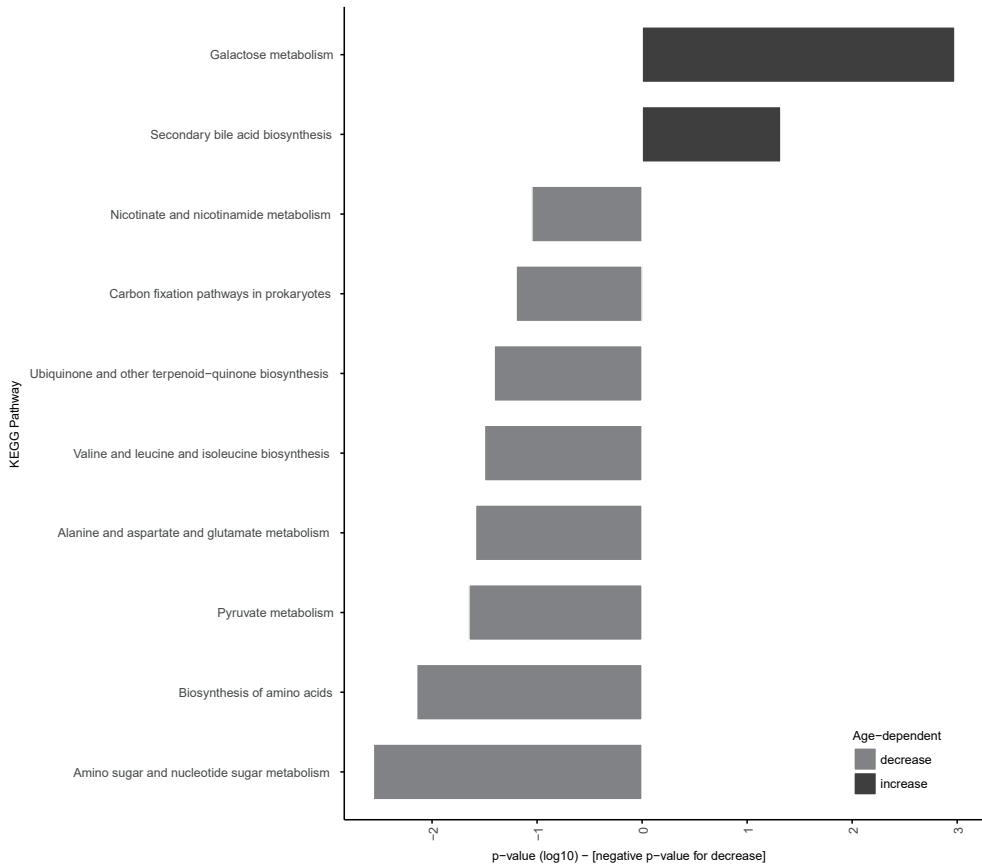


**Figure S4 - The contribution of metabolic groups to age-dependent shifts in the metabolome. (A and B)** Age-constrained ordination using redundancy analysis (RDA) of metabolomic groups colored according to age depicting **(A)** RDA1 and RDA2 and **(B)** RDA2 and RDA3. **(C-E)** Mean contribution of individual metabolic groups to explain the age-dependent variance of **(C)** RDA1, **(D)** RDA2 and **(E)** RDA3.

3



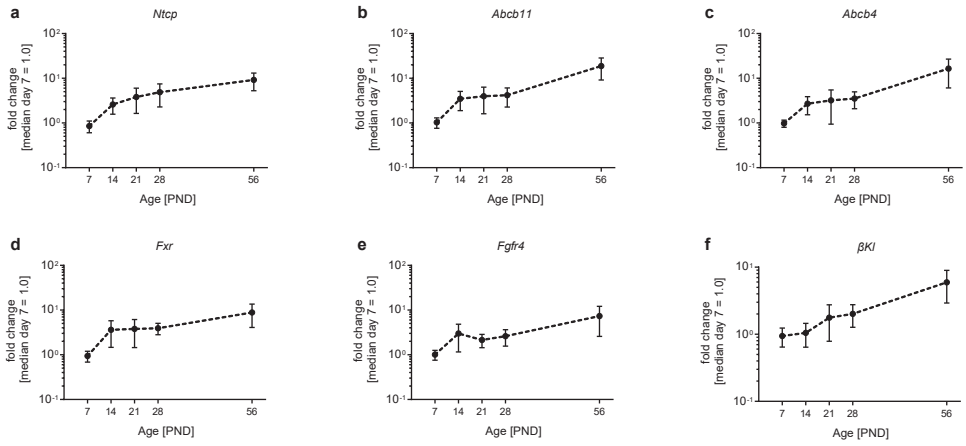
**Figure S5 - Compositional changes of metabolic groups during the postnatal period.** (A-L) Principal Component Analysis (PCA) illustrating changes in the composition of the indicated metabolic groups (A and B) bile acids, (C and D) biogenic amines, (E and F) acylcarnitines, (G and H) glycerophospholipids, (I and J) amino acids and (K and L) sphingolipids between PND 1 and 56. For each metabolic group, PC1 and PC2 (A, C, E, G, I and K) as well as PC2 and PC3 (B, D, F, H, J, L) are shown. Please note that since only the total sugar concentration was measured (single value/sample) we could not include sugars as a metabolic group. Due to the different number of metabolites in each group, the PCA's are not suitable for direct comparisons; see therefore Fig.2 and Fig. S4.



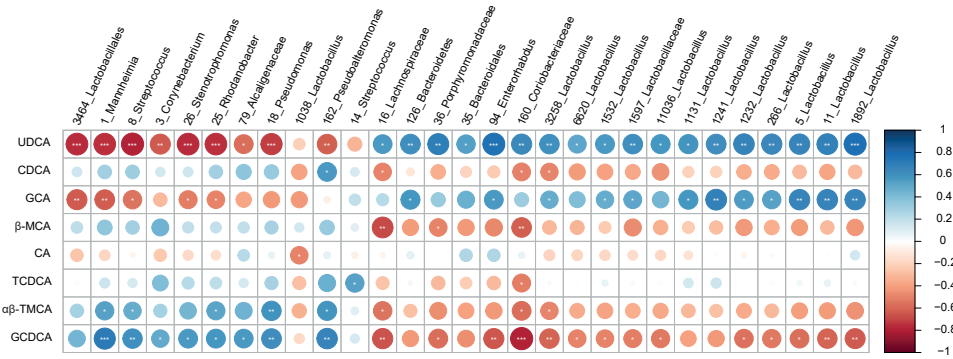
**Figure S6 - Age-dependent metagenomic enrichment analyses.** Mouse-PICRUSt based prediction of metagenomic information functionally annotated using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Enrichment analysis of the age-dependent increases and decreases with the corresponding p-values for the top 10 differentially expressed KEGG pathways (n=5 per PND, PND1 vs. PND56).

3

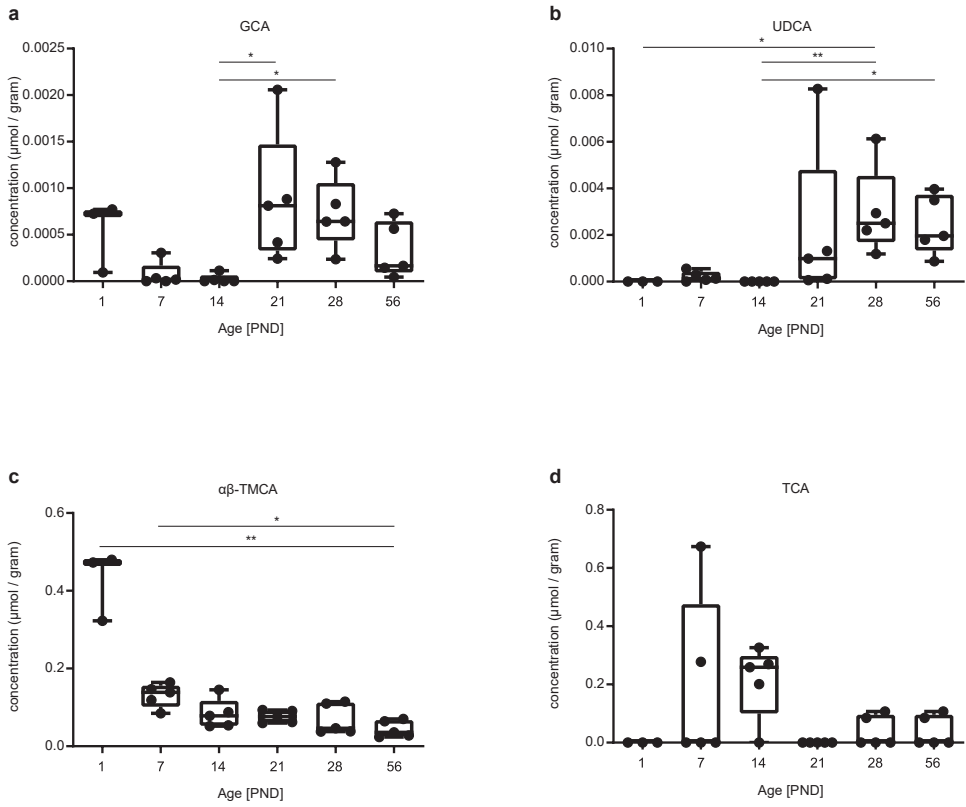




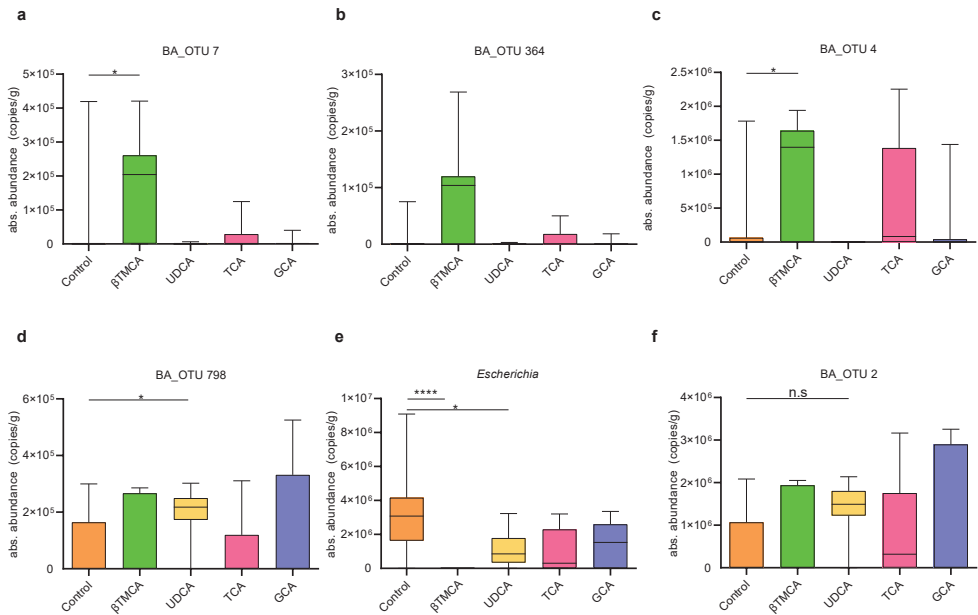
**Figure S7 - Maturation of the enterohepatic cycle.** (A-F) Expression analysis by RT-PCR for key factors in bile acid transport and signaling such as (A) the sodium/bile acid co-transporter Na<sup>+</sup>-taurocholate cotransporting polypeptide (Ntcp), (B) the bile salt export pump ATP-binding cassette, sub-family B member (Abcb)11, (C) the lipid transport molecule ATP-binding cassette, sub-family B member (Abcb)4, (D) the bile acid receptor farnesoid X receptor (Fxr), (E) the fibroblast growth factor receptor (Fgfr)4, and (F) the Fgfr associated protein  $\beta$ -Klotho ( $\beta$ Kl) in total liver tissue of mice at the indicated age (n = 5 per group, Kruskal-Wallis test, mean and SD).



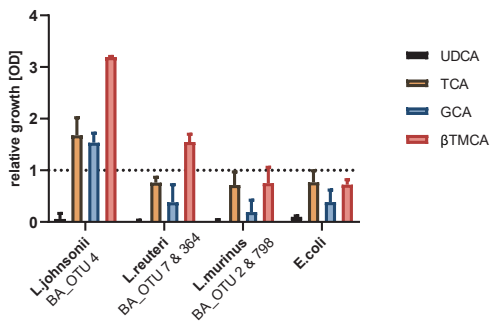
**Figure S8 - Correlation heatmap of primary bile acids and individual OTUs.** Correlation heatmap based on the rCCA with significant pairs of small intestinal OTUs/bile acids (Spearman, \*p<0.05, \*\*p<0.01, \*\*\*p<0.001).



**Figure S9 - Changes in bile acids according to age. A-D,** Concentration [ $\mu\text{mol}/\text{gram}$ ] of the four primary bile acids **(A)** UDCA, **(B)**  $\beta\text{TMCA}$ , **(C)** GCA, and **(D)** TCA in total liver tissue ( $n=3-5$  per group, (Kruskal-Wallis test to controls with Dunn's post-test and correction for multiple comparisons ,mean and SD; \*,  $p<0.05$ ; \*\*\*,  $p<0.001$ ).



**Figure S10 - Impact of orally administered bile acids on the absolute abundance of lactobacilli in the small intestinal microbiota. (A-F)** Absolute abundances of the dominant lactobacillus OTUs **(A)** BA\_OTU 7, **(B)** BA\_OTU 364, **(C)** BA\_OTU 4, **(D)** BA\_OTU 798, **(E)** *Escherichia* and **(F)** BA\_OTU 2 following oral administration of the indicated bile acids calculated based on the results of a quantitative analysis of the total 16S rDNA gene copy number (Kruskal-Wallis test to controls with Dunn's post-test and correction for multiple comparisons, mean and SD; \*,  $p < 0.05$ ; \*\*\*\*,  $p < 0.001$ ).



**Figure S11: Direct influence of bile acids on bacterial survival and growth.** Influence of UDCA, TCA, GCA and  $\beta$ TMCA at a concentration of 1% on the anaerobic growth of representative murine lactobacillus and *E. coli* isolates as illustrated by the relative growth in the presence/absence of bile acid supplementation over 24 h ( $\Delta OD_{570}$  in medium with 1% bile acid/ $\Delta OD_{570}$  in medium with 0% bile acid).





# Chapter 4

## **Gut colonization by methanogenic archaea is associated with organic dairy consumption in children**

**Niels van Best\***, Jeroen A.A. van de Pol\*, Catherine Akwi Mbakwa, Carel Thijs, Paul H.M. Savelkoul, Ilja C.W. Arts, Mathias W. Hornef, Monique Mommers‡, John Penders‡

\*equal contribution / ‡equal contribution

*Front Microbiol. 2017 Mar 10;8:355*

## Abstract

The gut microbiota represents a complex and diverse ecosystem with a profound impact on human health, promoting immune maturation, and host metabolism as well as colonization resistance. Important members that have often been disregarded are the methanogenic archaea. Methanogenic archaea reduce hydrogen levels via the production of methane, thereby stimulating food fermentation by saccharolytic bacteria. On the other hand, colonization by archaea has been suggested to promote a number of gastrointestinal and metabolic diseases such as colorectal cancer, inflammatory bowel disease, and obesity. Archaea have been shown to be absent during infancy while omnipresent in school-aged children, suggesting that colonization may result from environmental exposure during childhood. The factors that determine the acquisition of methanogenic archaea, however, have remained undefined. Therefore, we aimed to explore determinants associated with the acquisition of the two main gastrointestinal archaeal species, *Methanobrevibacter smithii* and *Methanosphaera stadtmanae*, in children. Within the context of the KOALA Birth Cohort Study, fecal samples from 472 children aged 6-10 years were analyzed for the abundance of *M. smithii* and *M. stadtmanae* using qPCR. Environmental factors such as diet, lifestyle, hygiene, child rearing and medication were recorded by repeated questionnaires. The relationship between these determinants and the presence and abundance of archaea was analyzed by logistic and linear regression respectively. 369 out of the 472 children (78.2%) were colonized by *M. smithii*, and 39 out of the 472 children (8.3%) by *M. stadtmanae*. The consumption of organic yogurt (odds ratio: 4.25, CI95: 1.51; 11.95) and the consumption of organic milk (odds ratio: 5.58, CI95: 1.83; 17.01) were positively associated with the presence of *M. smithii*. We subsequently screened raw milk, processed milk and yogurt samples for methanogens. We identified milk products as possible source for *M. smithii*, but not *M. stadtmanae*. In conclusion, *M. smithii* seems present in milk products and their consumption may determine archaeal gut colonization in children. For the first time, a large variety of determinants have been explored in association with gut colonization by methanogenic archaea. Although more information is needed to confirm and unravel the mechanisms in detail, it provides new insights on microbial colonization processes in early life.

## Introduction

The human gut contains a complex and diverse ecosystem consisting of hundreds of microbial species that are acquired during the first years of life [2]. Although a myriad of bacterial species has been studied within the human infant gut, important colonizers that are often disregarded are the methanogenic archaea [3]. At present, 5 methanogenic archaea species and two halophilic archaea have been isolated from human feces from which only the *Methanobrevibacter smithii* (*M. smithii*), *Methanosphaera stadtmanae* (*M. stadtmanae*) and *Methanomassiliicoccus luminyensis* (*M. luminyensis*) have been detected more than once [4-9]. A previous study from Dridi et al. showed that the most dominant archaeal gut inhabitant is *M. smithii* with a prevalence of 88% in children of 0 to 10 years of age. In contrast, *M. stadtmanae* and *M. luminyensis* tend to colonize the child's gut less frequently with a prevalence of 11% and 1%, respectively [10].

Upon colonization, methanogenic archaea are responsible for producing the majority of methane in the gut by reducing carbon dioxide into methane in the presence of hydrogen [11]. The hydrogen in the gut is mainly the result of bacterial fermentation, and accumulation of hydrogen subsequently inhibits this process of breaking down food components for energy. Therefore, reduction of hydrogen levels by methanogens stimulates food fermentation by saccharolytic bacteria [12, 13]. On the other hand, colonization by archaea has been suggested to be potentially detrimental for host health due to alterations in gut metabolism and syntrophic interactions with other microbes [11, 14-17]. For instance, in previous studies higher levels of archaea and excreted methane were found in patients with gastrointestinal and metabolic diseases such as colorectal cancer, inflammatory bowel disease, irritable bowel syndrome, constipation and obesity [18-24].

Methanogenic archaea have been shown to be absent during infancy while omnipresent in school-aged children and their presence seems to increase with age [10]. The latter suggests that colonization may result either through exposure to sources of archaea during childhood, or through factors shaping the gastrointestinal ecophysiology to make the gut more favorable for archaeal colonization during this time period. The factors that determine the acquisition of methanogenic archaea, however, have remained undefined. Although the rumen of beef cattle has been shown to be a carrier for *M. smithii* and *M. stadtmanae* [25], these human gut colonizers have not been identified in selected food products so far [26]. Moreover, no study has conducted a comprehensive analysis on potential lifestyle and dietary determinants



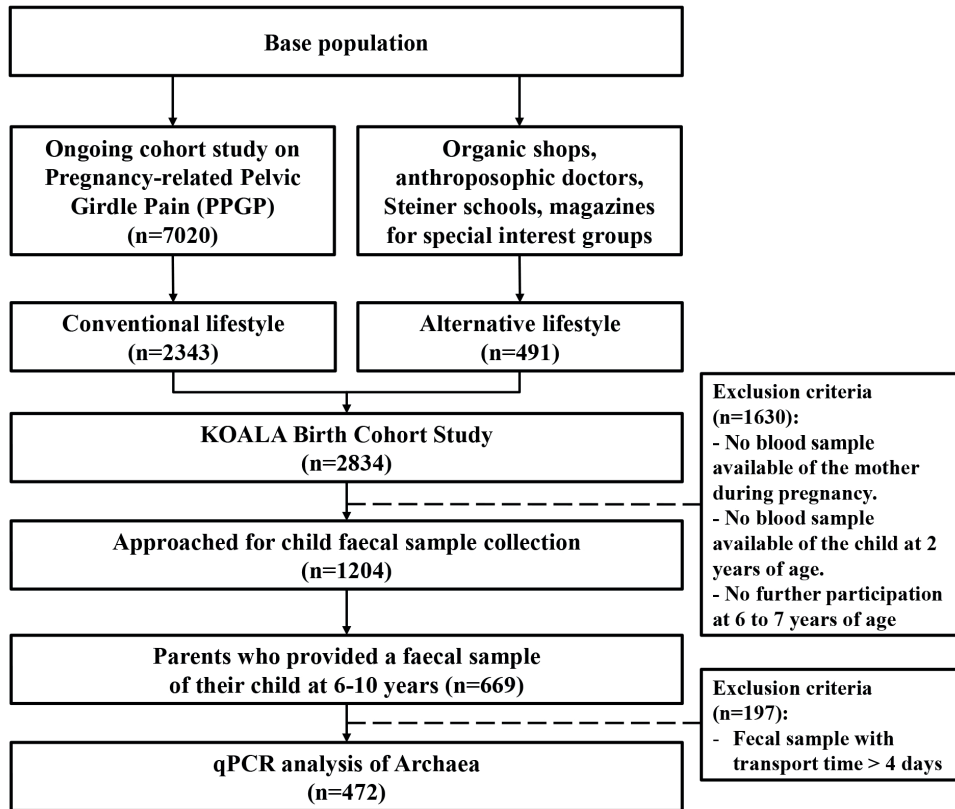
of human gut colonization by these methanogenic archaea. Therefore, we aimed to explore a wide variety of potential determinants associated with the acquisition of the two main archaeal species, *M. smithii* and *M. stadtmanae*, in children. To this end, we used extensive data on determinants prospectively gathered through repeated questionnaires within the KOALA Birth Cohort Study in combination with data on presence and abundance of *M. smithii* and *M. stadtmanae* obtained from fecal samples in children of 6 to 10 years.

## Material and Methods

### Study population

This study was conducted within the KOALA Birth Cohort Study in the Netherlands. The study design has been described in detail elsewhere [27]. In summary, 2834 pregnant women were recruited between October 2000 and December 2002. Pregnant women with a conventional lifestyle (N=2343) were recruited from an ongoing cohort study on pregnancy-related pelvic girdle pain in the Netherlands. In addition, pregnant women with an 'alternative' lifestyle (N=491) were recruited through organic shops, anthroposophical doctors and midwives, anthroposophical under-five clinics, Steiner schools and magazines for special interest groups. The 'alternative' lifestyle was expected to differ from the 'conventional' lifestyle in vaccination practices, use of antibiotics, dietary habits (breastfeeding, organic foods, and vegetarian diet) and child rearing practices [27].

Parents of a subgroup of children (N=1204) were approached for the collection of a fecal sample of their child at age 6-10 (Figure 1). Fecal samples were provided by 669 children. Transport time exceeded 4 days for 197 samples, which were therefore excluded from analyses. As a result, quantitative real-time PCR analysis was performed on 472 fecal samples. Informed consent was given by all parents, and the study was approved by the Medical Ethics Committee of Maastricht University and the National Ethical Committee for Medical Research.



**Figure 1** - Flowchart illustrating how the population under study (n=472) was obtained from the KOALA Birth Cohort Study population

### Fecal sampling

Parents received a stool sample kit consisting of a feces collection tube with a spoon attached to the lid (Sarstedt, Nürnberg, Germany) together with an instruction form on fecal collection. After collection, the fecal sample was sent to the Department of Medical Microbiology at the Maastricht University Medical Center+ by mail. Upon arrival fecal samples were 10-fold diluted in Peptone Water (Oxoid CM0009) containing 20% (vol/vol) glycerol (Merck, Darmstadt, Germany) and stored at  $-80^{\circ}\text{C}$  until further analysis.

### Fecal DNA isolation and quantitative real time PCR

Repeated-Bead-Beating (RBB) plus a column-based purification method was used to isolate DNA as described in detail elsewhere [28, 29]. Afterwards, DNA concentration and purity were

determined by the Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, USA). For the enumeration of *M. smithii* and *M. stadtmanae* all fecal samples were subjected to 5'-nuclease based real-time PCR assays. Primers and probes employed for PCR and the amplification and quantification process as performed on an Applied Biosystems Prism 7000 sequence detection system (Applied Biosystems) have been described in detail elsewhere [20]. In short, quantification of the abundance of *M. smithii* and *M. stadtmanae* in fecal samples was achieved by comparing the cycle threshold (Ct) values to a standard curve. The standard curves were constructed by subjecting serial tenfold dilutions of positive plasmid constructs containing the target sequences of *M. smithii* and *M. stadtmanae* to the same qPCR assays. The lower limit of detection was 3.81 log<sub>10</sub> copies/g feces for *M. smithii* and 4.82 log<sub>10</sub> copies/g feces for *M. stadtmanae*. Further details on qPCR assays and conditions can be found elsewhere [20].

### Determinants

Information on potential determinants was collected by repeated questionnaires in the KOALA Birth Cohort Study. Information was requested from 34 weeks of gestation until 6-7 years of age. Questionnaires covered multiple topics including dietary habits (food frequency questionnaires (FFQ)), lifestyle characteristics, hygiene, child rearing practices, diseases and medication use. Determinants of interest were selected based on findings from previous research on the establishment of the microbiota and the availability of information in repeated questionnaires.

Determinants selected for analysis were: age of the child at time of fecal sample (continuous in years), recruitment group (conventional cohort, alternative cohort), child's gender, maternal education (low: primary school, preparatory vocational or lower general secondary school; middle: vocational, higher general secondary or pre-university; high: higher vocational or academic education), maternal diet during pregnancy (conventional; organic/biodynamic), place and mode of delivery (vaginal delivery at home; vaginal delivery in hospital; artificially induced delivery in hospital; caesarean section in hospital), gestational age (continuous [weeks]), birth weight (continuous [grams]), regularly cleaning of pacifier in boiling water (no, yes), duration of pacifier use (never, short [until 7 months], long [at least until 12 months]), first time exposure to antibiotics during the first 2 years of life (Never, at 0-3 months, at 3-7 months, at 7-12 months, at 12-24 months), number of siblings at 1 year of age (0, 1, ≥2), regular stay at daycare during the first 2 years of life (never; at host parent or daycare; host parent and daycare combined), vitamin D/AD supplementation during the

first 2 years of life (no; yes), number of siblings at 4-5 years of age (0; 1;  $\geq 2$ ), type of pets kept during life of the child (none; dog; cat; other; combination), exposure to farm (animals) during the last 12 months at 6-7 years of age (never; pigs; chicken/pigeons; ox/goat/sheep/horse; children's farm visits; combination), antibiotic use within 1 year prior to fecal sampling (no; yes), child's diet at 6-7 years (conventional;  $>25\%$  organic/biodynamic), vegetarian diet of child (no; yes), total energy intake (continuous [kcal]), carbohydrate intake (continuous [en%]), protein intake (continuous [en%]), animal protein intake (continuous [en%]), fiber intake (continuous [grams]), regular intake of organic raw vegetables (no; yes), regular intake of organic fruit (no, yes), regular milk consumption (at least once per week; no, yes), regular organic milk consumption (at least once per week; no, yes), regular cheese consumption (at least once per week; no, yes), regular organic cheese consumption (at least once per week; no, yes), regular yogurt consumption (at least once per week; no, yes), regular organic yogurt consumption (at least once per week; no, yes).

### **Statistical analysis**

Characteristics of the present study population (N=472) and the total KOALA birth cohort (N=2834) are presented as median plus range for continuous variables and as frequency (n (%)) for categorical variables.

Multivariable logistic regression models were used to assess the association between potential determinants and the presence of *M. smithii* and *M. stadmanae*, respectively. Multivariable linear regression models were used to assess the association with *M. smithii* abundance. Variable selection for the regression models was based on purposeful selection as described by Hosmer & Lemeshow (Hosmer and Lemeshow, 2000). A multivariable model was fitted including all determinants identified from univariable models with a p-value of  $\leq 0.25$ . This model was further reduced to only include variables with either a p-value lower than 0.10 or variables that affected the parameter estimate ( $\beta$ ) of other variables in the model by more than 20%, as recommended by Bursac, Gauss [30]. This model was subsequently refitted with variables that were not statistically significant (i.e.  $p > 0.25$ ) in univariable analyses. Afterwards, the model was reduced as described before ( $p < 0.10$ ), but only for the added variables. As a result, all variables included in the main effects model were statistically significant or made an important contribution based on the presence of other variables in the model [30]. After completion of purposeful selection the resulting main effects model was additionally adjusted for the following covariables of prior interest: age at time of

fecal sampling (years), gender (male, female), total energy intake (kcal), recruitment group (conventional, alternative) and BMI (z-score). For determinants with more than 30 missing values a missing value group was constructed. Models were checked for multicollinearity through Variance Inflation Factor (VIF) scores. If VIF scores exceeded 10, determinants causing multicollinearity were separated into independent models to obtain effect estimates without multicollinearity for the affected determinants. To limit the number of spurious associations, results from multivariable analyses were corrected for multiple comparisons by adjusting  $\alpha$  by Benjamini-Hochberg procedure at a false discovery rate (FDR) level of  $q=0.05$  [31].

In secondary analyses, the impact of the number of consumed organic dairy products, and organic as well as raw vegetable and fruit products on the presence and abundance of *M. smithii* and *M. stadtmanae* was examined. For this purpose the individual variables for organic dairy (cheese, milk and yogurt) and organic fruit and raw vegetables were grouped into index variables as follows: number of different organic dairy products (0; 1; 2; 3), number of different organic raw vegetable and fruit products (0; 1; 2). Since organic dairy products were significantly associated with the presence of *M. smithii*, analyses were extended to yogurt, milk and cheese, irrespective of organic or conventional origin. For all statistical analyses IBM SPSS version 23 (SPSS Inc., Chicago, IL) was used.

### Archaeal presence in milk

To validate findings from this study, publicly available data from another study was used to identify archaea in 975 raw milk samples collected from 899 tanker trucks [1]. These individual trucks arrived at two dairy processors in San Joaquin Valley of California for product manufacturing during three seasons between October 2013 and September 2014. The sequences of the 16S rRNA V4 regions were obtained via the Qiita database (<https://qiita.ucsd.edu>) under study ID 10485, and further analyzed within Qiime [32]

Processed milk and yoghurt samples were collected from biodynamic, organic and non-organic brands from the supermarket (Supplementary Table 1). Raw unprocessed milk was obtained from a local farmer via a standard tap-procedure (Brunimat). Metagenomic DNA was extracted from dairy products with the PowerFood™ Microbial DNA Isolation kit (MoBio Laboratories Inc.) according to manufacturer's instructions as evaluated in detail elsewhere [33]. In short, cell lysis with chaotrophic agents and bead-beating plus a column-based extraction was used to isolate DNA.

## Results

Fecal samples from 472 children with a median age of 7.2 (6.0-12.0) years were analyzed. Of these, 369 (78.2%) out of 472 were colonized by *M. smithii*, while 39 (8.3%) out of 472 children were colonized by *M. stadtmanae*. Due to the low colonization rate of *M. stadtmanae*, analyses on its abundance were not performed in this study. Supplementary table 2 shows participant characteristics of the KOALA Birth Cohort Study (n=2834) and the study population (n=472) for all selected determinants. In general, the distribution of participant characteristics of the study population was similar to the population of the entire KOALA Birth Cohort Study.

### **Methanobrevibacter smithii presence and abundance**

The multivariable regression models (Table 1 and Supplementary Table 3) showed a positive association between organic yogurt and organic milk consumption and *M. smithii* colonization. Children who regularly consumed organic yogurt were more than four times as likely to be colonized by *M. smithii* as compared to children who did not consume organic yoghurt (OR<sub>adjusted</sub>: 4.25 95% Confidence Interval (CI95): 1.51; 11.95), whereas children consuming organic milk were even over 5 times more likely to be colonized by *M. smithii* as compared to children who did not consume organic milk (OR<sub>adjusted</sub>: 5.58, CI95: 1.83; 17.01). We subsequently performed secondary analyses to examine the effect of the number of different organic dairy products being consumed on *M. smithii* colonization (Table 2 and Supplementary Table 4). These analyses showed a statistically significant increasing trend between the number of dairy products consumed and the chance of colonization by *M. smithii* (Table 2, p=0.002).

**Table 1.** Final multivariable logistic regression model showing the association between potential determinants and the presence of *Methanobrevibacter smithii*. Adjusted main effects model (N=419)<sup>a</sup>

Determinant	OR (CI95%)	FDR crit. <sup>b</sup>	p-value
Regular intake of organic milk <sup>c, d</sup>			
No	Ref.		
Yes	5.58 (1.83; 17.01)	0.006	<b>0.003<sup>f</sup></b>
Regular intake organic yogurt <sup>c, d</sup>			
No	Ref.		
Yes	4.25 (1.51; 11.95)	0.006	0.006
Diet of child <sup>e</sup>			
Conventional (≤25% organic)	Ref.		
Organic (incl. biodynamic) (>25% organic)	0.36 (0.17; 0.79)	0.011	<b>0.010<sup>f</sup></b>
Fiber intake (g) <sup>e</sup>	0.95 (0.88; 1.03)	0.028	0.250

<sup>a</sup> Model adjusted for: age at fecal sampling (years), gender (male/female), recruitment group (conventional/alternative), total energy intake (kcal) and BMI (z-score); <sup>b</sup> critical FDR cut-off level as determined by Benjamini-Hochberg procedure; <sup>c</sup> Missing value category omitted (included in FDR correction); <sup>d</sup> Due to multicollinearity both regular intake of organic milk and regular intake of organic yogurt were included in separate models, which included all other variables as listed in this table; <sup>e</sup> Parameter estimates presented for the model containing organic milk intake. Parameter estimates did not change substantially when the variable 'organic milk intake' was replaced by the variable 'organic yoghurt intake' (data not shown); <sup>f</sup> Significant association after correction for FDR by Benjamini-Hochberg procedure.

**Table 2.** Secondary multivariable logistic regression model estimating the impact of the number of organic dairy products on the presence of *Methanobrevibacter smithii*.

<b>M. smithii presence</b>			
<b>Adjusted main effects model (N=406)<sup>a</sup></b>			
Determinant	OR (CI95%)	FDR crit. <sup>b</sup>	p-value
First exposure to antibiotics (during first 2 years of life)			
Never	Ref.		
At 0-7 months	0.61 (0.33; 1.13)	0.020	0.115
At 8-12 months	1.20 (0.58; 2.49)	0.035	0.631
At 13-24 months	0.78 (0.39; 1.57)	0.030	0.490
Regular intake of organic products (cheese, milk and yogurt)			
Trend (0, 1, 2, 3)	2.12 (1.31; 3.43)	0.005	<b>0.002<sup>c</sup></b>
Diet of child			
Conventional (≤25% organic)	Ref.		
Organic (incl. biodynamic) (>25% organic)	0.30 (0.14; 0.65)	0.010	<b>0.002<sup>c</sup></b>

<sup>a</sup> Model adjusted for: age at fecal sampling (years), gender (male/female), recruitment group (conventional/alternative), total energy intake (kcal) and BMI (z-score); <sup>b</sup> Critical FDR cut-off level as determined by Benjamini-Hochberg procedure; <sup>c</sup> Significant association after correction for FDR by Benjamini-Hochberg procedure.

For all other potential determinants that have been examined in the present study, no relationship with *M. smithii* colonization was found. Moreover, when examining the relationship between *M. smithii* colonization and the consumption of dairy products, irrespective of organic or conventional origin, we did neither find an association for milk nor for yogurt and nor for



cheese consumption. In contrast to the analyses on the presence of *M. smithii*, none of the determinants assessed in this study were positively associated with the abundance of *M. smithii* (Supplementary Tables 5 & 6). In models for both *M. smithii* presence and abundance the model adjustment by a priori selected covariates did not meaningfully alter parameter estimates. The minimal effect of these covariates on parameter estimates indicates that models for *M. smithii* were robust.

### **Methanosphaera stadtmanae presence**

Multivariable regression analyses on the presence of *M. stadtmanae* showed no significant results after FDR correction (Supplementary Table 7). However, after adjustment for a priori selected covariables birth by caesarean section (OR: 6.89, CI95: 2.09; 22.67), first exposure to antibiotics at 13-24 months (OR: 3.38, CI95: 1.34; 8.50) and organic fruit intake (OR: 4.73, CI95: 1.64; 13.62) were associated with an increased *M. stadtmanae* presence (Table 3). In secondary analyses to examine the effect of the number of organic fruit and vegetable products consumed on the presence of *M. stadtmanae* no significant trend was found ( $p=0.321$ ).

In models for *M. stadtmanae* presence minor differences in model estimates were found after adjustment for a priori selected covariates. These differences may indicate underlying relationships between potential determinants and a priori selected covariates or model instability of *M. stadtmanae* models.

### **Verification of archaeal presence in milk**

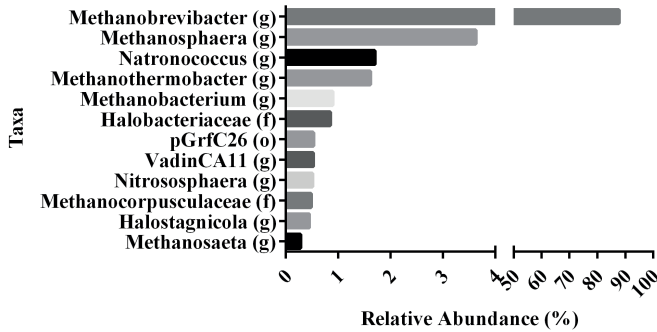
To evaluate the contribution of milk and yogurt to gut colonization by Archaea, we re-analyzed 16S sequence-data of raw milk from a previous study [1] as only the bacterial results of these samples from 899 tanker trucks were reported in the original publication. The relative abundances showed that the vast majority of all recovered archaeal sequences could be assigned to the genera *Methanobrevibacter* (87.70%) and *Methanosphaera* (3.63%) compared to the other 10 ( $\leq 1.70\%$ ) (Figure 2). In addition, 947 out of the 975 milk samples (97.13%) were positive for *Methanobrevibacter* while *Methanosphaera* sequences were present in 348 of the milk samples (35.69%). To further assess the absolute counts of methanogens in dairy products, we subsequently screened unprocessed raw milk, pasteurized milk, and pasteurized yogurt samples (Supplementary Table 1). *M. smithii* and *M. stadtmanae* were quantified by qPCR on isolated DNA of these samples. Although *M. smithii* could not be

detected in yogurt samples, we found significant levels of *M. smithii* in milk (Figure 3). The average absolute counts in processed milk was highly similar for biodynamic (2.63 log<sub>10</sub> DNA copies/ml), organic (2.88 log<sub>10</sub> DNA copies/ml) and conventional milk (2.94 log<sub>10</sub> DNA copies/ml). However, raw milk (3.73 log<sub>10</sub> DNA copies/ml) showed higher *M. smithii* counts compared to processed milk. For *M. stadtmanae*, no detectable levels could be measured in any of the samples. In conclusion, we identified milk as a possible source off *M. smithii*.

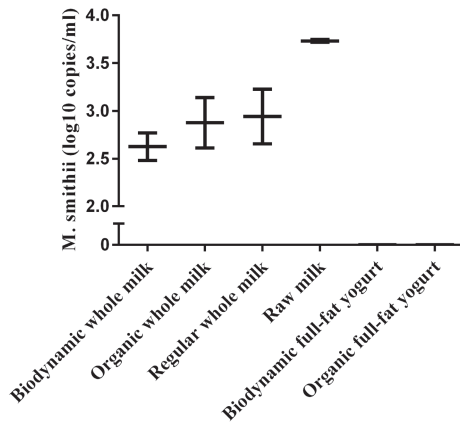
**Table 3.** Final multivariable logistic regression model showing the association between potential determinants and the presence of *Methanosphaera stadtmanae*.

<b><i>M. stadtmanae</i> presence</b>			
<b>Adjusted main effects model (N=420)<sup>a</sup></b>			
<b>Determinant</b>	<b>OR (CI95%)</b>	<b>FDR crit.<sup>b</sup></b>	<b>p-value</b>
<i>Birthweight (g)</i> <sup>c</sup>	1.001 (1.000; 1.002)	0.019	0.028
Place and mode of delivery <sup>c</sup>			
<i>Natural birth at home</i>	Ref.		
<i>Natural/artificial birth at hospital</i>	1.61 (0.67; 3.88)	0.034	0.288
<i>Caesarean section</i>	6.89 (2.09; 22.67)	0.003	<b>0.002<sup>f</sup></b>
First exposure to antibiotics (during first 2 years of life) <sup>c</sup>			
<i>Never</i>	Ref.		
<i>At 0-7 months</i>	0.54 (0.16; 1.83)	0.038	0.323
<i>At 8-12 months</i>	0.70 (0.21; 2.36)	0.050	0.566
<i>At 13-24 months</i>	3.38 (1.34; 8.50)	0.013	<b>0.010<sup>f</sup></b>
Regular intake of organic milk <sup>d, e</sup>			
<i>No</i>	Ref.		
<i>Yes</i>	3.52 (0.97; 12.73)	0.022	0.056
Regular intake of organic yogurt <sup>d, e</sup>			
<i>No</i>	Ref.		
<i>Yes</i>	1.690 (0.49; 5.86)	-	0.409
Regular intake of organic fruit <sup>c</sup>			
<i>No</i>	Ref.		
<i>Yes</i>	4.73 (1.64; 13.62)	0.006	<b>0.004<sup>f</sup></b>
<i>Animal protein intake (en%)<sup>f</sup></i>	0.77 (0.65; 0.92)	0.009	<b>0.004<sup>f</sup></b>

<sup>a</sup> Model adjusted for: age at fecal sampling (years), gender (male/female), recruitment group (conventional/alternative), total energy intake (kcal) and BMI (z-score); <sup>b</sup> Critical FDR cut-off level as determined by Benjamini-Hochberg procedure; <sup>c</sup>Parameter estimates presented for the model containing organic milk intake. Parameter estimates did not change substantially when the variable 'organic milk intake' was replaced by the variable 'organic yogurt intake' (data not shown); <sup>d</sup> missing value category omitted (included in FDR correction); <sup>e</sup> Due to multicollinearity both regular intake of organic milk and regular intake of organic yogurt were included in separate models, which included all other variables as listed in this table; <sup>f</sup> significant association after correction for FDR by Benjamini-Hochberg procedure.



**Figure 2** - Variation in the relative abundances of dominant Archaea taxa in raw milk of re-analyzed 16S sequence-data from Kable, Srisengfa [1]. Data represent the abundances of the various archaeal taxa (found at 0.1% or greater) as proportion of the total archaeal populations in raw tanker milk. Relative abundances of OTU's rarified at 15,000 sequences / sample are indicated at their highest identified taxa. (o) order. (f) family. (g) genus.



**Figure 3** - Absolute counts of *M. smithii* in whole or full-fat dairy products (n=2-3). Replicates (n) denotes repeated isolation and quantification of the product. Average counts were calculated from archaea positive samples only.

## Discussion

To the best of our knowledge, this is the first large-scale explorative study investigating the associations between a wide range of potential determinants and intestinal colonization by methanogenic archaea in school-aged children. We found that consumption of organic yogurt and milk products is related to colonization of *M. smithii* in particular. Moreover, we identified *M. smithii* in raw and processed milk, which could therefore be considered as potential sources for archaeal colonization.

The difference in metabolic activity might be a possible explanation for the large difference in the prevalence of the two main intestinal archaea. *M. smithii* is responsible for producing the majority of methane in the gut by reducing carbon dioxide using hydrogen as the primary electron donor. Alternatively, *M. smithii* can use formate as a direct substrate to produce methane [11]. In contrast, *M. stadtmanae* is only able to produce methane from hydrogen [12]. Moreover, colonization by *M. smithii* and *M. stadtmanae* might originate from different environmental sources, which could explain their different prevalence. Factors that play a role in the acquisition of methanogenic archaea are largely unknown. However, the environment is assumed to play a crucial role in gut colonization. This view is strengthened by a twin study that indicated that not genetic factors, but shared and unique environmental factors were of importance for the occurrence of methanogens in humans [34]. The prospective design and long-term follow-up allowed us to examine a wide range of environmental, dietary and life-style associated determinants throughout childhood. Out of the determinants present during infancy, caesarean section compared to vaginal delivery at home, first exposure to antibiotics in the 2<sup>nd</sup> year of life compared to no antibiotic treatment during the first two years and organic fruit consumption were associated with increased odds of colonization by *M. stadtmanae*. These findings on *M. stadtmanae* in the present study need however to be interpreted with great care. Due to the small group of children colonized by *M. stadtmanae* (n=39), models for *M. stadtmanae* were prone to instability. Therefore, these factors need to be verified in other studies with, preferably, larger sample sizes. In particular, the association with birth mode is in contrast with previous findings that intestinal colonization by methanogenic archaea in neonates and infants is very rare [35]. However, it could be postulated that the pioneer species associated with caesarian section delivery drive a subsequent colonization pattern that is more favorable for the establishment of *M. stadtmanae* at a later stage. As such, it is important to study the co-occurrence and co-exclusion of archaea and bacterial taxa in future studies, in order to identify microbial networks that might favor archaeal colonization.

Although one previous study found females to be more often colonized by methanogenic archaea than males [34], the current view is that there is no association between methanogenic archaea and gender [36]. In this study, we assessed fecal samples of 244 (51.7%) males and 228 (48.3%) females. Of all males under study 187 (76.6%) were colonized by *M. smithii*, while 182 (79.8%) females were colonized by *M. smithii*. Furthermore, 19 (7.8%) males were colonized by *M. stadtmanae*. This was similar to the colonization by *M. stadtmanae* in females, as 20 (9.6%) females were colonized. Moreover,

we also did not find an association between gender and either *M. smithii* or *M. stadtmanae*.

So far, only one study assessed potential carriers of archaea in food, but yoghurt and milk have not been assessed herein [26]. The present study suggests an association between dietary factors and archaeal colonization, as organic yogurt and milk were significantly associated with the presence of *M. smithii*. Due to the high correlation between organic milk and yoghurt intake, we were however not able to disentangle whether only organic milk or organic yoghurt or both were truly responsible for the association with *M. smithii* during multivariable analyses. Cheese was however not associated with *M. smithii* colonization, which is in line with previous findings where no *M. smithii* or *M. stadtmanae* were detected in different types of cheese [26]. Analyses on the number of dairy products consumed, used to address multicollinearity, indicated that the consumption of multiple different dairy products might be associated with an increased *M. smithii* presence.

A well-known carrier of *M. smithii* and *M. stadtmanae* is the rumen of beef and cows [25, 37]. Therefore, it is likely that products derived from cows, such as dairy products, may contain some of these taxa, which was reflected in our results. Moreover, these specific methanogenic archaea have also been found in soil which could be the route of origin to cows as revealed from the Integrated Microbial Next Generation Sequencing database [38]. The archaeal presence in soil might be the underlying reason that especially organic products were associated with colonization. Organic cows have more outdoor access compared to conventionally farmed cows and could therefore have increased archaea uptake. Additionally, drugs for organic cows are less prescribed whereas conventional cows more often receive antibiotics to prevent microbial infections [39]. Methanogens are susceptible to antibiotics such as bacitracin, a commonly used antibiotic in cattle, which might eliminate them [40]. Although the screening of multiple dairy products indicated the presence of *M. smithii* in both conventional and organic dairy products, the latter could alternatively explain the lack of association of consumption of conventional dairy products with *M. smithii* colonization within our cohort study.

As the culture-independent techniques applied in the present study do not distinguish viable microbes from cell-free DNA originating from lysed microorganisms, we cannot completely exclude that dairy products only contain archaeal DNA instead of viable archaea. However, we used the concentrated pellet of the microbial cells for downstream analyses of milk products, thereby minimizing the detection of circulating cell-free DNA. It is therefore likely that a living fraction of archaeal cells has been measured.

For milk, the typical neutral pH is viable for both *M. smithii* and *M. stadtmanae*,

which favor an optimal pH of 6.9-7.4 whereas yogurt has a more acidic environment [36, 41, 42], which might explain why we could not detect methanogenic archaea in yoghurt samples. In addition, methanogens have been discovered in a wide variety of extreme environments with temperatures until at least 100°C [43, 44]. The latter might indicate that these microbes could resist heat treatments and pasteurization performed for milk products. All in all, this strengthens the possibility that viable methanogenic archaea could indeed be present in milk products, even after thermal processing.

To get definite prove that milk products are a source of archaea, future studies using either culture-based methods or molecular methods that enable differentiation between intracellular and cell-free DNA, such as the use of propidium monoazide (PMA) as a membrane impermeable DNA intercalating dye [45], are warranted. Several studies already identified living bacteria with an implemented PMA-assay in both human feces and processed milk [46-49].

In conclusion, dairy products, in particular organic milk products, may play an influential role in the colonization of the gut by *M. smithii* in children. Moreover, *M. smithii* seems to be present in both raw and commonly consumed milk products. For the first time, a large variety of determinants have been explored in association with gut colonization by methanogenic archaea. Although more information is needed to confirm and unravel the mechanisms of archaeal colonization in more detail, it may provide new targets for prevention of diseases associated with the presence of methanogenic archaea.

---

## References

1. Kable, M.E., et al., *The Core and Seasonal Microbiota of Raw Bovine Milk in Tanker Trucks and the Impact of Transfer to a Milk Processing Facility*. MBio, 2016. 7(4).
2. van Best, N., et al., *On the origin of species: Factors shaping the establishment of infant's gut microbiota*. Birth Defects Res C Embryo Today, 2015. 105(4): p. 240-51.
3. Horz, H.P., *Archaeal Lineages within the Human Microbiome: Absent, Rare or Elusive?* Life (Basel), 2015. 5(2): p. 1333-45.
4. Miller, T.L. and M.J. Wolin, *Methanosphaera stadtmaniae* gen. nov., sp. nov.: a species that forms methane by reducing methanol with hydrogen. Arch Microbiol, 1985. 141(2): p. 116-22.
5. Miller, T.L., et al., *Isolation of Methanobrevibacter smithii from human feces*. Appl Environ Microbiol, 1982. 43(1): p. 227-32.
6. Dridi, B., et al., *Methanomassiliococcus luminyensis* gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. Int J Syst Evol Microbiol, 2012. 62(Pt 8): p. 1902-7.
7. Khelaifia, S., et al., *Draft genome sequence of a human-associated isolate of Haloferax alexandrinus strain Arc-hr, an extremely halophilic archaea*. New Microbes New Infect, 2017. 15: p. 44-45.
8. Khelaifia, S. and D. Raoult, *Haloferax massiliensis* sp. nov., the first human-associated halophilic archaea. New Microbes New Infect, 2016. 12: p. 96-8.
9. Khelaifia, S., D. Raoult, and M. Drancourt, *A versatile medium for cultivating methanogenic archaea*. PLoS One, 2013. 8(4): p. e61563.
10. Dridi, B., et al., *Age-related prevalence of Methanomassiliococcus luminyensis in the human gut microbiome*. APMIS, 2012. 120(10): p. 773-7.
11. Roccarina, D., et al., *The role of methane in intestinal diseases*. Am J Gastroenterol, 2010. 105(6): p. 1250-6.
12. Gaci, N., et al., *Archaea and the human gut: new beginning of an old story*. World J Gastroenterol, 2014. 20(43): p. 16062-78.
13. Horz, H.P. and G. Conrads, *The discussion goes on: What is the role of Euryarchaeota in humans?* Archaea, 2010. 2010: p. 967271.
14. Conway de Macario, E. and A.J. Macario, *Methanogenic archaea in health and disease: a novel paradigm of microbial pathogenesis*. Int J Med Microbiol, 2009. 299(2): p. 99-108.
15. Cavicchioli, R., et al., *Pathogenic archaea: do they exist?* Bioessays, 2003. 25(11): p. 1119-28.
16. Nakamura, N., et al., *Mechanisms of microbial hydrogen disposal in the human colon and implications for health and disease*. Annu Rev Food Sci Technol, 2010. 1: p. 363-95.
17. Gill, E.E. and F.S. Brinkman, *The proportional lack of archaeal pathogens: Do viruses/phages hold the key?* Bioessays, 2011. 33(4): p. 248-54.
18. Haines, A., et al., *Breath-methane in patients with cancer of the large bowel*. Lancet, 1977. 2(8036): p. 481-3.
19. Kim, G., et al., *Methanobrevibacter smithii is the predominant methanogen in patients with constipation-predominant IBS and methane on breath*. Dig Dis Sci, 2012. 57(12): p. 3213-8.
20. Mbakwa, C.A., et al., *Gut colonization with methanobrevibacter smithii is associated with childhood weight development*. Obesity (Silver Spring), 2015. 23(12): p. 2508-16.
21. Vandeputte, D., et al., *Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates*. Gut, 2016. 65(1): p. 57-62.

22. Triantafyllou, K., C. Chang, and M. Pimentel, *Methanogens, methane and gastrointestinal motility*. J Neurogastroenterol Motil, 2014. 20(1): p. 31-40.
23. Pimentel, M., et al., *Methane production during lactulose breath test is associated with gastrointestinal disease presentation*. Dig Dis Sci, 2003. 48(1): p. 86-92.
24. Blais Lecours, P., et al., *Increased prevalence of Methanosphaera stadtmanae in inflammatory bowel diseases*. PLoS One, 2014. 9(2): p. e87734.
25. Carberry, C.A., et al., *Rumen methanogenic genotypes differ in abundance according to host residual feed intake phenotype and diet type*. Appl Environ Microbiol, 2014. 80(2): p. 586-94.
26. Brusa, T., F. Ferrari, and E. Canzi, *Methanogenic bacteria: presence in foodstuffs*. J Basic Microbiol, 1998. 38(2): p. 79-84.
27. Kummeling, I., et al., *Etiology of atopy in infancy: the KOALA Birth Cohort Study*. Pediatr Allergy Immunol, 2005. 16(8): p. 679-84.
28. Zoetendal, E.G., et al., *Isolation of DNA from bacterial samples of the human gastrointestinal tract*. Nat Protoc, 2006. 1(2): p. 870-3.
29. Salonen, A., et al., *Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis*. J Microbiol Methods, 2010. 81(2): p. 127-34.
30. Bursac, Z., et al., *Purposeful selection of variables in logistic regression*. Source Code Biol Med, 2008. 3: p. 17.
31. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. Journal of the Royal Statistical Society. Series B (Methodological), 1995: p. 289-300.
32. Caporaso, J.G., et al., *QIIME allows analysis of high-throughput community sequencing data*. Nat Methods, 2010. 7(5): p. 335-6.
33. Quigley, L., et al., *A comparison of methods used to extract bacterial DNA from raw milk and raw milk cheese*. J Appl Microbiol, 2012. 113(1): p. 96-105.
34. Florin, T.H., et al., *Shared and unique environmental factors determine the ecology of methanogens in humans and rats*. Am J Gastroenterol, 2000. 95(10): p. 2872-9.
35. Palmer, C., et al., *Development of the human infant intestinal microbiota*. PLoS Biol, 2007. 5(7): p. e177.
36. Dridi, B., D. Raoult, and M. Drancourt, *Archaea as emerging organisms in complex human microbiomes*. Anaerobe, 2011. 17(2): p. 56-63.
37. Cersosimo, L.M., et al., *Influence of periparturient and postpartum diets on rumen methanogen communities in three breeds of primiparous dairy cows*. BMC Microbiol, 2016. 16(1): p. 78.
38. Lagkouvardos, I., et al., *IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies*. Sci Rep, 2016. 6: p. 33721.
39. Zwald, A.G., et al., *Management practices and reported antimicrobial usage on conventional and organic dairy farms*. J Dairy Sci, 2004. 87(1): p. 191-201.
40. Dridi, B., et al., *The antimicrobial resistance pattern of cultured human methanogens reflects the unique phylogenetic position of archaea*. J Antimicrob Chemother, 2011. 66(9): p. 2038-44.
41. Miller, T.L. and C. Lin, *Description of Methanobrevibacter gottschalkii sp. nov., Methanobrevibacter thaueri sp. nov., Methanobrevibacter woesei sp. nov. and Methanobrevibacter wolinii sp. nov.* Int J Syst Evol Microbiol, 2002. 52(Pt 3): p. 819-22.
42. Ledenbach, L.H. and R.T. Marshall, *Microbiological Spoilage of Dairy Products*, in *Compendium of the Microbiological Spoilage of Foods and Beverages*, W.H. Sperber and M.P. Doyle, Editors. 2009, Springer New



- York: New York, NY. p. 41-67.
43. Tung, H.C., N.E. Bramall, and P.B. Price, *Microbial origin of excess methane in glacial ice and implications for life on Mars*. Proc Natl Acad Sci U S A, 2005. 102(51): p. 18292-6.
  44. Elias, D.A., et al., *Estimation of methanogen biomass by quantitation of coenzyme M*. Appl Environ Microbiol, 1999. 65(12): p. 5541-5.
  45. Janssen, K.J., et al., *Viability-PCR Shows That NAAT Detects a High Proportion of DNA from Non-Viable Chlamydia trachomatis*. PLoS One, 2016. 11(11): p. e0165920.
  46. Cangelosi, G.A. and J.S. Meschke, *Dead or alive: molecular assessment of microbial viability*. Appl Environ Microbiol, 2014. 80(19): p. 5884-91.
  47. Bae, S. and S. Wuertz, *Discrimination of viable and dead fecal Bacteroidales bacteria by quantitative PCR with propidium monoazide*. Appl Environ Microbiol, 2009. 75(9): p. 2940-4.
  48. Quigley, L., et al., *The microbial content of raw and pasteurized cow milk as determined by molecular approaches*. J Dairy Sci, 2013. 96(8): p. 4928-37.
  49. Soejima, T., J. Minami, and K. Iwatsuki, *Rapid propidium monoazide PCR assay for the exclusive detection of viable Enterobacteriaceae cells in pasteurized milk*. J Dairy Sci, 2012. 95(7): p. 3634-42.

## Supplementary Material

**Supplementary Table 1.** Absolute counts of *M. smithii* and *M. stadmanae* in dairy whole/full-fat products (n=2-3). \*log<sub>10</sub> DNA copies/ml, average. Replicates (*n*) denotes repeated isolation and quantification of the product. Average counts were calculated from positive samples only.

Dairy Product	M.smithii counts*	M.stadmanae counts*	Thermal Processing	Expiration (days)	Origin
Biodynamic milk	2.63	-	Pasteurised	4	Zuiver Zuivel
Organic milk	2.88	-	Pasteurised	5	Weerribben Zuivel
Regular milk	2.94	-	Pasteurised	1-5	Campina & Plus
Raw milk	3.73	-	None	-	Local Farmer
Organic yogurt	-	-	Pasteurised	14	Weerribben Zuivel
Biodynamic yogurt	-	-	Pasteurised	15	Zuiver Zuivel

**Supplementary Table 2.** Patient characteristics of the KOALA birth cohort study and the population under study

<b>Determinants</b>	<b>KOALA Birth Cohort Study (n = 2834)</b>	<b>Study Population (n = 472)</b>
<b>Continuous variables</b>	<b>Median (range)</b>	<b>Median (range)</b>
<b>General characteristics</b>		
<i>Age at fecal sampling (years)<sup>b</sup></i>	-	7.2 (6.0 – 10.0)
<i>Birthweight (g)</i>	3500 (780 – 5455)	3570 (2140 – 5000)
<i>Gestational age (weeks)</i>	40 (26 – 43)	40 (35 – 42)
<i>Breastfeeding (months)</i>	4 (0 – 13)	5 (0 – 13)
<i>BMI (z-score)<sup>b</sup></i>	-	-0.2 (-2.7 – 2.8)
<b>Dietary intake (per day)</b>		
<i>Total energy intake (kcal)</i>	1437.5 (740.0 – 2844.1)	1419.5 (788.3 – 2415.2)
<i>Carbohydrate intake (en%)</i>	55.7 (37.1 – 74.5)	55.7 (42.3 – 69.5)
<i>Protein intake (en%)</i>	14.6 (8.0 – 21.9)	14.5 (8.8 – 20.4)
<i>Animal protein intake (en%)</i>	9.0 (1.6 – 16.3)	9.0 (2.2 – 15.3)
<i>Fiber intake (g)</i>	15.1 (5.4 – 31.4)	15.4 (6.1 – 28.1)
<b>Categorical variables</b>	<b>N (%)<sup>a</sup></b>	<b>N (%)<sup>a</sup></b>
<b>Gender</b>		
<i>Male</i>	1451 (51.3)	244 (51.7)
<i>Female</i>	1376 (48.7)	228 (48.3)
<b>Recruitment group</b>		
<i>Conventional</i>	2343 (82.7)	354 (75.2)
<i>Alternative</i>	491 (17.3)	117 (24.8)
<b>Maternal education<sup>c</sup></b>		
<i>Lower education</i>	289 (10.7)	34 (7.5)
<i>Medium education</i>	1060 (39.4)	171 (37.7)
<i>Higher education</i>	1341 (49.9)	248 (54.7)
<b>Older siblings (at 1 year of age)</b>		
<i>None</i>	1193 (44.1)	187 (39.7)
<i>One</i>	1110 (41.0)	206 (43.7)
<i>Two or more</i>	404 (14.9)	78 (16.6)
<b>Younger and older siblings (at 4-5 years of age)</b>		
<i>None</i>	199 (9.7)	46 (10.3)
<i>One</i>	1162 (56.8)	245 (54.9)
<i>Two or more</i>	685 (33.5)	155 (34.8)
<b>Regular visits to daycare or stay at a host parent (during the first 2 years of life)</b>		
<i>No</i>	960 (35.2)	161 (34.3)
<i>Host parent</i>	680 (24.9)	122 (26.0)
<i>Daycare or daycare and host parent</i>	1086 (39.8)	187 (39.8)
<b>Place and mode of delivery</b>		
<i>Natural birth at home</i>	1187 (44.9)	219 (47.1)
<i>Natural/artificial birth at hospital</i>	1147 (43.4)	198 (42.6)
<i>Caesarean section</i>	311 (11.8)	48 (10.3)
<b>First exposure to antibiotics (during first 2 years)</b>		
<i>Never</i>	1074 (43.5)	212 (46.7)
<i>At 0-7 months</i>	562 (22.8)	92 (20.3)
<i>At 8-12 months</i>	404 (16.4)	75 (16.5)
<i>At 13-24 months</i>	430 (17.4)	75 (16.5)
<b>Use of antibiotic treatments (1 year before fecal sampling)<sup>b</sup></b>		
<i>No</i>	-	396 (85.2)
<i>Yes</i>	-	69 (14.8)
<b>Duration of pacifier use by the child<sup>d</sup></b>		
<i>Never</i>	684 (27.5)	127 (27.6)
<i>Early</i>	587 (23.6)	108 (23.5)
<i>Late</i>	1214 (48.9)	225 (48.9)

<b>Determinants</b>	<b>KOALA Birth Cohort Study (n = 2834)</b>	<b>Study Population (n = 472)</b>
<b>Continuous variables</b>	<b>Median (range)</b>	<b>Median (range)</b>
Cleaning of pacifier in boiling water		
<i>Never/Sometimes</i>	458 (22.0)	84 (22.3)
<i>Always</i>	1624 (78.0)	292 (77.7)
Presence of pets at 6 to 7 years of life.		
<i>None</i>	928 (47.0)	217 (46.7)
<i>Dog</i>	227 (11.5)	44 (9.5)
<i>Cat</i>	242 (12.3)	58 (12.5)
<i>Other</i>	338 (17.1)	90 (19.4)
<i>Combination</i>	236 (12.0)	56 (12.0)
Monthly exposure to farm (animals) at 6-7 years		
<i>No</i>	1309 (72.7)	303 (72.8)
<i>Yes</i>	491 (27.3)	113 (27.2)
Diet of mother		
<i>Conventional</i>	1885 (67.1)	294 (62.8)
<i>Organic (incl. biodynamic)</i>	923 (32.9)	174 (37.2)
Diet of child:		
<i>Conventional (≤25% organic)</i>	1554 (81.4)	360 (80.2)
<i>Organic (incl. biodynamic) (&gt;25% organic)</i>	355 (18.6)	89 (19.8)
Vitamin D/AD supplementation during the first 2 years of life		
<i>No</i>	466 (18.8)	96 (20.6)
<i>Yes</i>	2011 (81.2)	370 (79.4)
Vegetarian diet of child		
<i>No</i>	1943 (96.0)	419 (94.4)
<i>Yes</i>	82 (4.0)	25 (5.6)
Regular intake of milk		
<i>No</i>	490 (25.3)	107 (26.5)
<i>Yes</i>	1445 (75.3)	347 (73.5)
Regular intake of organic milk		
<i>No</i>	641 (75.3)	152 (87.5)
<i>Yes</i>	210 (24.7)	59 (12.5)
Regular intake of yogurt		
<i>No</i>	464 (24.4)	110 (28.4)
<i>Yes</i>	1438 (75.6)	338 (71.6)
Regular intake of organic yogurt		
<i>No</i>	619 (71.5)	157 (87.5)
<i>Yes</i>	247 (28.5)	59 (12.5)
Regular intake of cheese		
<i>No</i>	510 (27.1)	114 (30.1)
<i>Yes</i>	1370 (72.9)	330 (69.9)
Regular intake of organic cheese		
<i>No</i>	740 (87.2)	186 (94.1)
<i>Yes</i>	109 (12.8)	28 (5.9)
Regular intake of organic fruit		
<i>No</i>	1738 (85.5)	376 (84.5)
<i>Yes</i>	295 (14.5)	69 (15.5)
Regular intake of organic raw vegetables:		
<i>No</i>	650 (74.8)	155 (71.4)
<i>Yes</i>	219 (25.2)	62 (28.6)

<sup>a</sup> Totals may not add up to the total group size due to missing values; <sup>b</sup> Data unavailable for the complete cohort; <sup>c</sup> Lower: primary school, preparatory vocational or lower general secondary school. Middle: vocational, higher general secondary or pre-university. High: higher vocational or academic education <sup>d</sup> Never: Child did not use a pacifier.

**Supplementary Table 3.** Multivariable logistic regression model showing associations between potential determinants and *Methanobrevibacter smithii* presence.

<b>M. smithii presence</b>			
<b>Main effects model (N=419)</b>			
<b>Determinant</b>	<b>OR (CI95%)</b>	<b>FDR crit.<sup>a</sup></b>	<b>p-value</b>
Regular intake of organic milk <sup>b,c</sup>			
No	Ref.		
Yes	5.73 (1.92; 17.03)	0.010	<b>0.002<sup>d</sup></b>
Regular intake of organic yogurt <sup>b,c</sup>			
No	Ref.		
Yes	4.40 (1.60; 12.08)	0.010	<b>0.004<sup>d</sup></b>
Diet of child <sup>e</sup>			
Conventional (≤25% organic)	Ref.		
Organic (incl. biodynamic) (>25% organic)	0.37 (0.18; 0.78)	0.020	<b>0.009<sup>d</sup></b>
Fiber intake (g) <sup>e</sup>	0.96 (0.90; 1.03)	0.040	0.256
Age at fecal sampling (years) <sup>e</sup>	0.83 (0.64; 1.09)	0.030	0.186

<sup>a</sup> Critical FDR cut-off level as determined by Benjamini-Hochberg procedure; <sup>b</sup> Missing value category omitted (included in FDR correction); <sup>c</sup> Due to multicollinearity both regular intake of organic milk and regular intake of organic yogurt are reported. Both were run in a separate model; <sup>d</sup> significant association after correction for FDR by Benjamini-Hochberg procedure; <sup>e</sup> Parameter estimates presented for the model containing organic milk intake. Parameter estimates did not change substantially when the variable 'organic milk intake' was replaced by the variable 'organic yoghurt intake' (data not shown).

**Supplementary Table 4.** Secondary multivariable logistic regression model showing associations between potential determinants and the presence of *Methanobrevibacter smithii*.

<b>M. smithii presence</b>			
<b>Main effects model (N=406)</b>			
<b>Determinant</b>	<b>OR (CI95%)</b>	<b>FDR crit.<sup>a</sup></b>	<b>p-value</b>
First exposure to antibiotics (during first 2 years of life)			
Never	Ref.		
At 0-7 months	0.57 (0.31; 1.04)	0.030	0.065
At 8-12 months	1.20 (0.58; 2.48)	0.050	0.623
At 13-24 months	0.80 (0.40; 1.59)	0.040	0.519
Regular intake of organic products (cheese, milk and yogurt)			
Trend (0, 1, 2, 3)	2.10 (1.32; 3.33)	0.010	<b>0.002<sup>b</sup></b>
Diet of child			
Conventional (≤25% organic)	Ref.		
Organic (incl. biodynamic) (>25% organic)	0.32 (0.16; 0.67)	0.020	<b>0.003<sup>b</sup></b>

<sup>a</sup> Critical FDR cut-off level as determined by Benjamini-Hochberg procedure; <sup>b</sup> Significant association after correction for FDR by Benjamini-Hochberg procedure.

**Supplementary Table 5.** Multivariable linear regression model showing associations between potential determinants and the presence of *Methanobrevibacter smithii* abundance.

<b>M. smithii abundance (log<sub>10</sub> DNA copies per gram feces)</b>			
<b>Main effects model (N=431)</b>			
<b>Determinant</b>	<b>β (CI95%)</b>	<b>FDR crit.<sup>a</sup></b>	<b>p-value</b>
Younger and older siblings (at 4-5 years of age)			
<i>None</i>	Ref.		
<i>One</i>	-0.88 (-1.91; 0.15)	0.029	0.094
<i>Two or more</i>	-0.99 (-2.07; 0.09)	0.021	0.072
First exposure to antibiotics (during first 2 years of life)			
<i>Never</i>	Ref.		
<i>At 0-7 months</i>	-0.81 (-1.65; 0.02)	0.017	0.056
<i>At 8-12 months</i>	0.16 (-0.71; 1.03)	0.042	0.717
<i>At 13-24 months</i>	0.27 (-0.61; 1.14)	0.038	0.551
Monthly exposure to farm (animals) at 6-7 years of age <sup>b</sup>			
<i>No</i>	Ref.		
<i>Yes</i>	-0.50 (-1.26; 0.25)	0.033	0.192
Presence of pets at 6 to 7 years of life			
<i>None</i>	Ref.		
<i>Dog</i>	0.12 (-0.99; 1.23)	0.050	0.829
<i>Cat</i>	0.14 (-0.84; 1.12)	0.046	0.774
<i>Other</i>	1.03 (0.20; 1.86)	0.004	0.016
<i>Combination</i>	0.91 (-0.10; 1.91)	0.025	0.076
<i>Breastfeeding (months)</i>	0.07 (0.00; 0.15)	0.013	0.051

<sup>a</sup> Critical FDR cut-off level as determined by Benjamini-Hochberg procedure; <sup>b</sup> Missing value category omitted (included in FDR correction)

**Supplementary Table 6.** Final multivariable linear regression model showing associations between potential determinants and abundance of *Methanobrevibacter smithii*.

<b>M. smithii abundance (log<sub>10</sub> DNA copies per gram feces)</b>			
<b>Adjusted main effects model (N=431)<sup>a</sup></b>			
<b>Determinant</b>	<b>β (CI95%)</b>	<b>FDR crit.<sup>b</sup></b>	<b>p-value</b>
Younger and older siblings (at 4-5 years of age)			
<i>None</i>	Ref.		
<i>One</i>	-0.89 (-1.94; 0.16)	0.021	0.095
<i>Two or more</i>	-1.04 (-2.14; 0.06)	0.011	0.065
First exposure to antibiotics (during first 2 years of life)			
<i>Never</i>	Ref.		
<i>At 0-7 months</i>	-0.81 (-1.65; 0.04)	0.009	0.062
<i>At 8-12 months</i>	0.17 (-0.70; 1.05)	0.035	0.697
<i>At 13-24 months</i>	0.26 (-0.63; 1.14)	0.029	0.568
Monthly exposure to farm (animals) at 6-7 years of age <sup>c</sup>			
<i>No</i>	Ref.		
<i>Yes</i>	-0.50 (-1.26; 0.27)	0.024	0.200
Presence of pets at 6 to 7 years of life			
<i>None</i>	Ref.		
<i>Dog</i>	0.13 (-0.99; 1.25)	0.044	0.821
<i>Cat</i>	0.15 (-0.84; 1.14)	0.041	0.767
<i>Other</i>	1.04 (0.20; 1.88)	0.003	0.015
<i>Combination</i>	0.94 (-0.08; 1.96)	0.018	0.070
<i>Breastfeeding (months)</i>	0.08 (-0.01; 0.15)	0.015	0.065

<sup>a</sup> Model adjusted for: age at fecal sampling (years), gender (male/female), recruitment group (conventional/alternative), total energy intake (kcal) and BMI (z-score); <sup>b</sup> Critical FDR cut-off level as determined by Benjamini-Hochberg procedure; <sup>c</sup> Missing value category omitted (included in FDR correction)

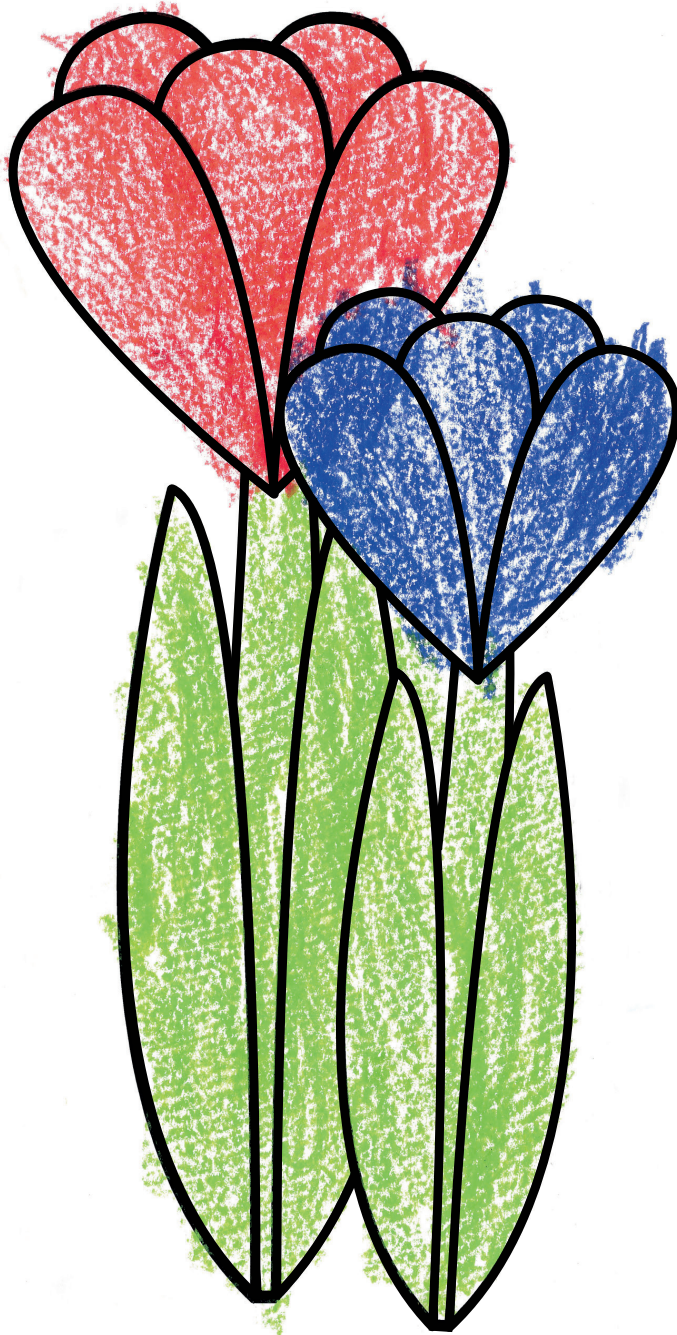
**Supplementary Table 7.** Main logistic regression model showing associations between potential determinants and presence of *Methanospira stadmanae*.

<b>M. stadmanae presence</b>			
<b>Main effects model (N=420)</b>			
<b>Determinant</b>	<b>OR (CI95%)</b>	<b>FDR crit.<sup>a</sup></b>	<b>p-value</b>
<i>Birthweight (g)</i> <sup>b</sup>	1.001 (1.000; 1.002)	0.027	0.061
Place and mode of delivery <sup>b</sup>			
<i>Natural birth at home</i>	Ref.		
<i>Natural/artificial birth at hospital</i>	1.49 (0.64; 3.46)	0.041	0.350
<i>Caesarean section</i>	5.14 (1.64; 16.07)	0.005	0.005
First exposure to antibiotics (during first 2 years of life) <sup>b</sup>			
<i>Never</i>	Ref.		
<i>At 0-7 months</i>	0.60 (0.19; 1.95)	0.046	0.395
<i>At 8-12 months</i>	0.70 (0.22; 2.28)	0.050	0.552
<i>At 13-24 months</i>	2.72 (1.13; 6.57)	0.023	0.026
Regular intake of organic milk <sup>c, d</sup>			
<i>No</i>	Ref.		
<i>Yes</i>	2.06 (0.66; 0.96)	0.029	0.214
Regular intake of organic yogurt <sup>c, d</sup>			
<i>No</i>	Ref.		
<i>Yes</i>	1.69 (0.49; 5.86)	-	0.680
<i>Breastfeeding (months)</i> <sup>b</sup>	1.06 (0.96; 1.16)	0.036	0.237
Regular intake of organic fruit <sup>b</sup>			
<i>No</i>	Ref.		
<i>Yes</i>	3.06 (1.17; 8.03)	0.018	0.023
<i>Animal protein intake (en%)</i> <sup>b</sup>	0.81 (0.69; 0.96)	0.009	0.014

<sup>a</sup> Critical FDR cut-off level as determined by Benjamini-Hochberg procedure; <sup>b</sup> Parameter estimates presented for the model containing organic milk intake. Parameter estimates did not change substantially when the variable 'organic milk intake' was replaced by the variable 'organic yoghurt intake' (data not shown); <sup>c</sup> Missing value category omitted (included in FDR correction); <sup>d</sup> Due to multicollinearity both regular intake of organic milk and regular intake of organic yogurt were included in separate models, which included all other variables as listed in this table.







Paul Hanl (5 years)

# Chapter 5

**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**

**Niels van Best\***, Gianluca Galazzo\*, Liene Bervoets, Isaac Oteng Dapaah, Paul H. Savelkoul, Mathias W. Hornef, the GI-MDH consortium, Susanne Lau‡, Eckard Hamelmann‡, John Penders‡

\*Equal contribution / ‡Equal contribution

*Gastroenterology. 2020 May;158(6):1584-1596*

## Abstract

Establishment of the gastrointestinal microbiota during infancy affects immune system development and oral tolerance induction. Perturbations in the microbiome during this period can contribute to development of immune-mediated diseases. We monitored microbiota maturation and associations with subsequent development of allergies in infants and children.

We collected 1453 stool samples, at 5, 13, 21 and 31 weeks post-partum (infants), and once at school-age (6–11 years), from 440 children (49.3% girls, 24.8% born by caesarean section; all children except for 6 were breastfed for varying durations; median 40 weeks; interquartile range, 30–53 weeks). Microbiota were analyzed by amplicon sequencing. Children were followed through 3 years of age for development of atopic dermatitis; data on allergic sensitization and asthma were collected when children were school age.

Diversity of fecal microbiota, assessed by Shannon index, did not differ significantly among children from 5 through 13 weeks after birth, but thereafter gradually increased to 21 and 31 weeks. Most bacteria within the Bacteroidetes and Proteobacteria phyla were already present at 5 weeks after birth, whereas many bacteria of the Firmicutes phylum were acquired at later times in infancy. At school age, many new Actinobacteria, Firmicutes, and Bacteroidetes bacterial taxa emerged. The largest increase in microbial diversity occurred after 31 weeks. Vaginal, compared with cesarean section delivery, was most strongly associated with an enrichment of *Bacteroides* species at 5 weeks through 31 weeks. From 13 weeks onwards, diet became most important determinant of microbiota composition—cessation of breastfeeding, rather than solid food introduction, was associated with changes. For example, Bifidobacteria, staphylococci, and streptococci significantly decreased upon cessation of breastfeeding, whereas bacteria within the Lachnospiraceae family (*Pseudobutyrvibrio*, *Lachnobacterium*, *Roseburia*, and *Blautia*) increased. When we adjusted for confounding factors, we found fecal microbiota composition to be associated with development of atopic dermatitis, allergic sensitization, and asthma. Members of the Lachnospiraceae family, as well as the genera *Faecalibacterium* and *Dialister*, were associated with a reduced risk of atopy.

In a longitudinal study of fecal microbiota of children from 5 weeks through 6–11 years, we tracked changes in diversity and composition associated with the development of allergies and asthma.

## Introduction

Colonization of the intestinal tract during the neonatal period is of crucial importance for the maturation of the mucosal immune system and the induction of oral tolerance [1-3]. Animal studies have provided compelling evidence to support a causal role of the intestinal microbiota and its metabolites, especially in early life, in the etiology of allergic diseases [3-6].

Numerous epidemiological studies [7-12] also suggest that the infant intestinal microbiota plays an important role in the manifestation of allergic diseases and asthma, although actual results vary considerably between studies. About half of the studies that examined intestinal microbial diversity in infancy and childhood reported a lower diversity among children with (subsequent) allergies, whereas the remaining studies found no evidence for such an association [13]. Moreover, despite many specific microbial taxa have been linked to allergies and asthma, it remains unclear which bacterial taxa prevent or promote disease onset [14].

Lack of early stool sampling and different ages of stool sample collection, different microbiological profiling methods, and an inadequate control for potential confounders have been suggested to contribute to the heterogeneity between study results [9, 13]. Additionally, cross-sectional studies are prone to reverse causality, i.e. changes in the microbiota composition as a result of the disease manifestation, and only very few studies have sufficient clinical follow-up to link infant microbiota maturation to the subsequent development of asthma [11].

Initial microbial colonization starts upon rupture of the amniotic membranes and subsequent passage through the birth canal when the infant is seeded by maternal microbial strains, a process which is impeded in case of a cesarean section delivery [15, 16]. Subsequently, microbial populations evolve as the diet changes and the host develops. Given the highly dynamic and complex process of microbial assembly, succession and maturation, repeated sampling is important to allow analysis of the overall development of the indigenous infant microbial ecology [17]. Moreover, many of the antenatal and postnatal factors that influence microbial community assembly during infancy, such as birth mode and the presence of older siblings and furry pets in the household, have also been associated with the development of allergic diseases and asthma [9, 13, 18, 19], highlighting the importance to account for potential confounding factors.

In the present study, we aimed to monitor microbial assembly, succession and maturation during the first year of life and identify hereditary, perinatal, environmental, lifestyle and dietary factors that drive microbiota development. Through the application

of various multivariable longitudinal models, including joint modelling, we next examined the dynamics in microbial diversity, composition and community structure in association with the subsequent risk of developing atopic dermatitis and asthma until to school-age.

Our findings indicate that alterations in microbial diversity and composition precede the onset of allergic manifestations, while emphasizing the importance of possible confounders.

## Results

### Design and clinical outcome measurements

Originally this study was designed as a randomized placebo-controlled clinical trial to examine the impact of a bacterial lysate, containing heat-killed *Escherichia coli* and *Enterococcus faecalis*, on the primary prevention of atopic dermatitis (AD) (registration no. ISRCTN60475069, ISRCTN registry) [20]. However, we did not find any evidence that the intervention affected the microbiota composition and therefore pooled both treatment arms in the downstream statistical analyses. Infants were clinically examined by a pediatrician on signs of AD at the ages of 1, 21 and 31 weeks and again at 1, 2 and 3 years of age as described previously [20]. The school-age follow-up of the study population (at 6-11 years) took place in 2013 including clinical examination, lung function testing, skin prick tests, and serum analyses of specific IgE to the most common aero-allergens (house dust mite, dog, cat, mold (*Alternaria*, *Cladosporium*), birch and grass pollen). Children were classified as having current asthma in case of a doctor's diagnosis in combination of any indicative symptoms in the last 12 months (wheezing, shortness of breath, nocturnal awakening due to shortness of breath and/or wheezing). Allergic sensitization was assessed by Skin Prick Test and serum sensitization for the above-mentioned allergens. The study and follow-up were approved by the hospitals local review board Charité Ethics Committee in 2002 and 2012. Parents and participants gave written informed consent.

### Microbial profiling of fecal samples

Fecal DNA was isolated by a combination of bead beating and column-based purification as described in detail previously [21]. Barcoded universal primers adapted from Bartram and colleagues [22] were used to amplify the variable 3 region of the 16S rRNA gene. Amplicons were sequenced using the Illumina MiSeq platform using 2x250 paired-end reads. The resulting sequencing data were processed using the short-read library 16S rRNA gene sequencing

pipeline (sl1p) [23] (for description see supplementary methods).

This resulted in a total of 93,475,612 reads from 1,468 samples that were clustered into 7,961 OTUs. Removal of OTUs that were observed in only a single sample and discarding OTUs with a fraction of the total number of sequences below 0.001, retained the majority of sequences (92,997,277) while significantly reducing the number of OTUs to 873. Finally, we eliminated 15 samples with a low coverage (<15,000 reads) and normalized the data using *Rhea* [24]. In order not to discard informative information, normalization in *Rhea* is performed by dividing OTU counts per sample for their total count (sample depth) and then multiplying the obtained relative abundance for the lowest sample depth (15,540 reads).

### Statistical Analysis

All the statistical analyses were performed two-sided using R, version 3.4.3.

#### ***GI microbiota richness and maturation***

The Chao1 index, as measure for the estimated microbial richness, and the Shannon index, as microbial diversity metric, were computed using the R package *vegan 2.5.3* [25]. In order to compare the microbial community structure of samples, we used the unweighted UniFrac which incorporates phylogenetic distances between observed organisms. The R package *GUniFrac 1.1* [26] was used to compute the unweighted UniFrac metric. Ordination of samples based upon the unweighted Unifrac dissimilarity was visualized using Principal Coordinate Analysis (PCoA) using the *capscale* function of the *vegan* package. The Friedman test was used to test for significant differences in Chao1, Shannon and loadings scores on the principal coordinates between all time-points, followed by Dunn's test for post hoc pairwise comparisons between individual time-points. The p-values were finally FDR adjusted for multiple comparisons. In order to track the dynamics of individual OTUs within the four main bacterial phyla, we created Sankey plots using Sankeymatic. For readability of the Sankey plots, the OTU table was further filtered to OTUs that were present in  $\geq 10\%$  of the samples during one or more time-points. We next investigated the maturity of the GI microbiota by computing the microbial-by-age z-score (MAZ) of the sample as described previously [27]. Shortly, we started feeding a random forest with a training set made of the microbial community composition of healthy children after filtering out all OTUs with a prevalence below 5%. Once the model was trained, we used it to predict the age of all the samples. Finally, the z-score was computed using the following formula:

$$MAZ = \frac{\text{Microbial age} - \text{median of microbial age of healthy children of same chronologic age}}{(\text{standard deviation of microbial age of healthy children of the same chronologic age})}$$

Dirichlet Multinomial Mixture (DMM) clustering, an unsupervised clustering method that uses Laplace approximation to identify groups of communities (enterotypes) with similar composition, was performed as previously described [28]. We then analysed the transition of infants through these DMM-clusters with age [29].

### ***Analysis of factors shaping the GI microbiota***

We examined which hereditary, perinatal, environmental, lifestyle and dietary factors were associated with the establishment of the microbiome during infancy (see supplementary methods for detailed description). In order to examine which of these factors were associated with the DMM clusters at baseline and/or with the transition of DMM clusters between the age of 5 and 31 weeks, multinomial logistic regression analysis were used. Only factors that were significantly associated with the (transition of) DMM clusters in the univariable analyses, were included in the final multivariable model. We next used multivariate association with linear models (MaAsLin) [30] to examine the association between these factors with and individual microbial taxa and multivariable linear regression models to examine the association with the microbial diversity and maturity. The effect size and significance of each of the covariates on the microbial community structure was determined using the *envfit* function in *vegan* [25]. Ordination was performed using the PCoA based on unweighted UniFrac metric obtained as described above. The significance value was determined based on 999 permutations. All *P* values derived from *envfit* were adjusted for multiple comparisons using FDR adjustment (Benjamini–Hochberg procedure). In order to understand which of the covariates had the strongest impact on the overall microbial community structure, we performed a Permutational Analysis of the Variance (PERMANOVA) based on unweighted UniFrac. Only covariates that were statistically significant in the *envfit* analyses were included in the PERMANOVA.

### ***Analysis on association between microbiota and allergic manifestations***

To examine how the longitudinal variation of the microbial diversity (Shannon index) and maturity (MAZ) of the GI microbiota affects the time to development of AD, we applied a joint model [31] using the JM function of the *JM* package [32] (for details see supplementary methods). To examine the impact of microbial diversity and maturity on asthma and allergic

sensitization at school-age, a Generalized Linear Model (GLM) was built using *lme4* 1.1.19. The same covariates as included in the JM were incorporated as potential confounding factors. Since both asthma and sensitization were binary outcomes, a binomial distribution was chosen for the GLM. To identify if specific bacterial genera were differentially abundant in children with and without allergic manifestations, we used the MetaLonDa package [33]. To ensure meaningful p-values we performed 999 permutations. To select only the significant associations, we choose a threshold of 0.05 for the p-values after FDR adjustment.

## Results

### Study population characteristics

The study, initially designed as a randomized, placebo-controlled trial on the primary prevention of AD by an orally applied lysate of heat-killed *Escherichia coli* and *Enterococcus faecalis*, consisted of healthy newborns (n = 606) with a single or double heredity for atopy [20]. During the first 3 years of life, children were deeply phenotyped by physical examination and the collection of detailed questionnaires at 7 clinical visits. At school-age, children were contacted again to determine the establishment of asthma and allergic sensitization.

For the present study, only children with at least three fecal samples collected during the first year and/or feces collected at school-age were included (n = 440). Of these children, 217 (49.3%) were girls, 187 (42.5%) had older siblings, 109 (24.8%) were born by caesarean section and 29 (6.6%) were reportedly treated with antibiotics in the first 31 weeks of life. All except 6 children received breastfeeding, although the duration of breastfeeding varied considerably with a median duration of 40 weeks [IQR: 30-53]. Solid food was introduced at a median age of 25 weeks [IQR: 22 – 27] (Supplementary Table S1).

### Development of the microbiota between early infancy and school-age

We first examined the compositional changes in the microbiota during infancy and compared this to the school-age microbiota composition. Samples collected at the age of 5 (n = 306), 13 (n = 287), 21 (n = 268) and 31 (n = 307) weeks post-partum and again at school-age (n = 300) were profiled by amplicon sequencing of the 16S rRNA hypervariable V3 gene region. Upon quality filtering and removal of samples with low sequencing depth (n = 15), 1,453 samples with a median sequencing depth of 62,420 reads/sample (range 15,540 – 168,848) were retained for downstream analysis and clustered into 873 operational taxonomic units (OTUs).

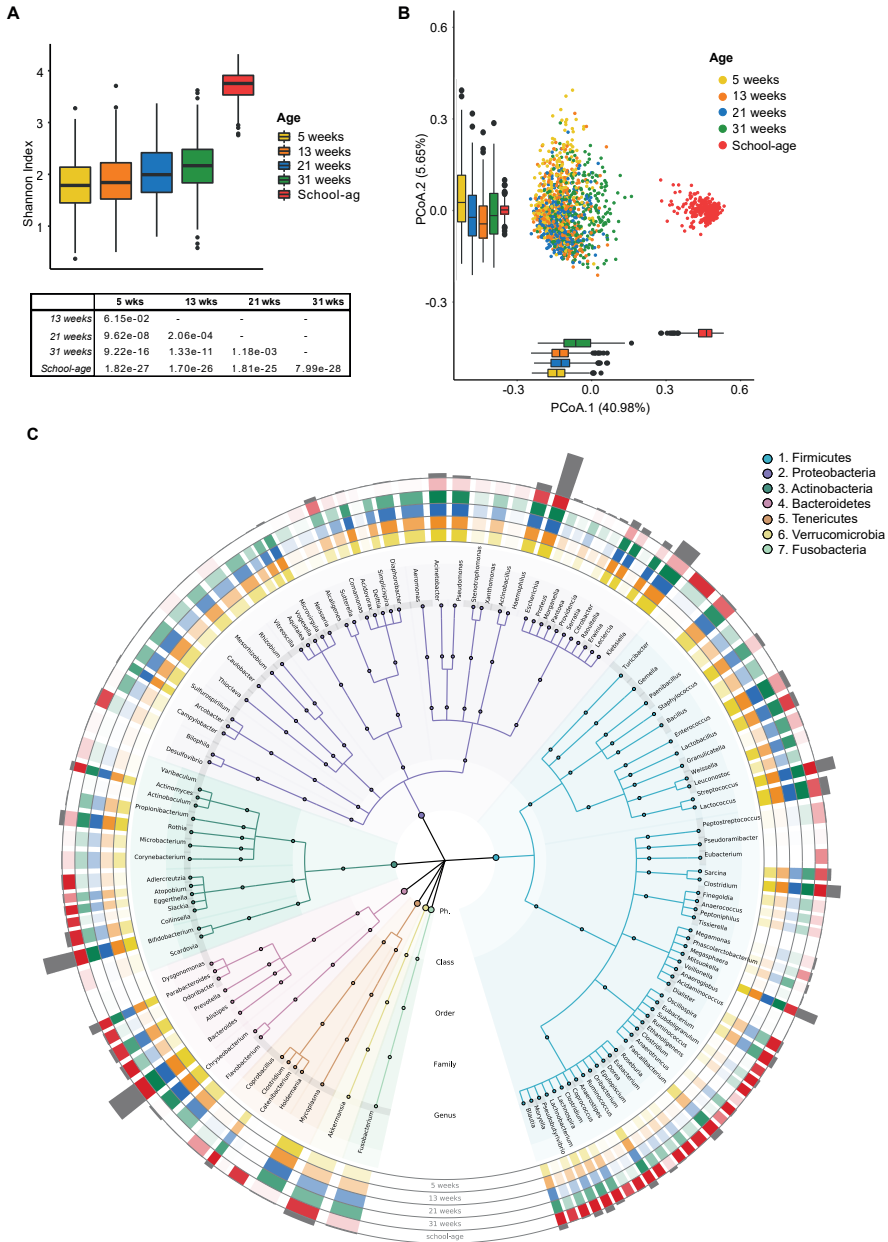


Microbial diversity, assessed by the Shannon index, was not significantly different between ages 5 and 13 weeks but thereafter gradually increased from 13 to 21 and 31 weeks after birth (Fig. 1A). The largest increase in microbial diversity occurred after the age of 31 weeks as indicated by the steep increase in the Shannon index at school-age ( $p = 7.99 \times 10^{-28}$ ). Similar findings were observed for the microbial richness as assessed by the Chao1 (Supplementary Table S2).

Principle coordinate analyses indicated that the microbial community structure as assessed by the unweighted Unifrac metric also gradually shifted during infancy with the most prominent shift between the age of 21 and 31 weeks (Fig 1B,  $p = 1.58 \times 10^{-27}$ , Supplementary Table S3). The school-age samples, however, clustered separately and showed less inter-individual variation as compared to the infant samples.

Tracking individual OTUs based on their presence or absence revealed different dynamics within the dominant phyla (Supplementary Fig 1). The majority of OTUs within the phyla of Actinobacteria, Bacteroidetes and Proteobacteria found during infancy were already present at 5 weeks after birth, whereas almost half of the OTUs within the phylum of Firmicutes were only acquired at later infant time-points. At school-age, many new Actinobacteria, Firmicutes, and Bacteroidetes OTUs emerged on top of the OTUs already present during infancy. In contrast, only few new OTUs emerged within the phylum Proteobacteria at school-age, while a significant portion of the infant OTUs were lost thereafter.

We next examined the bacterial genera that contribute most to the temporal dynamics in microbial diversity and community structure. Towards school-age the prevalence in many of the genera within the phylum of Proteobacteria dramatically decreased, whereas the prevalence of genera within the phylum of Firmicutes, and in particular within the Lachnospiraceae and Ruminococcaceae families strongly increased (Fig 1C). Moreover, with the exception of *Bifidobacterium*, the relative abundance of all of the major bacterial genera changed significantly (Friedman test, all P-values < 0.001, Supplementary Table 4) throughout infancy and childhood (Supplementary Fig. 2A-B). *Escherichia* was the most abundant genus at 5 weeks of age followed by *Bifidobacterium* and *Streptococcus*. *Escherichia* still remained the most abundant genus at 31 weeks of age but was now followed by *Bacteroides* and *Veillonella*. At school-age the most abundant genera were *Blautia*, *Faecalibacterium* and *Ruminococcus*.

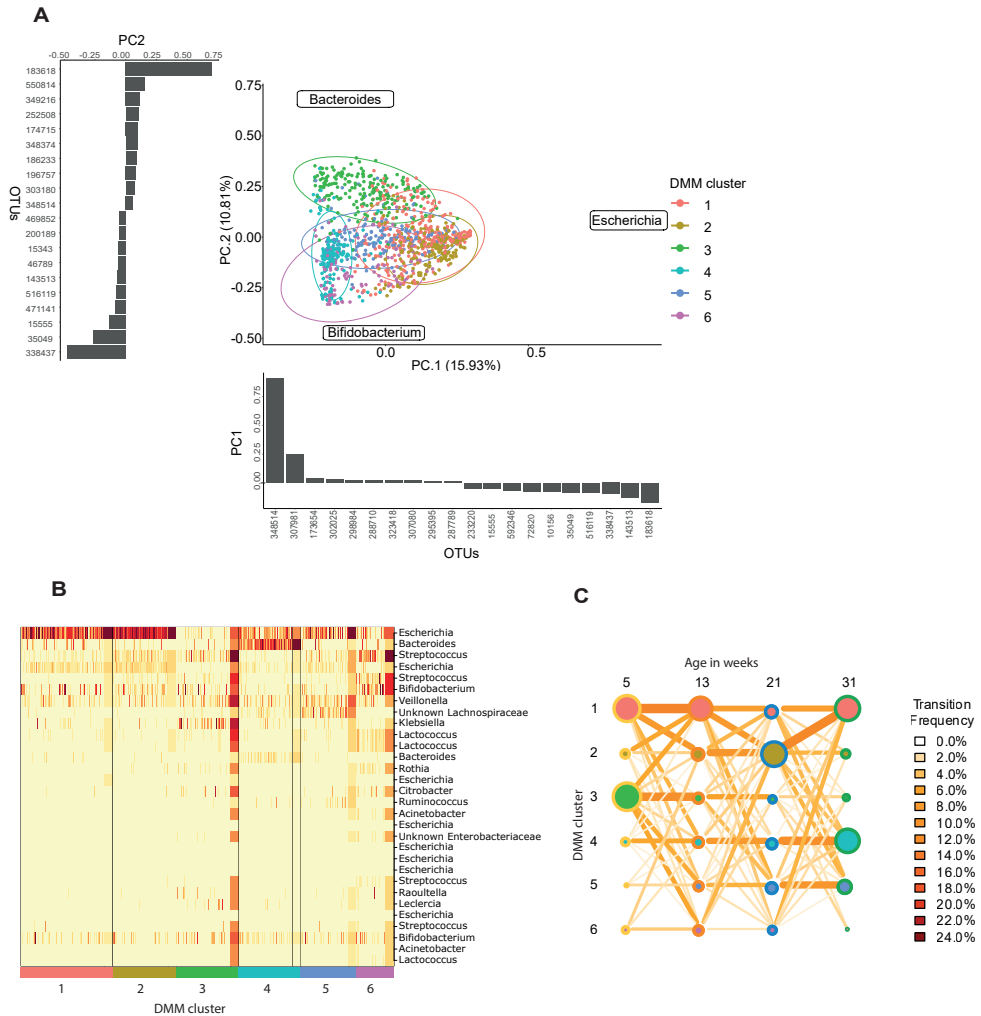


**Figure 1. Microbiota maturation throughout infancy and childhood (N= 1,453 stool samples from 440 children).** **A**, Microbial diversity (Shannon index) gradually increased throughout infancy and has markedly risen at school-age ( $p = 7.72 \times 10^{-53}$ , Friedman, p-values for post-hoc analyses using Dunn's test are depicted in the table). **B**, Principal Coordinate Analysis (PCoA) based on unweighted Unifrac dissimilarity indicates a gradual shift in microbial community structure along PC1 during infancy and a completely distinct structure at school-age ( $p = 6.0 \times 10^{-51}$ , Friedman, p-values for post-hoc analyses using Dunn's test are depicted in Supplementary Table 3). **C**, Cladogram depicting the bacterial genera detected in the children's fecal microbiota. Background and branch colors reflect the different phyla. The height of the outer ring reflects the average relative abundance of a genus across all infant time points, whereas the color density of the five inner rings reflects the prevalence of the genus at the individual time points (with opaque color indicating a prevalence of 100 percent and fully transparent indicating a prevalence of 0 percent).

**The length of breastfeeding represents the main driver of the infant's microbiota composition**

To identify covariates associated with the microbiota dynamics during infancy, we continued our analyses focusing on the infant samples. Dirichlet multinomial mixtures (DMM) modelling on OTU-level data formed six clusters (based on lowest Laplace approximation) (Fig. 2A-B).

To illustrate the progression of samples through each DMM cluster with age, we applied a transition model as described previously [29]. Clusters 1 and 3 were the most dominant at the age of 5 weeks and thereafter transitions were chaotic, consistent with the previously identified developmental phase of the microbiome during the first 14 months of life [29]. Although cluster 1 remained dominant until the age of 31 weeks, cluster 3 gradually disappeared in favor of clusters 4 and 5 (Fig 2C). Multinomial logistic regression analyses indicated that the initial microbiota cluster at 5 weeks of age was mainly determined by birth mode. The chance that a newborn's microbiota belonged to cluster 3 was strongly increased among infants born by cesarean section [Supplementary Table 5]. This cluster was remarkably different with respect to the abundance of several of the driving OTUs. In particular a *Klebsiella* OTU exhibited a high abundance at the expense of an *Escherichia* OTU that dominated many of the other clusters. Additionally, *Citrobacter*, *Leclercia* and *Raoultella* OTUs were characteristic for cluster 3 (Fig 2B).



**Figure 2. Community typing by Dirichlet Multinomial Mixtures of infant samples revealed six clusters (N= 1,154 stool samples from 312 children).** **A**, Heat map showing the relative abundance of the 30 most important/dominant OTUs per DMM cluster. **B**, Principal Component Analysis (PCA) on OTU-level data with samples colored according to DMM cluster. Ellipses indicate the 95% confidence interval. Vertical and horizontal bar charts depict OTUs with the highest loadings on PC1 and PC2 respectively. The OTUs with the highest positive and negative loadings on PC1 and PC2 are plotted in the PCA. **C**, Transition model showing the progression of samples through the six DMM clusters from one sampling time-point to the next time-point.

Analysis of the most common transition trajectories, revealed that for both children starting in cluster 1 as well as for children starting in cluster 3, transition towards clusters 4 and 5 significantly increased when breastfeeding was ceased [Supplementary Table 5]. These results were further supported by the overall bacterial profiles throughout infancy. At the age of 5 weeks,

the largest amount of variance was explained by birth mode (Fig. 3A). At the genus level, vaginal as compared to cesarean section delivery was most strongly associated with an enrichment of *Bacteroides spp.*, at 5 weeks and until the age of 31 weeks (Supplementary Table 6).

At the age of 13, 21 and 31 weeks, breastfeeding explained by far the greatest variance in bacterial community profiles (Fig 3B-D). Permutational multivariate analyses of variance confirmed that the duration of breastfeeding had a stronger impact than the introduction of solid foods (Fig 3E, Supplementary Table 7). Bifidobacteria, staphylococci and streptococci amongst others significantly decreased upon cessation of breastfeeding, whereas many bacteria within the Lachnospiraceae family (e.g. *Pseudobutyrvibrio*, *Lachnobacterium*, *Roseburia*, *Blautia*) increased (Supplementary Table 6).

A longer duration of breastfeeding was also associated with a lower microbial diversity (Supplementary Table 8) as well as with a lower microbial-by-age z-score (MAZ) (Supplementary Table 9). The MAZ is calculated by training a machine-learning algorithm on the microbiota composition of a dataset with known biological age, thereafter the age of samples is predicted based on its microbiota composition. A lower MAZ is thus indicative for a delayed microbial maturation.

Furthermore, the exposure to older siblings was associated with an increase in several genera within the phylum of Actinobacteria (*Bifidobacterium* and *Corynebacterium* at 5 weeks and *Egghertella* at 21 weeks, Supplementary Table 6) and a higher microbial diversity at 31 weeks of age (Supplementary Table 9). Finally, besides dietary factors, the microbial community structure was most strongly associated with the presence of AD at time of sample collection.

### **Alterations in microbial composition, diversity and maturity precede manifestations of atopy**

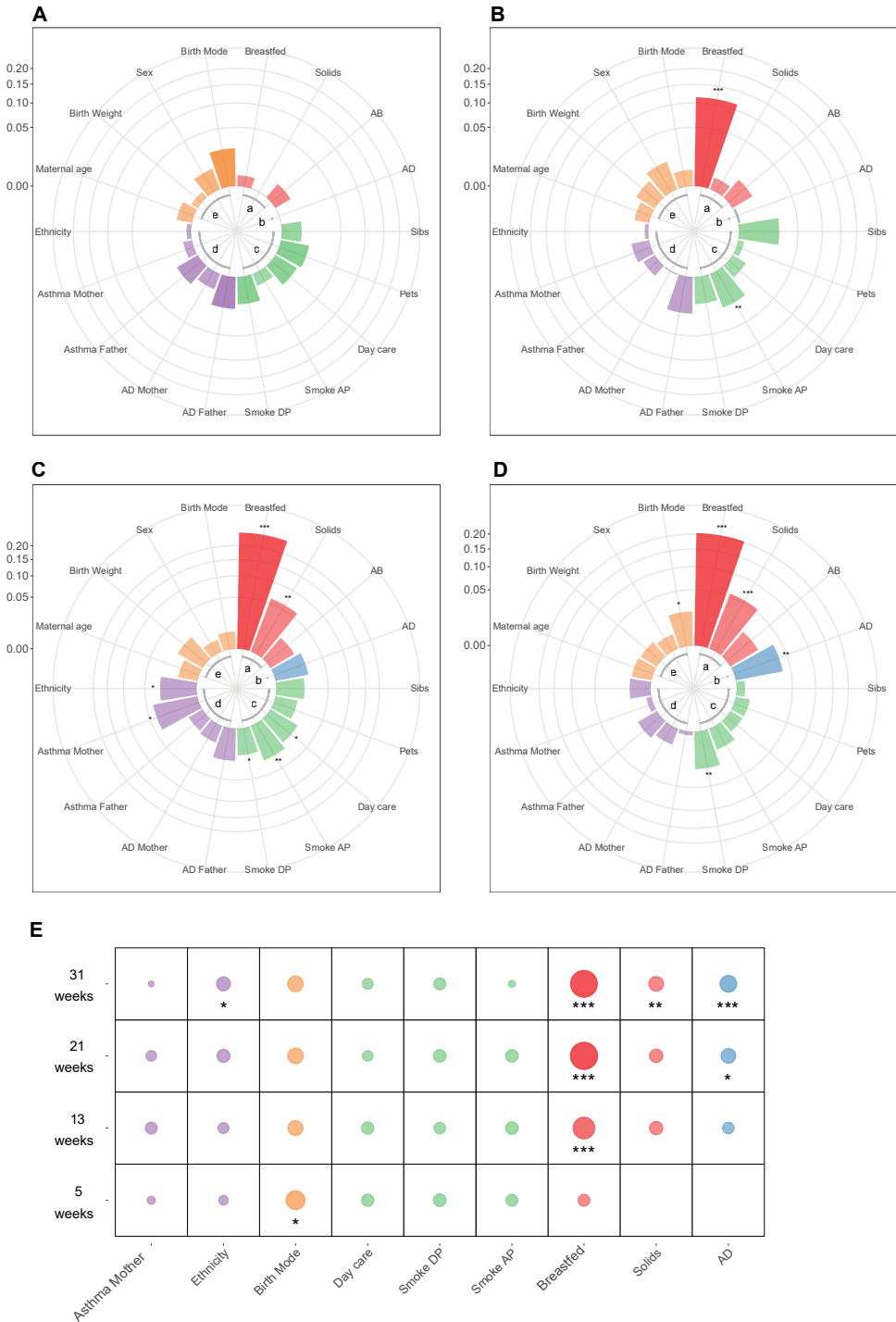
To further investigate whether differences in microbiota development precede the onset of atopic disease, we applied several longitudinal analyses while controlling for potential confounding factors by adjusting for other covariates.

We first applied multivariate joint models on the microbial diversity and maturity in association to AD. Joint models have become increasingly popular as a statistical framework to concurrently analyze longitudinal data (e.g. biomarker evolution) and survival data (e.g. time-to-disease onset). [31] They have, to our knowledge, not been applied in the microbiome research field so far. While accounting for known risk factors for AD, we found that the temporal pattern of microbial diversity was independently and inversely associated with

AD (Hazard rate (HR)= 0.21;  $p=1.15*10^{-4}$ , Figure 4A-B, Supplementary Table 10), indicating that a lower microbial diversity throughout infancy is associated with an increased risk of AD. For the temporal pattern of microbial maturity, expressed as microbial age z-scores, we found a statistically significant positive association with AD (HR=1.14,  $p=1.94*10^{-5}$ , Figure 4C-D, Supplementary Table 11). Next, we used the recently introduced metagenomics longitudinal differential abundance (MetaLonDa) method [33] to identify time intervals of differentially abundant bacterial genera between infants that did or did not develop AD.

Among children who did not develop AD during follow-up, the relative abundance of *Atopobium* (days 25.6 - 79.4,  $FDR_{\text{adjusted}} p = 7.65*10^{-3}$ ), *Corynebacterium* (days 126.1 - 151.2,  $FDR_{\text{adjusted}} p = 9.68*10^{-3}$ ), both members of the phylum Actinobacteria, and *Prevotella* (days 104.6 - 133.3,  $FDR_{\text{adjusted}} p < 0.001$ ) were temporarily enriched when compared to children who developed AD. Most pronounced were, however, the associations of *Lachnobacterium* and *Faecalibacterium*, which were significantly enriched during the entire period of fecal sampling among children who remained free from AD (Figure 4E-G, Supplementary Table 12).

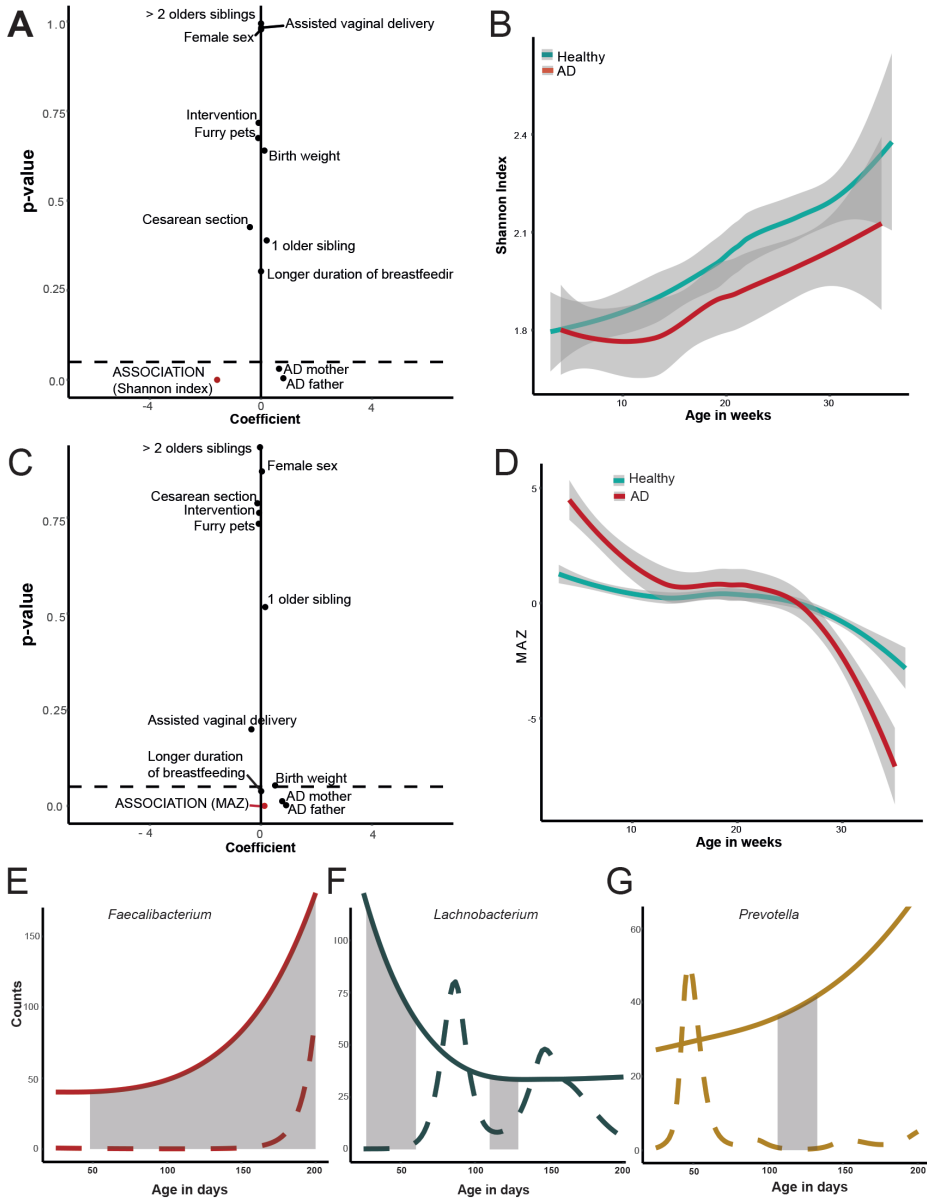
We next, examined whether the infant microbiota composition was also associated with allergic manifestations at school-age, including allergic sensitization and asthma. Blood samples for the determination of allergic sensitization at school age were available for 292 out of the 440 children included in the present study. Like for AD, we found a higher diversity of the infant microbiota to be associated with decreased risk of allergic sensitization at school-age (Shannon index at 31 weeks  $OR_{\text{adjusted}} = 0.19$ ,  $p = 7.33*10^{-3}$ , Supplementary Figure 3A, Supplementary Table 13). A higher microbial maturity very early in life was associated with an increased risk of allergic sensitization (MAZ at 5 weeks  $OR_{\text{adjusted}} = 1.46$ ,  $p = 5.01*10^{-3}$ , Supplementary Figure 3B, Supplementary Table 14), again in line with findings for AD. We could, however, not identify individual bacterial genera with differential abundance over a significant period of time between children who did or did not develop allergic sensitization.



**Figure 3. Microbiota community structure is most strongly influenced by breastfeeding (N= 961 stool samples from 312 children).** A-D, Polar plots visualizing the amount of variance of microbial communities at 5 weeks (A), 13 weeks (B), 21 weeks (C) and 31 weeks (D) of age that could be explained by 18 covariates as analyzed using EnvFit. The height of the bars reflects the amount of variance ( $r^2$ ) explained by each covariate. Covariates are colored to highlight parental health status and ethnicity (purple), perinatal covariates (orange), diet and medication (red) and environmental exposures (green). Asterisks indicate significant covariates (false discovery rate (FDR)  $P < 0.05$ ) at each time point. E, Permutational Multivariate Analysis of Variance (PERMANOVA) combining all covariates that were significantly associated with microbial community variation at any given time point in the EnvFit analyses. The size of the dots reflects the  $R^2$ . Only samples without missing data on the included covariates were included in PERMANOVA (5 weeks: N= 238, 13 weeks: N = 233, 21 weeks: N = 231, 31 weeks: N= 259). Asterisks indicate statistical significance with \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

A clear association between microbial diversity and asthma could not be detected. Yet, in line with allergic sensitization and AD, a higher microbial maturity at the age of 5 weeks was also associated with an increased risk for asthma (MAZ at 5 weeks  $OR_{\text{adjusted}} = 1.43$ ,  $p = 7.78 \times 10^{-3}$ , Supplementary Figure 3C-D, Supplementary Tables 15 and 16). Multiple bacterial genera were differentially abundant over time in children who did or did not develop asthma. The genera that were differentially abundant across the entire time-period during which the microbiota composition was monitored included *Lachnobacterium*, *Lachnospira* (both members of the Lachnospiraceae family) and *Dialister* (Veillonellaceae), which were all significantly enriched in healthy as compared to asthmatic children (Supplementary Figure 3E-G, Supplementary Table 17).





**Figure 4. Microbiota composition, diversity and maturity is linked to the subsequent development of atopic dermatitis (N= 961 stool samples from 312 children).** **A**, Volcano plot depicting the regression coefficients from the joint model on the association between the Shannon index and atopic dermatitis (AD). The dashed line depicts the threshold for statistical significance at  $p < 0.05$ . Variables depicted below the dashed line were statistically significantly associated with AD in the final model. Positive coefficients (variables to the right of the vertical line) were associated with an increased AD risk. Negative coefficients (variables to the left of the vertical line) were associated with a decreased AD risk. The Hazard Ratio is given by the exponent of the coefficient (e.g. for Shannon index:  $e^{-1.57}$  results in a HR of 0.21). **B**, Development of microbial diversity (Shannon index) throughout infancy among children that did (red line) or did not (green line) develop AD as modelled by Loess regression. Grey areas represent the 95%

confidence intervals. **C**, Volcano plot depicting the regression coefficients from the joint model on the association between the microbial age z-score (MAZ) and the development of atopic dermatitis (AD). The dashed line depicts the threshold for statistical significance at  $p < 0.05$ . **D**, Development of microbial maturity (MAZ-score) throughout infancy among children that did (red line) or did not (green line) develop AD as modelled by Loess regression. Grey areas represent the 95% confidence intervals. **E-G**, Time intervals of differential abundance in *Faecalibacterium* (E), *Lachnobacterium* (F), and *Prevotella* (G) between infants that did or did not develop AD as identified from MetaLonda analyses. Significantly different time-intervals (FDR-adjusted  $p < 0.05$ ) are depicted by gray shading.

## Discussion

This study aimed to longitudinally analyze the process of gastrointestinal microbial community assembly, succession and maturation throughout the most critical time-window of immune development and linked microbiota maturation during this time to the development of clinical signs of allergic disease, while carefully controlling for potential confounding factors.

Our results indicate a dynamic microbiota during infancy which is far from completely matured at 31 weeks of age. In early infancy the microbial composition was most strongly affected by birth mode, while from 13 weeks onwards diet became the most important factor. Our data support previous reports, showing that *Bacteroides* are most strongly affected by birth mode [29, 34-36]. The difference in microbial community structure and lower abundance of *Bacteroides* in cesarean section as compared to vaginal delivered infants persisted up to the age of 31 weeks of age and withstood mutual adjustment for other determinants, including breastfeeding. This suggests that the impact of cesarean section delivery could not be compensated by breastfeeding. Given the increased risk of future diseases, including allergies and asthma [37, 38], among children born by cesarean section, more research is warranted to elucidate the need for and efficacy of restoring the natural microbial colonization process upon cesarean delivery.

We furthermore showed that cessation of breastfeeding was more strongly associated with microbial composition and maturity than solid food introduction. In line with previous [11, 35, 39], these results suggest that the introduction of solid food does not appear to result in a profound shift in microbial community structure as long as breastfeeding is continued. Only when breastfeeding is ceased, maturation of the microbiota is accelerated with a decrease in degraders of human milk oligosaccharides and an increase in microbial diversity and compositional changes towards bacterial genera specialized in degrading complex dietary carbohydrates. The generally observed lower microbial diversity in infants during breastfeeding [40] seems at first contradictory with the concept that a “healthy” and resilient microbiome is highly diverse [41]. However, in line with most prospective studies [42], we did not find a direct

association between breastfeeding duration and the risk of AD. The fact that breastfeeding reduces the risk of several other diseases, including metabolic diseases, which on the other hand are also associated with a lower microbial diversity, suggests that the context is of crucial importance when considering microbial diversity. For example, loss of microbiota diversity generally opens up niches for opportunistic invaders [41], while the plethora of bioactive components transferred by breastfeeding protect against colonization by such opportunistic pathogens [43]. This further underscores the need for meticulous adjustment for diet as a confounding factor in the association between microbiota and disease outcomes.

Using various multivariable longitudinal analysis, we furthermore demonstrated that the microbial community structure, diversity and maturity as well as the relative abundance of several individual genera were associated with the subsequent development of allergic manifestation. We know from previous animal studies and large longitudinal human cohorts that intestinal microbial dysbiosis in allergic diseases is mainly observed within a critical window in early life [44]. The comparability between studies is, however, hampered by the highly dynamic microbial communities within this early time window, which likely results in different associations at different sampling time points. The main strengths of the present study are its prospective design, repeated sample collection and the deep clinical phenotyping.

The regular physical examinations of the children throughout the first 3 years of life in combination with the collection of detailed questionnaire data allowed not only deep clinical phenotyping, but also an accurate assessment of the time of disease onset. The follow-up into school-age further facilitated a reliable classification of children who developed allergic asthma as it is well-known that wheezing symptoms at an earlier age are often transient and caused by episodic viral infections [45].

We observed a lower microbial diversity to be associated with AD development and allergic sensitization, but not with asthma. This is consistent with previous studies that also reported a reduced microbial diversity in association to AD [46-49] and sensitization [7, 50]. In contrast, a link between microbial diversity and wheeze or asthma could often not be observed [51-53]. Although atopic manifestations are common comorbidities, these results support previous conclusions from COPSAC that extrapolation of risk factors between different atopic disorders may not always be justified [54].

The increased risk of AD, sensitization and asthma among children with a higher microbial maturity might at first seem in contrast with the above-mentioned results for microbial diversity and with findings of previous studies. Indeed recent results from COPSAC<sub>2010</sub> linked a

low microbial maturity with later onset of asthma in children born to asthmatic mothers [11]. In our study, the microbial maturity was however only significantly increased at 5 weeks of age in children who developed sensitization (as determined by Skin Prick Tests and serum IgE levels to the most common aero-allergens) and asthma. Also, for children with AD, we observed a microbial maturity (MAZ) that was significantly higher at the earliest recorded time-point but gradually decreased and became even lower at the age of 31 weeks when compared to the MAZ of children that did not develop AD. This temporally higher MAZ in very young infants might therefore suggest a dysregulated colonization process, e.g. with some bacterial taxa arriving (too) early, rather than a more mature overall microbial community structure.

Next to differences in microbial diversity and maturity, we were able to identify microbial taxa that were differentially abundant among infants who did or did not develop allergic disease manifestations. *Lachnobacterium* and *Faecalibacterium* were significantly decreased throughout infancy among children who developed AD. Also, *Lachnospira* and *Dialister*, next to *Lachnobacterium*, were significantly decreased among children who developed asthma.

The fact that these bacterial taxa were not only differentially abundant at a single time-point but throughout infancy strengthens the likelihood of a causal role in the protection against allergic disease. Altogether our results indicate that microbial perturbations in early life are also associated with asthma at school age, although perturbations are not identical to those observed in children that developed AD. In line with our findings, analysis on the microbiota composition at 3 months of age within the Canadian Healthy Infant Longitudinal Development (CHILD) Study revealed *Lachnospira* and *Faecalibacterium* to be significantly decreased among children at risk for allergic wheeze at the age of 1 year [52]. Moreover, a lower relative abundance of amongst others *Lachnospiraraceae incertae sedis*, *Faecalibacterium* and *Dialister* at the age of 1 year in children from COPSAC<sub>2010</sub> was associated with an increased risk of asthma at 5 years [11]. Fermentation products of these bacteria are a possible explanation for the protective effect of these bacteria. Acetate is one of the fermentation products of *Lachnospira* and to a lesser account *Lachnobacterium*. Animal studies have previously shown that acetate-feeding leads to a marked suppression of allergic airway disease in a mouse-model for human asthma. The underlying cellular mechanism was related to the effect of acetate on Treg cells, particularly through epigenetic modification of the *Foxp3* promoter [5].

*Faecalibacterium (prausnitzii)* is well-known for its anti-inflammatory effects, amongst others through the production of butyrate [55] and a microbial anti-inflammatory molecule

(MAM) that inhibits the NF- $\kappa$ B pathway [56]. Two recent studies have identified another lactate-consuming butyrate-producing genus, *Anaerostipes*, associated with a decreased risk of food allergy [57] and eczema [58]. The very low abundance of this genus in our population could potentially explain the lack of association in our study.

The application of several types of longitudinal data-analysis, including the joint modelling of longitudinal and survival data which had previously not been used for microbiota data analyses, enabled us to demonstrate that alterations in microbial diversity, maturity and composition preceded the clinical manifestations of atopic diseases. Although this statistical framework reveals the temporality of associations thereby suggesting causal relationships, causality can never be proven in an observational study. For example, microbial perturbations could be an epiphenomenon of exposure to yet another unknown risk factor for allergy. Also, it cannot be ruled out that early preclinical manifestation of allergies or genetic predisposition for allergy might impact the microbiota composition. Moreover, our findings on fecal collections might not fully reflect alterations in the microbiome on allergy development at the level of (small) intestinal mucosa.

It is therefore of importance not only to replicate findings in similar cohorts, but also to conduct future experimental studies building upon these findings in order to reveal the underlying biological mechanisms and prove causality.

In conclusion, our results demonstrate the importance of birth mode and diet on the early maturation of the infant microbiota and demonstrate that, upon careful adjustment of important confounding factors, alterations in the microbial colonization process of the infant intestinal tract precede the development of AD, sensitization and asthma. In particular members of the Lachnospiraceae family, as well as the genera *Faecalibacterium* and *Dialister* appear to protect against allergies. These findings further support the future development of evidence-based intervention strategies targeting the microbiota to prevent or treat allergic diseases in early life.

## Acknowledgments

We thank Christel Driessen for lab analyses and Mayk Lucchesi for designing the graphical abstract.

## References

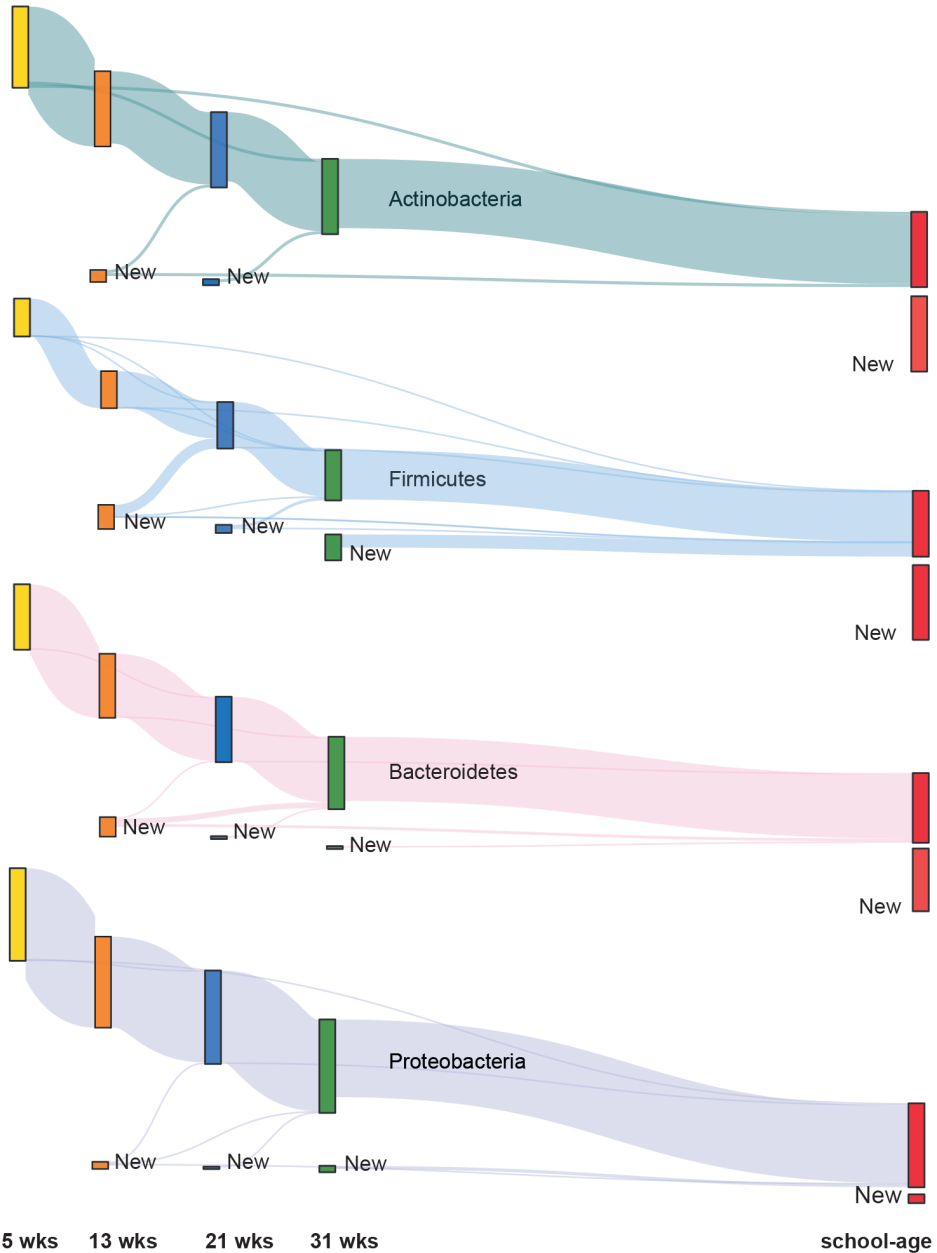
1. Bennek, E., et al., *Subcellular antigen localization in commensal E. coli is critical for T cell activation and induction of specific tolerance*. *Mucosal Immunol*, 2019. 12(1): p. 97-107.
2. Cahenzli, J., et al., *Intestinal microbial diversity during early-life colonization shapes long-term IgE levels*. *Cell Host Microbe*, 2013. 14(5): p. 559-70.
3. Stefka, A.T., et al., *Commensal bacteria protect against food allergen sensitization*. *Proc Natl Acad Sci U S A*, 2014. 111(36): p. 13145-50.
4. Olszak, T., et al., *Microbial exposure during early life has persistent effects on natural killer T cell function*. *Science*, 2012. 336(6080): p. 489-93.
5. Thorburn, A.N., et al., *Evidence that asthma is a developmental origin disease influenced by maternal diet and bacterial metabolites*. *Nat Commun*, 2015. 6: p. 7320.
6. Trompette, A., et al., *Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis*. *Nat Med*, 2014. 20(2): p. 159-66.
7. Azad, M.B., et al., *Infant gut microbiota and food sensitization: associations in the first year of life*. *Clin Exp Allergy*, 2015. 45(3): p. 632-43.
8. Dzidic, M., et al., *Aberrant IgA responses to the gut microbiota during infancy precede asthma and allergy development*. *J Allergy Clin Immunol*, 2017. 139(3): p. 1017-1025 e14.
9. Penders, J., et al., *The role of the intestinal microbiota in the development of atopic disorders*. *Allergy*, 2007. 62(11): p. 1223-36.
10. Simonyte Sjodin, K., et al., *Temporal and long-term gut microbiota variation in allergic disease: A prospective study from infancy to school age*. *Allergy*, 2019. 74(1): p. 176-185.
11. Stokholm, J., et al., *Maturation of the gut microbiome and risk of asthma in childhood*. *Nat Commun*, 2018. 9(1): p. 141.
12. Fujimura, K.E., et al., *Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation*. *Nat Med*, 2016. 22(10): p. 1187-1191.
13. Zimmermann, P., et al., *Association between the intestinal microbiota and allergic sensitization, eczema, and asthma: A systematic review*. *J Allergy Clin Immunol*, 2019. 143(2): p. 467-485.
14. Simonyte Sjodin, K., et al., *Emerging evidence of the role of gut microbiota in the development of allergic diseases*. *Curr Opin Allergy Clin Immunol*, 2016. 16(4): p. 390-5.
15. Korpela, K., et al., *Selective maternal seeding and environment shape the human gut microbiome*. *Genome Res*, 2018. 28(4): p. 561-568.
16. Wampach, L., et al., *Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential*. *Nat Commun*, 2018. 9(1): p. 5091.
17. Gilbert, J.A. and S.V. Lynch, *Community ecology as a framework for human microbiome research*. *Nat Med*, 2019. 25(6): p. 884-889.
18. Davidson, R., et al., *Influence of maternal and perinatal factors on subsequent hospitalisation for asthma in children: evidence from the Oxford record linkage study*. *BMC Pulm Med*, 2010. 10: p. 14.
19. van Nimwegen, F.A., et al., *Mode and place of delivery, gastrointestinal microbiota, and their influence on asthma and atopy*. *J Allergy Clin Immunol*, 2011. 128(5): p. 948-55 e1-3.
20. Lau, S., et al., *Oral application of bacterial lysate in infancy decreases the risk of atopic dermatitis in children*

- with 1 atopic parent in a randomized, placebo-controlled trial. *J Allergy Clin Immunol*, 2012. 129(4): p. 1040-7.
21. Penders, J., et al., *Establishment of the intestinal microbiota and its role for atopic dermatitis in early childhood*. *J Allergy Clin Immunol*, 2013. 132(3): p. 601-607 e8.
  22. Bartram, A.K., et al., *Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads*. *Appl Environ Microbiol*, 2011. 77(11): p. 3846-52.
  23. Whelan, F.J. and M.G. Surette, *A comprehensive evaluation of the sl1p pipeline for 16S rRNA gene sequencing analysis*. *Microbiome*, 2017. 5(1): p. 100.
  24. Lagkouvardos, I., et al., *Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons*. *PeerJ*, 2017. 5: p. e2836.
  25. Oksanen, J., et al., *Vegan: Community Ecology Package. R Package Version. 2.5-3*. CRAN. 2013.
  26. Chen, J., *GUniFrac: Generalized UniFrac Distances. R Package Version. 1.1*. CRAN. 2018.
  27. Subramanian, S., et al., *Persistent gut microbiota immaturity in malnourished Bangladeshi children*. *Nature*, 2014. 510(7505): p. 417-21.
  28. Holmes, I., K. Harris, and C. Quince, *Dirichlet multinomial mixtures: generative models for microbial metagenomics*. *PLoS One*, 2012. 7(2): p. e30126.
  29. Stewart, C.J., et al., *Temporal development of the gut microbiome in early childhood from the TEDDY study*. *Nature*, 2018. 562(7728): p. 583-588.
  30. Morgan, X.C., et al., *Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment*. *Genome Biol*, 2012. 13(9): p. R79.
  31. Cadarso Suarez, C., et al., *Editorial "Joint modeling of longitudinal and time-to-event data and beyond"*. *Biom J*, 2017. 59(6): p. 1101-1103.
  32. Rizopoulos, D., *JM: An R Package for the Joint Modelling of Longitudinal and Time-to-event Data*. *Journal of Statistical Software*, 2010. 35(9).
  33. Metwally, A.A., et al., *MetaLonDA: a flexible R package for identifying time intervals of differentially abundant features in metagenomic longitudinal studies*. *Microbiome*, 2018. 6(1): p. 32.
  34. Azad, M.B., et al., *Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months*. *CMAJ*, 2013. 185(5): p. 385-94.
  35. Backhed, F., et al., *Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life*. *Cell Host Microbe*, 2015. 17(5): p. 690-703.
  36. Penders, J., et al., *Factors influencing the composition of the intestinal microbiota in early infancy*. *Pediatrics*, 2006. 118(2): p. 511-21.
  37. Huang, L., et al., *Is elective cesarean section associated with a higher risk of asthma? A meta-analysis*. *J Asthma*, 2015. 52(1): p. 16-25.
  38. Thavagnanam, S., et al., *A meta-analysis of the association between Caesarean section and childhood asthma*. *Clin Exp Allergy*, 2008. 38(4): p. 629-33.
  39. Pannaraj, P.S., et al., *Association Between Breast Milk Bacterial Communities and Establishment and Development of the Infant Gut Microbiome*. *JAMA Pediatr*, 2017. 171(7): p. 647-654.
  40. Ho, N.T., et al., *Meta-analysis of effects of exclusive breastfeeding on infant gut microbiota across populations*. *Nat Commun*, 2018. 9(1): p. 4169.
  41. Bello, M.G.D., et al., *Preserving microbial diversity*. *Science*, 2018. 362(6410): p. 33-34.
  42. Lin, B., et al., *Breastfeeding and Atopic Dermatitis Risk: A Systematic Review and Meta-Analysis of Prospective Cohort Studies*. *Dermatology*, 2019: p. 1-16.

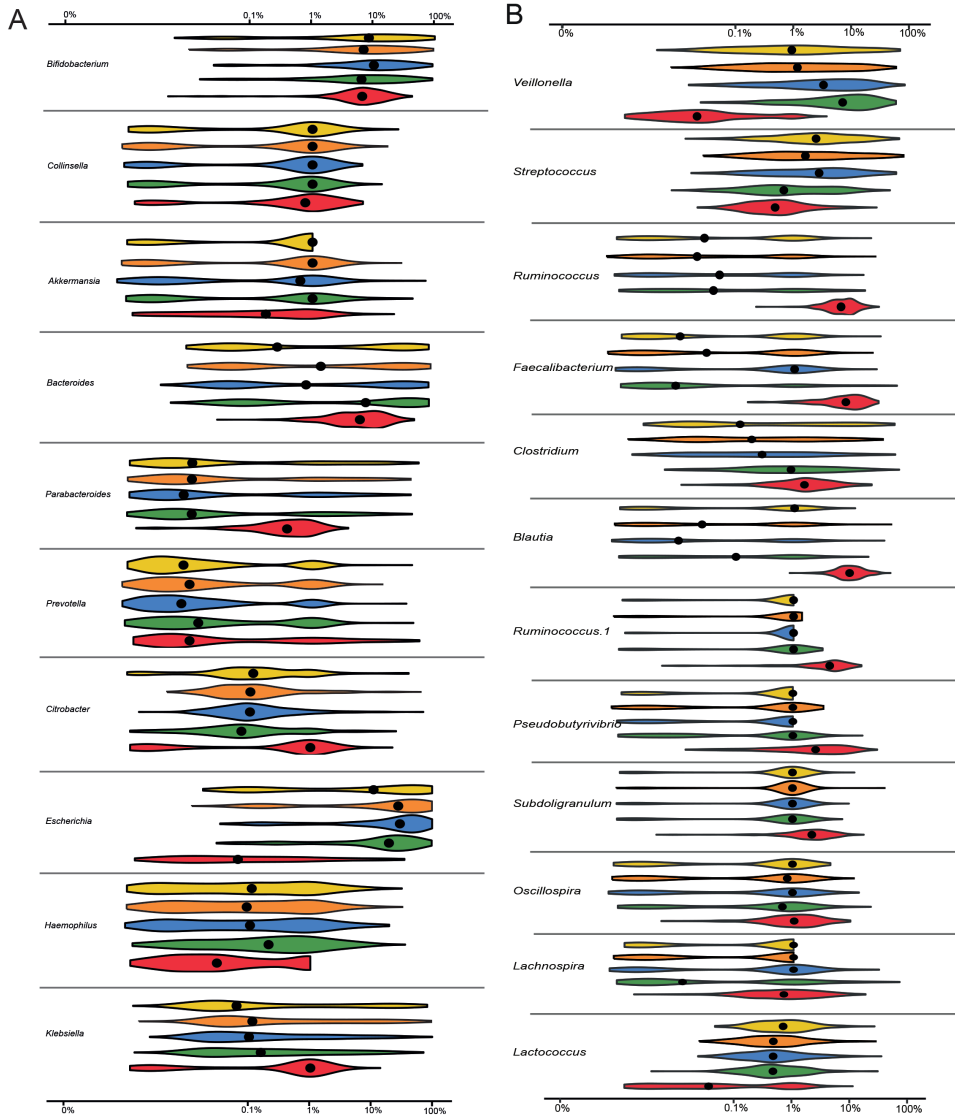
43. van den Elsen, L.W.J., et al., *Shaping the Gut Microbiota by Breastfeeding: The Gateway to Allergy Prevention?* Front Pediatr, 2019. 7: p. 47.
44. Stiemsma, L.T. and S.E. Turvey, *Asthma and the microbiome: defining the critical window in early life.* Allergy Asthma Clin Immunol, 2017. 13: p. 3.
45. Townshend, J., S. Hails, and M. McKean, *Diagnosis of asthma in children.* BMJ, 2007. 335(7612): p. 198-202.
46. Abrahamsson, T.R., et al., *Low diversity of the gut microbiota in infants with atopic eczema.* J Allergy Clin Immunol, 2012. 129(2): p. 434-40, 440 e1-2.
47. Forno, E., et al., *Diversity of the gut microbiota and eczema in early life.* Clin Mol Allergy, 2008. 6: p. 11.
48. Ismail, I.H., et al., *Reduced gut microbial diversity in early life is associated with later development of eczema but not atopy in high-risk infants.* Pediatr Allergy Immunol, 2012. 23(7): p. 674-81.
49. Wang, M., et al., *Reduced diversity in the early fecal microbiota of infants with atopic eczema.* J Allergy Clin Immunol, 2008. 121(1): p. 129-34.
50. Chen, C.C., et al., *Alterations in the gut microbiotas of children with food sensitization in early life.* Pediatr Allergy Immunol, 2016. 27(3): p. 254-62.
51. Arrieta, M.C., et al., *Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting.* J Allergy Clin Immunol, 2018. 142(2): p. 424-434 e10.
52. Arrieta, M.C., et al., *Early infancy microbial and metabolic alterations affect risk of childhood asthma.* Sci Transl Med, 2015. 7(307): p. 307ra152.
53. Bisgaard, H., et al., *Reduced diversity of the intestinal microbiota during infancy is associated with increased risk of allergic disease at school age.* J Allergy Clin Immunol, 2011. 128(3): p. 646-52 e1-5.
54. Bisgaard, H., et al., *Risk analysis of early childhood eczema.* J Allergy Clin Immunol, 2009. 123(6): p. 1355-60 e5.
55. Canani, R.B., et al., *Potential beneficial effects of butyrate in intestinal and extraintestinal diseases.* World J Gastroenterol, 2011. 17(12): p. 1519-28.
56. Quevrain, E., et al., *Identification of an anti-inflammatory protein from Faecalibacterium prausnitzii, a commensal bacterium deficient in Crohn's disease.* Gut, 2016. 65(3): p. 415-425.
57. Feehley, T., et al., *Healthy infants harbor intestinal bacteria that protect against food allergy.* Nat Med, 2019. 25(3): p. 448-453.
58. Wopereis, H., et al., *Intestinal microbiota in infants at high risk for allergy: Effects of prebiotics and role in eczema development.* J Allergy Clin Immunol, 2018. 141(4): p. 1334-1342 e5.



## Supplementary Data



**Supplementary Figure 1. Tracking the flow of OTUs within the four main phyla throughout infancy and childhood (N= 1,453 stool samples from 440 children).** OTUs that were shared by at least 10 percent of the population during one or more time points were tracked using Sankey plots in the four major phyla. The rectangle height indicates the relative number of OTUs and the rectangle color reflects the children's age. The lines represent the transfer of OTUs between time points. At school-age many new OTUs within the Actinobacteria, Firmicutes and Bacteroidetes were gained, whereas a major loss of OTUs in Proteobacteria was observed.



**Supplementary Figure 2. Relative abundance of the main bacterial genera at different time-points during infancy and at school-age (N= 1,453 stool samples from 440 children).** Comparison were made among the 20 most abundant genera and stratified according to age. The black dots represent the median values and the violins are colored according to age. **A**, relative abundance of the main genera within the phyla of Proteobacteria, Actinobacteria, Bacteroidetes and Verrucomicrobia. **B**, relative abundances of the main genera within the Firmicutes phylum. To test for significant changes in relative abundances with age, the Friedman test was used followed by the Dunn's test for post-hoc analyses. FDR-adjusted p-values are presented in Supplementary Table 4.

**Supplementary Table S1 - Demographics of the study population**

<b>Variable</b>	<b>n (%)*</b>
<b>Sex</b>	
Boys	223 (50.7)
Girls	217 (49.3)
<b>Birth mode</b>	
Natural vaginal delivery	304 (69.1)
Assisted vaginal delivery	26 (5.9)
Cesarean section	109 (24.8)
<b>Number of older siblings</b>	
No older siblings	253 (57.5)
1 older sibling	150 (34.1)
≥ 2 older siblings	37 (8.4)
<b>Eczema mother</b>	
Yes	161 (36.2)
No	279 (63.4)
<b>Eczema father</b>	
Yes	92 (20.9)
No	346 (78.6)
<b>Asthma mother</b>	
Yes	142 (32.3)
No	298 (67.6)
<b>Asthma father</b>	
Yes	102 (23.2)
No	337 (76.6)
<b>Maternal smoking in pregnancy</b>	
No	348 (79.1)
Yes	92 (20.9)
<b>Maternal smoking after pregnancy</b>	
No	341 (77.5)
Yes	99 (22.5)
<b>Households with furry pets</b>	
No	138 (31.4)
Yes	277 (63.0)
<b>Ethnicity</b>	
Caucasian	429 (97.5)
Other	11 (2.5)
<b>Antibiotics before age 31 weeks</b>	
No	411 (93.4)
Yes	29 (6.6%)

\*Numbers may not add up to 440 (100%) due to missing data

<b>Variable</b>	<b>Median (IQR)</b>
<b>Maternal age at delivery (in years)</b>	33 (29-36)
<b>Birth Weight (in grams)</b>	3497 (3202-3800)
<b>Breast feeding in weeks</b>	40 (30-53)
<b>Age at introduction of solids (in weeks)</b>	25 (22-27)
<b>Gestational age (in weeks)</b>	40 (38-40)
<b>Start at daycare (in months)</b>	13 (11-22)

**Supplementary Table S2 - Microbial richness as assessed by Chao1 index across time-points**

Age	Median Chao1	Interquartile range
5 weeks	147.23	41.6 - 188.64
13 weeks	170.8	65.13 - 214.3
21 weeks	182.71	66.11 - 220.88
31 weeks	189.83	74.13 - 239.52
School-age	379.83	216.52 - 432.17

Test	P-value
Friedman test	7.724E-53

Variable 1	Variable 2	Test	P-value
Chao1 at 5 weeks	Chao1 at 13 weeks	Dunn's test	<b>1.37E-02</b>
Chao1 at 5 weeks	Chao1 at 21 weeks	Dunn's test	<b>1.52E-05</b>
Chao1 at 5 weeks	Chao1 at 31 weeks	Dunn's test	<b>1.65E-07</b>
Chao1 at 5 weeks	Chao1 at school-age	Dunn's test	<b>7.55E-21</b>
Chao1 at 13 weeks	Chao1 at 21 weeks	Dunn's test	<b>4.09E-03</b>
Chao1 at 13 weeks	Chao1 at 31 weeks	Dunn's test	<b>4.09E-03</b>
Chao1 at 13 weeks	Chao1 at school-age	Dunn's test	<b>7.68E-21</b>
Chao1 at 21 weeks	Chao1 at 31 weeks	Dunn's test	9.09E-01
Chao1 at 21 weeks	Chao1 at school-age	Dunn's test	<b>7.55E-21</b>
Chao1 at 31 weeks	Chao1 at school-age	Dunn's test	<b>8.56E-21</b>

*Related samples statistics based only on children with samples collected at every single time-point (N = 606 samples of 121 children)*

**Supplementary Table S3 - Testing for statistically significant associations between age of sample collection and principal coordinates of PCo1 as depicted in Figure 1**

Test	P-value
Friedman	6.002E-51

Variable 1	Variable 2	Test	P-value
PCo1 at 5 weeks	PCo1 at 13 weeks	Dunn's test	<b>2.32E-04</b>
PCo1 at 5 weeks	PCo1 at 21 weeks	Dunn's test	<b>2.12E-03</b>
PCo1 at 5 weeks	PCo1 at 31 weeks	Dunn's test	<b>5.65E-34</b>
PCo1 at 5 weeks	PCo1 at school-age	Dunn's test	<b>1.21E-27</b>
PCo1 at 13 weeks	PCo1 at 21 weeks	Dunn's test	3.27E-01
PCo1 at 13 weeks	PCo1 at 31 weeks	Dunn's test	<b>1.21E-27</b>
PCo1 at 13 weeks	PCo1 at school-age	Dunn's test	<b>8.53E-27</b>
PCo1 at 21 weeks	PCo1 at 31 weeks	Dunn's test	<b>1.58E-27</b>
PCo1 at 21 weeks	PCo1 at school-age	Dunn's test	<b>1.03E-25</b>
PCo1 at 31 weeks	PCo1 at school-age	Dunn's test	<b>3.99E-28</b>

*Related samples statistics based only on children with samples collected at every single time-point (N = 606 samples of 121 children)*

Supplementary Table S4 - Testing for statistically significant associations between age of sample collection and relative abundance of the main bacterial genera as depicted in Supplementary Figures ZA-B

	<i>Akkermansia</i>	<i>Prevotella</i>	<i>Oscillospira</i>	<i>Subdoligranulum</i>	<i>Haemophilus</i>	<i>Lachnospira</i>	<i>Citrobacter</i>
†	7.69E-37	4.60E-05	3.76E-45	5.91E-75	1.48E-11	8.79E-57	4.37E-59
§							
Time-point 1	Time-point 2						
5 weeks	13 weeks	NS	NS	NS	NS	NS	NS
5 weeks	21 weeks	<0.001	NS	NS	NS	NS	0.05
5 weeks	31 weeks	0.002	<0.001	NS	<0.001	<0.001	NS
5 weeks	school-age	<0.001	<0.001	<0.001	0.012	<0.001	<0.001
13 weeks	21 weeks	0.012	NS	NS	NS	NS	NS
13 weeks	31 weeks	0.039	NS	NS	0.012	<0.001	NS
13 weeks	school-age	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
21 weeks	31 weeks	NS	NS	NS	0.042	<0.001	NS
21 weeks	school-age	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
31 weeks	school-age	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
†							
	<i>Pseudobutyrvibrio</i>	<i>Ruminococcus</i>	<i>Lactococcus</i>	<i>Parabacteroides</i>	<i>Ruminococcus</i>	<i>Faecalibacterium</i>	<i>Blautia</i>
†	4.96E-64	6.30E-55	7.75E-42	1.54E-11	3.75E-78	8.95E-59	2.16E-59
§							
Time-point 1	Time-point 2						
5 weeks	13 weeks	NS	NS	NS	NS	NS	NS
5 weeks	21 weeks	NS	0.005	NS	NS	NS	0.012
5 weeks	31 weeks	0.028	<0.001	NS	<0.001	0.002	<0.001
5 weeks	school-age	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
13 weeks	21 weeks	NS	NS	NS	NS	NS	NS
13 weeks	31 weeks	NS	0.023	NS	0.023	<0.001	0.003
13 weeks	school-age	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
21 weeks	31 weeks	NS	NS	NS	NS	<0.001	NS
21 weeks	school-age	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
31 weeks	school-age	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

	<i>Clostridium</i>	<i>Klebsiella</i>	<i>Streptococcus</i>	<i>Veillonella</i>	<i>Bacteroides</i>	<i>Bifidobacterium</i>
†	2.55E-10	3.42E-39	1.61E-14	6.00E-53	0.000067	3.00E-01
§						
<b>Time-point 1</b>	<b>Time-point 2</b>					
5 weeks	13 weeks	NS	NS	NS	NS	-
5 weeks	21 weeks	NS	NS	NS	NS	-
5 weeks	31 weeks	<0.001	<0.001	<0.001	<0.001	-
5 weeks	school-age	<0.001	<0.001	<0.001	NS	-
13 weeks	21 weeks	NS	NS	NS	NS	-
13 weeks	31 weeks	0.005	<0.001	<0.001	0.011	-
13 weeks	school-age	<0.001	<0.001	<0.001	NS	-
21 weeks	31 weeks	0.007	<0.001	0.007	<0.001	-
21 weeks	school-age	<0.001	<0.001	<0.001	NS	-
31 weeks	school-age	<0.001	NS	<0.001	NS	-

Related samples statistics based only on children with samples collected at every single time-point (N = 606 samples of 121 children)

†Friedman-test p-values & § Post-hoc analyses Dunn's test Bonferroni adjusted p-values

Supplementary Table S5 - Multinomial logistic regression analyses on the association between demographic, lifestyle and medical factors in association with DMM clusters and trajectories

FINAL MULTIVARIABLE MULTINOMIAL LOGISTIC REGRESSION MODEL ON COVARIATES IN ASSOCIATION TO DMM CLUSTERS AT 5 WEEKS OF AGE.				
	Variable	Exp(B)	95% Confidence Interval for Exp (B)	P-value
DMM Cluster 1 (n = 75) - Reference category	-	-	-	-
DMM Cluster 2 (n = 29)	Breastfed (in weeks)	0.998	0.983-1.014	8.27E-01
	C-section delivery	1.610	0.520 - 4.974	4.10E-01
	Eczema mother	1.574	0.476 - 5.204	4.57E-01
	Asthma father	0.685	0.188 - 2.497	5.67E-01
	1 older sibling	2.097	0.815 - 5.398	1.25E-01
	≥ 2 older siblings	2.931	0.645 - 13.315	1.64E-01
DMM Cluster 3 (n = 76)	Breastfed (in weeks)	0.997	0.985 - 1.010	6.81E-01
	C-section delivery	<b>3.346</b>	<b>1.488 - 7.528</b>	<b>3.00E-03</b>
	Eczema mother	2.238	0.895 - 5.596	8.50E-02
	Asthma father	0.700	0.274 - 1.787	4.56E-01
	1 older sibling	1.019	0.487 - 2.133	9.60E-01
	≥ 2 older siblings	0.765	0.180 - 3.254	7.17E-01
DMM Cluster 4 (n = 20)	Breastfed (in weeks)	<b>0.965</b>	<b>0.938 - 0.992</b>	<b>1.20E-02</b>
	C-section delivery	0.242	0.028 - 2.109	1.99E-01
	Eczema mother	2.564	0.713 - 9,220	1.49E-01
	Asthma father	<b>3.338</b>	<b>1.006 - 11.075</b>	<b>4.90E-02</b>
	1 older sibling	2.527	0.802 - 7.956	1.13E-01
	≥ 2 older siblings	2.815	0.443 - 17.903	2.73E-01
DMM Cluster 5 (n = 11)	Breastfed (in weeks)	0.990	0.964 - 1.017	4.57E-01
	C-section delivery	<b>6.347</b>	<b>1.467 - 27.456</b>	<b>1.30E-02</b>
	Eczema mother	0,808	0.105 - 6.212	8.37E-01
	Asthma father	0.931	0.140 - 6.175	9.41E-01
	1 older sibling	1.635	0.323 - 8.277	5.52E-01
	≥ 2 older siblings	<b>16.669</b>	<b>2.535 - 109.587</b>	<b>3.00E-03</b>
DMM Cluster 6 (n = 14)	Breastfed (in weeks)	1.013	0.993 - 1.034	1.98E-01
	C-section delivery	1.371	0.310 - 6.065	6.78E-01
	Eczema mother	2.233	0.514 - 9.703	2.84E-01
	Asthma father	3.455	0.937 - 12.748	6.30E-02
	1 older sibling	0.343	0.066 - 1.783	2.03E-01
	≥ 2 older siblings	0.502	0.045 - 5.555	5.74E-01

FINAL MULTIVARIABLE MULTINOMIAL LOGISTIC REGRESSION MODEL ON COVARIATES IN ASSOCIATION TO MOST COMMON DMM CLUSTER TRAJECTORIES				
	Variable	Exp(B)	95% Confidence Interval for Exp (B)	P-value
DMM Cluster trajectory 1 -> 1 (n =26) - Reference category	-	-	-	-
DMM Cluster trajectory 1 -> 4 (n = 22)	Breastfed (in weeks)	<b>0.959</b>	<b>0.928 - 0.992</b>	<b>1.40E-02</b>
	Age introduction solids (weeks)	0.943	0.774-1.148	5.57E-01
	Age at start day care (months)	0.958	0.876 - 1.048	3.53E-01
DMM Cluster trajectory 1 -> 5 (n = 14)	Breastfed (in weeks)	<b>0.948</b>	<b>0.909 - 0.989</b>	<b>1.20E-02</b>
	Age introduction solids (weeks)	0.818	0.660 - 1.014	6.70E-02
	Age at start day care (months)	1.004	0.909 - 1.108	0./944
DMM Cluster trajectory 3 -> 1 (n = 16)	Breastfed (in weeks)	0.980	0.950 - 1.010	1.92E-01
	Age introduction solids (weeks)	1.074	0.876 - 1.315	4.93E-01
	Age at start day care (months)	0.981	0.890 - 1.081	6.95E-01
DMM Cluster trajectory 3 -> 3 (n = 11)	Breastfed (in weeks)	0.997	0.958 - 1.037	8.69E-01
	Age introduction solids (weeks)	0.836	0.636 - 1.100	2.01E-01
	Age at start day care (months)	<b>1.130</b>	<b>1.016 - 1.257</b>	<b>2.50E-02</b>
DMM Cluster trajectory 3 -> 4 (n = 13)	Breastfed (in weeks)	<b>0.939</b>	<b>0.896 - 0.985</b>	<b>1.00E-02</b>
	Age introduction solids (weeks)	<b>0.776</b>	<b>0.618 - 0.974</b>	<b>2.80E-02</b>
	Age at start day care (months)	0.970	0.865 - 1.087	5.97E-01
DMM Cluster trajectory 3 -> 5 (n = 25)	Breastfed (in weeks)	<b>0.949</b>	<b>0.916 - 0.983</b>	<b>4.00E-03</b>
	Age introduction solids (weeks)	0.905	0.744 - 1.102	3.21E-01
	Age at start day care (months)	1.009	0.925 - 1.099	8.44E-01



Supplementary Table S6 - Multivariate Analysis by Linear Models (MaAsLin) on demographic, lifestyle and medical factors in association with genus level taxa

Multivariate Analysis by Linear Models (MaAsLin) on demographic, lifestyle and medical factors associated with genus level taxa (only results with statistically significant associations upon FDR-correction are presented)									
Age	Variable	Feature (genus)	Value	Coefficient	N	N. not.0	P-value	FDR-adjusted P-value (Q)	
5 weeks	Birth Mode	Bacteroides	Vaginal delivery	0.175837645	297	297	1.38E-05	8.04E-03	
5 weeks	Older siblings	Bifidobacterium	Older siblings in households	0.166120378	297	297	1.55E-05	8.04E-03	
5 weeks	Older siblings	Corynebacterium	Older siblings in households	0.002526116	297	138	2.06E-05	8.04E-03	
13 weeks	Cessation breastfeeding prior to sample collection	Pantoea	Breastfeeding ceased	0.00131711	286	79	2.96E-05	4.16E-02	
21 weeks	Birth Mode	Bacteroides	Vaginal delivery	0.14714048	267	267	7.16E-04	6.28E-02	
21 weeks	Birth Mode	Campylobacter	Vaginal delivery	-0.001529738	267	77	1.11E-03	8.25E-02	
21 weeks	Cessation breastfeeding prior to sample collection	Other	Breastfeeding ceased	0.139999237	267	267	1.51E-09	1.72E-06	
21 weeks	Cessation breastfeeding prior to sample collection	Lachnobacterium	Breastfeeding ceased	0.003857002	267	145	2.31E-09	1.72E-06	
21 weeks	Cessation breastfeeding prior to sample collection	Pseudobutyrvibrio	Breastfeeding ceased	0.002832593	267	95	8.08E-08	4.02E-05	
21 weeks	Cessation breastfeeding prior to sample collection	Bifidobacterium	Breastfeeding ceased	-0.197143826	267	267	7.22E-06	1.79E-03	
21 weeks	Cessation breastfeeding prior to sample collection	Clostridium_C	Breastfeeding ceased	0.009197433	267	154	1.47E-05	3.14E-03	
21 weeks	Cessation breastfeeding prior to sample collection	Oscillospira	Breastfeeding ceased	0.002240703	267	155	2.20E-05	4.10E-03	
21 weeks	Cessation breastfeeding prior to sample collection	Eubacterium_A	Breastfeeding ceased	0.002266881	267	81	5.18E-05	8.59E-03	
21 weeks	Cessation breastfeeding prior to sample collection	Ruminococcus	Breastfeeding ceased	0.007854753	267	194	6.28E-05	9.37E-03	
21 weeks	Cessation breastfeeding prior to sample collection	Blautia	Breastfeeding ceased	0.002264146	267	203	2.04E-04	2.76E-02	
21 weeks	Cessation breastfeeding prior to sample collection	Staphylococcus	Breastfeeding ceased	-0.004207764	267	195	2.93E-04	3.36E-02	
21 weeks	Cessation breastfeeding prior to sample collection	Clostridium	Breastfeeding ceased	0.040367347	267	265	6.21E-04	6.28E-02	
21 weeks	Cessation breastfeeding prior to sample collection	Megasphaera	Breastfeeding ceased	0.001435393	267	85	7.13E-04	6.28E-02	
21 weeks	Cessation breastfeeding prior to sample collection	Enterococcus	Breastfeeding ceased	-0.015797733	267	263	1.11E-03	8.25E-02	
21 weeks	Ethnicity	Proteus	non Caucasian	0.033164297	267	49	8.31E-04	6.88E-02	
21 weeks	Older siblings	Eggerthella	Older siblings in households	0.006674972	267	124	4.98E-06	1.49E-03	
21 weeks	Smoking_in_Pregnancy	Veillonella	Smoking_in_Pregnancy2	-0.088597495	267	267	2.60E-04	3.23E-02	
21 weeks	Solids introduced prior to sample collection	Campylobacter	Solids introduced	0.002285898	267	77	2.94E-06	1.10E-03	
21 weeks	Solids introduced prior to sample collection	Clostridium_C	Solids introduced	0.006088278	267	154	6.57E-04	6.28E-02	
31 weeks	AD status at time of sample collection	Citrobacter	AD present	-0.0083787	304	288	1.19E-03	8.39E-02	
31 weeks	Birth Mode	Bacteroides	Vaginal delivery	0.1860148	304	304	4.30E-05	5.14E-03	
31 weeks	Birth Mode	Raoultella	Vaginal delivery	-0.0064418	304	294	1.05E-03	7.75E-02	

Multivariate Analysis by Linear Models (MaAsLin) on demographic, lifestyle and medical factors associated with genus level taxa (only results with statistically significant associations upon FDR-correction are presented)

Age	Variable	Feature (genus)	Value	Coefficient	N	N. not.0	P-value	FDR-adjusted P-value (Q)
31 weeks	Birth weight (in grams)	Corynebacterium	Birth weight (in grams)	1.22E-06	304	93	4.53E-04	4.14E-02
31 weeks	Cessation breastfeeding prior to sample collection	Oscillospira	Breastfeeding ceased	0.04809638	304	209	1.18E-17	1.83E-14
31 weeks	Cessation breastfeeding prior to sample collection	Ruminococcus	Breastfeeding ceased	0.06942166	304	259	1.04E-15	8.05E-13
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Other	Breastfeeding ceased	0.1276393	304	304	1.13E-11	5.86E-09
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Pseudobutyrvibrio	Breastfeeding ceased	0.00440686	304	147	1.53E-10	5.95E-08
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Roseburia	Breastfeeding ceased	0.00412616	304	60	6.19E-10	1.92E-07
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Lachnobacterium	Breastfeeding ceased	0.00304302	304	182	2.20E-09	5.71E-07
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Coprobacillus	Breastfeeding ceased	0.00542691	304	31	4.50E-09	9.99E-07
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Eubacterium_A	Breastfeeding ceased	0.00411655	304	156	6.38E-09	1.24E-06
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Bifidobacterium	Breastfeeding ceased	-0.1679934	304	304	1.06E-07	1.82E-05
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Bacteroides	Breastfeeding ceased	0.20501827	304	304	6.47E-06	1.00E-03
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Clostridium_C	Breastfeeding ceased	0.01510415	304	223	8.10E-06	1.14E-03
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Staphylococcus	Breastfeeding ceased	-0.0034376	304	216	3.92E-05	5.07E-03
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Actinomyces	Breastfeeding ceased	-0.0055536	304	276	1.61E-04	1.67E-02
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Streptococcus	Breastfeeding ceased	-0.043623	304	304	3.06E-04	2.97E-02
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Blautia	Breastfeeding ceased	0.00395886	304	231	8.53E-04	6.98E-02
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Escherichia	Breastfeeding ceased	-0.135552	304	304	9.71E-04	7.54E-02
31 weeks	Cessation_breastfeeding_prior_to_sample_collection	Veillonella	Breastfeeding ceased	-0.0786019	304	304	1.41E-03	9.50E-02
31 weeks	Smoking during pregnancy	Bliphila	Yes	-0.0073827	304	49	5.82E-04	5.02E-02
31 weeks	Solids introduced prior to sample collection	Veillonella	Solids introduced	0.14597588	304	304	4.96E-05	5.50E-03

Supplementary Table S7 - Permutational analyses of variance on demographic, dietary, lifestyle and medical factors associated with microbial community structure

Permutational analysis of the variance at age 5 weeks (N = 238)					
Variable	Df	SumOfSqs	R2	F	P-value
Asthma Mother	1	0.049	0.002	0.556	9.59E-01
Ethnicity	1	0.063	0.003	0.716	8.13E-01
Birth Mode	2	0.291	0.014	1.654	<b>1.70E-02</b>
Day care	1	0.095	0.005	1.079	3.22E-01
Smoke DP	1	0.112	0.005	1.276	1.69E-01
Smoke AP	1	0.102	0.005	1.157	2.56E-01
Breastfed	1	0.094	0.004	1.066	3.90E-01
Residual	229	20.169	0.962		
Total	237	20.976	1.000		

Permutational analysis of the variance at age 13 weeks (N = 233)					
Variable	Df	SumOfSqs	R2	F	P-value
Asthma Mother	1	0.082	0.004	0.997	4.47E-01
Ethnicity	1	0.074	0.004	0.903	5.69E-01
Birth Mode	2	0.169	0.009	1.025	4.07E-01
Day care	1	0.099	0.005	1.202	2.25E-01
Smoke DP	1	0.107	0.005	1.299	1.74E-01
Smoke AP	1	0.122	0.006	1.486	7.39E-02
Breastfed	1	0.305	0.016	3.717	<b>9.99E-04</b>
Solids	1	0.119	0.006	1.448	9.19E-02
AD	1	0.085	0.004	1.038	3.81E-01
Residual	222	18.241	0.940		
Total	232	19.403	1.000		

Permutational analysis of the variance at age 21 weeks (N = 231)					
Variable	Df	SumOfSqs	R2	F	P-value
Asthma Mother	1	0.103	0.006	1.360	1.25E-01
Ethnicity	1	0.113	0.006	1.498	8.19E-02
Birth Mode	2	0.171	0.009	1.132	2.36E-01
Day care	1	0.099	0.006	1.319	1.33E-01
Smoke DP	1	0.111	0.006	1.466	7.79E-02
Smoke AP	1	0.099	0.005	1.312	1.27E-01
Breastfed	1	0.468	0.026	6.212	<b>9.99E-04</b>
Solids	1	0.103	0.006	1.370	1.15E-01
AD	1	0.135	0.007	1.785	<b>2.00E-02</b>
Residual	220	16.589	0.922		
Total	230	17.991	1.000		

Permutational analysis of the variance at age 31 weeks (N = 259)					
Variable	Df	SumOfSqs	R2	F	P-value
Asthma Mother	1	0.046	0.002	0.516	9.84E-01
Ethnicity	1	0.148	0.006	1.647	<b>4.80E-02</b>
Birth Mode	2	0.244	0.010	1.359	<b>4.90E-02</b>
Day care	1	0.098	0.004	1.093	2.93E-01
Smoke DP	1	0.152	0.006	1.694	3.70E-02
Smoke AP	1	0.078	0.003	0.871	6.28E-01
Breastfed	1	0.663	0.027	7.377	<b>9.99E-04</b>
Solids	1	0.180	0.007	1.999	<b>2.00E-03</b>
AD	1	0.241	0.010	2.679	<b>9.99E-04</b>
Residual	248	22.287	0.923		
Total	258	24.137	1.000		

Supplementary Table S8 - Multivariable linear regression analyses on demographic, dietary, lifestyle and medical factors associated with microbial maturity (Microbial Age z-scores)

MAZ index at age 5 weeks (N = 235)							
Unstandardized Coefficients	Std. Error	Standardized Coefficients		t	P-value	95,0% Confidence Interval for B	
		Beta				Lower Bound	Upper Bound
(Constant)	-2.378	5.265		-0.452	6.520E-01	-12.757	8.000
Birhmode assisted vaginal	-0.476	0.680	-0.040	-0.701	4.840E-01	-1.816	0.863
Birhmode cesarean section	0.711	0.367	0.111	1.937	5.400E-02	-0.013	1.434
1 older sibling	-0.362	0.371	-0.060	-0.976	3.300E-01	-1.092	0.369
≥ 2 older siblings	0.763	0.620	0.078	1.230	2.200E-01	-0.459	1.984
Sex	-0.287	0.320	-0.051	-0.897	3.710E-01	-0.918	0.344
Birth Weight	-0.001	0.000	-0.077	-1.220	2.240E-01	-0.001	0.000
Gestational age	0.175	0.136	0.082	1.284	2.010E-01	-0.094	0.444
Breastfeeding duration	-0.010	0.007	-0.097	-1.507	1.330E-01	-0.024	0.003
Ethnicity	0.104	1.022	0.006	0.102	9.190E-01	-1.910	2.119
Maternal age delivery	-0.035	0.035	-0.060	-1.003	3.170E-01	-0.104	0.034
AD father	0.279	0.405	0.043	0.689	4.920E-01	-0.519	1.077
AD mother	-0.694	0.436	-0.103	-1.592	1.130E-01	-1.553	0.165
Asthma father	0.417	0.407	0.060	1.023	3.080E-01	-0.386	1.220
Asthma mother	<b>0.894</b>	<b>0.366</b>	<b>0.152</b>	<b>2.444</b>	<b>1.500E-02</b>	<b>0.173</b>	<b>1.615</b>
Smoking in pregnancy	0.364	0.393	0.056	0.928	3.540E-01	-0.410	1.138
Smoking after pregnancy	-0.168	0.401	-0.026	-0.418	6.760E-01	-0.959	0.623
AD	<b>3.203</b>	<b>0.356</b>	<b>0.536</b>	<b>8.986</b>	<b>0.000E+00</b>	<b>2.500</b>	<b>3.905</b>
Household with furry pets	0.179	0.339	0.030	0.528	5.980E-01	-0.490	0.848
Start of daycare in months	-0.014	0.021	-0.039	-0.647	5.180E-01	-0.055	0.028
Age of introduction solids (months)	-0.020	0.043	-0.030	-0.477	6.340E-01	-0.105	0.064
antibiotic use prior to sample collection	-0.116	1.717	-0.004	-0.068	9.460E-01	-3.500	3.268

MAZ index at age 13 weeks (N = 231)							
Unstandardized Coefficients	Std. Error	Standardized Coefficients		t	P-value	95,0% Confidence Interval for B	
		Beta				Lower Bound	Upper Bound
(Constant)	5.589	5.543		1.008	3.140E-01	-5.338	16.516
Birhmode assisted vaginal	0.262	0.734	0.025	0.357	7.220E-01	-1.185	1.709
Birhmode cesarean section	0.337	0.384	0.060	0.878	3.810E-01	-0.420	1.094
1 older sibling	0.172	0.382	0.033	0.449	6.540E-01	-0.581	0.925
≥ 2 older siblings	0.598	0.674	0.071	0.887	3.760E-01	-0.731	1.927
Sex	-0.138	0.336	-0.028	-0.410	6.820E-01	-0.799	0.524
Birth Weight	-0.001	0.000	-0.147	-1.937	5.400E-02	-0.002	0.000
Gestational age	-0.059	0.143	-0.031	-0.409	6.830E-01	-0.341	0.224
Breastfeeding duration	-0.008	0.007	-0.087	-1.111	2.680E-01	-0.022	0.006
Ethnicity	-0.764	1.009	-0.054	-0.757	4.500E-01	-2.753	1.225
Maternal age delivery	0.019	0.037	0.038	0.523	6.010E-01	-0.054	0.092
AD father	-0.043	0.410	-0.008	-0.105	9.160E-01	-0.852	0.766
AD mother	0.211	0.469	0.036	0.451	6.530E-01	-0.714	1.137
Asthma father	0.724	0.450	0.115	1.610	1.090E-01	-0.162	1.610
Asthma mother	0.293	0.379	0.057	0.774	4.400E-01	-0.453	1.039

Supplementary Table S8 - Continued

MAZ index at age 13 weeks (N = 231)								
Unstandardized Coefficients	Std. Error	Standardized Coefficients		t	P-value	95,0% Confidence Interval for B		
		Beta				Lower Bound	Upper Bound	
Smoking in pregnancy	0.296	0.400	0.054	0.740	4.600E-01	-0.492	1.084	
Smoking after pregnancy	-0.349	0.403	-0.064	-0.867	3.870E-01	-1.144	0.445	
AD	0.052	0.366	0.010	0.142	8.870E-01	-0.669	0.773	
Household with furry pets	-0.136	0.347	-0.027	-0.391	6.960E-01	-0.820	0.549	
Start of daycare in months	0.023	0.018	0.085	1.229	2.210E-01	-0.014	0.059	
Age of introduction solids (months)	-0.005	0.045	-0.009	-0.113	9.100E-01	-0.093	0.083	
antibiotic use prior to sample collection	-0.021	0.054	-0.026	-0.388	6.990E-01	-0.128	0.086	

MAZ index at age 21 weeks (N - 228)								
Unstandardized Coefficients	Std. Error	Standardized Coefficients		t	P-value	95,0% Confidence Interval for B		
		Beta				Lower Bound	Upper Bound	
(Constant)	8.762	5.692		1.539	1.250E-01	-2.463	19.986	
Birhmode assisted vaginal	-0.628	0.734	-0.056	-0.856	3.930E-01	-2.076	0.820	
Birhmode cesarean section	-0.652	0.395	-0.110	-1.651	1.000E-01	-1.430	0.126	
1 older sibling	-0.134	0.384	-0.025	-0.349	7.270E-01	-0.891	0.623	
≥ 2 older siblings	0.985	0.632	0.113	1.558	1.210E-01	-0.262	2.232	
Sex	<b>0.727</b>	<b>0.335</b>	<b>0.143</b>	<b>2.171</b>	<b>3.100E-02</b>	<b>0.067</b>	<b>1.386</b>	
Birth Weight	0.000	0.000	0.004	0.055	9.560E-01	-0.001	0.001	
Gestational age	-0.105	0.150	-0.052	-0.700	4.850E-01	-0.400	0.190	
Breastfeeding duration	-0.004	0.007	-0.040	-0.542	5.880E-01	-0.018	0.010	
Ethnicity	<b>-4.363</b>	<b>1.117</b>	<b>-0.255</b>	<b>-3.907</b>	<b>0.000E+00</b>	<b>-6.565</b>	<b>-2.161</b>	
Maternal age delivery	-0.047	0.036	-0.091	-1.289	1.990E-01	-0.118	0.025	
AD father	-0.437	0.417	-0.076	-1.049	2.960E-01	-1.260	0.385	
AD mother	-0.243	0.465	-0.039	-0.522	6.020E-01	-1.160	0.674	
Asthma father	0.072	0.429	0.012	0.167	8.680E-01	-0.775	0.918	
Asthma mother	<b>0.838</b>	<b>0.382</b>	<b>0.156</b>	<b>2.191</b>	<b>3.000E-02</b>	<b>0.084</b>	<b>1.592</b>	
Smoking in pregnancy	-0.504	0.421	-0.083	-1.198	2.320E-01	-1.333	0.325	
Smoking after pregnancy	<b>1.106</b>	<b>0.419</b>	<b>0.189</b>	<b>2.637</b>	<b>9.000E-03</b>	<b>0.279</b>	<b>1.933</b>	
AD**	<b>0.917</b>	<b>0.358</b>	<b>0.172</b>	<b>2.559</b>	<b>1.100E-02</b>	<b>0.210</b>	<b>1.623</b>	
Household with furry pets	0.215	0.362	0.039	0.593	5.540E-01	-0.499	0.929	
Start of daycare in months	0.002	0.019	0.008	0.114	9.090E-01	-0.035	0.039	
Age of introduction solids (months)	-0.018	0.044	-0.029	-0.411	6.820E-01	-0.104	0.068	
antibiotic use prior to sample collection	-1.055	0.676	-0.107	-1.561	1.200E-01	-2.388	0.278	

## MAZ index at age 31 weeks (N = 253)

	Unstandardized Coefficients	Std. Error	Standardized Coefficients		t	P-value	95,0% Confidence Interval for B	
			Beta				Lower Bound	Upper Bound
(Constant)	-2.944	4.998			-0.589	5.560E-01	-12.792	6.904
Birhmode assisted vaginal	1.245	0.686	0.113		1.814	7.100E-02	-0.107	2.597
Birhmode cesarean section	0.254	0.354	0.045		0.719	4.730E-01	-0.443	0.952
1 older sibling	-0.361	0.353	-0.069		-1.024	3.070E-01	-1.056	0.334
≥ 2 older siblings	-0.273	0.581	-0.033		-0.470	6.390E-01	-1.418	0.871
Sex	-0.072	0.306	-0.015		-0.236	8.130E-01	-0.676	0.531
Birth Weight	-0.001	0.000	-0.097		-1.408	1.600E-01	-0.001	0.000
Gestational age	0.077	0.130	0.041		0.592	5.540E-01	-0.179	0.332
Breastfeeding duration**	<b>-0.018</b>	<b>0.007</b>	<b>-0.192</b>		<b>-2.724</b>	<b>7.000E-03</b>	<b>-0.031</b>	<b>-0.005</b>
Ethnicity	1.271	0.947	0.086		1.342	1.810E-01	-0.594	3.136
Maternal age delivery*	<b>0.072</b>	<b>0.034</b>	<b>0.141</b>		<b>2.130</b>	<b>3.400E-02</b>	<b>0.005</b>	<b>0.139</b>
AD father	-0.551	0.384	-0.099		-1.434	1.530E-01	-1.309	0.206
AD mother	-0.384	0.421	-0.064		-0.912	3.630E-01	-1.214	0.446
Asthma father	-0.329	0.398	-0.054		-0.827	4.090E-01	-1.114	0.455
Asthma mother	-0.526	0.348	-0.101		-1.511	1.320E-01	-1.211	0.160
Smoking in pregnancy	-0.195	0.373	-0.034		-0.523	6.020E-01	-0.930	0.540
Smoking after pregnancy	0.131	0.373	0.023		0.351	7.260E-01	-0.605	0.866
AD**	<b>-0.977</b>	<b>0.332</b>	<b>-0.188</b>		<b>-2.945</b>	<b>4.000E-03</b>	<b>-1.630</b>	<b>-0.323</b>
Household with furry pets	0.287	0.322	0.055		0.890	3.740E-01	-0.348	0.921
Start of daycare in months	-0.030	0.017	-0.113		-1.798	7.400E-02	-0.063	0.003
Age of introduction solids (months)	-0.012	0.040	-0.021		-0.308	7.590E-01	-0.092	0.067
antibiotic use prior to sample collection	0.431	0.502	0.055		0.860	3.910E-01	-0.557	1.420

Supplementary Table S9 - Multivariable linear regression analyses on demographic, dietary, lifestyle and medical factors associated with microbial diversity (Shannon)

Shannon index at age 5 weeks (N = 235)							
Unstandardized Coefficients	Std. Error		Standardized	t	P-value	95,0% Confidence Interval for B	
			Beta			Lower Bound	Upper Bound
(Constant)	0.984	1.142		0.861	3.900E-01	-1.267	3.234
Birhmode assisted vaginal	0.075	0.153	0.032	0.487	6.270E-01	-0.227	0.376
Birhmode cesarean section	-0.025	0.081	-0.020	-0.314	7.540E-01	-0.184	0.134
1 older sibling	0.057	0.080	0.050	0.722	4.710E-01	-0.099	0.214
≥ 2 older siblings	0.006	0.134	0.003	0.047	9.620E-01	-0.257	0.270
Sex	0.045	0.070	0.042	0.636	5.250E-01	-0.093	0.183
Birth Weight	0.000	0.000	0.080	1.110	2.680E-01	0.000	0.000
Gestational age	0.023	0.030	0.056	0.765	4.450E-01	-0.036	0.082
Breastfeeding duration	-0.001	0.002	-0.065	-0.875	3.820E-01	-0.004	0.002
Ethnicity	0.017	0.220	0.005	0.078	9.380E-01	-0.417	0.451
Maternal age delivery	0.003	0.008	0.030	0.441	6.590E-01	-0.012	0.018
AD father	0.019	0.088	0.015	0.212	8.320E-01	-0.154	0.191
AD mother	-0.183	0.097	-0.140	-1.888	6.000E-02	-0.374	0.008
Asthma father	0.083	0.091	0.062	0.917	3.600E-01	-0.095	0.261
Asthma mother	-0.068	0.079	-0.060	-0.860	3.900E-01	-0.224	0.088
Smoking in pregnancy	-0.029	0.087	-0.023	-0.331	7.410E-01	-0.199	0.142
Smoking after pregnancy	0.014	0.084	0.012	0.166	8.680E-01	-0.152	0.180
AD	-0.004	0.076	-0.004	-0.055	9.560E-01	-0.154	0.145
Household with furry pets	-0.146	0.075	-0.127	-1.953	5.200E-02	-0.293	0.001
Start of daycare in months	0.002	0.004	0.027	0.413	6.800E-01	-0.006	0.010
Age of introduction solids (months)	-0.010	0.009	-0.078	-1.084	2.800E-01	-0.028	0.008
antibiotic use prior to sample collection	-0.025	0.281	-0.006	-0.090	9.280E-01	-0.580	0.529

Shannon index at age 13 weeks (N = 231)							
Unstandardized Coefficients	Std. Error		Standardized	t	P-value	95,0% Confidence Interval for B	
			Beta			Lower Bound	Upper Bound
(Constant)	4.139	1.268		3.264	1.000E-03	1.639	6.639
Birhmode assisted vaginal	0.184	0.168	0.076	1.098	2.730E-01	-0.147	0.515
Birhmode cesarean section	0.031	0.088	0.025	0.358	7.210E-01	-0.142	0.205
1 older sibling	0.005	0.087	0.004	0.054	9.570E-01	-0.168	0.177
≥ 2 older siblings	0.157	0.154	0.081	1.019	3.090E-01	-0.147	0.461
Sex	0.018	0.077	0.016	0.236	8.140E-01	-0.133	0.170
Birth Weight	0.000	0.000	-0.119	-1.573	1.170E-01	0.000	0.000
Gestational age	-0.034	0.033	-0.079	-1.035	3.020E-01	-0.099	0.031
Breastfeeding duration	-0.001	0.002	-0.046	-0.587	5.580E-01	-0.004	0.002
Ethnicity	0.100	0.231	0.031	0.431	6.670E-01	-0.356	0.555
Maternal age delivery	-0.004	0.008	-0.036	-0.493	6.230E-01	-0.021	0.013
AD father	0.020	0.094	0.016	0.208	8.350E-01	-0.166	0.205
AD mother	0.124	0.107	0.092	1.154	2.500E-01	-0.088	0.336
Asthma father	-0.012	0.103	-0.008	-0.113	9.100E-01	-0.214	0.191

Shannon index at age 13 weeks (N = 231)							
Unstandardized Coefficients	Std. Error		Standardized	t	P-value	95,0% Confidence Interval for B	
			Beta			Lower Bound	Upper Bound
Asthma mother	-0.006	0.087	-0.005	-0.064	9.490E-01	-0.176	0.165
Smoking in pregnancy	-0.053	0.091	-0.042	-0.583	5.610E-01	-0.234	0.127
Smoking after pregnancy	-0.035	0.092	-0.028	-0.380	7.040E-01	-0.217	0.147
AD	-0.074	0.084	-0.062	-0.888	3.760E-01	-0.239	0.091
Household with furry pets	-0.034	0.079	-0.029	-0.429	6.680E-01	-0.191	0.123
Start of daycare in months	-0.006	0.004	-0.101	-1.452	1.480E-01	-0.014	0.002
Age of introduction solids (months)	-0.006	0.010	-0.047	-0.605	5.460E-01	-0.026	0.014
antibiotic use prior to sample collection	0.001	0.012	0.006	0.094	9.250E-01	-0.023	0.026

Shannon index at age 21 weeks (N = 228)							
Unstandardized Coefficients	Std. Error		Standardized	t	P-value	95,0% Confidence Interval for B	
			Beta			Lower Bound	Upper Bound
(Constant)	3.346	1.192		2.807	5.000E-03	0.996	5.696
Birhmode assisted vaginal	-0.156	0.154	-0.067	-1.009	3.140E-01	-0.460	0.148
Birhmode cesarean section	-0.074	0.082	-0.061	-0.902	3.680E-01	-0.236	0.088
1 older sibling	0.042	0.081	0.038	0.526	6.000E-01	-0.117	0.201
≥ 2 older siblings	0.005	0.133	0.003	0.037	9.710E-01	-0.257	0.267
Sex	0.083	0.070	0.079	1.183	2.380E-01	-0.055	0.222
Birth Weight	0.000	0.000	-0.056	-0.757	4.500E-01	0.000	0.000
Gestational age	0.004	0.031	0.009	0.122	9.030E-01	-0.058	0.065
Breastfeeding duration**	<b>-0.004</b>	<b>0.001</b>	<b>-0.199</b>	<b>-2.685</b>	<b>8.000E-03</b>	<b>-0.007</b>	<b>-0.001</b>
Ethnicity	-0.207	0.235	-0.058	-0.882	3.790E-01	-0.669	0.255
Maternal age delivery	0.000	0.008	-0.002	-0.024	9.810E-01	-0.015	0.015
AD father	-0.111	0.088	-0.094	-1.265	2.070E-01	-0.284	0.062
AD mother	-0.069	0.097	-0.054	-0.712	4.780E-01	-0.260	0.122
Asthma father	-0.012	0.089	-0.010	-0.137	8.910E-01	-0.188	0.164
Asthma mother	0.059	0.080	0.053	0.729	4.670E-01	-0.100	0.217
Smoking in pregnancy	-0.058	0.087	-0.046	-0.661	5.090E-01	-0.230	0.115
Smoking after pregnancy	-0.060	0.088	-0.050	-0.686	4.930E-01	-0.233	0.113
AD	-0.100	0.075	-0.091	-1.334	1.840E-01	-0.248	0.048
Household with furry pets	0.036	0.076	0.032	0.477	6.340E-01	-0.113	0.186
Start of daycare in months	0.003	0.004	0.060	0.882	3.790E-01	-0.004	0.011
Age of introduction solids (months)**	<b>-0.028</b>	<b>0.009</b>	<b>-0.224</b>	<b>-3.086</b>	<b>2.000E-03</b>	<b>-0.046</b>	<b>-0.010</b>
antibiotic use prior to sample collection	0.109	0.138	0.055	0.793	4.290E-01	-0.162	0.380



Supplementary Table S9 - Continued

Shannon index at age 31 weeks (N = 253)							
Unstandardized Coefficients	Std. Error		Standardized	t	P-value	95,0% Confidence Interval for B	
			Beta			Lower Bound	Upper Bound
(Constant)	2.551	1.064		2.399	1.700E-02	0.456	4.647
Birhmode assisted vaginal	-0.012	0.142	-0.005	-0.082	9.340E-01	-0.292	0.269
Birhmode cesarean section	0.009	0.076	0.008	0.122	9.030E-01	-0.140	0.158
1 older sibling	0.104	0.075	0.091	1.386	1.670E-01	-0.044	0.251
≥ 2 older siblings*	<b>0.246</b>	<b>0.124</b>	<b>0.136</b>	<b>1.978</b>	<b>4.900E-02</b>	<b>0.001</b>	<b>0.491</b>
Sex	-0.029	0.065	-0.028	-0.448	6.550E-01	-0.157	0.099
Birth Weight	0.000	0.000	0.017	0.248	8.040E-01	0.000	0.000
Gestational age	0.021	0.028	0.051	0.757	4.500E-01	-0.033	0.075
Breastfeeding duration**	<b>-0.005</b>	<b>0.001</b>	<b>-0.238</b>	<b>-3.449</b>	<b>1.000E-03</b>	<b>-0.008</b>	<b>-0.002</b>
Ethnicity	-0.202	0.203	-0.063	-0.996	3.200E-01	-0.602	0.198
Maternal age delivery	-0.007	0.007	-0.060	-0.915	3.610E-01	-0.021	0.008
AD father	-0.073	0.082	-0.060	-0.885	3.770E-01	-0.235	0.089
AD mother	0.072	0.089	0.056	0.810	4.190E-01	-0.104	0.249
Asthma father	0.092	0.084	0.071	1.106	2.700E-01	-0.072	0.257
Asthma mother	0.068	0.074	0.061	0.918	3.600E-01	-0.078	0.215
Smoking in pregnancy	0.031	0.079	0.025	0.391	6.960E-01	-0.125	0.187
Smoking after pregnancy	0.002	0.079	0.002	0.025	9.800E-01	-0.155	0.158
AD	-0.178	0.070	-0.158	-2.525	1.200E-02	-0.317	-0.039
Household with furry pets	-0.010	0.069	-0.009	-0.148	8.820E-01	-0.145	0.125
Start of daycare in months	0.000	0.004	0.002	0.028	9.780E-01	-0.007	0.007
Age of introduction solids (months)**	<b>-0.029</b>	<b>0.008</b>	<b>-0.232</b>	<b>-3.459</b>	<b>1.000E-03</b>	<b>-0.046</b>	<b>-0.013</b>
antibiotic use prior to sample collection	-0.001	0.106	-0.001	-0.010	9.920E-01	-0.209	0.207

Supplementary Table S10 - Joint modelling on the association between longitudinal microbial diversity (Shannon index) and the time to development of atopic dermatitis (AD)

Joint modelling on MAZ scores and atopic dermatitis (N = 312)				
Effect	Value	Std. Err.	z-value	p-value
Treatment (reference = placebo)	-0.06407407	0.211695794	-0.30267049	7.62E-01
Gender (reference = boy)	0.034598927	0.211291403	0.16374981	8.70E-01
AD father (reference = no)	0.906944883	0.290099079	3.126328035	1.77E-03
Breastfeeding duration (weeks)	0.007557587	0.00364736	2.072070356	3.83E-02
1 older sibling (reference = no sibs)	0.154546605	0.238565208	0.647817031	5.17E-01
> 2 older siblings	-0.034512157	0.375489343	-0.091912482	9.27E-01
Furry pets in household (reference = no)	-0.078684706	0.23136654	-0.340086798	7.34E-01
Cesarean section birth mode (reference = natural vaginal delivery)	-0.12821753	0.474555339	-0.27018457	7.87E-01
Assisted vaginal delivery	-0.343478873	0.267486101	-1.284099886	1.99E-01
Birth weight (in grams)	0.510627325	0.264140586	1.933164954	5.32E-02
AD mother (reference = no)	0.766246518	0.305654581	2.506903427	1.22E-02
ASSOCIATION (MAZ scores)	<b>0.127799617</b>	<b>0.029913897</b>	<b>4.272249057</b>	<b>1.94E-05</b>

log(xi.1)	-4.595666187	1.097884853	-4.185927309	2.84E-05
log(xi.2)	-4.671636806	1.098344847	-4.253342489	2.11E-05
log(xi.3)	-2.512362383	1.302212452	-1.929302995	5.37E-02
log(xi.4)	-7.065394817	8240.266544	-0.000857423	9.99E-01

**Supplementary Table S11 - Joint modelling on the association between longitudinal microbial maturity (MAZ score) and the time to development of atopic dermatitis (AD)**

Joint modelling on MAZ scores and atopic dermatitis (N = 312)				
Effect	Value	Std. Err.	z-value	p-value
Treatment (reference = placebo)	-0.06407407	0.211695794	-0.30267049	7.62E-01
Gender (reference = boy)	0.034598927	0.211291403	0.16374981	8.70E-01
AD father (reference = no)	0.906944883	0.290099079	3.126328035	1.77E-03
Breastfeeding duration (weeks)	0.007557587	0.00364736	2.072070356	3.83E-02
1 older sibling (reference = no sibs)	0.154546605	0.238565208	0.647817031	5.17E-01
<u>&gt; 2 older siblings</u>	-0.034512157	0.375489343	-0.091912482	9.27E-01
Furry pets in household (reference = no)	-0.078684706	0.23136654	-0.340086798	7.34E-01
Cesarean section birth mode (reference = natural vaginal delivery)	-0.12821753	0.474555339	-0.27018457	7.87E-01
Assisted vaginal delivery	-0.343478873	0.267486101	-1.284099886	1.99E-01
Birth weight (in grams)	0.510627325	0.264140586	1.933164954	5.32E-02
AD mother (reference = no)	0.766246518	0.305654581	2.506903427	1.22E-02
ASSOCIATION (MAZ scores)	<b>0.127799617</b>	<b>0.029913897</b>	<b>4.272249057</b>	<b>1.94E-05</b>
log(xi.1)	-4.595666187	1.097884853	-4.185927309	2.84E-05
log(xi.2)	-4.671636806	1.098344847	-4.253342489	2.11E-05
log(xi.3)	-2.512362383	1.302212452	-1.929302995	5.37E-02
log(xi.4)	-7.065394817	8240.266544	-0.000857423	9.99E-01

**Supplementary Table S12 - METALONDA analyses on microbial genera in association to atopic dermatitis**

MetaLonDa Analysis Output in association to Atopic Dermatitis				
Bacterial Genera	start (days)	end (days)	dominant	FDR adjusted p-value
<b>Corynebacterium</b>	126.110	151.240	HC	9.68E-03
<b>Atopobium</b>	25.590	79.440	HC	7.65E-03
<b>Prevotella</b>	104.570	133.290	HC	0.00E+00
<b>Lachnobacterium</b>	25.590	61.490	HC	1.18E-02
<b>Lachnobacterium</b>	108.160	129.700	HC	6.49E-03
<b>Lachnobacterium</b>	208.680	269.710	HC	5.00E-03
<b>Faecalibacterium</b>	47.130	201.500	HC	5.08E-03

**Supplementary Table S13 - Logistic regression model of microbial diversity (Shannon index) in association to allergic sensitization at school-age**

<b>Logistic regression model of microbial diversity (Shannon index) in association to allergic sensitization at school-age</b>				
<b>Coefficients</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>P-value</b>
(Intercept)	5.627791598	3.674612075	1.531533529	1.26E-01
Shannon at 5 weeks	-0.410285084	0.454296315	-0.903122193	3.66E-01
Shannon at 13 weeks	-0.208920785	0.527504064	-0.396055309	6.92E-01
Shannon at 21 weeks	0.279310503	0.601239738	0.464557623	6.42E-01
Shannon at 31 weeks	-1.620866258	0.604467106	-2.681479675	7.33E-03
Breastfeeding duration (weeks)	-3.030233351	0.873149718	-3.470462497	5.20E-04
Solid food introduction (week)	6.294484319	3.381347875	1.861531126	6.27E-02
Cesarean section	-0.238619825	1.081714761	-0.220594036	8.25E-01
Assisted vaginal delivery	-0.94531093	0.626875226	-1.507973023	1.32E-01
Birth Weight (g)	-0.000366139	0.00065891	-0.555673949	5.78E-01
Sex Female (reference = male)	-0.527185935	0.500811149	-1.052664135	2.92E-01
Treatment Control (reference = yes)	0.742100392	0.503761986	1.473117091	1.41E-01
No AD Father (reference=Yes)	-1.635728509	0.645015214	-2.535953375	1.12E-02
No AD Mother (reference=Yes)	-0.655147341	0.663647229	-0.987192159	3.24E-01
1 older sibling (reference = no sibs)	0.45591163	0.568647631	0.801747172	4.23E-01
≥ 2 olders siblings	-6.44E-05	0.999108354	-6.44E-05	1.00E+00
Furry pets in household (reference = no)	-0.72774705	0.548034524	-1.327921908	1.84E-01

**Supplementary Table S14 - Logistic regression model of microbial maturity (MAZ score) in association to allergic sensitization at school-age**

<b>Logistic regression model of microbial maturity (MAZ-score) in association to allergic sensitization at school-age</b>				
<b>Coefficients</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>P-value</b>
(Intercept)	-0.087464637	2.640818493	-0.033120276	9.74E-01
MAZ at 5 weeks	0.382210056	0.136202238	2.806195127	5.01E-03
MAZ at 13 weeks	-0.164301672	0.140623631	-1.16837882	2.43E-01
MAZ at 21 weeks	0.050549523	0.182113301	0.277571833	7.81E-01
MAZ at 31 weeks	-0.016145267	0.133545263	-0.120897342	9.04E-01
Breastfeeding duration (weeks)	-3.026775739	0.884920175	-3.420394091	6.25E-04
Solid food introduction (week)	8.847895913	3.301639509	2.679849175	7.37E-03
Birth Weight (g)	-0.00039128	0.000649369	-0.602553926	5.47E-01
Treatment Control (reference = yes)	0.957059287	0.534603349	1.790223143	7.34E-02
Sex Female (reference = male)	-0.402875364	0.508750997	-0.791891054	4.28E-01
Cesarean section	-0.236701605	1.019008403	-0.232286215	8.16E-01
Assisted vaginal delivery	-0.842345125	0.656241797	-1.283589569	1.99E-01
No AD Father (reference=Yes)	-1.43835435	0.64026914	-2.246483953	2.47E-02
No AD Mother (reference=Yes)	-0.52709629	0.640499938	-0.822945108	4.11E-01
1 older sibling (reference = no sibs)	0.119864913	0.552340617	0.217012672	8.28E-01
≥ 2 olders siblings	-0.745760306	1.048383644	-0.711342942	4.77E-01
Furry pets in household (reference = no)	-0.840188796	0.565073526	-1.486866323	1.37E-01

**Supplementary Table S15 - Logistic regression model of microbial diversity (Shannon index) in association to asthma at school-age**

<b>Logistic regression model of microbial diversity (Shannon index) in association to asthma at school-age</b>				
<b>Coefficients</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>P-value</b>
(Intercept)	-9.519296394	5.319877537	-1.789382618	7.36E-02
Shannon at 5 weeks	-0.431558541	0.692669049	-0.623037137	5.33E-01
Shannon at 13 weeks	0.065089728	0.751687765	0.086591443	9.31E-01
Shannon at 21 weeks	-0.02811964	0.661640302	-0.04249989	9.66E-01
Shannon at 31 weeks	0.164692463	0.801999013	0.205352451	8.37E-01
Breastfeeding duration (weeks)	-0.254586729	0.764690582	-0.332927768	7.39E-01
Solid food introduction (week)	6.194863693	5.019759796	1.234095643	2.17E-01
Cesarean section	0.344889489	1.456595944	0.236777735	8.13E-01
Birth Weight (g)	0.001270255	0.00090951	1.396637011	1.63E-01
Assisted vaginal delivery	-0.825554535	0.948850266	-0.870057758	3.84E-01
Treatment Control (reference = yes)	-0.166071147	0.657007741	-0.252768936	8.00E-01
Sex Female (reference = male)	-2.051878298	0.859550668	-2.387152234	1.70E-02
No AD Father (reference=Yes)	0.085116272	0.823378904	0.103374365	9.18E-01
No AD Mother (reference=Yes)	-0.313860614	1.003653848	-0.31271799	7.54E-01
1 older sibling (reference = no sibs)	0.643148333	0.754705643	0.852184343	3.94E-01
≥2 olders siblings	-15.68707851	1926.712049	-0.00814189	9.94E-01
Furry pets in household (reference = no)	1.483880523	0.934270057	1.588277941	1.12E-01

**Supplementary Table S16 - Logistic regression model of microbial maturity (MAZ score) in association to asthma at school-age**

<b>Logistic regression model of microbial maturity (MAZ-score) in association to asthma at school-age</b>				
<b>Coefficients</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>P-value</b>
(Intercept)	-12.40821269	5.470392851	-2.268248923	2.33E-02
MAZ at 5 weeks	0.363442479	0.136558324	2.661445086	7.78E-03
MAZ at 13 weeks	-0.415956399	0.187305674	-2.220735705	2.64E-02
MAZ at 21 weeks	0.254280302	0.228921442	1.110775385	2.67E-01
MAZ at 31 weeks	-0.104036691	0.157184716	-0.661875362	5.08E-01
Breastfeeding duration (weeks)	-0.486254724	0.839838341	-0.578986099	5.63E-01
Solid food introduction (week)	10.08291775	5.955334706	1.69309002	9.04E-02
Cesarean section	1.480313854	1.561995355	0.94770695	3.43E-01
Assisted vaginal delivery	-1.1528046	1.0910663	-1.056585287	2.91E-01
Birth Weight (g)	0.001070665	0.000970668	1.103018401	2.70E-01
Treatment Control (reference = yes)	0.126012384	0.750777261	0.167842569	8.67E-01
Sex Female (reference = male)	-3.129532598	1.16397764	-2.688653536	7.17E-03
No AD Father (reference=Yes)	0.731989105	0.882583537	0.829370903	4.07E-01
No AD Mother (reference=Yes)	-1.282273155	1.392172865	-0.921058862	3.57E-01
1 older sibling (reference = no sibs)	0.961383693	0.768278099	1.25134856	2.11E-01
≥2 olders siblings	-14.55804809	1838.156717	-0.007919917	9.94E-01
Furry pets in household (reference = no)	1.829645625	1.082872207	1.68962285	9.11E-02

**Supplementary Table S17 - METALONDA analyses on microbial genera in association to asthma**

<b>MetaLonDa Analysis Output in association to Asthma</b>				
<b>Bacterial Genera</b>	<b>start (days)</b>	<b>end (days)</b>	<b>dominant</b>	<b>FDR adjusted p-value</b>
Actinomyces	183.040	192.000	HC	1.809E-02
Enterococcus	167.360	194.240	HC	3.307E-02
Streptococcus	44.160	129.280	Asthma	1.741E-02
Streptococcus	183.040	236.800	HC	3.156E-03
Sarcina	28.480	57.600	HC	8.062E-03
Sarcina	84.480	86.720	HC	4.844E-02
Sarcina	88.960	93.440	HC	4.521E-02
Lachnobacterium	28.480	100.160	HC	1.365E-02
Lachnobacterium	138.240	176.320	HC	4.618E-03
Lachnobacterium	221.120	236.800	HC	2.783E-02
Lachnospira	48.640	80.000	HC	1.717E-02
Lachnospira	91.200	200.960	HC	2.878E-03
Lachnospira	218.880	236.800	HC	1.040E-03
Dialister	28.480	122.560	HC	6.209E-04
Dialister	131.520	144.960	Asthma	1.378E-02
Dialister	151.680	176.320	HC	1.200E-02
Dialister	203.200	207.680	HC	4.823E-02
Dialister	216.640	236.800	HC	1.935E-02
Klebsiella	28.480	102.400	HC	6.922E-03
Leclercia	28.480	187.520	HC	2.746E-03
Haemophilus	115.840	236.800	HC	1.038E-02





Nicolas Jongejan (5 years)

# Chapter 6

## **Gut microbiota in wheezing preschool children and the association with childhood asthma**

**Niels van Best\***, Michiel A. G. E. Bannier\*, Liene Bervoets, Paul H.M. Savelkoul, Mathias W. Hornef, Kim D. G. van de Kant, Quirijn Jöbsis, Edward Dompeling‡, John Penders‡

\*equal contribution / ‡equal contribution

*Allergy. 2019 Dec 15.*



## Abstract

Reliable biomarkers to predict asthma in wheezing preschool children are lacking. We examined whether gut microbial composition in preschool wheezers was associated with asthma, and investigated associations between gut microbiota and atopic sensitisation, regulatory T-cells (Tregs), and *Foxp3* gene expression.

In the Asthma Detection and Monitoring study, 202 wheezing children aged 2-4 years and 50 healthy controls were followed until the age of 6 years, when a diagnosis of asthma was made. At age 2-4 years, faecal microbial composition was analysed by sequencing of the 16S rRNA V3-V4 gene region. In blood, atopic sensitisation (Phadiatop Infant test), proportion of Tregs (CD4+CD25<sup>high</sup>CD127<sup>-</sup>), and *Foxp3* gene expression were assessed.

We found no difference in gut microbial diversity, richness, or overall microbial community structure between transient wheezers and true asthmatics, or between preschool wheezers and healthy controls. However, the relative abundance of the genera *Collinsella* ( $p=0.01$ ) and *Dorea* ( $p=0.02$ ) was significantly lower in wheezing compared to healthy children, whereas *Gemmiger* ( $p=0.03$ ) and *Escherichia* ( $p=0.02$ ) were significantly higher in preschool wheezers who developed asthma. In particular, a high relative abundance of *Escherichia* was associated with a 4.6-fold increased odds of asthma ( $p=0.02$ ). No significant correlations between abundance of specific bacterial genera and atopic sensitisation, Tregs, and *Foxp3* gene expression were detected.

In conclusion, gut microbial diversity and overall gut microbial community structure at preschool age were not associated with wheezing or future asthma. However, some bacterial genera were associated with wheezing or subsequent asthma, suggesting some microbial dysbiosis in children prone for developing asthma.

## Introduction

Wheezing in preschool children is common, but only one third of wheezing preschool children will eventually develop asthma [1]. Currently, a reliable biomarker to predict which wheezing children will develop asthma is lacking [2]. Consequently, both underdiagnosis of asthma and overtreatment of children with episodic viral wheeze still frequently occur.

In the development of childhood asthma, genetic susceptibility and early life environmental exposures play a major role, leading to an immunological dysbalance favouring a T-helper type 2 (Th2) response [3]. In the past years, the impact of gut microbiota perturbations on this immunological dysbalance has gained widespread attention. Data from birth cohort studies showed that gut microbial dysbiosis in the first year of life was associated with asthma development [4-8]. Recent studies have shown the influence of short-chain fatty acids (SCFAs), originating from bacterial fermentation of dietary fibres in the gut, on regulatory T-cells (Tregs) which are crucial in maintaining an adequate immune balance [9-11]. Moreover, a high-fibre diet increased the levels of SCFAs in mice, which markedly suppressed allergic airways disease by enhancing Treg numbers and function [12]. This protective effect was partly explained by increasing acetylation at the *Foxp3* promoter region [12].

These studies show our increased understanding regarding the role of the gut microbiota in asthma development. However, the majority of studies have been performed in animals, whereas most human studies have focused on the association between infant gut microbiota and asthma-like symptoms at an age when a reliable diagnosis of asthma cannot yet be made. Finally, no studies have been performed that investigated gut microbial composition in wheezing children, and its association with subsequent development of asthma.

In the current study, we profiled the gut microbiota in a cohort of wheezing and healthy preschool children, which were prospectively followed until the age of 6 years when a definitive diagnosis of asthma was made. Our aims were to: (1) investigate differences in gut microbial composition between wheezing preschool children and healthy controls, (2) assess whether the gut microbiota in wheezing preschool children were associated with the development of childhood asthma, (3) investigate associations between the gut microbial profiles and atopic sensitisation, proportion of Tregs, and *Foxp3* gene expression.

## Methods

### Study Design

In the Asthma Detection and Monitoring (ADEM) study (clinicaltrials.gov: NCT 00422747), 202 wheezing children and 50 healthy controls aged 2 to 4 years were prospectively followed until 6 years of age [13, 14]. The primary goal of the ADEM study was to assess the diagnostic accuracy of various tests. Moreover, important pathogenic pathways for the early development of asthma were studied. A detailed study protocol with in- and exclusion criteria was previously published [15]. The study was approved by the Dutch national medical ethical committee, and written informed consent was given by all parents.

At inclusion, clinical characteristics, and faecal and blood samples were collected. Clinical characteristics were collected using validated questionnaires [16, 17].

At the age of 6 years, a clinical diagnosis (healthy, transient wheeze or true asthma) was made by two experienced paediatric pulmonologists and a computer based algorithm [18].

### DNA isolation and amplicon sequencing of the 16S rRNA V3-V4 gene region

By using a standardised protocol, faeces of all participating children was collected at inclusion [19]. After reception, samples were immediately frozen at -80 °C until further analysis. DNA isolation was performed using a combination of mechanical (repeated bead beating) and chemical lysis followed by column-based purification as described previously [20]. For each DNA isolation batch, additional isolation was performed on PCR-grade water as a negative control. Amplicon libraries of the 16S rRNA V3-V4 gene region and subsequent sequencing (Illumina, MiSeq) was performed by Baseclear [Leiden, the Netherlands], see supplementary material for more details. Data demultiplexing, length and quality filtering, pairing of reads and clustering of reads into Operational Taxonomic Units (OTUs) at 97% sequence identity was done using the online Integrated Microbial Next Generation Sequencing (IMNGS) platform [21], using default settings except for minimum and maximum length for amplicons which were set at 300 and 600 bp respectively. IMNGS is based upon the UPARSE algorithm [22].

Data normalisation, diversity, taxonomical binning and group comparisons were performed using Rhea package [23]. In order not to discard informative data, normalisation in Rhea is performed by dividing OTU counts per sample for their total count (sample depth) followed by multiplying all of the obtained relative abundance for the lowest sample depth

(8,278 reads/sample). Alpha- (observed species, Chao1 and Shannon index) and beta-diversity (generalised Unifrac, Bray-Curtis dissimilarity) estimates were computed in R [24].

### Atopy

Atopy was defined positive when a specific IgE concentration against a mixture of inhalant and food allergens was  $\geq 0.35$  kU/L (Phadiatop Infant test; Phadia, Uppsala, Sweden). IgE-mediated asthma was defined as children with a diagnosis of asthma at age 6 years, and specific IgE levels  $\geq 0.35$  kU/l against one or more aeroallergens as determined by the Phadiatop Infant test.

### Regulatory T cells

Tregs ( $CD4^+CD25^{\text{high}}CD127^-$ ) were determined by flow cytometry. As published before, the  $CD4^+CD25^{\text{high}}CD127^-$  phenotype was shown to be a valid surrogate for *Foxp3* positive Tregs in our study population [25].

### Foxp3 gene expression

Total RNA was extracted from peripheral blood mononuclear cells (PBMC), and expression of *Foxp3* gene was assessed. A detailed study procedure was previously published [14].

### Data analysis

IBM SPSS Statistics 20, SAS 7.30 and R statistical software were used for data analysis. To test for statistical significant differences in the overall microbial community structure between study groups, permutational multivariate analysis of variance (PERMANOVA) using distance matrices (vegan::adonis) were performed. Enterotyping, using Dirichlet Multinomial Mixture (DMM) clustering, was performed as described previously [26], in order to cluster faecal samples based upon specific microbial community structures. To examine whether the study group associated with specific enterotypes, multivariable logistic regression analyses were performed while adjusting for potential confounders (sex, breastfeeding, birth season, parental atopy, siblings, parental smoking status, day care). To test whether specific microbial taxa and/or microbial diversity differed between study groups (e.g. healthy controls vs. wheeze or transient wheeze vs. asthma), Kruskal-Wallis tests and Mann-Whitney U tests were used. For those bacteria that appeared to be statistically significantly different between the study groups in univariable analyses, multivariable logistic regression analyses were performed to examine whether these associations remained upon adjustment for potential confounders

(see above). To identify possible relations between microbial richness, diversity and genus abundance and continuous outcomes (i.e. *Foxp3* gene-expression and Tregs) spearman correlations were used.

## Results

Faecal and blood samples were collected. Faecal microbial composition was analysed by sequencing of the 16S rRNA V3-V4 gene region (see flow chart, Figure S1). A total of 12,230,374 V3-V4 reads were generated. After trimming, quality filtering, removal of potential chimeric reads, de-multiplexing, and removal of low abundant operational taxonomic units (OTUs), 6,885,888 paired-end sequences belonging to 336 OTUs were retained for downstream analysis. Samples of three participants failed during sequencing and were excluded from subsequent analysis. For the remaining samples of 230 children (70 asthmatics, 114 transient wheezers, and 46 healthy controls), the number of sequences per sample ranged from 2,362 to 44,357 (median 30,471). In blood, atopic sensitisation (Phadiatop Infant test), proportion of Tregs by flow cytometry (CD4+CD25<sup>high</sup>CD127<sup>-</sup>), and *Foxp3* gene expression were assessed. The baseline characteristics are displayed in Table S1.

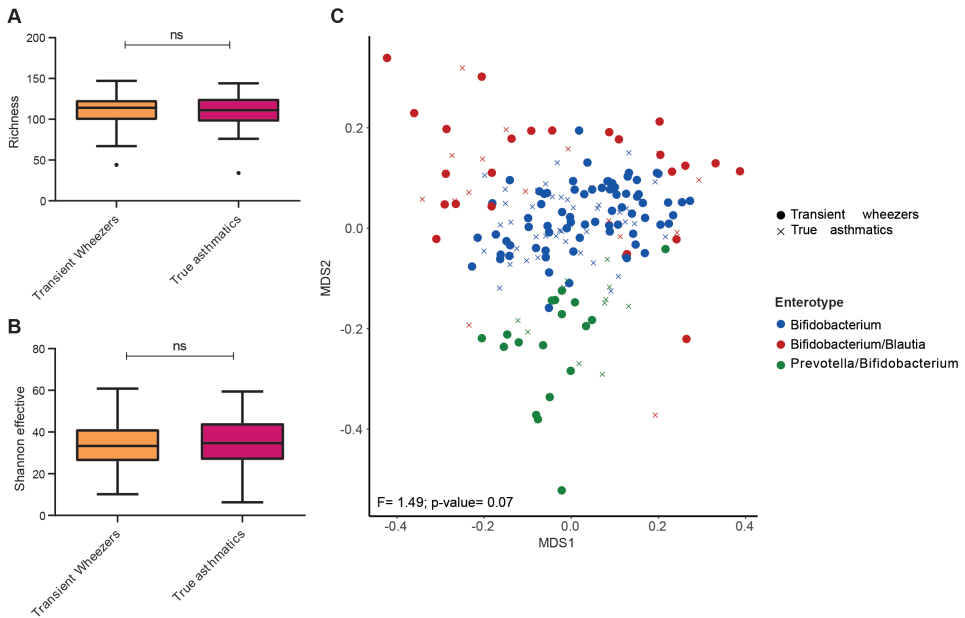
### Microbiota richness and diversity

First, we examined whether microbial richness and diversity at preschool age were predictive for future asthma development. Neither the microbial richness (OR 0.99 [95%CI 0.98-1.01];  $P = .46$ ), nor the microbial diversity as assessed by the Shannon index (OR 1.01 [0.98-1.04];  $P = .53$ ) were significantly different between transient wheezing children and true asthmatics while adjusting for potential confounders (sex, breastfeeding, birth season, atopy parents, siblings, parental smoking status, day care attendance; Figure 1A-B). At preschool age, these indices were also not different between wheezers and healthy controls, while adjusting for potential confounders (Figure S2).

### Enterotypes

Next, we examined the overall microbial community structure (as assessed by the Bray-Curtis dissimilarity), which was neither significantly different between transient wheezers and true asthmatics (PERMANOVA  $P = .07$ , Figure 1C), nor between preschool wheezers and healthy controls (PERMANOVA  $P = .22$ , Figure S2). Microbial profiles of all children were clustered using

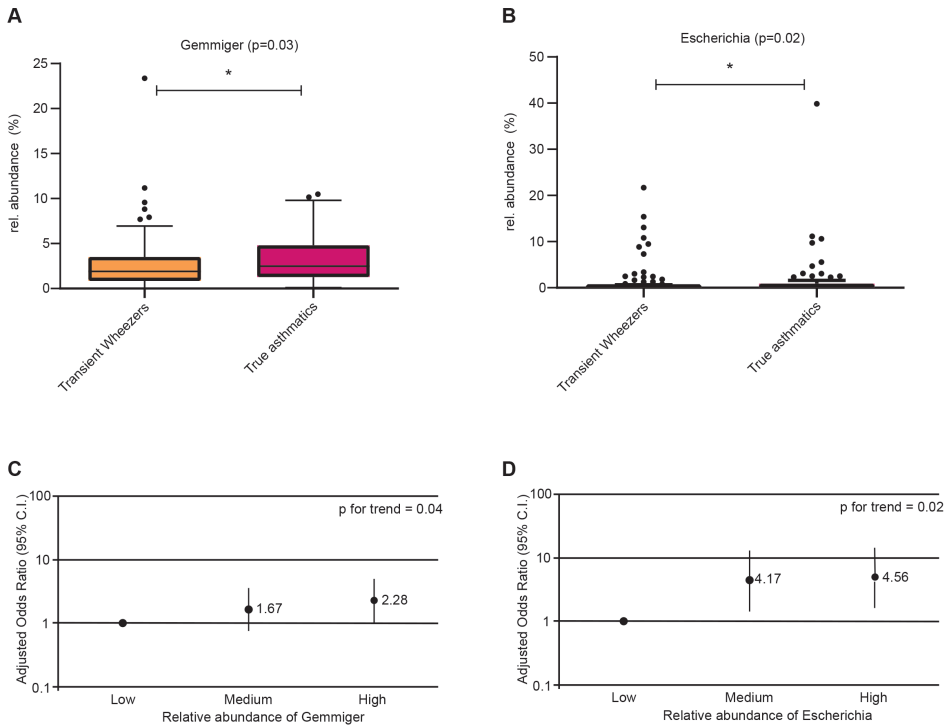
Dirichlet Multinomial Mixture (DMM) modelling. Three distinct clusters (enterotypes) were identified, that is those that were driven by a relatively high abundance of Bifidobacterium, Bifidobacterium combined with Blautia, and Prevotella combined with Bifidobacterium, respectively. Neither the number of wheezing children who developed asthma, nor the proportion of children with preschool wheeze were significantly different among the three enterotypes, while adjusting for multiple confounders in multivariable logistic regression analyses (Table S2 and S3). Altogether these results indicate that microbial diversity and overall microbial community structure are not predictive for subsequent asthma development among preschool wheezing children.



**Figure 1** - Microbial richness, diversity and community structure among preschool wheezing children who did (true asthmatics) or did not (transient wheezers) subsequently develop asthma. Microbial richness (observed species) (A) and diversity (Shannon effector index) (B) are not significantly different between transient wheezers and true asthmatics (ns: nonsignificant; Kruskal-Wallis). C, Multidimensional scaling (MDS) based on Bray-Curtis dissimilarity indicates three different enterotypes driven by Bifidobacterium, Bifidobacterium/Blautia and Prevotella/Bifidobacterium. Overall, microbial community structure is not statistically significantly different between transient wheezers and true asthmatics (permutational analysis of variance [PERMANOVA])

### Specific microbial taxa

Furthermore, we examined whether the relative abundance of specific bacterial genera was predictive for future asthma development. Using multivariable logistic regression, we found that the relative abundance of the genera *Gemmiger* ( $P = .03$ ) and *Escherichia* ( $P = .02$ ) was significantly higher in wheezing children who developed asthma at age 6 years (Figure 2A-B). The risk of developing asthma was highest in those children who harboured the highest relative abundance of these two bacterial genera (Figure 2C-D). In particular, a high relative abundance of *Escherichia* was associated with 4.6-fold increased odds of asthma ( $P = .02$ , Figure 2D). When comparing preschool wheezers with healthy controls, the relative abundance of *Collinsella* ( $P = .01$ ) and *Dorea* ( $P = .02$ ) was significantly lower in wheezing children (Figure S3).



**Figure 2** - The relative abundance of specific microbial taxa increases the risk of subsequent asthma. The relative abundance of bacterial genera *Gemmiger* ( $P = .03$ ) and *Escherichia* ( $P = .02$ ) is higher in true asthmatic children compared with transient wheezers (A, B). Multiple logistic regression analyses show that the higher risk to develop asthma among children with a high abundance of *Gemmiger* and *Escherichia* remained statistically significant upon adjustment for sex, breastfeeding, birth season, atopy parents, siblings, parental smoking status and day care attendance (C, D)

### Gut microbiota and sensitization, Tregs and *Foxp3* gene expression

Finally, we examined whether gut microbial profiles were related to atopic sensitisation, Tregs and *Foxp3* gene expression. Besides a weak, yet statistically significant, positive correlation between *Foxp3* gene expression and bacterial diversity (Shannon index) within the entire study population (Spearman's  $\rho = 0.16$ ;  $P = .02$ ), atopic sensitisation, Tregs and *Foxp3* gene expression were neither associated with the overall microbial community structure (Bray-Curtis dissimilarity) nor with the abundance of specific bacterial genera.

### Discussion

In this study, we investigated the gut microbiota in wheezing preschool children and healthy controls at 2-4 years of age and assessed its association with asthma at age 6 years. To the best of our knowledge, no previous studies have examined the microbiota in wheezing preschool children and its association with asthma progression. We did not observe a difference in gut microbial diversity, richness, or overall microbial community structure between transient wheezers and true asthmatics, or between preschool wheezers and healthy controls. However, on a genus level, some bacterial genera were associated with wheezing (*Collinsella* and *Dorea*) or with the subsequent development of asthma (*Gemmiger* and *Escherichia*). No significant correlations between the abundance of specific bacterial genera and atopic sensitisation, Tregs, and *Foxp3* gene expression were detected.

The ADEM study was designed to assess the diagnostic accuracy of various tests in the early prediction of asthma in wheezing preschool children and to investigate basic pathogenetic mechanisms underlying an early asthma development. Previously, we reported on a high diagnostic accuracy for an early asthma diagnosis (95% area under the receiver operating characteristic curve) by combining exhaled volatile organic compounds in exhaled breath with expression of inflammation genes, and the Asthma Predictive Index [14].

The present study revealed that gut microbial richness, diversity and overall community structure at the age of 2-4 years were, however, not suitable to predict asthma. These findings suggest that at an age beyond infancy, when the gut microbiome is largely stabilised, microbiota perturbations might only be modest. This is in line with the proposed window-of-opportunity, the first months of life, during which the microbiome is thought to have its strongest impact on immune maturation and tolerance development [27, 28], and underscores the importance of timing of faecal sampling. However, this early time-



window might not be a suitable age to identify biomarkers for asthma prediction as manifestation of wheezing symptoms may not have yet occurred. Moreover, the bacterial genera *Gemmiger* and *Escherichia* were significantly associated with asthma, suggesting that some microbial dysbiosis might still exist at preschool age among wheezing children prone for developing asthma. In particular, a high relative abundance of *Escherichia* in wheezing children was associated with a 4.6-fold increased odds of asthma at the age of six years.

Given the lack of similar studies in this age group, we can only compare our findings with those of birth cohort studies assessing the relationship between the infant microbiota and subsequent asthma development. The design of these birth cohort studies differ, however, with respect to timing of faecal sampling, geographical influences, differences in environmental exposures (e.g. mode of delivery, diet, exposure to antibiotics), and heterogeneity in the definitions of asthma. Only in two Scandinavian birth cohorts, asthma diagnosis was robust and in accordance with international guidelines [6, 29]. In other studies, asthma was diagnosed at a younger age than a reliable diagnosis is normally made, or diagnosis was based on parentally reported symptoms [4, 5, 7, 8].

Regarding gut microbial diversity, a lower diversity at age 1 week and 1 month was associated with asthma in a small Swedish birth cohort [29]. However, this could not be replicated in two other studies [4, 6]. In the COPSAC<sub>2010</sub> study, beta-diversity (but not alpha-diversity) at age 12 months was significantly related to asthma development [6]. With respect to gut microbial composition, alterations in early life were associated with asthma development [4-8]. Microbial alterations at age 1 and 3 months were associated with asthma at age 4 in two North-American birth cohorts [4, 7, 8]. Gut microbial alterations at 12 months were associated with asthma at age 5 in the COPSAC<sub>2010</sub> study, predominantly in children born to asthmatic mothers [6]. In contrast, no significant association between gut microbial changes in the first month of life and asthma were found [6]. Beside these differences in timing, some bacterial genera were repeatedly and significantly associated with asthma in these birth cohort studies. In particular, decreased relative abundances of *Faecalibacterium* [4, 6, 7], *Bifidobacterium* [6, 7], and the family *Lachnospiraceae* [4, 6, 8], showed a relationship with asthma.

Although the genus *Escherichia* was not associated with asthma in the COPSAC<sub>2010</sub> study, children who were still in a relatively immature microbiome cluster at the age of 1 year did show a higher asthma risk [6]. This cluster was amongst others driven by a higher abundance of *Enterobacteriaceae*. Moreover, in a recent paediatric study, *Escherichia* was one of three significantly predominant genera in children with asthma (mean age 5.6 years)

compared to healthy controls [30]. Furthermore, in a recent adult study, *Escherichia* was one of two genera that discriminated asthmatics with fixed airway obstruction from those with no airway obstruction [31]. These results are in line with our findings, potentially indicating that the abundance of *Escherichia* may play a role in the early development of asthma.

The increased *Escherichia* abundance might be indicative for delayed microbiota maturation as it has been previously shown that *Escherichia* dominate the microbiota composition in early infancy after which their abundance gradually decreases [6, 32]. On the other hand, an increased number of aerotolerant bacteria, in particular the family of *Enterobacteriaceae*, has also been associated with other types of inflammatory diseases, such as Inflammatory Bowel Disease [33], and it has been suggested that an increased abundance of aerotolerant bacteria might be a nonspecific response to inflammatory conditions [34]. Finally, in a recent study an increase in *Escherichia* appeared to reduce butyrate production, which was associated with asthma [30]. Potentially, *Escherichia* abundance has a wider influence on SCFAs production, thereby causing an immunological dysbalance leading to a Th2-response.

In this study, we tried to unravel pathophysiological mechanisms by combining gut microbiota assessment with atopic sensitisation, Tregs, and *Foxp3* gene expression. However, no significant associations between either of these parameters and genera abundance were observed. The association between atopic sensitisation and gut microbiota was recently reviewed to be inconsistent [35]. In our study, the number of sensitised preschool children might have been too low for a significant result. Regarding Tregs, a possible explanation might be that we assessed the number of Tregs, but not their function. Another potential explaining mechanism comes from two recent murine studies where SCFAs, microbial metabolites of dietary fibre fermentation, influenced lung inflammation by direct or indirect effects on T cells and dendritic cells [36, 37]. These data suggest that gut microbial influences on the development of asthma are not purely restricted to the number and function of Tregs. This might also explain the lack of a positive association with *Foxp3* gene expression, which is a key regulator gene of Tregs.

The strength of our study is that we assessed the gut microbial profiles of a large group of wheezing preschool children, using modern sequencing techniques. To our knowledge no similar study in wheezing preschool children, at an age that the gut microbiome is more mature, has been performed. Another strength is a reliable asthma diagnosis at age 6 based on symptoms, medication use, and lung function measurements.

In conclusion, gut microbial diversity and overall gut microbial community structure at age 2-4 years was not associated with preschool wheezing or future asthma development at

age 6. When compared to microbial perturbations during infancy, microbial perturbations at preschool age appear to be only modestly associated with asthma. On a genus level some bacterial genera were associated with wheezing (*Collinsella* and *Dorea*) or subsequent development of asthma (*Gemmiger* and *Escherichia*), suggesting some microbial dysbiosis in children prone for developing asthma. The role of these genera in the development of asthma warrants further investigation.

## References

1. Martinez, F.D., et al., Asthma and wheezing in the first six years of life. The Group Health Medical Associates. *N Engl J Med*, 1995. 332(3): p. 133-8.
2. Bannier, M.A., et al., Biomarkers to predict asthma in wheezing preschool children. *Clin Exp Allergy*, 2015. 45(6): p. 1040-50.
3. Martinez, F.D. and S. Guerra, Early Origins of Asthma. Role of Microbial Dysbiosis and Metabolic Dysfunction. *Am J Respir Crit Care Med*, 2018. 197(5): p. 573-579.
4. Arrieta, M.C., et al., Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med*, 2015. 7(307): p. 307ra152.
5. Arrieta, M.C., et al., Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting. *J Allergy Clin Immunol*, 2018. 142(2): p. 424-434 e10.
6. Stokholm, J., et al., Maturation of the gut microbiome and risk of asthma in childhood. *Nat Commun*, 2018. 9(1): p. 141.
7. Fujimura, K.E., et al., Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med*, 2016. 22(10): p. 1187-1191.
8. Stiemsma, L.T., et al., Shifts in *Lachnospira* and *Clostridium* sp. in the 3-month stool microbiome are associated with preschool age asthma. *Clin Sci (Lond)*, 2016. 130(23): p. 2199-2207.
9. Smith, P.M., et al., The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science*, 2013. 341(6145): p. 569-73.
10. Furusawa, Y., et al., Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature*, 2013. 504(7480): p. 446-50.
11. Arpaia, N., et al., Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature*, 2013. 504(7480): p. 451-5.
12. Thorburn, A.N., et al., Evidence that asthma is a developmental origin disease influenced by maternal diet and bacterial metabolites. *Nat Commun*, 2015. 6: p. 7320.
13. van de Kant, K.D., et al., Early diagnosis of asthma in young children by using non-invasive biomarkers of airway inflammation and early lung function measurements: study protocol of a case-control study. *BMC Public Health*, 2009. 9(1): p. 210.
14. Klaassen, E.M., et al., Exhaled biomarkers and gene expression at preschool age improve asthma prediction at 6 years of age. *Am J Respir Crit Care Med*, 2015. 191(2): p. 201-7.
15. van de Kant, K.D., et al., Early diagnosis of asthma in young children by using non-invasive biomarkers of airway inflammation and early lung function measurements: study protocol of a case-control study. *BMC Public Health*, 2009. 9: p. 210.
16. Powell, C.V., et al., A parent completed questionnaire to describe the patterns of wheezing and other respiratory symptoms in infants and preschool children. *Arch Dis Child*, 2002. 87(5): p. 376-9.
17. Asher, M.I., et al., Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. *Lancet*, 2006. 368(9537): p. 733-43.
18. Klaassen, E.M., et al., Symptoms, but not a biomarker response to inhaled corticosteroids, predict asthma in preschool children with recurrent wheeze. *Mediators Inflamm*, 2012. 2012: p. 162571.

19. Scheepers, L.E., et al., The intestinal microbiota composition and weight development in children: the KOALA Birth Cohort Study. *Int J Obes (Lond)*, 2015. 39(1): p. 16-25.
20. Costea, P.I., et al., Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol*, 2017. 35(11): p. 1069-1076.
21. Lagkouvardos, I., et al., IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. *Sci Rep*, 2016. 6: p. 33721.
22. Edgar, R.C., UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods*, 2013. 10(10): p. 996-8.
23. Lagkouvardos, I., et al., Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. *PeerJ*, 2017. 5: p. e2836.
24. Chen, J., et al., Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics*, 2012. 28(16): p. 2106-13.
25. Klaassen, E.M., et al., CD14/Toll-like receptors interact with bacteria and regulatory T-cells in the development of childhood asthma. *Eur Respir J*, 2014. 44(3): p. 799-802.
26. Holmes, I., K. Harris, and C. Quince, Dirichlet multinomial mixtures: generative models for microbial metagenomics. *PLoS One*, 2012. 7(2): p. e30126.
27. Yatsunenکو, T., et al., Human gut microbiome viewed across age and geography. *Nature*, 2012. 486(7402): p. 222-7.
28. Stiemsma, L.T. and K.B. Michels, The Role of the Microbiome in the Developmental Origins of Health and Disease. *Pediatrics*, 2018. 141(4).
29. Abrahamsson, T.R., et al., Low gut microbiota diversity in early infancy precedes asthma at school age. *Clin Exp Allergy*, 2014. 44(6): p. 842-50.
30. Chiu, C.Y., et al., Gut microbial-derived butyrate is inversely associated with IgE responses to allergens in childhood asthma. *Pediatr Allergy Immunol*, 2019.
31. Buendia, E., et al., Gut microbiota components are associated with fixed airway obstruction in asthmatic patients living in the tropics. *Sci Rep*, 2018. 8(1): p. 9582.
32. Stewart, C.J., et al., Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature*, 2018. 562(7728): p. 583-588.
33. Gophna, U., et al., Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis. *J Clin Microbiol*, 2006. 44(11): p. 4136-41.
34. Rawls, J.F., Enteric infection and inflammation alter gut microbial ecology. *Cell Host Microbe*, 2007. 2(2): p. 73-4.
35. Zimmermann, P., et al., Association between the intestinal microbiota and allergic sensitization, eczema, and asthma: A systematic review. *J Allergy Clin Immunol*, 2019. 143(2): p. 467-485.
36. Trompette, A., et al., Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat Med*, 2014. 20(2): p. 159-66.
37. Cait, A., et al., Microbiome-driven allergic lung inflammation is ameliorated by short-chain fatty acids. *Mucosal Immunol*, 2018. 11(3): p. 785-795.

## Supplementary Data

**Table S1** - Baseline characteristics of the study population.

Subject Characteristics	Controls (N=46)	Transient wheeze (N=114)	Asthma (N=70)
Age (years), mean (SEM)	3.5 (0.12)	3.3 (0.07)	3.3 (0.08)
Sex: male / female (n/n)	22/24	60/54	42/28
Atopy*	22	18	33
Allergic rhinitis	2	2	7
Eczema**	24	34	44
Use of inhaled corticosteroids***	0	12	30
Use of short acting $\beta_2$ -agonists***	0	28	51
Use of long acting $\beta_2$ -agonists**	0	2	6
Day-care attendance	65	75	73
Exposure to furry pets	57	46	46
Passive smoking	35	28	30
Breast feeding	65	61	73
First degree relative with atopic constitution†	63	72	74

Results are displayed as %, unless stated otherwise. Abbreviation SEM: standard error of the mean. \*Significantly different between children with transient wheeze and children with asthma ( $p < 0.05$ ); \*\*Significantly different between healthy controls and children with asthma ( $p < 0.05$ ); \*\*\*Significantly different between all groups ( $p < 0.05$ ); †First-degree relative with eczema, asthma or allergic rhinitis.

**Table S2.** Association between enterotype status and childhood asthma.

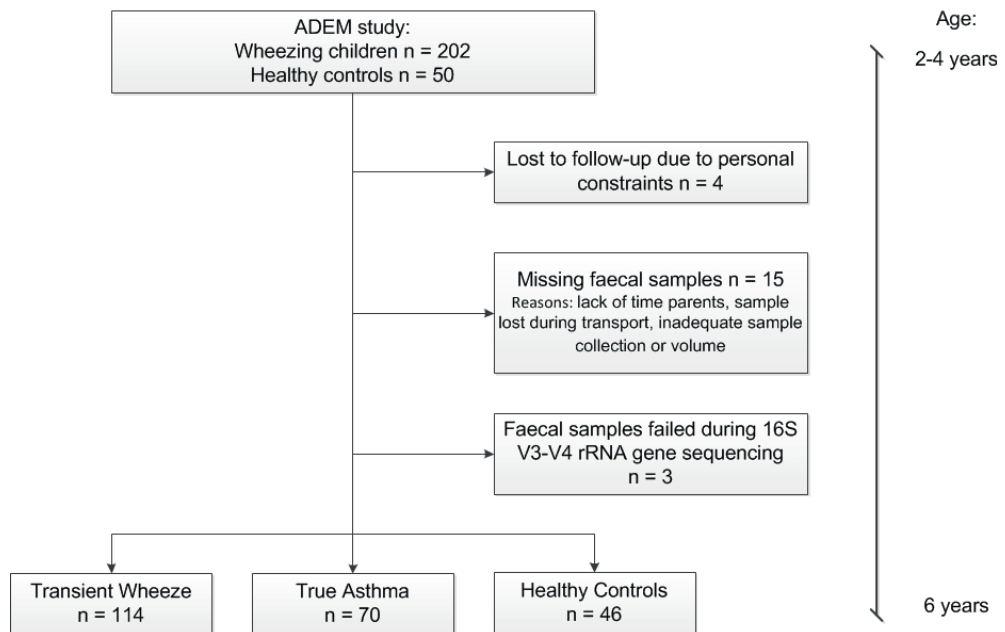
Enterotypes	n/N (%) <sup>§</sup>	OR <sub>adj</sub> <sup>*</sup>	95% CI	p-value
<i>Bifidobacterium</i> ‡	38/105 (36.2)	1.0	-	-
<i>Bifidobacterium-Blautia</i>	22/51 (43.1)	1.454	0.707-2.994	0.31
<i>Prevotella</i>	10/28 (35.7)	0.707	0.276-1.809	0.47

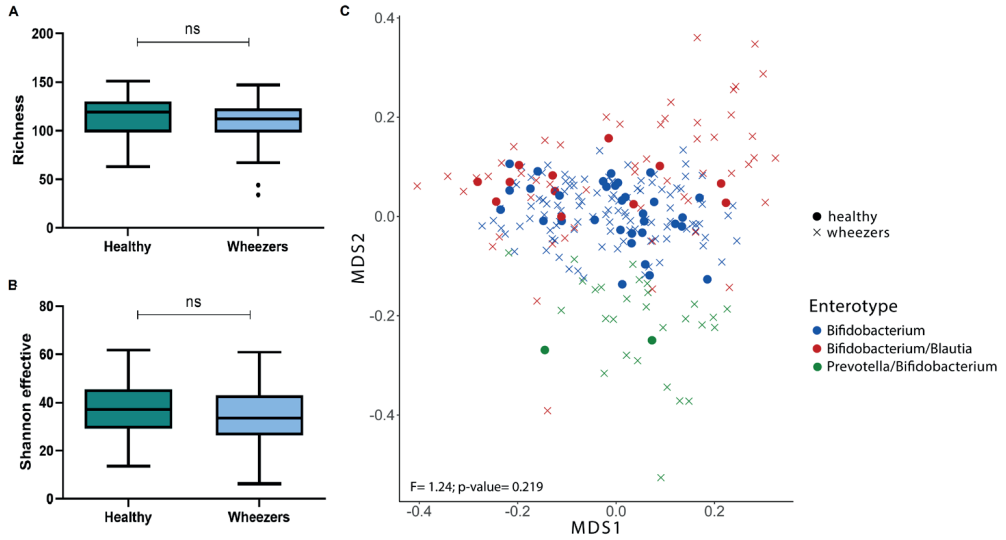
<sup>§</sup>Number of children with asthma/total number of children in the given enterotype (proportion of children with asthma in enterotype). ‡ Reference category. \*Adjusted for sex, breastfeeding, birth season, atopy parents, siblings, parental smoking status, day care attendance.

**Table S3.** Association between enterotype status and childhood wheeze.

Enterotypes	n/N (%) <sup>§</sup>	OR <sub>adj</sub> <sup>*</sup>	95% CI	p-value
<i>Bifidobacterium</i> ‡	105/137 (76.6)	1.0	-	-
<i>Bifidobacterium-Blautia</i>	51/63 (80.1)	1.167	0.537-2.540	0.70
<i>Prevotella</i>	28/30 (93.3)	4.282	0.940-19.51	0.06

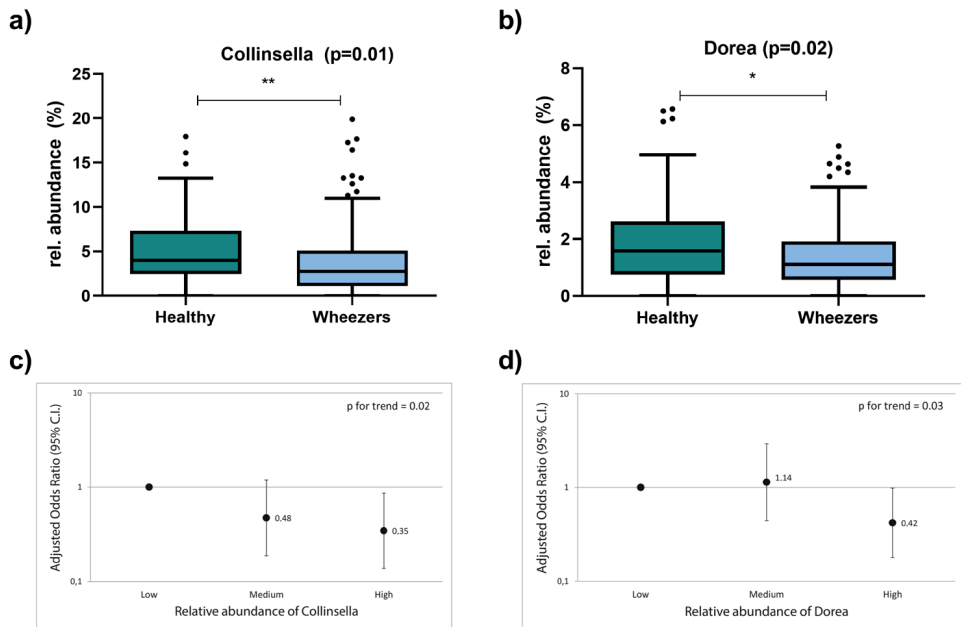
<sup>§</sup>Number of children with wheeze/total number of children in the given enterotype (proportion of children with wheeze in enterotype). ‡ Reference category. \*Logistic regression model Adjusted for sex, breastfeeding, birth season, atopy parents, siblings, parental smoking status, day care attendance.

**Figure S1.** Study flow chart.



**Figure S2. Microbial richness (observed species), microbial diversity (Shannon effector index), and microbial profiles of all children clustered into different enterotypes.** Microbial richness (A) and diversity (B) are not significantly different between healthy and wheezing children (ns: non-significant; Kruskal-Wallis). C: Multidimensional scaling (MDS) based on Bray-Curtis dissimilarity indicates three different enterotypes of Bifidobacterium, Bifidobacterium/Blautia and Prevotella/Bifidobacterium in the general microbial community structure. Microbial community structure is not statistically significantly different between healthy and wheezing children (permutational analysis of variance (PERMANOVA)).





**Figure S3. The relative abundance of specific microbial taxa decreases the risk of wheeze.** The relative abundance of bacterial genera *Collinsella* ( $p=0.01$ ) and *Dorea* ( $p=0.02$ ) is lower in wheezing children compared to healthy controls (A, B). The adjusted odds ratio for wheeze according to the relative abundance of the genera as derived from multiple logistic regression analysis show a higher risk to develop wheeze with a low abundance of *Collinsella* and *Dorea* (C, D). Adjusted for sex, breastfeeding, birth season, atopy parents, siblings, parental smoking status, day care attendance.





Julia Penders (9 years)

# Chapter 7

## Probiotic-mediated changes in the developing neonatal microbiota

Niels van Best\*, Sonja Trepels-Kottek, Paul H.M. Savelkoul, Thorsten Orlikowsky, Mathias W. Hornef‡, John Penders‡

\*equal contribution

*Gut Microbes. 2020 Nov 9;12(1):1-16*

## Abstract

Oral administration of probiotic bacteria to preterm neonates to prevent the development of necrotizing enterocolitis (NEC) has been recommended. The influence of probiotics on the endogenous microbiome has however remained incompletely understood. We therefore aimed to characterize the persistence of probiotic bacteria and their influence on the microbial ecosystem during and after the intervention and their association with the development of NEC.

We conducted an observational study including 80 preterm neonates, and collected fecal material of preterm infants hospitalized at the NICU of RWTH Aachen University Hospital weekly between January 1, 2016, and December 31, 2017. Neonates born at a gestational age <32-weeks received daily probiotic supplementation with *Bifidobacterium* (*B.*) *longum* and *Lactobacillus* (*L.*) *acidophilus* as a standard procedure (probiotic1 group). Supplementation had to be ceased due to a supply limitations (control group). Another probiotic formula consisting of *B. longum*, *L. acidophilus*, *B. lactis*, and *L. casei* was subsequently introduced (probiotic2 group). We subsequently profiled the fecal samples by 16S rRNA sequencing and monitored the presence of the probiotic bacteria by quantitative PCR.

80 neonates with a median gestational age of 26-30 weeks were included. Microbiota profiles differed between the control group and both probiotic groups. Probiotic supplementation was associated with lower temporal variation as well as higher relative abundance of *Bifidobacterium* and *Enterobacter* combined with reduced abundance of *Escherichia*, *Enterococcus* and *Klebsiella*. Colonization by probiotic bifidobacteria was observed in approximately 50% of infants although it remained transient in the majority of cases. Few children exhibited a more persistent colonization of up to 13 weeks. A significantly reduced monthly incidence of NEC was observed in neonates supplemented with probiotics.

Our results demonstrate successful transient colonization of preterm neonates by probiotic bacteria influencing the endogenous microbiota. The abundance of individual bacterial taxa associated with the development of NEC was reduced by the treatment. These results emphasizes that early infancy represents a critical period during which probiotic supplementation may allow targeted manipulation of the enteric microbiota and confer a clinical benefit.

## Introduction

Following birth, establishment of the early intestinal microbiota is characterized by major fluctuations in the relative abundance of the prominent bacterial taxa associated with an increase in the bacterial diversity over time [1, 2].

Preterm birth increases the inter-individual variation in the early fecal microbial ecosystem and enhances the colonization by opportunistic pathogens due to delivery by cesarean section, delayed onset and a reduced rate of enteral feeding by breast milk, the need for antibiotic therapy and prolonged exposure to the hospital environment [3, 4]. Together with the immature immune system and mucosal tissue, these microbiota alterations may enhance the susceptibility for necrotizing enterocolitis (NEC), an inflammatory disease of the preterm neonate's mucosal tissue associated with high morbidity and mortality [5].

Oral administration of probiotic bacteria to preterm neonates was reported to decrease the incidence of NEC in a number of studies[6]. Despite these promising results and the subsequent routine probiotic administration to preterm neonates in some institutions [7, 8], the precise influence of the administered probiotic bacteria on the existing microbial ecosystem during and following the intervention has remained incompletely understood. In particular, this relates to the selection and comparative analysis of probiotic strains and formulations, the persistence of the administered probiotic bacteria as well as their influence on the preexisting colonizing bacteria [9-11]. To gain more insight into the underlying mechanisms, we compared the microbiota in 174 fecal samples of 80 preterm neonates in the absence of as well as prior to, during and after oral administration of two different probiotic consortia within the context of a natural experiment.

## Material & Methods

### Study design

Preterm infants of gestational age <32 weeks were included in this natural experiment at the NICU of RWTH Aachen University Hospital between January 2016 and 2018. Initially, from birth until 36 weeks of gestational age, infants received daily a supplemented milk regimen (Infloran) as a standard procedure (P1 group). The probiotic contained the bacteria *Bifidobacterium longum subsp. infantis* and *Lactobacillus acidophilus* at a 1:1 ratio at  $10^9$  CFU/g. This standard supplementation with probiotic bacteria had to be temporarily suspended due to the unavailability of this probiotic on the market. During this time between May and December 2016, preterm neonates only received milk without probiotic supplement (control group). After December 2016, probiotic administration to all preterm (<36 weeks) neonates was reinstalled with another probiotic supplement (DarmfloraPlus) containing *Bifidobacterium longum*, *Bifidobacterium animalis subsp. lactis*, *Lactobacillus acidophilus*, *Lactobacillus casei* at a 1:1 ratio at  $10^9$  CFU/g. There were no exclusion criteria for the probiotic supplementation and the initial antibiotic standard regimen for neonates consisted of ampicillin (100mg/kg) and tobramycin (3 mg/kg) in 3-4 daily doses. Infants were sampled weekly during their hospitalization on the neonatal intensive care unit (NICU). The final samples were selected based on the first available sample within the timepoints described in Figure 1a. In an additional group of 16 neonates we longitudinally sampled weekly from supplementation up to 20 weeks after cessation of the probiotic to assess the persistence. Moreover, we also included longitudinal samples of three patients that developed NEC closely to the diagnosed onset of the disease. The study was recorded in the German Register Clinical Studies under accession number DRKS/GCTR 00021034 and was approved by the ethics committee of the medical faculty of the RWTH University of Aachen (approval number EK307/15).

### Microbial profiling of fecal samples

Metagenomic DNA was isolated from fecal samples thawed on ice and resuspended in 600  $\mu$ l DNA stabilization buffer (Stratec Biomedical) following a previously described protocol [12]. Briefly, 500  $\mu$ l 5% (w/v) N-lauroylsarcosine and 250  $\mu$ l 4 M guanidinethiocyanat were added to the samples and cells were mechanically lysed by repeated bead beating (3 x 40 s) with 500 mg 0.1 mm glass beads (Roth). Next, samples were vortexed and centrifuged (15.000  $\times$  g for 3 min at 4° C) after adding 15 mg poly-vinylpolypyrrolidone. Subsequently, the supernatant

was incubated with 1% ribonuclease (10 mg/mL) for 40 min at 37 °C. Finally, column-based purification was performed using the NucleoSpin gDNA Clean-up Kit (Macherey-Nagel). The bacteria-specific primers 341F and 785R were used to amplify the variable 3-4 region of the 16S rRNA gene from 12 ng metagenomic DNA by a two-step PCR (2x 15 cycles) [13, 14]. PCR amplicons were purified using the AMPure XP system (Beckmann), and sequenced using 2x275 paired-end reads spiked with 25% PhiX on a Miseq platform (Illumina Inc.). All generated sequences from this study have been deposited to the Qiita and ENA databases and can be accessed under No. 13071.

### **Microbiome data processing and statistics**

A total of 8,335,421 generated sequences were processed using the DADA2 package in R to create single amplicon sequence variants (ASV's) [15]. Quality filtering, removing of sequence errors and chimeras was performed using default settings and taxonomy was assigned with SILVA on a species level. For the removal of potential contamination, we used the Decontam package in R with the 'combined' setting [16]. Here, all samples with a DNA concentration below the detection limit (<0.10 ng/ul) were set at half the lowest detected concentration of 0.05 ng/ul and extraction controls were used as negative controls. Subsequently, we omitted samples with a low sequencing depth (<8000 sequences) and removed ASVs that were low abundant (<0.25% in all samples). This resulted in a total of 7,708,120 sequences ranging from 7,727 to 72,816 reads per sample (median 21,113) that belonged to 4,981 ASVs for downstream analyses. Lastly, we normalized the retained ASV counts by dividing each value to the total sum per sample and multiplied by the lowest sample depth [17].

We computed the alpha-diversity (Shannon index and observed species) and the Bray-Curtis metrics in order to determine the microbial community structure using the R package Phyloseq 1.30.0 [18]. The ordination of samples based on the Hellinger transformed genus composition or Bray-Curtis dissimilarity was visualized using Principal Component Analysis (PCA) or Principal Coordinate Analysis (PCoA), respectively. The Mann-Whitney U or Kruskal-Wallis tests were used to test for differences in relative abundances, Bray-Curtis distances, microbial richness and diversity between two or multiple groups respectively. Following the Kruskal-Wallis test, the Dunn's test was performed for pairwise post hoc comparisons. Subsequently, the alpha-diversity was used as a continuous variable in a linear regression model to test for a significant trend across timepoints. In order to evaluate significant separations between groups of samples, we performed a permutational multivariate



analysis of variance (PERMANOVA) based on Bray-Curtis distances. We examined factors that were significantly associated with the microbial community variation at the different timepoints by PERMANOVA to determine potential confounders: timepoint, feeding-type, sepsis, antibiotics, mode of delivery, birthweight, gestational age, sex, and PROM. We next combined the identified potential confounding factors with the variable probiotic group in a multivariate analyses and re-evaluated the significance. Moreover, we also tested the dispersion of the samples, a required assumption for the PERMANOVA test. All these statistical analyses were performed two-sided with R 3.6.1 or GraphPad Prism 8.

SparCC correlation analyses were performed to identify specific associations taking the compositional nature of the data into account by MicrobiomeAnalyst with default settings [19]. Only correlations of >0.3 that were statistically significant were included in the network analyses.

Differentially abundant bacterial genera were identified using Linear discriminant analysis Effect Size (LEfSe) with default settings [20].

### **Detection of probiotic bacteria & antimicrobial resistance genes (qPCR)**

For the quantification of *Lactobacillus acidophilus* (5' AACTGCAATTTAAGATTATGAGT TTC / GGTACCGTCTTGATTATTAGTGTA 3'), *Bifidobacterium lactis* (5' CATCGCAACTTCACCCACATTG 3' / 5' ATGCCGTACCCCTGAATGAAG 3'), *Bifidobacterium longum* (5' CGGCGTYGTGACCGTTGAAGAC 3' / 5' TGYTTCGCCRTCGACGTCCTCA 3'), *Lactobacillus casei* (5' TGCCATTAGCATACTGGACC 3' / 5' ACCCGAGCCTTTGCCAA 3') all fecal samples were subjected to real-time PCRs [21, 22]. We validated the existing assays by using amplifying the DNA of the cultured probiotic strains. Upon confirmation of the PCR amplicons by Sanger sequencing, products were cloned into a pGEM-T easy vector (Promega Corporation, Madison, WI, USA) for construction of standard curves. Thereafter, the amount of plasmid copies normalized to the amount of DNA were calculated for each fecal sample from the cycle threshold values using the standard curves. For this assay 2 µl DNA and 0.4 µM primer was used as input, employing SYBR Green chemistry (Supermix™ SYBR®, Bio-Rad) and a CFX96 System (Bio-Rad). All targets were amplified at an annealing temperature of 65 °C except for the *Lactobacillus acidophilus* (60 °C). CFX Manager 3.1 (Bio-Rad) was used for analysis of the real-time qPCR data. Samples that were below the detection limit were included with zero plasmid copies.

The qPCR of antimicrobial resistance genes *vanA*, *vanB*, *CTXM-1*, *CTXM-2* was performed in a multiplex reaction and *CTXM-9* in a simplex reaction as described earlier [23, 24]. Primer-probe concentrations for *vanA* (GCCGAAAAAGGCTCTGAA / TCCTCGTCTCTGCTGAA) and

vanB (CGCAGCTTGCATGGACAA / GGCGATGCCCGCATT) were 800 nM of each primer and 200 nM probe (vanA\_probe: FAM-ACGCAGTTATAACCGTTCGCCGAGACC-BHQ1 / vanB\_probe: VIC-TCACTGGCCTACATTC-MGB-NFQ). The qPCR of CTX-M (ATGTGCAGYACCAGTAARGTKATGGC / ATCACKCGGRTCGCCNGGRAT) contained 500 nM of each primer and 100 nM probe (CTX-M-1: JOE-CCCGACAGCTGGGAGACGAAACGT-BHQ1 / CTX-M-2: FAM-CAGGTGCTTATCGCTCTCGCTCTGTT-BHQ1 / CTX-M-9: JOE-CTGGATCGCACTGAACCTACGCTGA-BHQ1). Each PCR was performed in a total volume of 25  $\mu$ L consisting of 12.5  $\mu$ L Absolute QPCR ROX Mix (Thermo Scientific, Waltham, MA, USA), the primer-probe mixture and template DNA. The amplification was achieved on a 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) according to the standard conditions.

## Results

### Study population characteristics

As part of a longitudinal observational study, we collected fecal material of preterm infants hospitalized at the NICU of RWTH Aachen University Hospital weekly between January 2016 and 2018. Between January 2016 and May 2016, neonates born at <32-weeks gestational age received daily probiotic supplementation consisting of *Bifidobacterium longum* and *Lactobacillus acidophilus* as a standard clinical procedure (probiotic 1 [P1] group). In May 2016, this probiotic was no longer commercially available and supplementation had to be ceased (control group). From January 2017 onwards, another probiotic mixture consisting of *B. longum*, *L. acidophilus*, *Bifidobacterium lactis*, and *Lactobacillus casei* (probiotic 2 [P2] group) was introduced and administered to preterm neonates. Supplementation of both probiotic products continued until 36 weeks postmenstrual age. In total, we selected 174 fecal samples from 80 preterm neonates. The population characteristics, including gestational age, birthweight, sex and feeding-mode, were comparable between all three groups (Table S1). The median gestational age ranged from 26 to 30 weeks and most neonates were born by caesarian section. For the comparative analysis between groups (P1, P2 and control group), samples were clustered in four age-windows as follows: week 1 after birth, week 3-4 after birth, week 7-9 after birth and week 10-12 after birth corresponding to “before” (T0), “during” (T1), “shortly after” (T2) and “long after” (T3) probiotic supplementation, respectively (Figure 1a).

### Supplementation of probiotic bacteria affects the microbiota composition in early life

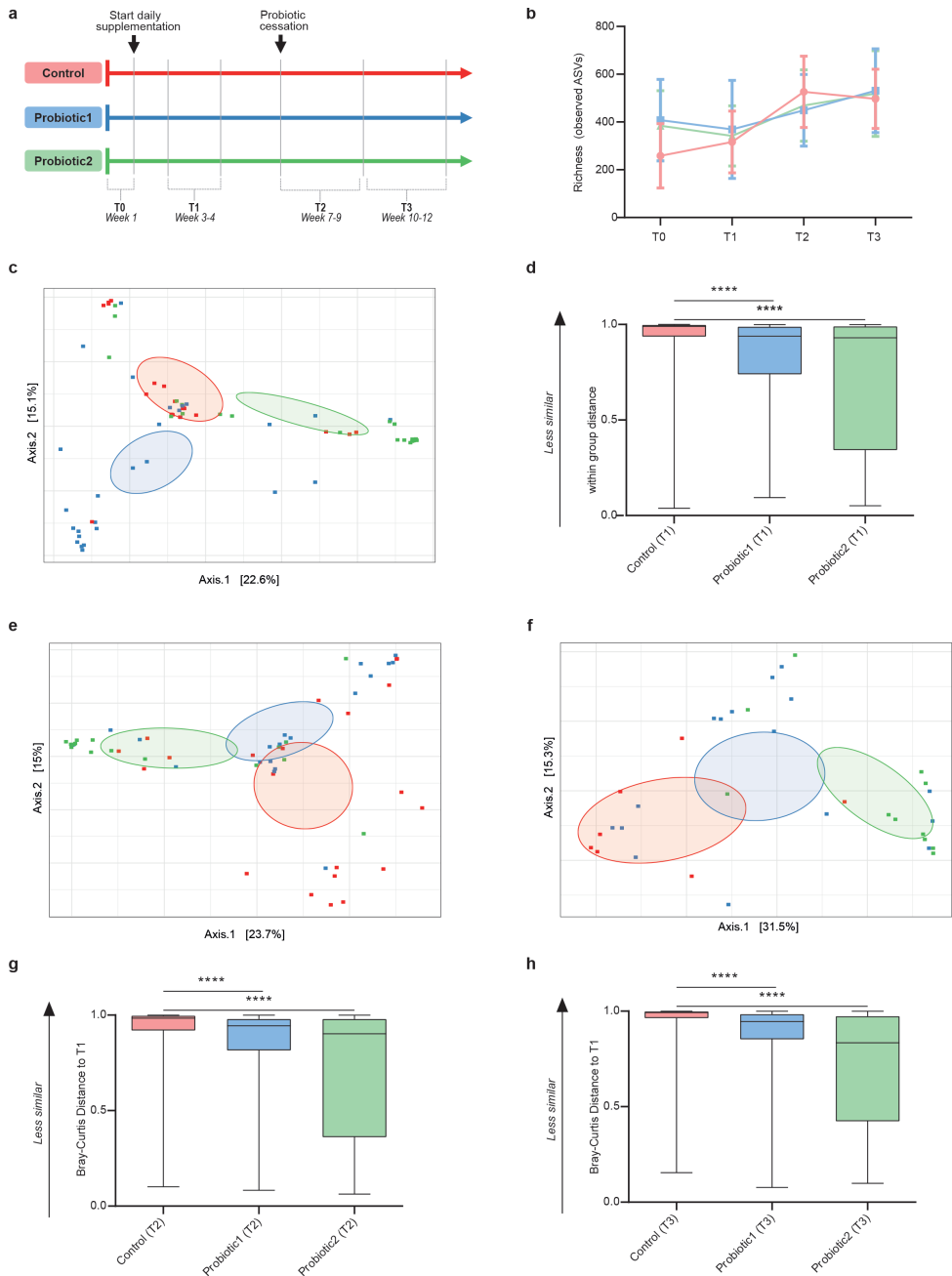
To assess global compositional differences in the gut microbiome of neonates, we profiled fecal samples by V3-V4 16S rRNA gene amplicon sequencing. A total of 7,708,120 sequences with a median of 21,113 [range 7,727 - 72,816] reads per sample were retained after quality filtering and assigned to 4,981 amplicon sequence variants (ASVs) for subsequent analyses.

The fecal microbial richness gradually increased significantly with age (linear regression  $p < 0.0001$ ,  $p = 0.01$ , and  $p = 0.0021$  for control, P1, and P2 groups, respectively), but was not significantly different between the control and probiotic groups (Figure 1b, Figure S1a-d). Similar results were obtained for the microbial diversity as examined by the Shannon index (Figure S1e-h).

As expected, no significant difference in the general microbial community composition was observed between the three groups prior to the start of probiotic supplementation (T0) (Figure S2a). In contrast, significant differences in the microbial composition between the

control group and both probiotic groups at the time of probiotic supplementation was observed (Figure 1c). The changes in the microbial community structure remained after identification and correction for potential confounders such as feeding-type, sepsis, administration of antibiotics and gestational age (Table S2, Table S3). Notably, significant differences between the control group and both probiotic groups were also detected after removal of all probiotic-specific ASV's from the analysis (Figure S3). This suggests that the observed changes are not solely a direct consequence of the bacteria added to the enteric community by oral probiotic supplementation.

Additionally, neonates receiving probiotic supplementation at T1 showed less inter-individual variation and thus a more homogenous microbiota composition compared to infants in the control group (Figure 1d). Although the latter effect was not detected at subsequent time-points (Figure S2b-f), the significant difference in the global microbial community structure remained also shortly (T2) and long (T3) after cessation of probiotic supplementation (Figure 1e & f). The Bray-Curtis dissimilarity between samples collected during probiotic supplementation (T1) and both follow-up time-points (T2 and T3) was significantly lower for both probiotic groups as compared to the control group (Figure 1g & h). This suggests that probiotic supplementation reduces the temporal compositional variation and results in a more stable microbial community structure.



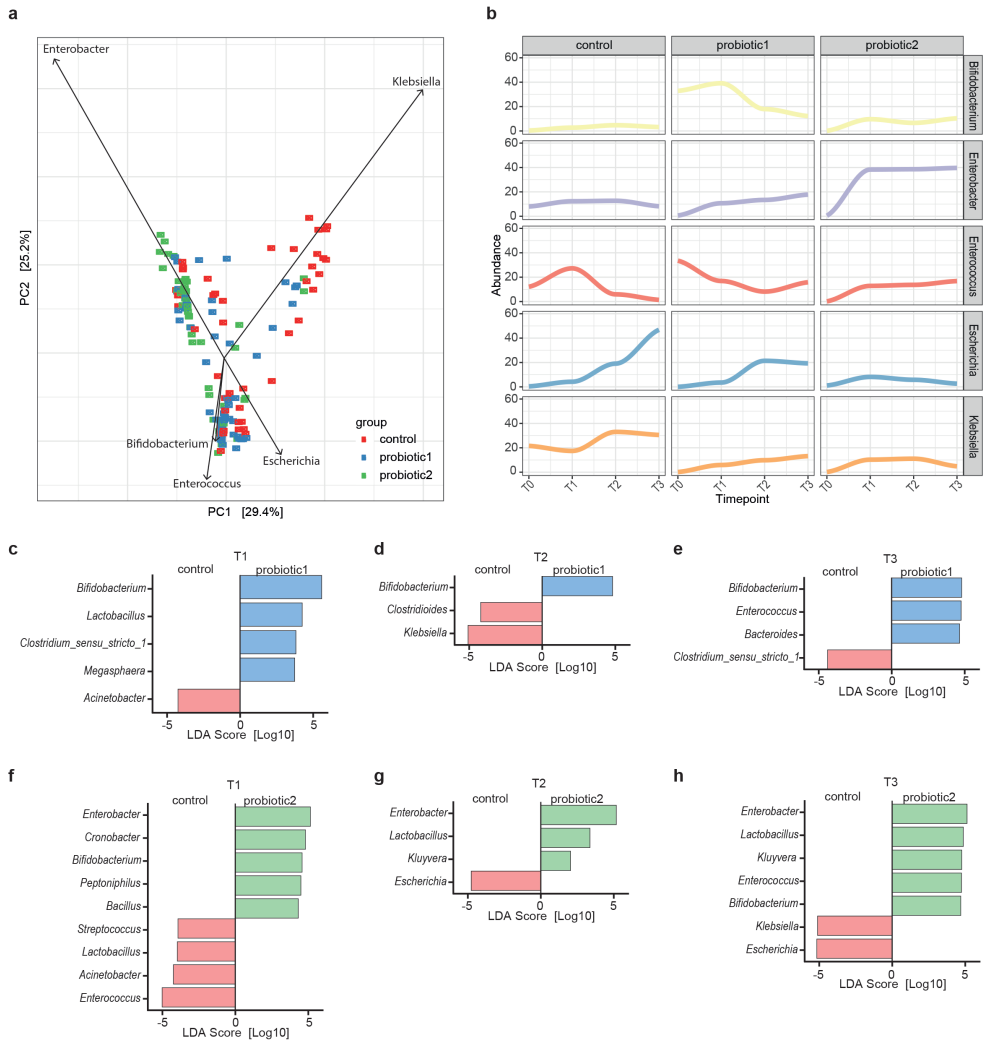
**Figure 1 - Microbiota composition and stability in neonates with or without probiotic supplementation. (A)** Infographic of study outline (n=51-66 per group for all subsequent cohort-analyses) **(B)** Microbial richness (number of ASVs) shows a gradual increase with age for all groups (linear regression: pfor trend<0.0001, pfor trend = 0.01, pfor trend =0.0021 for controls, probiotic1, probiotic2, respectively). **(C)** Principal Coordinate Analysis (PCoA) based on ASV-derived Bray-Curtis dissimilarity exhibiting differences in microbial community structure between control and probiotic groups at T1 (p<0.001, Permanova). **(D)** Bray-Curtis distance compared to T1 at T1 (within group distance). **(E,F)** PCoA between control and probiotic groups at T2 (p<0.001, Permanova) **(E)**, and at T3 (p<0.001, Permanova) **(F)**. **(G,H)** Bray-Curtis distance compared to T1 at T2 **(G)**, and at T3 **(H)**. (Kruskal-Wallis test with Dun's test for post-hoc comparisons (control-group is reference), mean and SD; \*\*\*\*, p<0.0001).

**Probiotic administration influences specific taxa of the endogenous microbiota**

We subsequently assessed how probiotic administration influenced specific endogenous bacterial taxa over time (T1-T3). First, we determined how the genera that explained most of the variation in the overall microbiota community structure were associated with probiotic administration. *Klebsiella*, *Escherichia*, *Bifidobacterium*, *Enterococcus* and *Enterobacter* represented the main genera driving the separation in microbial community structure between the three groups. Enhanced abundance of *Klebsiella* was typically found in neonates without probiotic supplementation. In contrast, neonates receiving P2 were characterized by higher abundance of *Enterobacter* and reduced abundance of *Escherichia* (Figure 2a). Increased relative numbers of *Bifidobacterium* and *Enterococcus* were most commonly observed among neonates that received P1.

Secondly, the taxonomic changes of these five genera over time indicated that probiotic supplementation reduced the temporal fluctuations in microbiota composition. Whereas the abundance of *Enterococcus*, *Escherichia* and *Klebsiella* were highly dynamic in the control group, a more stable pattern could be observed in probiotic supplemented neonates (Figure 2b). This more stable pattern was accompanied by an expansion of the genus *Bifidobacterium* in P1 group and an expansion of the genus *Enterobacter* in the P2 group.

Lastly, linear discriminant analysis revealed that *Bifidobacterium* was the most strongly enriched genus among neonates receiving P1 at all time-points, while lactobacilli were only significantly enriched at T1 (Figure 2c-e). In addition, after probiotic supplementation neonates harbored less *Klebsiella* and two members of the Clostridiales order when compared to control neonates. Moreover, although the second probiotic similarly to P1 was associated with a higher abundance of bifidobacteria and decreased levels of *Klebsiella* at T3, these neonates could most profoundly be differentiated from the control group by a higher abundance of *Enterobacter* at all time points (Figure 2f-h). Remarkably, the neonates receiving P2 showed a strong decrease in *Enterococcus* and *Escherichia* at T1 and T2-T3, respectively. Furthermore, in both probiotic groups the abundance of bifidobacterial species other than those included in the probiotic formulas were enhanced when compared to the control neonates (Figure S4a-f).



**Figure 2 - Association of probiotic administration with specific taxa. (A)** Principal Component Analysis (PCA) illustrating the separation between the control and probiotic groups at T1-T3. Vectors depict the five genera that explained most of the variation in microbial community structure **(B)** Relative abundances of the five genera over time for the three different groups. **(C)** Linear discriminant analyses with Effect Size (LEfSe) were employed to identify differentially abundant bacterial genera between probiotic 1 supplemented and control neonates at T1, **(D)** at T2, **(E)** and at T3. **(F)** LefSe between probiotic 2 supplemented and control neonates at T1, **(G)** at T2, **(H)** and at T3 (LDA-score>0.2 and p<0.05, Wilcoxon signed-rank test).

### Intestinal colonization by probiotic bacteria after cessation of supplementation

Next, we established qPCRs specific for the administered probiotic bacteria in order to quantitatively monitor their ability to colonize and persist after cessation of supplementation. None of the probiotic bacteria were detected prior to supplementation (T0) (Figure S4g-j). In contrast, qPCR detection confirmed the markedly higher abundance of *B. longum* and *L. acidophilus* in neonates receiving P1 and *B. longum*, *B. lactis*, *L. acidophilus* and *L. casei* in neonates receiving P2 as compared to controls (Figure 3a-l). Notably, *B. longum* was still enriched at T2 i.e. after cessation of supplementation with P1, whereas *B. lactis* and *L. acidophilus* remained increased after cessation (T2) of supplementation with P2. Importantly, probiotic bacteria did not permanently colonize since no significant differences between probiotic and control groups were observed at T3. However, *B. lactis* was still detectable in some (4/11) infants that received P2 at T3 (Figure 3f).

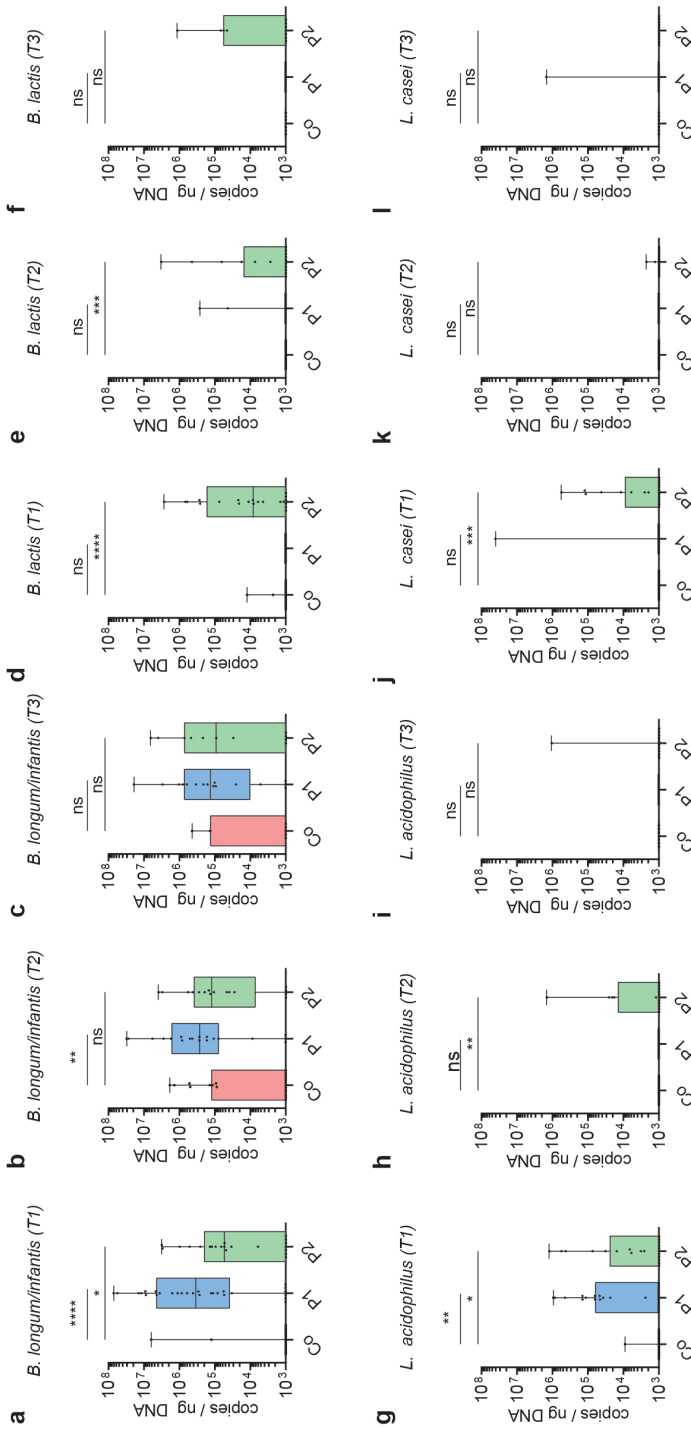
To further investigate the nature and duration of probiotic persistence we analyzed fecal samples of 16 neonates collected weekly during and until 20 weeks after cessation of probiotic supplementation. We tracked probiotic-specific *Bifidobacterium* ASVs (i.e. ASVs present in the probiotic formulas and absent in any of the control neonates) (Figure S5a-b) and performed the strain-specific qPCRs. Since the prevalence pattern of the different probiotic bacteria varied considerably (Figure S5c-f), we categorized the neonates into ‘persistently’ colonized (first two samples after cessation positive for probiotic strains), ‘transiently’ colonized (at least one sample collected in the first two weeks after cessation positive for probiotic strains, but subsequent samples negative) and not colonized (1st and 2nd sample upon probiotic cessation negative for probiotic strains). Thereby, the probiotic bifidobacteria were shown to persistently but mostly transiently colonize the neonatal gut after cessation (Figure 4 a-b). On the other hand, *L. acidophilus* was absent in almost all neonates that had received P1 whereas it remained present (persistently or transiently) in the majority of neonates supplemented with P2 (Figure 4c-d). Neonates supplemented with P2 also showed persistence of *B. lactis* and *L. casei* in some neonates, but most neonates had become negative for these probiotic strains (Figure 4e-f). In some children, the colonization of *Bifidobacterium* persisted up to 13 weeks and *L. casei* up to 15 weeks after cessation of probiotic supplementation (Figure S5c-f).



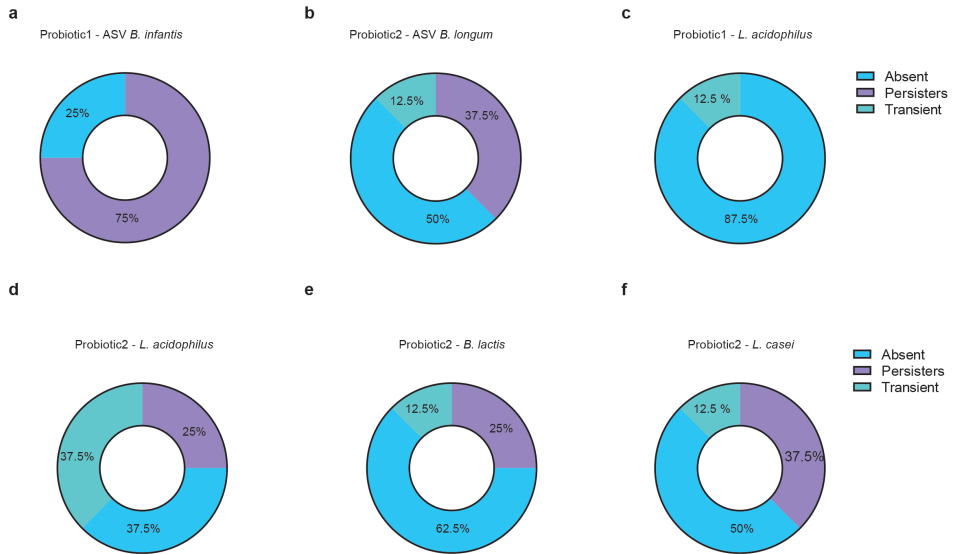
### **Abundance of pathobionts and prevention of necrotizing enterocolitis**

Probiotic supplementation both increased the stability in the overall community structure of the neonatal gut microbiota as well as reduced the abundance of pathobionts known to induce strong proinflammatory responses. Since these changes might protect against the adverse effects of an altered microbiota, we next investigated the influence of probiotic supplementation on the incidence of necrotizing enterocolitis (NEC) and the presence of clinically relevant antimicrobial resistance (AMR) genes. Despite the low overall occurrence of NEC during our study period (n=10), a significantly reduced incidence was observed in neonates supplemented with P1 or 2 as compared to control neonates (Figure 5a). We therefore analyzed the composition of the fecal microbiota of three NEC-patients prior to the clinical onset of the disease. Interestingly, a dramatic bloom of enterococci was observed, representing the most abundant taxa before the onset of the disease (Figure S6a). Notably, *Enterococcus* is only low abundant in neonates without NEC as depicted from the control group at T2 accompanied with *Klebsiella* as the most dominant taxa (Figure 2b). To further elucidate the role of the probiotic bacteria herein, we performed SparCC correlation network analyses taking the compositional nature of the data into account [25]. Interestingly, probiotic supplementation with bifidobacteria and lactobacilli both correlated with the abundance of the NEC-pathobiont *Enterococcus* [26, 27]. In neonates receiving P1, bifidobacteria and lactobacilli were positively correlated. Lactobacilli in turn exhibited a direct negative network correlation with enterococci (Figure 5b, Figure S7a). In neonates receiving P2, bifidobacteria and lactobacilli positively correlated with *Enterobacter* that in turn exhibited a negative correlation with the NEC-pathobiont *Enterococcus* (Figure 5c, Figure S7b).

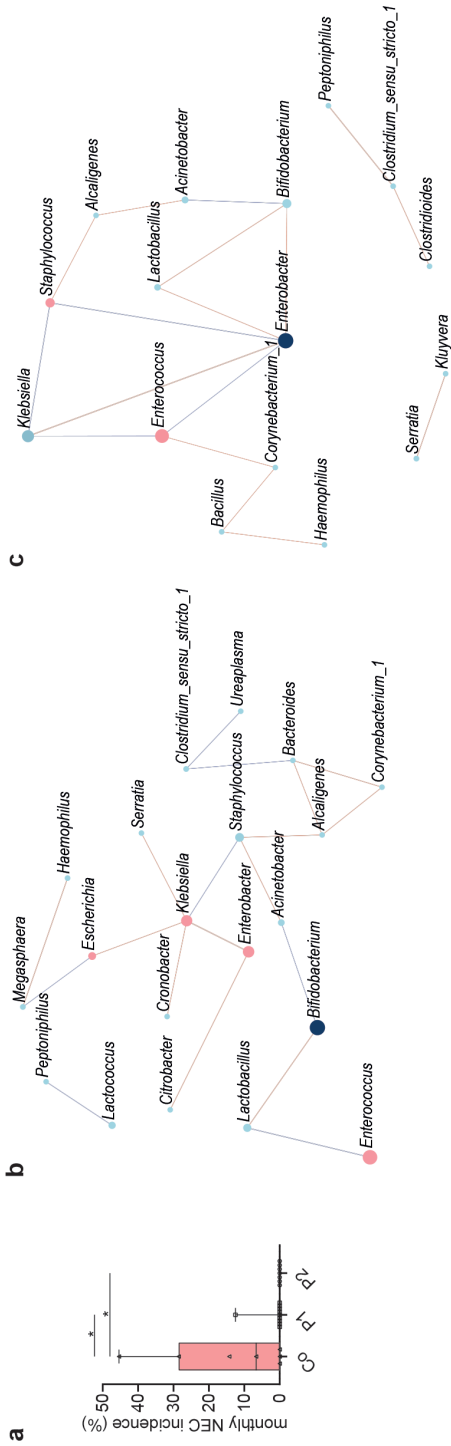
As probiotic administration was associated with a decreased abundance of *Enterococcus*, *Escherichia* and *Klebsiella* in preterm infants (Figure 2c-h), probiotics might also reduce the prevalence of antibiotic resistant strains. The presence of resistant strains significantly hampers an effective antimicrobial therapy in the event of infection. Enterococci can harbor the resistance genes *vanA* and *vanB* conferring resistance to glycopeptide antibiotics such as vancomycin and teicoplanin. Enterobacteriaceae, such as *Escherichia* and *Klebsiella*, can encode for extended-spectrum beta-lactamases (ESBLs) conferring resistance to most beta-lactam antibiotics. We therefore investigated the presence of *vanA* and *vanB* as well as the most prevalent ESBL (*CTX-M* group 1, 2, and 9) genes. However, the prevalence of these antibiotic resistance genes was not significantly altered by probiotic supplementation, possibly due to the low overall presence of antibiotic resistance genes in the enteric microbiota of the population analyzed (Figure S6b-c).



**Figure 3 - Association of probiotic administration with probiotic species.** (A-L) Quantitative abundance (log<sub>10</sub> copies/ng DNA) during probiotic supplementation at T1, after cessation of probiotic supplementation at T2 and T3 respectively of the probiotic bacteria *Bifidobacterium longum* (A-C), *Lactobacterium lactis* (D-F), *Lactobacillus acidophilus* (G-I), *Lactobacillus casei* (J-L). (Kruskal-Wallis test with Dun's test for post-test comparisons (control-group is reference), mean and SD; \*, p<0.05; \*\*, p<0.01; \*\*\*, p<0.001, \*\*\*\*, p<0.0001; ns, not significant).



**Figure 4 - Colonization by probiotic bacteria after cessation of supplementation. (A)** Colonization pattern from longitudinally sampled neonates (n=8 / group) after cessation of probiotic supplementation for ASV-tracked Bifidobacterium in probiotic 1 supplemented neonates, **(B)** Bifidobacterium in probiotic 2 supplemented neonates, **(C)** qPCR-tracked Lactobacillus acidophilus in probiotic 1 supplemented neonates, **(D)** Lactobacillus acidophilus, **(E)** Bifidobacterium Lactis, **(F)** and Lactobacillus casei in probiotic 2 supplemented neonates. Categorized into 'persistent' colonizers (first two weekly timepoints after cessation were positive for probiotic strains), 'transient' (sample collected in the first two weeks after cessation was positive, but subsequent sample was negative) and 'absent' (first timepoint upon probiotic cessation was negative for probiotic bacteria).



**Figure 5 - Probiotic intake may prevent necrotizing enterocolitis by interfering with pathobionts.** (A) Monthly NEC-incidence during the 2-year study period among all preterm born neonates (n=179, gestational age<32 weeks) stratified according to supplementation group. (B) SparCC correlation network analyses with genera at T1 in the probiotic 1 supplemented, (C) and probiotic 2 supplemented neonates as compared to control neonates (correlation>0.3, p<0.05). Blue circles indicate genera with an increased abundance in probiotic groups, while red circles represent genera enriched in control neonates. Red lines indicate negative correlations, while blue lines indicate positive correlations.

## Discussion

With weekly fecal sampling and clinical monitoring, our study represents a longitudinal observational study with three groups, a group of children that was left untreated (control group) and two groups of children that received two different probiotic supplements (P1, and P2). Although the different groups were sampled during consecutive time periods, the standard medical care regimen (except probiotic supplementation), the medical and nursing staff as well as the neonatal intensive care unit (NICU) environment and medical equipment were identical for all patients included in the study and no cross colonization between patients could occur [28, 29].

In contrast to many previous studies among adults and infants [10, 30, 31], we showed persistence of several probiotic strains upon cessation of supplementation. Administered bifidobacteria persisted for 2-12 weeks following cessation in approximately a quarter of the neonates, whereas persistence of lactobacillus strains was observed less frequently. This superior persistence of bifidobacterial strains when compared to lactobacilli has also been observed in several other studies [9, 11, 29]. The mechanisms underlying the differential persistence are largely unclear, but the gut luminal milieu, the characteristics of the administered bacterial strains as well as their competitive fitness may all contribute. Interestingly, we detected differences in colonization of the probiotic bacteria between our two supplemented probiotic formulas which suggests strain differences or competition. For example, *B. lactis* but not *B. longum* remained significantly enriched after cessation of supplementation in neonates that received P2 when compared to neonates in the control group. In addition, *L. acidophilus* persisted in neonates that received P2 in contrast to P1 supplemented neonates.

Probiotic administration significantly altered the enteric microbiota composition even after removing probiotic ASVs from the sequencing data. Thus, the administered probiotic bacteria directly or indirectly influence the endogenous microbiota consistent with the results of other studies [9-11, 29]. Most notably, although bacterial richness and diversity remained unaffected, probiotic administration significantly reduced the inter-individual variation and temporal microbiota changes. Neonates, in particular preterm neonates, have been shown to exhibit high inter-individual variation and major fluctuations in the enteric microbiota composition during the early postnatal period [1, 4, 32, 33]. High individual variation and compositional fluctuations may contribute to a decreased resilience and

colonization resistance, and thus a more stable microbiota composition might protect from disease-promoting compositional changes. Moreover, the preterm microbiota is typically dominated by pathobionts such as *Escherichia*, *Enterobacter* and *Klebsiella* that produce potent immunomodulatory molecules such as endotoxin and fimbriae and frequently carry genes conferring antibiotic resistance [3, 26, 34]. Reducing such bacteria might lower the risk of inappropriate inflammation or systemic infection. In this respect, *Klebsiella*, one of the major bacteria driving the global community structure in the preterm neonates without probiotic supplementations did not contribute to the overall community structure in children supplemented with probiotics administration. An observation which is in line with the results of the study by Alcon-Giner and colleagues [29].

As expected probiotic supplementation led to enrichment of *Bifidobacterium spp.* and *Lactobacillus spp.* However, this was not limited to an enrichment of the probiotic bifidobacterial and lactobacillus species. In accordance with another recent report, the administration of probiotic bifidobacterial species fostered the growth of endogenous bifidobacterial species such as *B. breve*, *B. bifidum* and *B. animalis* [29]. Cross-feeding between species within a genus or induced adaptation of the local gut environmental might contribute to this. In contrast, pathobionts such as *Klebsiella* or *Escherichia* as well as *Clostridioides* and enterococci were reduced in probiotic treated infants. These results are largely consistent albeit not identical with observations described in other studies [9, 29]. Despite the reduced abundance of enterococci and Enterobacteriaceae (*Klebsiella* and *Escherichia*), the prevalence of genes encoding vancomycin resistance and ESBLs was not reduced among neonates supplemented with probiotics. These findings contrast a previous study, which observed ESBL-genes in stools of non-probiotic very preterm infants but not in probiotic-treated extremely preterm infants [35]. Previous clinical and animal studies suggested that the enteric microbiota together with the immature mucosal tissue contribute to the etiology of NEC [34]. Changes in the microbiota composition would therefore be expected prior to disease onset. A number of human cohort studies with regular postnatal sampling of preterm neonates reported on the enteric microbiota composition in healthy infants and infants with NEC [26, 36]. However, no specific pathogen, pathobiont or microbiota signature was detected that unequivocally predicted disease onset. Notably, almost all cases of NEC documented during the study period occurred in neonates of the control group confirming a significant protective effect of probiotic supplementation. For three patients with NEC, consecutive samples were available prior to disease onset. All three exhibited a significantly enhanced

abundance of enterococci prior to disease onset. An increase in enterococci or enterococcus-associated secondary metabolite gene clusters was also noted in other studies [26, 37]. Additionally, certain strains of the *Enterococcus faecalis* have been shown to increase the tissue pathology in a rodent NEC model [27]. Two of the patients in the present study also exhibited an enhanced abundance of *Klebsiella*, an organism associated with an increased risk of NEC in several studies [26, 36]. Moreover, an increased abundance of Enterobacteriaceae prior to NEC onset have been found [36, 38, 39]. As expected, the abundance of members of the Enterobacteriaceae family (*Escherichia* and *Klebsiella*) also increased in abundance during the NEC-induced intestinal inflammation [36]. Although the mechanism by which probiotics protect against NEC is unclear, our results are consistent with the idea that it may include enhanced resilience of the enteric microbiota to exogenous stimuli, the reduction of pathobionts or indirect immunomodulatory effects on the host's developing intestinal tissue and mucosal immune system [40, 41].

The daily dose of  $10^9$  supplemented bacteria is substantial given that the total bacterial number within the newborn gut is estimated at  $10^6$ - $10^{12}$  [42, 43]. Given the immature intestinal tissue and mucosal immune system and high susceptibility to systemic bacterial infection of preterm neonates a careful consideration of potential risks of oral administration of viable bacteria is warranted. Cases of sporadic systemic infections with orally applied probiotic bacteria in extremely preterm neonates have indeed been reported sporadically [44]. In the present study no adverse effects or infections by the administered probiotic were observed during the study period.

Nevertheless, the observational character represents a limitation of our study precluding any blinded and randomized study design. The consecutive "enrollment" of the three patient groups also harbors the risk of unidentified changes in the standard patient care during the study period that may influence the clinical outcome. In addition, the relatively small sample size limits the statistical power of our analysis in particular with respect to the analysis of the enteric microbiota prior to the clinical onset of NEC. Altogether our study provides important insights that warrant confirmation, preferably in experimental studies. Sampling after discharge would ideally be included in such future studies in order to assess the long-term consequences of probiotic supplementation.

In conclusion, our results show a significant influence of oral probiotic administration on the global microbial community structure and the abundance of specific taxa. The data emphasize the unique period of early infancy for the establishment of the enteric microbiota

and suggest a beneficial effect of probiotic supplementation to preterm neonates. It is where probiotics seem to be successful to influence the gut microbiota.

## **Acknowledgment**

We would like to thank Martina Bernecker for excellent technical support.



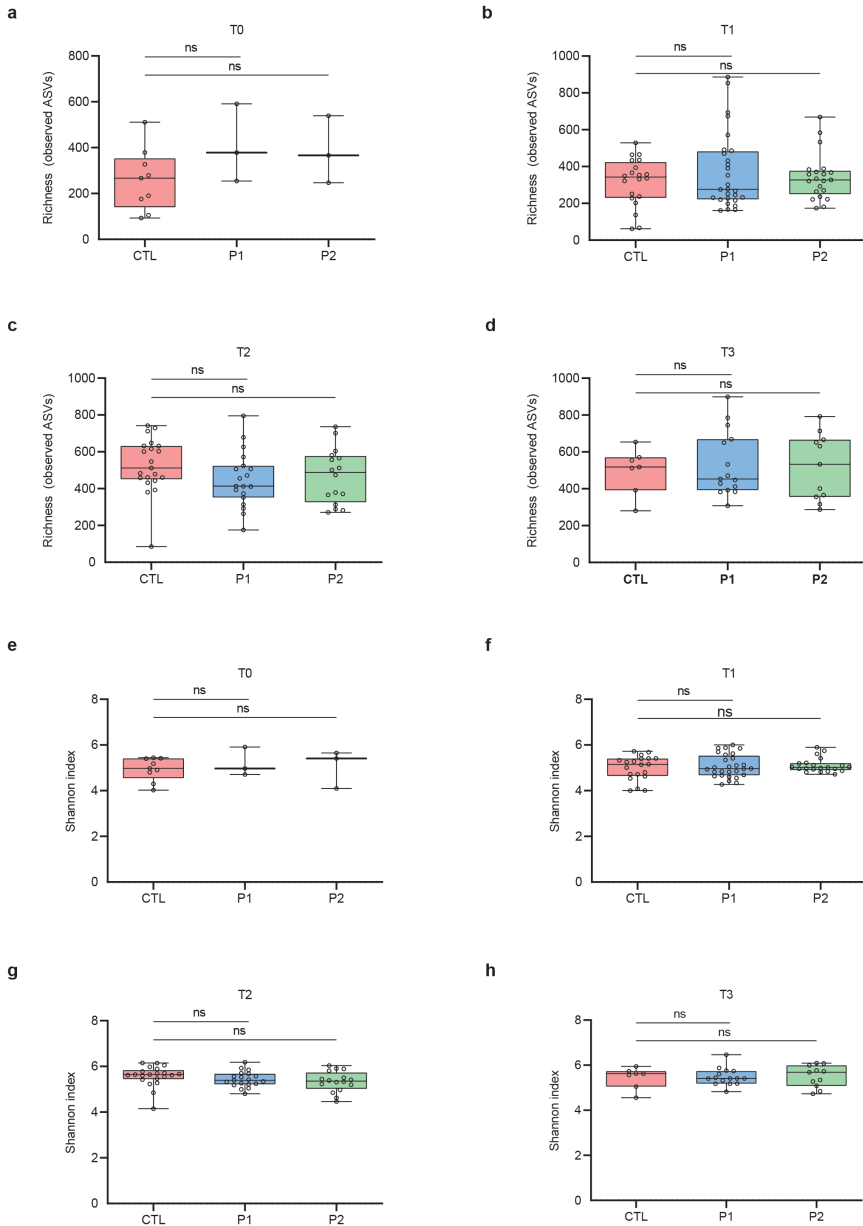
## References

1. Backhed, F., et al., *Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life*. Cell Host Microbe, 2015. 17(5): p. 690-703.
2. 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.
3. Stewart, C.J., et al., *Temporal bacterial and metabolic development of the preterm gut reveals specific signatures in health and disease*. Microbiome, 2016. 4(1): p. 67.
4. Hill, C.J., et al., *Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort*. Microbiome, 2017. 5(1): p. 4.
5. Gopalakrishna, K.P., et al., *Maternal IgA protects against the development of necrotizing enterocolitis in preterm infants*. Nat Med, 2019. 25(7): p. 1110-1115.
6. Dermyshe, E., et al., *The "Golden Age" of Probiotics: A Systematic Review and Meta-Analysis of Randomized and Observational Studies in Preterm Infants*. Neonatology, 2017. 112(1): p. 9-23.
7. Denkel, L.A., et al., *Protective Effect of Dual-Strain Probiotics in Preterm Infants: A Multi-Center Time Series Analysis*. PLoS One, 2016. 11(6): p. e0158136.
8. Guthmann, F. and C. Buhner, *Routine probiotics in preterm infants?* Arch Dis Child Fetal Neonatal Ed, 2011. 96(4): p. F311-2.
9. Abdulkadir, B., et al., *Routine Use of Probiotics in Preterm Infants: Longitudinal Impact on the Microbiome and Metabolome*. Neonatology, 2016. 109(4): p. 239-47.
10. Bazanella, M., et al., *Randomized controlled trial on the impact of early-life intervention with bifidobacteria on the healthy infant fecal microbiota and metabolome*. Am J Clin Nutr, 2017. 106(5): p. 1274-1286.
11. Korpela, K., et al., *Probiotic supplementation restores normal microbiota composition and function in antibiotic-treated and in caesarean-born infants*. Microbiome, 2018. 6(1): p. 182.
12. Godon, J.J., et al., *Molecular microbial diversity of an anaerobic digester as determined by small-subunit rDNA sequence analysis*. Appl Environ Microbiol, 1997. 63(7): p. 2802-13.
13. 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): p. e1.
14. Berry, D., et al., *Barcoded primers used in multiplex amplicon pyrosequencing bias amplification*. Appl Environ Microbiol, 2011. 77(21): p. 7846-9.
15. Callahan, B.J., P.J. McMurdie, and S.P. Holmes, *Exact sequence variants should replace operational taxonomic units in marker-gene data analysis*. ISME J, 2017. 11(12): p. 2639-2643.
16. Davis, N.M., et al., *Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data*. Microbiome, 2018. 6(1): p. 226.
17. Lagkouvardos, I., et al., *Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons*. PeerJ, 2017. 5: p. e2836.
18. McMurdie, P.J. and S. Holmes, *phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data*. PLoS One, 2013. 8(4): p. e61217.
19. Chong, J., et al., *Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data*. Nat Protoc, 2020. 15(3): p. 799-821.

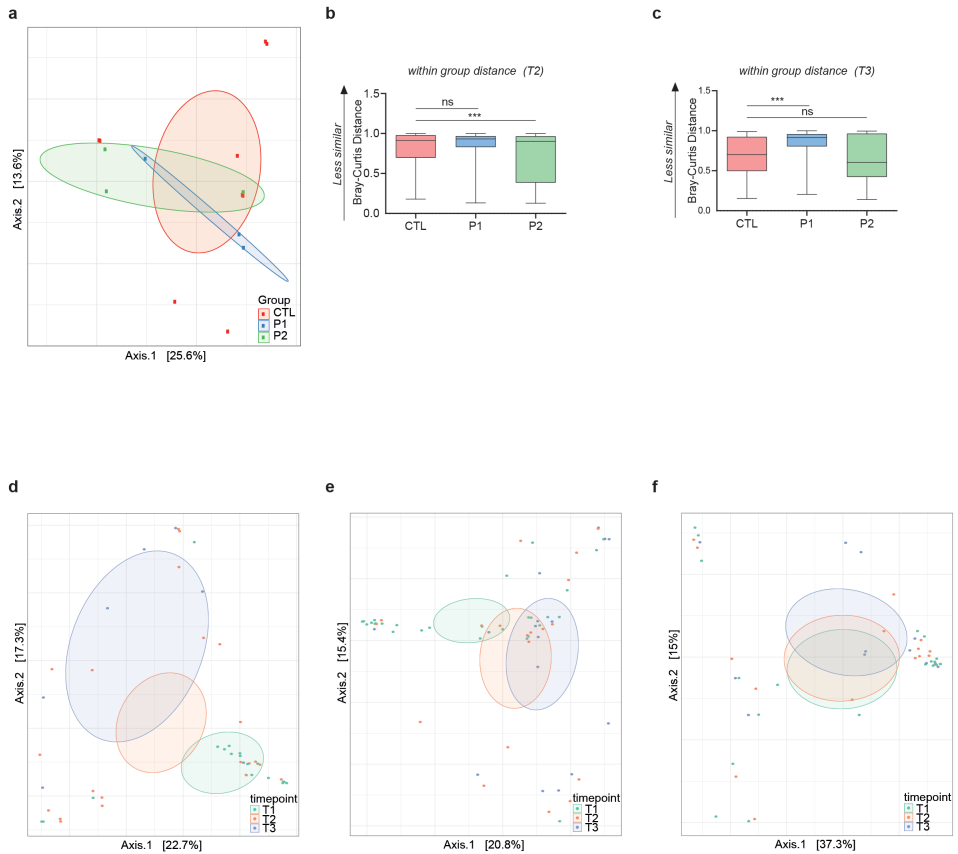
20. Segata, N., et al., *Metagenomic biomarker discovery and explanation*. *Genome Biol*, 2011. 12(6): p. R60.
21. Junick, J. and M. Blaut, *Quantification of human fecal bifidobacterium species by use of quantitative real-time PCR analysis targeting the groEL gene*. *Appl Environ Microbiol*, 2012. 78(8): p. 2613-22.
22. *United States Pharmacopeial Convention, in Food Chemicals Codex*. (2015): Rockville, MD: United States.
23. Flipse, J., et al., *Appearance of vanD-positive Enterococcus faecium in a tertiary hospital in the Netherlands: prevalence of vanC and vanD in hospitalized patients*. *Sci Rep*, 2019. 9(1): p. 6949.
24. von Wintersdorff, C.J., et al., *High rates of antimicrobial drug resistance gene acquisition after international travel, The Netherlands*. *Emerg Infect Dis*, 2014. 20(4): p. 649-57.
25. Friedman, J. and E.J. Alm, *Inferring correlation networks from genomic survey data*. *PLoS Comput Biol*, 2012. 8(9): p. e1002687.
26. Olm, M.R., et al., *Necrotizing enterocolitis is preceded by increased gut bacterial replication, Klebsiella, and fimbriae-encoding bacteria*. *Sci Adv*, 2019. 5(12): p. eaax5727.
27. Delaplain, P.T., et al., *Effects of artificially introduced Enterococcus faecalis strains in experimental necrotizing enterocolitis*. *PLoS One*, 2019. 14(11): p. e0216762.
28. Costeloe, K., et al., *Bifidobacterium breve BBG-001 in very preterm infants: a randomised controlled phase 3 trial*. *Lancet*, 2016. 387(10019): p. 649-60.
29. Alcon-Giner, C., et al., *Microbiota supplementation with Bifidobacterium and Lactobacillus modifies the preterm infant gut microbiota and metabolome*. *bioRxiv*, 2019: p. 698092.
30. Kristensen, N.B., et al., *Alterations in fecal microbiota composition by probiotic supplementation in healthy adults: a systematic review of randomized controlled trials*. *Genome Med*, 2016. 8(1): p. 52.
31. Zmora, N., et al., *Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features*. *Cell*, 2018. 174(6): p. 1388-1405 e21.
32. Yatsunenکو, T., et al., *Human gut microbiome viewed across age and geography*. *Nature*, 2012. 486(7402): p. 222-7.
33. Chu, D.M., 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): p. 314-326.
34. Nino, D.F., C.P. Sodhi, and D.J. Hackam, *Necrotizing enterocolitis: new insights into pathogenesis and mechanisms*. *Nat Rev Gastroenterol Hepatol*, 2016. 13(10): p. 590-600.
35. Esaiassen, E., et al., *Effects of Probiotic Supplementation on the Gut Microbiota and Antibiotic Resistome Development in Preterm Infants*. *Front Pediatr*, 2018. 6: p. 347.
36. Pammi, M., et al., *Intestinal dysbiosis in preterm infants preceding necrotizing enterocolitis: a systematic review and meta-analysis*. *Microbiome*, 2017. 5(1): p. 31.
37. Mshvildadze, M., et al., *Intestinal microbial ecology in premature infants assessed with non-culture-based techniques*. *J Pediatr*, 2010. 156(1): p. 20-5.
38. Zhou, Y., et al., *Longitudinal analysis of the premature infant intestinal microbiome prior to necrotizing enterocolitis: a case-control study*. *PLoS One*, 2015. 10(3): p. e0118632.
39. Claud, E.C., et al., *Bacterial community structure and functional contributions to emergence of health or necrotizing enterocolitis in preterm infants*. *Microbiome*, 2013. 1(1): p. 20.
40. Fouhy, F., et al., *Perinatal factors affect the gut microbiota up to four years after birth*. *Nat Commun*, 2019. 10(1): p. 1517.
41. Fujimura, K.E., et al., *Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation*. *Nat Med*, 2016. 22(10): p. 1187-1191.

42. Hansen, R., et al., *First-Pass Meconium Samples from Healthy Term Vaginally-Delivered Neonates: An Analysis of the Microbiota*. PLoS One, 2015. 10(7): p. e0133320.
43. Sender, R., S. Fuchs, and R. Milo, *Revised Estimates for the Number of Human and Bacteria Cells in the Body*. PLoS Biol, 2016. 14(8): p. e1002533.
44. Esaiassen, E., et al., *Bifidobacterium longum Subspecies infantis Bacteremia in 3 Extremely Preterm Infants Receiving Probiotics*. Emerg Infect Dis, 2016. 22(9): p. 1664-6.

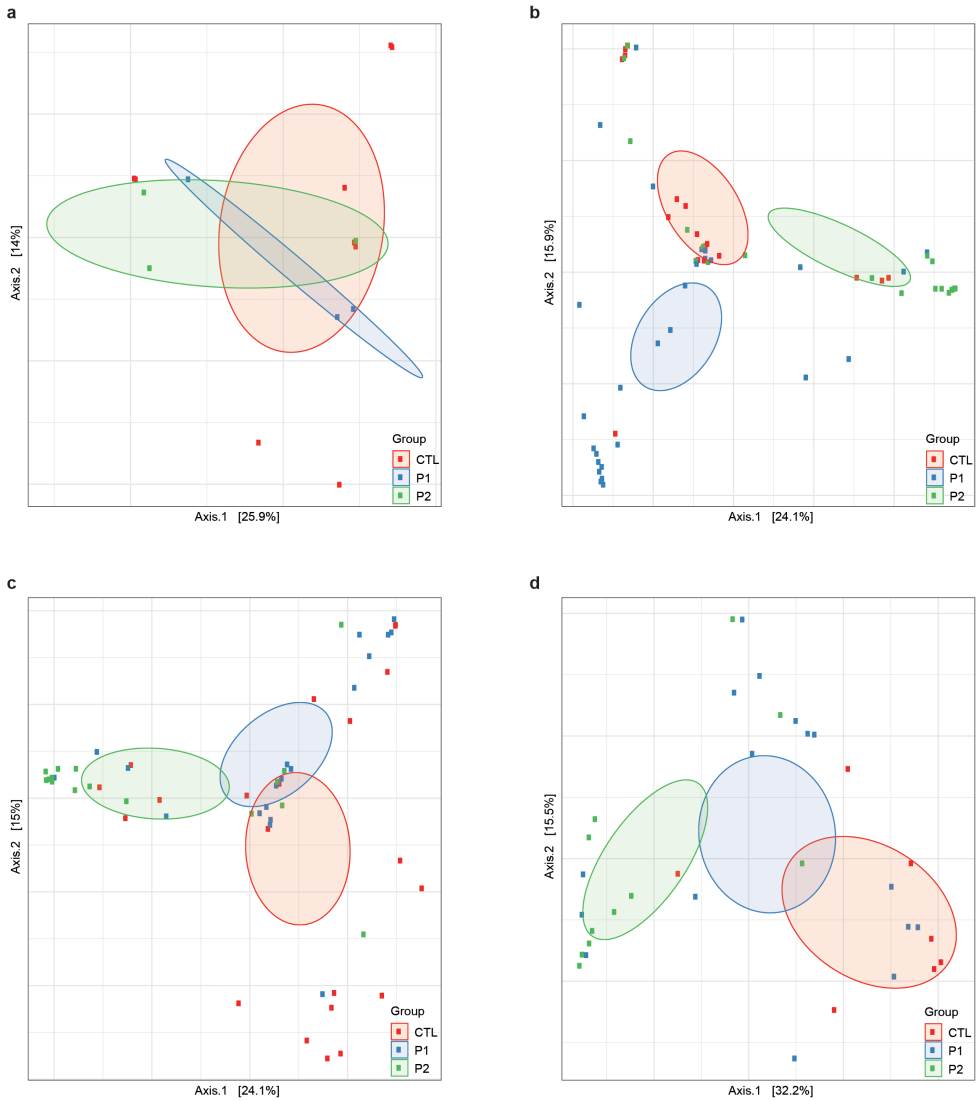
## Supplementary Data



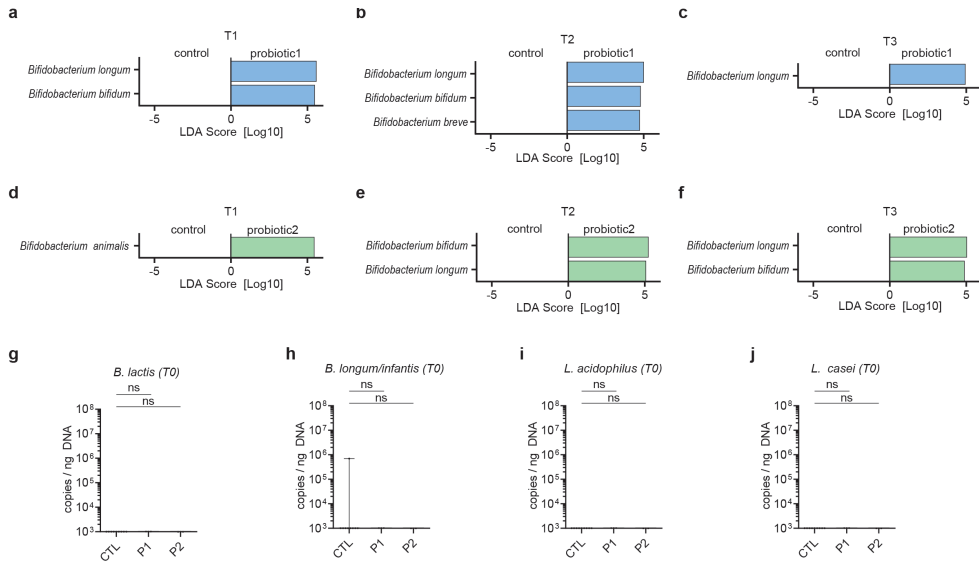
**Figure S1 - Microbiota diversity in preterms with or without probiotic supplementation. (A-D)** Richness (number of ASVs) of the three different groups at T0 (A), T1 (B), T2 (C) and T3 (D). (E-H) Shannon index for the microbiota of the three different groups at T0 (E), T1 (F), T2 (G) and T3 (H). Kruskal-Wallis test to controls with Dun's post-test (control-group is reference), mean and SD; ns, not significant).



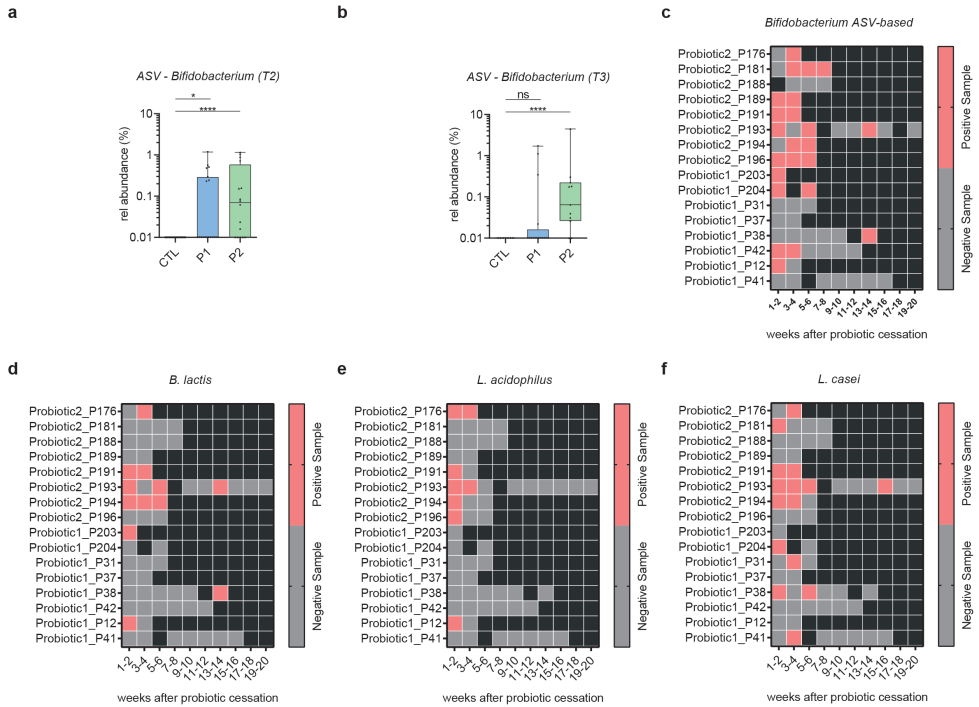
**Figure S2 - Microbial composition and stability in preterms with or without probiotic supplementation. (A)** PCoA based on ASV-derived Bray-Curtis dissimilarity does not show significant differences in microbial community structure between control and probiotic groups at T0 (not significant, Permanova). **(B-C)** Bray-Curtis within group (control-group is reference) distance at T2 **(B)**, and T3 **(C)**. **(D-F)** PCoA based on ASV-derived Bray-Curtis dissimilarity in microbial community structure between T1-T3 for the control group **(D)**, probiotic 1 group **(E)**, and probiotic 2 group **(F)**. (Kruskal-Wallis test with Dun's test for post-hoc comparisons, mean and SD; \*\*\*,  $p < 0.001$ ; ns, not significant).



**Figure S3 - Microbial composition in preterms with or without probiotic supplementation upon removal of probiotic ASVs. (A-D) PCoA based on Bray-Curtis dissimilarity at T0 (ns: not significant, Permanova) (A), T1 ( $p < 0.001$ , Permanova) (B), T2 ( $p < 0.001$ , Permanova) (C) and T3 ( $p < 0.001$ , Permanova) (D).**



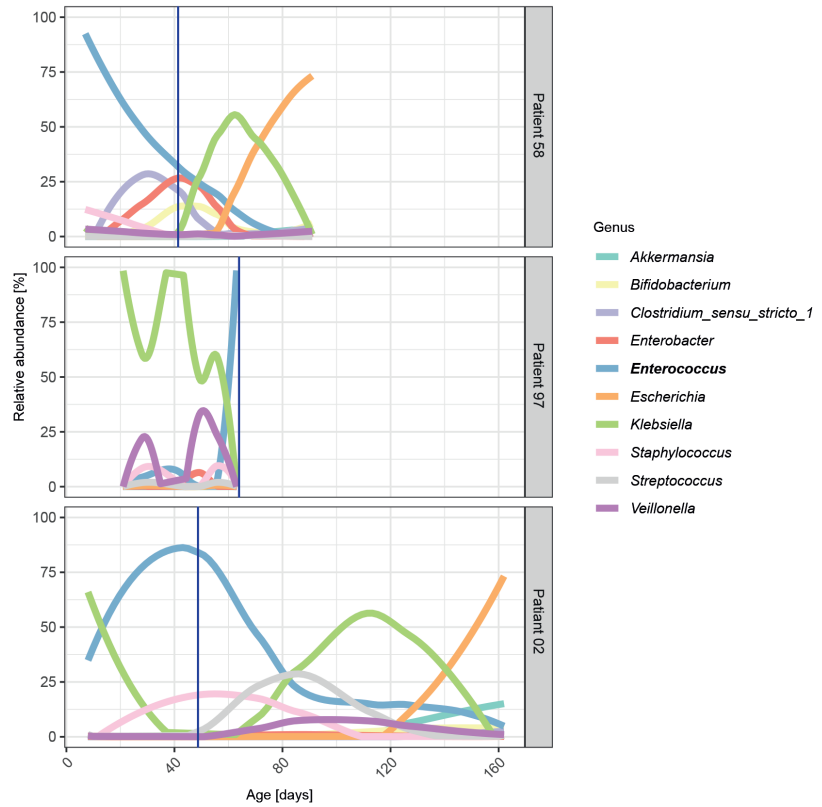
**Figure S4 - Association of probiotic administration with specific taxa. (A-C)** LefSe with bifidobacterial species between probiotic 1 supplemented and control neonates at T1, **(A)** at T2, **(B)** and at T3 **(C)**. **(D-F)** LefSe with bifidobacterial species between probiotic 2 supplemented and control neonates at T1, **(D)** at T2, **(E)** and at T3 **(F)** (LDA-score>0.2 and  $p < 0.05$ , Wilcoxon signed-rank test). **(G-J)** Quantitative abundance before probiotic supplementation at T0 of the probiotic bacteria *Bifidobacterium lactis* **(G)**, *Bifidobacterium longum* **(H)**, *Lactobacillus acidophilus* **(I)**, *Lactobacillus casei* **(J)** (Kruskal-Wallis test with Dun's test for post-hoc comparisons (control-group is reference), mean and SD; ns, not significant).



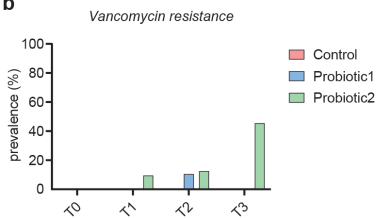
**Figure S5 - Colonization by probiotic bacteria after cessation of supplementation. (A-B)** Relative abundance of probiotic-specific *Bifidobacterium*-ASVs that are absent in control neonates after cessation of probiotic supplementation at T2 (A) and T3 (B) (Kruskal-Wallis test with Dun’s test for post-hoc comparisons (control-group is reference), mean and SD; \*,  $p < 0.05$ ; \*\*\*\*,  $p < 0.0001$ ; ns, not significant). (C-F) Colonization prevalence from longitudinally sampled neonates ( $n = 8$  / group) after cessation of probiotic supplementation for ASV-tracked *Bifidobacterium* (C), qPCR-tracked *Bifidobacterium lactis* (D), *Lactobacillus acidophilus* (E), and *Lactobacillus casei* (F). No necrotizing enterocolitis, sepsis or antibiotic-treatment at the time of sample collection occurred in the included neonates directly after probiotic cessation except for P41 and P188 (both antibiotics provided 1-2 weeks after probiotic cessation for sepsis treatment).



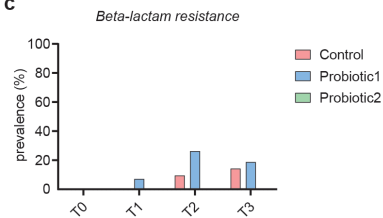
a



b

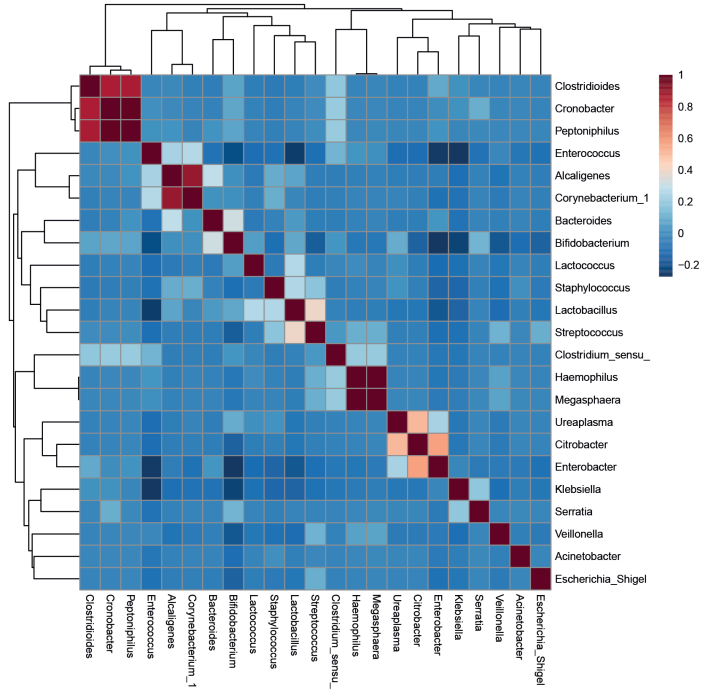


c

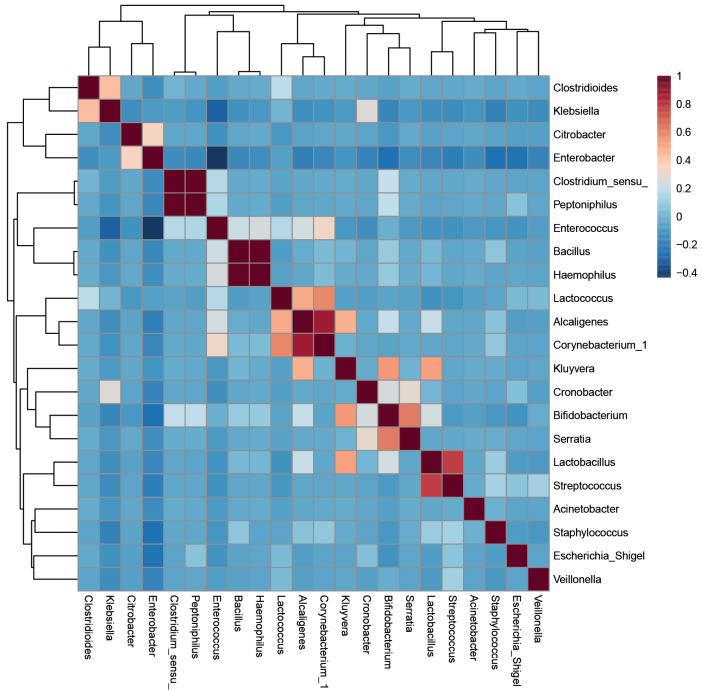


**Figure S6 - Microbial dynamics in preterm neonates developing necrotizing enterocolitis. (A)** The relative abundances of the 10 most abundant genera over time around the diagnosed onset of necrotizing enterocolitis (blue line) for three patients. **(B)** Prevalence of vanA/vanB genes conferring resistance to vancomycin. **(C)** Prevalence of extended-spectrum beta-lactamase encoding genes (CTX-M genogroups 1, 2 and/or 9) conferring resistance to most beta-lactam antibiotics.

a



b



**Figure S7 - : Correlations between genera upon probiotic administration. (A-B)** Correlation heatmap based on the coefficients of the SparCC analyses between relative genus abundances of control and probiotic 1 administered neonates at T1 (A), or between control and probiotic 2 administered neonates at T1 (correlation>0.3, p<0.05) (B).



Supplementary Table 1 – Demographics of the study population

Variable*	Controls			Probiotic1			Probiotic2					
	T0 [n (%)]	T1 [n (%)]	T2 [n (%)]	T3 [n (%)]	T0 [n (%)]	T1 [n (%)]	T2 [n (%)]	T3 [n (%)]	T0 [n (%)]	T1 [n (%)]	T2 [n (%)]	T3 [n (%)]
<b>Birthmode</b>												
Vaginal delivery	0 (0)	1 (5)	2 (10)	0 (0)	1 (33)	2 (7)	1 (5)	2 (12)	1 (33)	2 (10)	2 (13)	1 (9)
Cesarean section	9 (100)	19 (95)	19 (90)	7 (100)	2 (66)	26 (93)	18 (95)	14 (88)	2 (67)	19 (90)	14 (88)	10 (91)
<b>Antibiotics at T after age of 3 days</b>												
Yes	2 (22)	1 (5)	1 (5)	0 (0)	1 (33)	3 (11)	0 (0)	1 (6)	2 (67)	2 (10)	0 (0)	0 (0)
No	7 (78)	14 (70)	18 (86)	6 (86)	2 (66)	25 (89)	19 (100)	15 (94)	0 (0)	19 (90)	16 (100)	11 (100)
<b>Birth weight</b>												
<1000 g	3 (33)	9 (45)	9 (43)	4 (57)	2 (66)	10 (36)	13 (68)	11 (69)	3 (100)	15 (71)	14 (87)	10 (91)
≥1000 g	6 (67)	11 (55)	12 (57)	3 (43)	1 (33)	18 (64)	6 (32)	5 (31)	0 (0)	6 (29)	2 (13)	1 (10)
<b>Gestational age</b>												
<28	2 (22)	8 (40)	8 (38)	3 (43)	1 (33)	11 (39)	8 (42)	11 (69)	3 (100)	13 (62)	12 (75)	9 (82)
≥28	7 (78)	12 (60)	13 (62)	4 (57)	1 (33)	17 (61)	11 (58)	5 (31)	0 (0)	8 (38)	4 (25)	2 (18)
<b>PROM</b>												
at delivery	6 (67)	12 (60)	13 (62)	4 (57)	3 (100)	19 (68)	12 (63)	7 (44)	3 (100)	17 (81)	14 (87)	10 (91)
before delivery	3 (33)	8 (40)	8 (38)	3 (43)	0 (0)	9 (32)	7 (37)	9 (56)	0 (0)	4 (19)	2 (13)	1 (9)
<b>Sepsis</b>												
Yes	2 (22)	1 (5)	1 (5)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
No	7 (78)	14 (70)	18 (86)	6 (86)	3 (100)	28 (100)	19 (100)	16 (100)	3 (100)	21 (100)	16 (100)	11 (100)
<b>Milktype</b>												
Breastfeeding	3 (33)	7 (35)	7 (33)	2 (29)	0 (0)	17 (61)	10 (53)	8 (50)	3 (100)	15 (71)	13 (81)	8 (73)
Formula	5 (56)	6 (40)	10 (48)	3 (43)	3 (100)	8 (29)	6 (32)	4 (25)	0 (0)	6 (29)	3 (19)	3 (27)
Mixed	1 (11)	3 (15)	2 (10)	1 (14)	0 (0)	3 (11)	3 (16)	4 (25)	0 (0)	0 (0)	0 (0)	0 (0)
<b>Sex</b>												
Male	3 (33)	7 (35)	8 (38)	2 (29)	1 (33)	13 (46)	11 (58)	7 (44)	2 (67)	11 (52)	9 (56)	8 (73)
Female	6 (67)	13 (65)	13 (62)	5 (71)	2 (66)	15 (54)	8 (42)	9 (56)	1 (33)	10 (48)	7 (44)	3 (27)
<b>Variable - Median [IQR]</b>												
Birth Weight (in grams)	1040 [985-1085]	1038 [824-1123]	1040 [835-1130]	990 [748-1128]	970 [760-1450]	943 [830-2171]	900 [738-1150]	825 [611-1113]	880 700 [725-680]	880 700 [680-1030]	870 870 [695-940]	780 780 [670-943]
Gestational age (in weeks)	29 [28-29]	28 [27-29]	28 [27-29]	29 [26-29]	30 [29-30]	29 [26-30]	29 [26-29]	26 [24-28]	26 [25-26]	27 [26-28]	27 [26-27]	27 [26-27]

\*Numbers may not add up to 100% due to missing data

Supplementary Table 2 – Permutational analysis of the variance

Variable	Timepoint	P-value
Milktype	T0	0.211
	T1	<b>0.034</b>
	T2	0.074
	T3	<b>0.005</b>
Sepsis	T0	0.437
	T1	0.465
	T2	0.231
	T3	NA
Antibiotics	T0	0.506
	T1	0.344
	T2	0.247
	T3	0.167
Birthmode	T0	0.17
	T1	0.972
	T2	0.607
	T3	0.756
Birthweight	T0	0.204
	T1	0.893
	T2	0.139
	T3	0.52
Gestational age	T0	<b>0.044</b>
	T1	0.963
	T2	0.247
	T3	0.096
Sex	T0	0.77
	T1	0.732
	T2	0.228
	T3	0.513
PROM	T0	0.97
	T1	0.732
	T2	<b>0.02</b>
	T3	0.514

Supplementary Table 3 – PERMANOVA on P groups corrected by significant covariates

Adjusted for	Timepoint	P-value for probiotic group
Milktype	T1	<b>0.001</b>
	T3	<b>0.002</b>
Gestational Age	T0	<b>0.125</b>
PROM	T2	<b>0.001</b>



Noa Penders (12 years)

# **Chapter 8**

**General discussion and summary**

### **The role of the gut microbiome and how it unfolds in early life**

Our body harbors numerous microorganisms, including bacteria, archaea, eukaryotic microbes and viruses, which together exert a wide variety of vital biological and physiological processes. The majority of these microbes consist of bacteria that help with dietary digestion (e.g. resistant carbohydrates), produce essential micronutrients and signaling molecules in the gut (i.e. vitamins and serotonin) [1] and defend against infections by providing colonization resistance [2]. The intestinal microbiome plays also a critical role in the maturation of our immune system, particularly in the early phases of life, during infant development and growth [3-5]. In contrast to the microbiota of adults, which is relatively stable, this early age-window is characterized by major changes in the gut microbiota composition. Given the instability of the microbiome during this early time-window and its strong and non-redundant priming effect on the host's immune system, a comprehensive insight into the ecological factors that shape the microbiota during early life is of particular importance. Not in the least, because this co-development of the microbiome and the host immune system during infancy may have life-long consequences on the susceptibility to inflammatory and immune-mediated diseases [4-6]. To this end, my thesis aimed to unravel environmental and host-related factors that determine the ultimate composition and explain inter-individual variation in the microbiota during infancy, which is pivotal in order to better understand the role of the microbiota in health and disease.

### **The development of the gut microbiota during the postnatal period**

In **chapter 2**, we provide an overview of the origin and importance of pioneer bacteria and the known evolutionary factors that influence the development of the infant gut microbiota. The microbiota establishes from the initial inoculum to a complex and diverse ecosystem during the first years of life and is influenced by a large number of perinatal environmental and maternal factors, e.g. mode of delivery, antibiotic treatment, breast milk *versus* formula and other dietary factors. Thus, microbial maturation relies on the transmission of microbes (dispersal) and is further influenced by factors that provide a selective advantage to certain bacteria (habitat filtering) or impact succession of initial bacteria (historical contingency) during infancy. However, the perinatal, environmental and maternal factors studied so far together can only partly explain the individual variation in the microbiota development. One explanation is the lack of detailed insight in the precise dynamics of the microbiome development in early life as the majority of birth cohort studies

only collected fecal samples at a few time-points during infancy or focused on a limited number of determinants. In addition, many other unidentified factors, such as internal host factors, might also explain why our understanding of inter-individual microbiota variation is still limited.

In **chapter 3**, we therefore closely monitored the development and succession of bacteria in both the small intestine as well as colon of mice and examined whether host metabolic factors are involved in this maturation process of the microbiota. Careful microbiota profiling throughout the postnatal period until adulthood revealed major age-dependent compositional changes. The bacterial richness and diversity initially decreased in both small intestine and colon within the first 24 hours after birth. This reduction immediately after birth most likely illustrates the inability of many of the ingested bacterial species to permanently colonize the neonatal intestine during or shortly after birth consistent with previous reports [7-9]. In addition, the general microbiota composition, dominated by facultative anaerobes such as *Staphylococcus*, *Streptococcus* and *Escherichia*, was highly similar between the colon and small intestine early after birth. The most profound changes could be observed upon weaning – the complete transition to solid food and cessation of breastfeeding. After weaning, the composition diverged into a site-specific intestinal microbiota. In the small intestine the relative increase of *Lactobacillus* continued, whereas a more complex composition developed in the colon, characterized by an expansion of obligate anaerobes such as *Bacteroides* among others. Altogether, this suggests a late emergence of an organ-specific microbiota that occurs around weaning.

Next to these dietary changes, the development of digestive organs is accompanied by increased metabolic capacity during the postnatal period [10, 11]. These metabolic and dietary changes may together contribute to intestinal substrate availability and could thereby influence the observed bacterial succession and overall microbial ecosystem maturation [12]. We therefore screened hepatic bile acids, amino acids, biogenic amines, acylcarnitines, glycerophospholipids, sphingolipids and sugar at different ages. The switch in diet during weaning most likely induced the observed moderate but significant decrease of mainly amino acids and biogenic amines towards adulthood. These metabolites are present in breast milk and the majority of the amino acids in the gut is indeed suggested to derive from nitrogen sources such as dietary proteins while only a fraction is produced by intestinal bacteria [13, 14]. Moreover, we identified bile acids as potential drivers of the site-specific microbiota development with causality being confirmed by functional *in vitro* tests and oral bile acid administration. The selected bile acids (e.g. taurine-conjugated  $\beta$ -muricholic acid



and ursodeoxycholic acid) particularly fostered the succession of *Lactobacillus* within the small intestine. Lactobacilli are well-known for their health-beneficial properties [15, 16], and the rapid expansion of *Lactobacillus* along with the increase in bacterial richness upon bile acid supplementation could also promote colonization resistance. Hence, the capability of the microbiota community to protect our gut against pathogens might be accelerated by bile acids, preventing adverse effects on the infant's health [chapter 7]. The concept that bile acids can mediate colonization resistance is further supported by data demonstrating that specific bile acids prevent infections by the opportunistic pathogen *Clostridioides difficile* [17-19]. Therefore, bile acids might contribute to the outcompetition of pathogens. Moreover, chapter 3 clearly demonstrated that host factors, often omitted in microbial ecology studies, are important in the maturation of the microbiota in general. The use of animal studies is crucial to decipher the complete puzzle of determinants that contribute to microbial homeostasis. The controlled environment and shared lifestyle of these animal models are suitable for subsequent studies discovering the local host aspects associated with the infant microbiota. All in all, future research should take the ontogeny of other organs into account in combination with the focus on the detailed host-microbe interactions herein.

In chapter 3, we provide further results that support the idea that the development and maturation of the intestinal microbiota continues towards adulthood as reported previously by other studies [20, 21], chapter 5]. The early colonization phase was characterized by a high interindividual variation and low diversity. Towards adult-age microbial diversity continuously increased and the interindividual variation decreased. Interestingly, during the postnatal period the abundance of the most dominant genera such as *Escherichia* gradually decreased in favor of a relative enrichment in *Lactobacillus* in both small intestine and colon [chapter 3]. This dramatic reduction of Proteobacteria is in line with findings of our human birth cohort towards school-age [chapter 5] where *Escherichia* was the most abundant genus until 31 weeks after birth.

In line with chapter 3 and a few other studies [12, 22, 23], our data from the birth cohort presented in chapter 5 show that the dietary transition with cessation of breastmilk was more strongly related with microbial compositional changes and maturity than the introduction of solid food. The latter suggests that the change in diet to solid food does not result in major shifts in microbial community structure as long as breastfeeding is continued. The microbial transformation upon cessation of breastfeeding induces powerful responses of the mucosal immune system for later in life, the so-called 'weaning reaction' [24]. This

reaction is characterized by enhanced expression of intestinal interferon (IFN)- $\gamma$  and tumor necrosis factor (TNF)- $\alpha$ , which is caused by the omission of epidermal growth factor (EGF) present in breast milk. Interestingly, this response led to protection to inflammatory diseases in adults associated with induction of specific regulatory T cells. However, this ‘weaning reaction’ seems to be highly time-dependent as it does not occur in an effort to introduce this change a week later. An intriguing phenomenon that emphasizes the combined effect of the microbiota, diet, and ontogeny in priming of the immune system and lifelong gut homeostasis.

The dietary changes continue during infancy, impacting the aforementioned substrate availability and associated microbial successions. In particular, a wide range of bacteria consume dietary fibers and generate hydrogen and beneficial short-chain fatty acids (SCFA’s). The methanogenic archaea, often neglected as members of the gut microbiota, stimulate the constant fermentation of these dietary fibers by metabolizing excessive hydrogen generated during carbohydrate fermentation into methane. In **chapter 4**, we showed that among environmental factors such as hygiene, medication, lifestyle and diet the occurrence of methanogenic archaea (*Methanobrevibacter smithii*) in children positively correlated with the consumption of organic yogurt and milk. We subsequently confirmed that these archaea are indeed present in milk products, suggesting that dairy consumption may facilitate archaeal gut colonization in children.

Colonization of methanogens is of importance in facilitating the fermentation capacity of the microbiome and thereby the production of the immune-modulatory SCFAs which in turn might protect against the development of inflammatory or immune-mediated diseases such as allergies. The methanogenic archaea directly stimulate the induction of adaptive and innate immune responses and expansion of specific B and T cells. Moreover, they are able to regulate the secretion of antimicrobial peptides, activate the inflammasome, and drive the production of certain cytokines such as TNF (Th1) [25]. The beneficial events are supported by a first study on the association of intestinal colonization by methanogenic archaea and asthma in the same pediatric cohort [[26], **chapter 4**]. A lower risk of asthma was associated with higher abundance of archaea in these children. Moreover, a trend for a decreased risk for atopic dermatitis and sensitization to allergens was also observed in the presence of archaea. In addition, there is evidence that organic dairy consumption as such is protective against eczema [27]. Thus, these data again shed light on the complex functional interactions between dietary, microbial and developmental facets of the immune system, which need to be elucidated together to unravel the link between

the presence of methanogenic archaea and the onset of allergic diseases in early life.

Mode of delivery is perhaps one of the factors most commonly linked to changes in the microbiota establishment in early life [21, 28]. C-section numbers are rising worldwide [29]. Nowadays one out of five babies is delivered by c-section with numbers having almost doubled since 2000. Notably, the Netherlands has one of the lowest c-section rates (17%) among developed countries, especially when compared to neighboring countries such as Germany (30%) [29]. The latter might be because pregnant women at low risk of birth complications get unique midwife-led care in the Netherlands which is associated with lower intervention rates [30, 31]. Although in some cases it might be a life-saving procedure, the high number of c-section without medical reason (called elective c-section) are worrisome. The microbiota of c-section delivered infants is characterized by delayed colonization of important members like *Bacteroides* species compared to vaginally delivered infants [12, 21, 28, 32-36]. In **chapter 5**, we studied the establishment of the human infant's gut microbiota longitudinally and further support that *Bacteroides* are indeed most strongly affected by the birth mode. The decreased levels of *Bacteroides* in infants born by c-section remained significant in our study after careful adjustment for other confounders such as feeding type and persisted for up to 31 weeks. Although recent findings further demonstrated that the compositional effect of these pioneer bacteria could last until 4-years postpartum in the infant gut [33, 37], we could no longer identify differences at school-age (6-11 years) between children born by c-section as compared to those born vaginally. In **chapter 7**, we investigated the microbiota of preterm infants during the initial months of life. These preterms were most often delivered by c-section, and *Bacteroides* were indeed absent among the abundant taxa. Although c-section is accompanied by prophylactic antibiotic administration to the mother, the effect of delivery mode on the infant's gut microbiota has recently been demonstrated to be independent of this antibiotic exposure [32]. In this study, the intrapartum antibiotic dose was administered only after clamping of the umbilical cord in mothers undergoing c-section, enabling the researchers to examine the impact of the birth mode on the neonatal microbiota in the absence of antibiotic exposure to the baby. This study confirmed the lower levels of *Bacteroides spp.* and *Bifidobacterium spp.* in c-section delivered infants, suggesting that maternal transfer during vaginal delivery is crucial to acquire certain microbial species in early life. Moreover, new evidence suggests that the microbiota seeding from mother to infant not only occurs by vaginal transmission but rather via fecal microbial transmission during delivery [32, 36]. The latter has been investigated by comparing both maternal rectal and vaginal

samples to the newborn's fecal microbiota demonstrating more identical bacterial strains in infant samples and the maternal samples obtained rectally. Altogether, our (**Chapter 5**) and previous work show that *Bacteroides* represent the main maternal strains being transferred to the intestine of neonates born by vaginal delivery [28, 36, 38]. The lack of transmission of these strains during c-section delivery is related to lower levels of TNF- $\alpha$  and IL-18 in both the neonatal blood as well as the culture medium of primary immune cells stimulated with endotoxin isolated from their fecal material [38, 39]. This decrease in endotoxin-induced immune system priming has the potential for profound effects in later life.

### **The gut microbiota in the pathogenesis of (allergic) diseases**

The alterations in the gut microbiota have been implicated in the pathogenesis of several inflammatory and immune-mediated disorders in adults, including Crohn's disease [40, 41], ulcerative colitis [42] and non-alcoholic fatty liver disease [43] as well as other chronic disorders such as obesity [44], autism [45] and allergies [46]. There is increased concern that disruption of the early life microbiota might play a causal role in the onset of these conditions, especially in allergic diseases like asthma and atopic dermatitis. For instance, as compared to babies born vaginally, infants born by c-section more frequently experience infections, adverse effects on immune development and inflammatory disorders, including allergies [47, 48]. In the context of two birth cohort studies, mode of delivery was indeed shown to strongly affect both the risk of allergies and asthma and the infant microbiota [49]. Nowadays, these atopic diseases are the most common chronic pediatric disorders affecting more than 300 million children in the (western) world [50]. Although numerous epidemiological studies suggest that the infant intestinal microbiota plays an important role in the manifestation of allergic diseases and asthma, these are mostly based on cross-sectional studies and thus lack early and additional samples prior to disease onset. Therefore, the timing and nature of the influence of intestinal microbes in the etiology of allergies remains largely unknown and has been incompletely elucidated.

In **chapter 5**, we therefore studied the development of the gut microbiota during infancy at several time points during the first year of life and once at school-age, and related microbiota dynamics to the onset of allergies and asthma. We showed that the microbiota in general, its developmental state as well as the abundance of specific bacterial taxa were associated with the subsequent development of allergic disorders. For instance, using longitudinal analyses and careful correction of potential confounders, we found the abundance of *Lachnobacterium*

to be significantly lower throughout infancy both in children who developed atopic dermatitis and asthma as compared to children that remained free from allergic symptoms whereas a lower abundance of *Faecalibacterium* was indicative for atopic dermatitis.

The protective property of *Faecalibacterium (prausnitzii)* against allergies could be due to its immunomodulatory activities [51]. The bacterial metabolite butyrate and a specific microbial anti-inflammatory molecule (MAM) of *Faecalibacterium* might contribute to these effects [52, 53]. MAM inhibits nuclear factor  $\kappa$ B (NF- $\kappa$ B) activity in immune cells of the lamina propria and thereby reduces the interferon (IFN)- $\gamma$  and interleukin (IL)-17 production. Together with the epigenetic regulatory effect of butyrate this may induce differentiation and expansion of regulatory T cells. Moreover, butyrate also suppresses allergen-induced histamine release by mast cells via histone deacetylases [54]. Our reported association between *Faecalibacterium* abundance and the risk of subsequent allergy development confirmed previous findings. The CHILD study, for instance, revealed *Faecalibacterium* to be decreased in children at 3 months of age who were at risk for allergic wheeze at one year of age [55]. Furthermore, a lower relative *Faecalibacterium* abundance at one year of age was associated with an increased risk of asthma at five years in the COPSAC study [22].

To date, we seem to be the first that link substantial perturbations of *Lachnobacterium* to the development of allergic manifestations. The latter might be due to the late discovery of this microbe [56] and so far its depletion has only been associated with primary biliary cirrhosis [57] and ulcerative colitis[58]. *Lachnobacterium* strains are capable to ferment sugars such as lactose and glucose into primarily lactic acid and minor amounts of butyric and acetic acids. Although these strains have characteristics similar to those of *Lachnospira*, they are metabolically highly equivalent to the common lactic acid-producing bacteria, i.e. *Lactobacillus*. The acidic environment that these bacteria create might on one hand avoid colonization by harmful invaders. On the other hand, lactic acid bacteria can modulate the dendritic cell and natural-killer cell interactions, shifting T-cell maturation towards Th1 by induction of IL-12 and IFN- $\gamma$  production, respectively [59-61]. Alternatively, lactic acid bacteria are able to modify the DNA methylation patterns of CD4+ T cells prompting enhanced immune activation at birth [61]. Moreover, an increased abundance of *Lachnobacterium* in infants at three to six months of age was linked to higher vitamin D levels directly after birth [62]. Although causality should be warranted, this may provide another explanation as vitamin D is an immune modulator that can induce synthesis of antimicrobial peptides such as cathelicidins and activate toll-like receptors (TLRs) [63, 64]. Asthma might be accompanied

by periods of acute worsening of the symptoms, named exacerbations, which are frequently triggered by viral infections. The latter is also related to subsequent allergic responses by increased production of pro-inflammatory cytokines like IL-17A [65]. The anti-viral actions of vitamin D together with the inhibition of IL-17A is thereby able to decrease the susceptibility to microbial infections and has been linked to reduced asthma exacerbations [66]. All in all, our study provides important insight to stimulate further analysis on this possibly pivotal bacterium.

Although the pathogenesis of asthma may start in early in life [**chapter 5**], the actual diagnosis in the vast majority of children can only be made at 6 years of age and above. Wheezing can be recognized earlier, but only one third of wheezing preschool children will finally develop asthma [67]. Currently, a reliable biomarker to predict, which wheezing children will develop asthma is lacking and little is known about the association between the gut microbiota and asthma development in wheezing preschool children [68]. In **chapter 6**, we therefore examined the gut microbiota in 202 wheezing children aged 2-4 years and 50 healthy controls and the link with asthma at the age of 6 years. We observed that gut microbial diversity and overall community structure at this age were not associated with preschool wheezing or future childhood asthma. However, some bacteria were linked with wheezing and subsequent development of asthma. In particular, a high relative abundance of *Escherichia* was associated with an increased risk of asthma.

**Chapter 6** describes the first study that examined the microbiota in wheezing preschool children and its association with disease progression to asthma. In studies taking earlier time points into account, a general link between microbial diversity and asthma occurrence could indeed often not be observed [[55, 69, 70], **chapter 5**]. Moreover, despite the discrepancy between individual studies some earlier investigations had also reported an association of *Escherichia* with asthma risk [22, 71, 72]. In a large birth cohort study, infants with increased abundance of *Escherichia (coli)* developed IgE-mediated eczema [46]. Members of the genus *Escherichia* are able to produce compounds such as flagellae, fimbriae and endotoxin, which are potent stimuli of TLR receptors, increasing the risk of inappropriate inflammation which may in turn precede allergy development [73-76]. Alternatively, the higher abundance of *Escherichia* might indicate a delayed maturation of the microbiota as it has been shown that this bacterium dominates the composition in early infancy, and decreases thereafter gradually [**chapter 3, chapter 5**]. A delay in microbiota maturation has also recently been reported in children born to asthmatic mothers with later onset of asthma [22]. This delay might however

be partly influenced by decreased amounts of other bacteria such as *Lactobacillus*, which seems to be inversely related to *Escherichia* and thereby to exert a beneficial homeostatic effect in the gut [chapter 3, chapter 7]. Altogether, changes in the microbiota during infancy precede the development of allergic diseases and both the role of these bacteria as well as factors driving these bacteria warrant further investigation.

In both chapter 5 and chapter 6, we attempted to link alterations in the microbiota composition to asthma development. The differences in the results of these studies not only rely on the fact that the asthmatic children were compared to different groups, but perhaps also on the age-window in which the microbiota composition was measured. Although in both studies the infants were clinically examined at school age (from 6 years), the preceding time points for microbiota assessment were mainly during the first year in the PAPS-study (chapter 5) and around 3 years of age in the ADEM-study (chapter 6). The modest microbiota perturbations described in chapter 6 suggest that the early microbial priming for life might have already been completed at this age. These findings are in line with the concept of the window of opportunity that takes place within the first months of life where the microbiota is suggested to have its strongest impact on tolerance development and immune maturation [77, 78]. Previous large longitudinal human cohort and animal studies observed that gut microbial alterations in individuals with allergic diseases indeed mainly occurred within this critical window in early life. This again underscores the importance of a good timing of biosampling and calls for studies with regular clinical examinations throughout infancy in combination with deep immunological phenotyping.

### **Targeted manipulation of the gut microbiota**

Oral administration of specific beneficial bacteria, named probiotics, might be a way for targeted manipulation of the gut microbiota to achieve clinical benefits, i.e. to decrease the incidence of infections and prevent the development of immune-mediated and inflammatory diseases such as allergies, asthma and in particular necrotizing enterocolitis (NEC). In chapter 7, we describe the influence of different interventions with *Lactobacillus* and *Bifidobacterium* strains on the microbiota in preterm neonates. We observed that the probiotics were linked to alterations of the microbial ecosystem and could in part persistently colonize after administration. The administration of probiotic bacteria, especially members of the genus *Bifidobacterium*, were inversely correlated with the abundance of pathobionts associated with NEC and with NEC incidence. Although this might not be unexpected at

first sight, it should be noted that most probiotics do not colonize in adults [79]. Therefore, our study emphasizes that early infancy represents a critical period during which targeted manipulation of the gut microbiota might be possible and could confer a clinical benefit.

The use of bacteria-rich fecal material to cure diseases has already been implemented in our society for thousands of years [80]. In the last decade, the essence of this ancient treatment has encountered a massive revival with a global increase in awareness. The use of these so-called ‘fecal transplantations’ however needs to be tightly controlled as it is still unclear what exactly is introduced into the recipient and decent evidence of the health consequences on the long-term are lacking. For example, case reports described the transmission of multi-resistant bacteria, which subsequently caused life-threatening infections [81]. Therefore, more personalized and precision medicine would be recommended rather than transplanting the complete fecal mass. The studies of this thesis, addressing the involvement of the microbiota in various diseases [**chapter 5, chapter 6**] and highlighting the valuable properties that certain pioneer bacteria exert [**chapter 2, chapter 3, chapter 4, chapter 7**], may contribute to further increase the awareness for precision medicine. For instance, specific members of the genus *Lachnobacterium* or *Bacteroides* could be of potential use to prevent allergic manifestations or delayed maturation after c-section delivery, respectively [**chapter 5**]. Moreover, many efforts in this respect have recently been made in which neonates born by c-section are directly exposed to the maternal vaginal microbiota upon delivery, named ‘vaginal seeding’ or ‘bacterial baptism’. Recent data however challenges the biological feasibility of this vaginal seeding to manipulate the neonatal gut microbiota as the key members of *Bacteroides* that differ by delivery mode seem to originate from the gastrointestinal tract rather than the vagina [36]. Although fecal transplantation might be an alternative strategy for vaginal seeding, concerns remain on the risks and uncertainties of this treatment. It is therefore crucial to first obtain detailed insight into the exact mechanisms, transmission routes and disturbing determinants before continuing with these attempts in restoring the baby’s intestinal microbiota. In accordance with this, the examination of specific bacteria, or perhaps more strongly, the identification of specific neonatal populations that will benefit most from certain microbiota manipulations is pivotal in order to support the development of more targeted interventions. This relates not only to those infants born by c-section but also to those infants that received antibiotic treatment or were susceptible to infections or necrotizing colitis for instance.



**Concluding remarks & future perspectives**

There is a rainforest inside us, consisting of various microbial species that make up an entire complex ecosystem in the gut. The disturbance of such ecosystems could have dramatic implications and for the human intestinal microbiota this might impact health and disease. Since van Leeuwenhoek's discovery 300 years ago, we are only starting to understand the impact and role of our little friends in this forest. The gut microbiota is an intrinsic element of our physiology, which starts to root and grow with birth. The microbial species that colonize first take a crucial part in the development of this ecosystem, possibly influencing the ultimate composition and functionality of the microbiota for life. The (lack of) establishment and expansion of these pioneer species are partly affected by maternal factors, medication, host characteristics, diet and environment. A better understanding of the interplay between the gut microbiota, our body's physiology and the determinants interfering with the microbial development is important given its essential role in metabolic, immune and other processes. This thesis provided novel insight into specific factors driving the intestinal microbial community composition and inter-individual microbiota variations in early life and how changes in the microbiota composition and manipulation thereof relate to the onset of inflammatory diseases such as allergies.

Together with previous studies, this thesis showed that different factors contribute to the variation of the gut microbiota composition between individuals and finally shape a stable microbial community that is as unique as one's fingerprint. However, so far, we can only explain a minor part, and even identical twins sharing the exact similar genetics and most likely the same diet and environment exhibit major differences in their enteric microbiota. Therefore, host genetics appears to exert a low contribution to the microbiota variation as taken from twin studies [82, 83], whereas diet seems to contribute for only about 6% to the overall microbiota variation [84]. In total, the known determinants such as diet, drugs and lifestyle related factors account for less than 36% of the variation [85, 86], and although diet has the biggest relative proportional influence of all known determinants a large part of inter-individual microbiota remains still unexplained. This suggests that we either have overlooked important determinants or that other ecological processes such as historical contingency or stochastic effects play a major role in this unexplained inter-individual microbiota variation. For instance, it has been suggested that the order in which microbes arrive (historical contingency) impacts the community assembly and interactions as well [87]. This 'first come, first serve' principle could occur in early life and hence potentially alter

the development of the gut microbiota in babies including identical twins. Further research should take this ecological phenomenon into account, but also focus on unravelling other so far unknown factors that significantly contribute to the variation and acquisition of the microbiota, and thereby the vital biological processes in life.

The use of big data might be useful in the search for these unknown factors. This comprises the combination of genomics, transcriptomics, and proteomics accompanied by extensive information on environmental exposure to not only unravel novel determinants but also to provide a more detailed functional insight. The latter calls for both large human birth cohort studies as well for well-controlled animal studies and perhaps *in vitro* experiments, preferably paralleling these to overcome the possible drawbacks of each. The emphasis on the link between the different biological systems, the metabolic pathways, immune responses and microbial interactions, which is until now often neglected, is crucial to further increase our understanding. Furthermore, the consideration and eventual proof of causality is of importance herein, i.e. causal inferences via mediation analyses that detect potential directions of association pathways. The archaea are for instance correlated with the asthma risk and influenced by consumption of milk products [chapter 4], but if and how this is exactly interconnected with each other remains unclear. A recommendation would be that this research field with many association studies should attempt to move towards more mechanistic and causative approaches albeit relationships may generate new hypotheses.

We amongst others showed that cessation of breast milk profoundly influences the microbiota to a more mature composition during the initial years of life [chapter 3, chapter 5]. The fact that breast milk is beneficial for the infant's health has been widely accepted and the WHO recommends therefore to breastfeed up to 2 years. Is this however still the right duration in terms of the cooperative microbial, dietary and immune-mediated effects? Which signals contribute to finally complete the ultimate weaning transitions? Why is this time period non-redundant in light of the window of opportunity? Hence, studies are necessary to address these critical aspects also in order to support future targeted microbial therapies to prevent the risk of immune-mediated diseases in early life. In line with this, the selection and understanding of microbial manipulation and appropriate probiotics is pivotal and should rely on essential host-microbe interactions. In particular, increased knowledge on the microbes that are crucial for specific health outcomes may also promote the usage of certain drugs or diets to manipulate and "optimize" the microbiota. The potency of recovering the altered microbial colonization upon disturbance, especially upon antibiotic administration or

c-section delivery needs to be investigated. Lastly, further research is necessary to identify not only beneficial bacteria, but also their lasting long-term effect after administration and the components that impact these pioneer species.

Moreover, the lack of certain microbes could be partly attributed to the lack of environmental exposure of children, spending more time in urban surroundings. Changes in lifestyle thereby could contribute to the increase in chronic inflammatory diseases such as obesity, asthma and allergies. The reduced exposure to nature might impede the beneficial health effects including the acquisition of bacteria, the maturation and balancing of immunological responses, and elevation of the vitamin D level. For instance, favorable ( $\gamma$ -proteo)bacteria were detected in higher abundance in children living closely to non-urban environments such as farms and forests [88]. Moreover, a 'farm-like' microbiota at home, originating from the outdoor environment and green areas around homes, decreased the incidence of asthma and allergy risk [89, 90]. Outdoor activity in a natural biodiverse environment may improve the microbial colonization, decrease the risk of allergies and asthma, and improve children's general well-being. This concept of 'microbiota re-wilding' has recently been explored in a few animal studies. For instance, it has been demonstrated that laboratory mice released into an outdoor environment altered both the microbiota and enhanced the immune activation via the differentiation of T cell populations and circulating granulocytes [91]. In another study, the transplantation of embryos of lab mice into 'wild' mothers generated genetically defined offspring that harbor a microbiota that closely resembles that of wild mice [92]. The microbiota of these 'wildlings' was more resistant to environmental exposures such as diet and antibiotics and improved the immune regulation compared to the normal isolated lab mice. Conversely, the lack of microbes as reflected by germ-free mice results in a higher susceptibility to inflammatory disorders like allergies. Altogether, these studies support that a diverse microbiota derived from the wild outdoor contributes to greater stimulation of the immune system in free-living mammals. The mechanisms of the natural microbiota to induce these effects must be further investigated as it might explain the inter-individual differences in the predisposition to develop certain diseases. Future studies should also address the human interventional strategies if strengthening connections to nature as part of everyday's life during infancy indeed exerts a protective effect. Perhaps we need to go back to our basic roots from ancient times, the original home of *Homo sapiens*, back to nature and into the wilds.

In conclusion, this thesis sheds further light on the relevance of specific host and environmental factors and first settlers in the maturing early life microbiota, seeing more and more the forest for the trees. By contributing to the growing knowledge on the nature and opportunities of the microbiota in early life, our research has provided new leads to further explore the precise mechanisms and determinants of microbiota development and its impact on the host. In order to support the development of personalized and targeted microbial therapies, I would recommend to further unravel these leads between diet, host factors, immune maturation and pioneer species. Additional functional studies integrating animal models or co-culture systems with findings or samples from extensive longitudinal human birth cohorts will aid in understanding the inter-individual microbiome variation and the onset of non-communicable diseases herein. The field of microbiology has long been focusing on methodologies within its own boundaries, but this thesis highlights that the microbial developmental interactions are complex and can only be determined by cross-border research. The different fields of microbiology, immunology, and molecular biology should join their forces to move from compositional associations to functional relations and expand their horizons.

## References

1. Koropatkin, N.M., E.A. Cameron, and E.C. Martens, *How glycan metabolism shapes the human gut microbiota*. Nat Rev Microbiol, 2012. 10(5): p. 323-35.
2. Buffie, C.G. and E.G. Pamer, *Microbiota-mediated colonization resistance against intestinal pathogens*. Nat Rev Immunol, 2013. 13(11): p. 790-801.
3. Gensollen, T., et al., *How colonization by microbiota in early life shapes the immune system*. Science, 2016. 352(6285): p. 539-44.
4. Cahenzli, J., et al., *Intestinal microbial diversity during early-life colonization shapes long-term IgE levels*. Cell Host Microbe, 2013. 14(5): p. 559-70.
5. Olszak, T., et al., *Microbial exposure during early life has persistent effects on natural killer T cell function*. Science, 2012. 336(6080): p. 489-93.
6. Tamburini, S., et al., *The microbiome in early life: implications for health outcomes*. Nat Med, 2016. 22(7): p. 713-22.
7. Wampach, L., et al., *Colonization and Succession within the Human Gut Microbiome by Archaea, Bacteria, and Microeukaryotes during the First Year of Life*. Front Microbiol, 2017. 8: p. 738.
8. Palmer, C., et al., *Development of the human infant intestinal microbiota*. PLoS Biol, 2007. 5(7): p. e177.
9. Mueller, N.T., et al., *Delivery Mode and the Transition of Pioneering Gut-Microbiota Structure, Composition and Predicted Metabolic Function*. Genes (Basel), 2017. 8(12).
10. Henning, S.J., *Ontogeny of enzymes in the small intestine*. Annu Rev Physiol, 1985. 47: p. 231-45.
11. Grijalva, J. and K. Vakili, *Neonatal liver physiology*. Semin Pediatr Surg, 2013. 22(4): p. 185-9.
12. Backhed, F., et al., *Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life*. Cell Host Microbe, 2015. 17(5): p. 690-703.
13. Metges, C.C., *Contribution of microbial amino acids to amino acid homeostasis of the host*. J Nutr, 2000. 130(7): p. 1857S-64S.
14. Lonnerdal, B., et al., *Longitudinal evolution of true protein, amino acids and bioactive proteins in breast milk: a developmental perspective*. J Nutr Biochem, 2017. 41: p. 1-11.
15. Heeney, D.D., M.G. Gareau, and M.L. Marco, *Intestinal Lactobacillus in health and disease, a driver or just along for the ride?* Curr Opin Biotechnol, 2018. 49: p. 140-147.
16. Wypych, T.P., L.C. Wickramasinghe, and B.J. Marsland, *The influence of the microbiome on respiratory health*. Nat Immunol, 2019. 20(10): p. 1279-1290.
17. Weingarden, A.R., et al., *Microbiota transplantation restores normal fecal bile acid composition in recurrent Clostridium difficile infection*. Am J Physiol Gastrointest Liver Physiol, 2014. 306(4): p. G310-9.
18. Buffie, C.G., et al., *Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile*. Nature, 2015. 517(7533): p. 205-8.
19. Kang, J.D., et al., *Bile Acid 7alpha-Dehydroxylating Gut Bacteria Secrete Antibiotics that Inhibit Clostridium difficile: Role of Secondary Bile Acids*. Cell Chem Biol, 2019. 26(1): p. 27-34 e4.
20. Yatsunenko, T., et al., *Human gut microbiome viewed across age and geography*. Nature, 2012. 486(7402): p. 222-7.
21. Bokulich, N.A., et al., *Antibiotics, birth mode, and diet shape microbiome maturation during early life*. Sci Transl Med, 2016. 8(343): p. 343ra82.

22. Stokholm, J., et al., *Maturation of the gut microbiome and risk of asthma in childhood*. Nat Commun, 2018. 9(1): p. 141.
23. Pannaraj, P.S., et al., *Association Between Breast Milk Bacterial Communities and Establishment and Development of the Infant Gut Microbiome*. JAMA Pediatr, 2017. 171(7): p. 647-654.
24. Al Nabhani, Z., et al., *A Weaning Reaction to Microbiota Is Required for Resistance to Immunopathologies in the Adult*. Immunity, 2019. 50(5): p. 1276-1288 e5.
25. Sereme, Y., et al., *Methanogenic Archaea: Emerging Partners in the Field of Allergic Diseases*. Clin Rev Allergy Immunol, 2019. 57(3): p. 456-466.
26. Barnett, D.J.M., et al., *Intestinal archaea inversely associated with childhood asthma*. J Allergy Clin Immunol, 2019. 143(6): p. 2305-2307.
27. Kummeling, I., et al., *Consumption of organic foods and risk of atopic disease during the first 2 years of life in the Netherlands*. Br J Nutr, 2008. 99(3): p. 598-605.
28. Shao, Y., et al., *Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth*. Nature, 2019. 574(7776): p. 117-121.
29. Boerma, T., et al., *Global epidemiology of use of and disparities in caesarean sections*. Lancet, 2018. 392(10155): p. 1341-1348.
30. Offerhaus, P.M., et al., *Change in primary midwife-led care in the Netherlands in 2000-2008: A descriptive study of caesarean sections and other interventions among 807,437 low-risk births*. Midwifery, 2015. 31(6): p. 648-54.
31. Zhao, Y., et al., *Modest Rise in Caesarean Section from 2000-2010: The Dutch Experience*. PLoS One, 2016. 11(5): p. e0155565.
32. Reyman, M., et al., *Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life*. Nat Commun, 2019. 10(1): p. 4997.
33. Stewart, C.J., et al., *Temporal development of the gut microbiome in early childhood from the TEDDY study*. Nature, 2018. 562(7728): p. 583-588.
34. 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): p. 343ra81.
35. Penders, J., et al., *Factors influencing the composition of the intestinal microbiota in early infancy*. Pediatrics, 2006. 118(2): p. 511-21.
36. Mitchell, C., et al., *Delivery mode impacts newborn gut colonization efficiency*. bioRxiv, 2020: p. 2020.01.29.919993.
37. Fouhy, F., et al., *Perinatal factors affect the gut microbiota up to four years after birth*. Nat Commun, 2019. 10(1): p. 1517.
38. Wampach, L., et al., *Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential*. Nat Commun, 2018. 9(1): p. 5091.
39. Vatanen, T., et al., *Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans*. Cell, 2016. 165(4): p. 842-53.
40. Gevers, D., et al., *The treatment-naive microbiome in new-onset Crohn's disease*. Cell Host Microbe, 2014. 15(3): p. 382-392.
41. Galazzo, G., et al., *Faecal Microbiota Dynamics and their Relation to Disease Course in Crohn's Disease*. J Crohns Colitis, 2019. 13(10): p. 1273-1282.
42. Lavelle, A., et al., *Spatial variation of the colonic microbiota in patients with ulcerative colitis and control*

- volunteers. *Gut*, 2015. 64(10): p. 1553-61.
43. van Best, N., P.L. Jansen, and S.S. Rensen, *The gut microbiota of nonalcoholic fatty liver disease: current methods and their interpretation*. *Hepato Int*, 2015. 9(3): p. 406-15.
  44. Turnbaugh, P.J., et al., *An obesity-associated gut microbiome with increased capacity for energy harvest*. *Nature*, 2006. 444(7122): p. 1027-31.
  45. Mayer, E.A., K. Tillisch, and A. Gupta, *Gut/brain axis and the microbiota*. *J Clin Invest*, 2015. 125(3): p. 926-38.
  46. Penders, J., et al., *Gut microbiota composition and development of atopic manifestations in infancy: the KOALA Birth Cohort Study*. *Gut*, 2007. 56(5): p. 661-7.
  47. Eggesbo, M., et al., *Is delivery by cesarean section a risk factor for food allergy?* *J Allergy Clin Immunol*, 2003. 112(2): p. 420-6.
  48. Sevelsted, A., et al., *Cesarean section and chronic immune disorders*. *Pediatrics*, 2015. 135(1): p. e92-8.
  49. Penders, J., et al., *New insights into the hygiene hypothesis in allergic diseases: mediation of sibling and birth mode effects by the gut microbiota*. *Gut Microbes*, 2014. 5(2): p. 239-44.
  50. Collaborators, G.B.D.C.R.D., *Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015*. *Lancet Respir Med*, 2017. 5(9): p. 691-706.
  51. Hornef, M.W. and O. Pabst, *Real friends: Faecalibacterium prausnitzii supports mucosal immune homeostasis*. *Gut*, 2016. 65(3): p. 365-7.
  52. Canani, R.B., et al., *Potential beneficial effects of butyrate in intestinal and extraintestinal diseases*. *World J Gastroenterol*, 2011. 17(12): p. 1519-28.
  53. Quevrain, E., et al., *Identification of an anti-inflammatory protein from Faecalibacterium prausnitzii, a commensal bacterium deficient in Crohn's disease*. *Gut*, 2016. 65(3): p. 415-425.
  54. Folkerts, J., et al., *Butyrate inhibits human mast cell activation via epigenetic regulation of FcepsilonRI-mediated signaling*. *Allergy*, 2020.
  55. Arrieta, M.C., et al., *Early infancy microbial and metabolic alterations affect risk of childhood asthma*. *Sci Transl Med*, 2015. 7(307): p. 307ra152.
  56. Whitford, M.F., et al., *Lachnobacterium bovis gen. nov., sp. nov., a novel bacterium isolated from the rumen and faeces of cattle*. *Int J Syst Evol Microbiol*, 2001. 51(Pt 6): p. 1977-81.
  57. Lv, L.X., et al., *Alterations and correlations of the gut microbiome, metabolism and immunity in patients with primary biliary cirrhosis*. *Environ Microbiol*, 2016. 18(7): p. 2272-86.
  58. Yilmaz, B., et al., *The presence of genetic risk variants within PTPN2 and PTPN22 is associated with intestinal microbiota alterations in Swiss IBD cohort patients*. *PLoS One*, 2018. 13(7): p. e0199664.
  59. Perdigon, G., et al., *Interaction of lactic acid bacteria with the gut immune system*. *Eur J Clin Nutr*, 2002. 56 Suppl 4: p. S21-6.
  60. Rizzello, V., et al., *Role of natural killer and dendritic cell crosstalk in immunomodulation by commensal bacteria probiotics*. *J Biomed Biotechnol*, 2011. 2011: p. 473097.
  61. Fink, L.N., et al., *Distinct gut-derived lactic acid bacteria elicit divergent dendritic cell-mediated NK cell responses*. *Int Immunol*, 2007. 19(12): p. 1319-27.
  62. Sordillo, J.E., et al., *Factors influencing the infant gut microbiome at age 3-6 months: Findings from the ethnically diverse Vitamin D Antenatal Asthma Reduction Trial (VDAART)*. *J Allergy Clin Immunol*, 2017. 139(2): p. 482-491 e14.
  63. Wang, T.T., et al., *Cutting edge: 1,25-dihydroxyvitamin D3 is a direct inducer of antimicrobial peptide gene*

- expression. *J Immunol*, 2004. 173(5): p. 2909-12.
64. Liu, P.T., et al., *Toll-like receptor triggering of a vitamin D-mediated human antimicrobial response*. *Science*, 2006. 311(5768): p. 1770-3.
  65. Jackson, D.J. and S.L. Johnston, *The role of viruses in acute exacerbations of asthma*. *J Allergy Clin Immunol*, 2010. 125(6): p. 1178-87; quiz 1188-9.
  66. Jolliffe, D.A., et al., *Vitamin D supplementation to prevent asthma exacerbations: a systematic review and meta-analysis of individual participant data*. *Lancet Respir Med*, 2017. 5(11): p. 881-890.
  67. Martinez, F.D., et al., *Asthma and wheezing in the first six years of life*. *The Group Health Medical Associates*. *N Engl J Med*, 1995. 332(3): p. 133-8.
  68. Bannier, M.A., et al., *Biomarkers to predict asthma in wheezing preschool children*. *Clin Exp Allergy*, 2015. 45(6): p. 1040-50.
  69. Arrieta, M.C., et al., *Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting*. *J Allergy Clin Immunol*, 2018. 142(2): p. 424-434 e10.
  70. Bisgaard, H., et al., *Reduced diversity of the intestinal microbiota during infancy is associated with increased risk of allergic disease at school age*. *J Allergy Clin Immunol*, 2011. 128(3): p. 646-52 e1-5.
  71. Chiu, C.Y., et al., *Gut microbial-derived butyrate is inversely associated with IgE responses to allergens in childhood asthma*. *Pediatr Allergy Immunol*, 2019. 30(7): p. 689-697.
  72. Buendia, E., et al., *Gut microbiota components are associated with fixed airway obstruction in asthmatic patients living in the tropics*. *Sci Rep*, 2018. 8(1): p. 9582.
  73. Wiles, T.J., R.R. Kulesus, and M.A. Mulvey, *Origins and virulence mechanisms of uropathogenic Escherichia coli*. *Exp Mol Pathol*, 2008. 85(1): p. 11-9.
  74. Fischer, H., et al., *Mechanism of pathogen-specific TLR4 activation in the mucosa: fimbriae, recognition receptors and adaptor protein selection*. *Eur J Immunol*, 2006. 36(2): p. 267-77.
  75. Ren, Y., et al., *Co-exposure to lipopolysaccharide and desert dust causes exacerbation of ovalbumin-induced allergic lung inflammation in mice via TLR4/MyD88-dependent and -independent pathways*. *Allergy Asthma Clin Immunol*, 2019. 15: p. 82.
  76. Li, J., et al., *Pollen/TLR4 Innate Immunity Signaling Initiates IL-33/ST2/Th2 Pathways in Allergic Inflammation*. *Sci Rep*, 2016. 6: p. 36150.
  77. Stiemsma, L.T. and K.B. Michels, *The Role of the Microbiome in the Developmental Origins of Health and Disease*. *Pediatrics*, 2018. 141(4).
  78. Renz, H., et al., *The neonatal window of opportunity-early priming for life*. *J Allergy Clin Immunol*, 2018. 141(4): p. 1212-1214.
  79. Zmora, N., et al., *Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features*. *Cell*, 2018. 174(6): p. 1388-1405 e21.
  80. Zhang, F., et al., *Should we standardize the 1,700-year-old fecal microbiota transplantation?* *Am J Gastroenterol*, 2012. 107(11): p. 1755; author reply p 1755-6.
  81. DeFilipp, Z., et al., *Drug-Resistant E. coli Bacteremia Transmitted by Fecal Microbiota Transplant*. *N Engl J Med*, 2019. 381(21): p. 2043-2050.
  82. Goodrich, J.K., et al., *Genetic Determinants of the Gut Microbiome in UK Twins*. *Cell Host Microbe*, 2016. 19(5): p. 731-43.
  83. Goodrich, J.K., et al., *Human genetics shape the gut microbiome*. *Cell*, 2014. 159(4): p. 789-99.
  84. Wang, J., et al., *Genome-wide association analysis identifies variation in vitamin D receptor and other host*



- factors influencing the gut microbiota*. Nat Genet, 2016. 48(11): p. 1396-1406.
85. Rothschild, D., et al., *Environment dominates over host genetics in shaping human gut microbiota*. Nature, 2018. 555(7695): p. 210-215.
  86. Falony, G., et al., *Population-level analysis of gut microbiome variation*. Science, 2016. 352(6285): p. 560-4.
  87. Litvak, Y. and A.J. Baumler, *The founder hypothesis: A basis for microbiota resistance, diversity in taxa carriage, and colonization resistance against pathogens*. PLoS Pathog, 2019. 15(2): p. e1007563.
  88. Hanski, I., et al., *Environmental biodiversity, human microbiota, and allergy are interrelated*. Proc Natl Acad Sci U S A, 2012. 109(21): p. 8334-9.
  89. Kirjavainen, P.V., et al., *Farm-like indoor microbiota in non-farm homes protects children from asthma development*. Nat Med, 2019. 25(7): p. 1089-1095.
  90. Ruokolainen, L., et al., *Green areas around homes reduce atopic sensitization in children*. Allergy, 2015. 70(2): p. 195-202.
  91. Yeung, F., et al., *Altered Immunity of Laboratory Mice in the Natural Environment Is Associated with Fungal Colonization*. Cell Host Microbe, 2020. 27(5): p. 809-822 e6.
  92. Rosshart, S.P., et al., *Laboratory mice born to wild mice have natural microbiota and model human immune responses*. Science, 2019. 365(6452).





# **Addendum**

**Samenvatting**

**Zusammenfassung**

**Impact**

**Acknowledgements**

**About the author**

**List of Publications**

## Samenvatting

De samenstelling van de darmmicrobiota in de vroege kindertijd is zeer dynamisch en heeft een sterke invloed op het immuunsysteem van de gastheer. Verstoringen in de totstandkoming van de microbiota tijdens deze kindertijd gaan gepaard met een inadequante ontwikkeling van het immuunsysteem en kunnen daarmee levenslange gevolgen hebben voor de gevoeligheid voor ontstekings- en immuun-gemedieerde ziekten. Daarom is een gedetailleerd inzicht in de ecologische factoren die de microbiota tijdens het vroege leven vormgeven van bijzonder belang.

In **hoofdstuk 2** geven we een overzicht van de oorsprong en het belang van pionier bacteriën en de bekende evolutionaire factoren die de ontwikkeling van de darmmicrobiota in kinderen beïnvloeden. De microbiota ontplooit zich gedurende de eerste levensjaren tot een complex en divers ecosysteem onder invloed van een groot aantal omgevings- en maternale factoren, bijvoorbeeld de manier van bevalling, medicatiegebruik, moedermelk *versus* flesvoeding en andere voedingsfactoren. Deze microbiële volgroeïing is dus mede afhankelijk van de verspreiding van microben (dispersie) vanuit de omgeving en wordt verder beïnvloed door de volgorde waarin verschillende bacteriesoorten de darmen koloniseren (historische contingentie) en door factoren die een selectief voordeel bieden aan bepaalde bacteriën (habitat-filtering).

In **hoofdstuk 3** hebben we de ontwikkeling en successie van bacteriën naar volwassenheid in zowel de dunne als de dikke darm van muizen nauwlettend gevolgd en hebben we onderzocht of er metabole factoren van de gastheer betrokken zijn bij dit maturatieproces. De bacteriële soortenrijkheid en diversiteit nam aanvankelijk af in zowel de dunne als de dikke darm binnen de eerste 24 uur na de geboorte. Dit illustreert het onvermogen om de neonatale darm permanent te koloniseren op dit tijdpunt. Verder was de algehele microbiota samenstelling zeer vergelijkbaar tussen de dikke darm en de dunne darm kort na de geboorte. De meest ingrijpende veranderingen, gekenmerkt door de totstandkoming van een orgaan-specifieke microbiota, werden waargenomen tijdens het stoppen van borstvoeding. Deze successie van bacteriën en de algehele volgroeïing van het microbiële ecosysteem zijn mede afhankelijk van variaties in de beschikbare substraten in de darm als gevolg van metabole veranderingen. We hebben om die reden de concentraties van galzuren, aminozuren, biogene amines,

acylcarnitines, glycerofosfolipiden, sfingolipiden en suikers op verschillende leeftijden gemeten. We identificeerden de galzuren als potentiële drijfveren voor de ontwikkeling van de microbiota samenstelling. De causaliteit van deze bevinding werd bevestigd door orale toediening van galzuren. De geselecteerde galzuren (bijv. tauro- $\beta$ -muricholzuur en ursodeoxycholzuur) bevorderden in het bijzonder de groei van lactobacilli in de dunne darm. Deze snelle toename van lactobacilli na toediening van galzuren zou wellicht ook de bescherming tegen infecties, de zogenaamde kolonisatieresistentie, kunnen bevorderen.

In **hoofdstuk 4** toonden we aan dat onder omgevingsfactoren zoals hygiëne, medicatie, leefstijl en voeding de aanwezigheid van methanogene archaea (*Methanobrevibacter smithii*) in de ontlasting van kinderen van 6-10 jaar positief gecorreleerd is met de consumptie van biologische yoghurt en melk. We hebben vervolgens bevestigd dat deze archaea aanwezig zijn in melkproducten hetgeen suggereert dat zuivelconsumptie de darmkolonisatie met archaea bij kinderen kan bevorderen. De kolonisatie van methanogene archaea is van belang vanwege hun vele gunstige (immune-modulerende) effecten en vanwege hun mogelijk beschermend effect tegen astma, atopische dermatitis en sensibilisatie voor allergenen zoals aangetoond in hetzelfde cohort.

In **hoofdstuk 5** hebben we de ontwikkeling van de darmmicrobiota bestudeerd in een cohort van pasgeboren kinderen op verschillende tijdstippen in het eerste levensjaar en eenmaal op de schoolleeftijd. We hebben aangetoond dat het stoppen met borstvoeding sterker samenhangt met veranderingen in de microbiële samenstelling dan de introductie van vaste voeding. We bevestigden verder dat *Bacteroides* het sterkst worden beïnvloed door de manier van geboorte. De lagere aantallen *Bacteroides* onder zuigelingen die via een keizersnede werden geboren bleef ook na een zorgvuldige aanpassing voor andere confounders, zoals het voedingstype, statistisch significant in onze studie tot op de leeftijd van 31 weken post-partum. Uit ons [**hoofdstuk 5**] en eerder werk blijkt dat het voornamelijk deze *Bacteroides*-stammen zijn die van moeder op kind worden overgebracht tijdens een vaginale bevalling. De verstoring van de microbiota in het vroege leven kan een oorzakelijke rol spelen bij het ontstaan van aandoeningen zoals chronische darmontsteking, obesitas en allergieën. De manier van bevalling heeft niet alleen een belangrijk effect op de darmmicrobiota van kinderen, maar verhoogd ook het risico op allergieën en astma, de meest voorkomende chronische kinderziekten. Deze bevindingen zijn echter meestal gebaseerd op dwarsdoorsnede onderzoek,

er zijn maar weinig studies die de ontwikkeling van de darmmicrobiota voorafgaand aan de openbaring van allergische klachten hebben bestudeerd. In **hoofdstuk 5** hebben we vervolgens de veranderingen in de microbiota gerelateerd aan het ontstaan van allergieën en astma. We toonden aan dat de algehele microbiota samenstelling, de ontwikkelingstoestand en de hoeveelheid van specifieke bacteriesoorten geassocieerd waren met het risico op allergische aandoeningen op latere leeftijd. We vonden bijvoorbeeld met behulp van longitudinale analyses en zorgvuldige correctie van potentiële confounders dat de hoeveelheid van *Lachnobacterium* significant lager was gedurende de hele kindertijd in kinderen die atopische dermatitis of astma ontwikkelden. Een lagere hoeveelheid van *Faecalibacterium* was daarentegen indicatief voor een verhoogd risico op de ontwikkeling van atopisch eczeem.

Ondanks dat het ontstaan van astma in het begin van het leven kan beginnen [**hoofdstuk 5**], kan een klinische diagnose bij de overgrote meerderheid van de kinderen pas gesteld worden vanaf 6 jaar en ouder. Het zogenoemde ‘piepen’ kan al op jongere leeftijd worden herkend, maar slechts een derde van de piepende kleuters ontwikkeld uiteindelijk astma. Een betrouwbare biomarker om te voorspellen welke piepende kinderen astma zullen ontwikkelen ontbreekt op dit moment. In **hoofdstuk 6** hebben we daarom de darmmicrobiota bij 202 piepende kinderen en 50 gezonde controle kinderen van 2-4 jaar en het verband met astma op 6 jarige leeftijd onderzocht. De microbiële diversiteit en de algehele microbiële samenstelling op deze leeftijd bleken niet geassocieerd met een piepende ademhaling noch voorspellend voor de ontwikkeling van astma. Specifieke bacteriën, zoals *Escherichia*, bleken daarentegen wel geassocieerd met een verhoogd risico op de ontwikkeling van astma.

Zowel in **hoofdstuk 5** als in **hoofdstuk 6** hebben we geprobeerd om veranderingen in de samenstelling van de microbiota te koppelen aan de ontwikkeling van astma. In beide studies werden de kinderen klinisch onderzocht op schoolleeftijd (vanaf 6 jaar). Het belangrijkste verschil tussen deze studies was de leeftijd waarop de microbiota werd gekarakteriseerd. In **hoofdstuk 5** bestudeerden we de darmmicrobiota in het eerste levensjaar, terwijl in **hoofdstuk 6** de microbiota op kleuterleeftijd werd bestudeerd. De matige microbiële verstoringen die zijn beschreven in **hoofdstuk 6** suggereren dat de vroege microbiële stimulatie mogelijk op deze leeftijd al is voltooid. Deze bevindingen zijn in lijn met de zogenaamde ‘*window of opportunity*’, de kritische periode in de eerste levensmaanden waarin de microbiota zijn sterkste invloed heeft op de ontwikkeling van het immuunsysteem. Dit onderstreept

nogmaals het belang van een goede timing van monsterverzameling en vraagt om studies met regelmatige klinische onderzoeken gedurende de kindertijd in combinatie met gedetailleerde immunologische fenotypering.

De orale toediening van specifieke gunstige bacteriën, genaamd probiotica, kan een manier zijn om de darmmicrobiota gericht te manipuleren om zo klinische voordelen te bereiken. In **hoofdstuk 7** beschrijven we hoe het toedienen van verschillende stammen lactobacillen en bifidobacteriën de microbiota van prematuren beïnvloedt. We toonden aan dat het toedienen van probiotica geassocieerd was met veranderingen in het microbiële ecosysteem en dat ook na beëindiging van de toediening de probiotica-stammen deels blijvend koloniseren. De toediening van probiotische bacteriestammen, vooral bifidobacteriën, was gecorreleerd met een verlaging van bacteriën die geassocieerd zijn met necrotiserende enterocolitis (NEC) alsmede de incidentie van deze dodelijke ziekte. Onze studie benadrukt dat de vroege kindertijd een kritieke periode is waarin manipulatie van de microbiota mogelijk is en gezondheid bevorderd kan worden.

De studies in dit proefschrift, waarin de betrokkenheid van de microbiota in verschillende ziekten wordt beschreven [**hoofdstuk 5, hoofdstuk 6**] en waarin de waardevolle eigenschappen van bepaalde pionier-bacteriën worden benadrukt [**hoofdstuk 2, hoofdstuk 3, hoofdstuk 4, hoofdstuk 7**], kunnen het bewustzijn vergroten om bacteriën te gebruiken als gerichte medicatie. Specifieke stammen van bijvoorbeeld *Lachnobacterium* of *Bacteroides* kunnen potentieel van nut zijn om allergische aandoeningen of vertraagde microbiota ontwikkeling van niet-vaginaal geboren kinderen te voorkomen [**hoofdstuk 5**].

Er is een regenwoud in ons dat bestaat uit verschillende microbiële soorten die een heel complex ecosysteem vormen in de darmen. De verstoring van dergelijke ecosystemen kan dramatische gevolgen hebben en wat betreft de menselijke darmmicrobiota kan dit onze gezondheid sterk beïnvloeden. Sinds de ontdekking van microbiëel leven door van Leeuwenhoek 300 jaar geleden beginnen we pas net de impact en de rol van onze kleine vrienden in dit bos te begrijpen. De darmmicrobiota is een intrinsiek onderdeel van onze fysiologie, die begint te wortelen en te groeien vanaf de geboorte. De microbiële soorten die als eerste koloniseren spelen een cruciale rol in de ontwikkeling van dit ecosysteem en beïnvloeden mogelijk de uiteindelijke samenstelling en functionaliteit van de microbiota voor



de rest van ons leven. De vestiging van deze pioniers in het darmkanaal van zuigelingen wordt mede beïnvloed door maternale factoren, medicatie, gastheercharacteristieken, voeding en omgeving. Een beter inzicht in het samenspel tussen de darmmicrobiota, de fysiologie van ons lichaam en de factoren die de microbiële ontwikkeling beïnvloeden is belangrijk gezien de essentiële rol van de microbiota op ons metabolisme, immuunsysteem en andere processen. Dit proefschrift gaf vernieuwend inzicht in specifieke factoren die de samenstelling van de microbiota in de darm beïnvloeden en bijdragen aan de variatie tussen individuen in het vroege leven. Tevens laat het zien hoe veranderingen in de microbiële samenstelling en de manipulatie daarvan verband houden tot het ontstaan van ontstekingsziekten zoals allergieën.

Het gebrek aan bepaalde microben zou deels kunnen worden toegeschreven aan het gebrek aan omgevingsblootstelling van kinderen die steeds meer tijd in de stedelijke omgeving doorbrengen. Buitenactiviteiten in een natuurlijke omgeving met een grote biodiversiteit kunnen de microbiële kolonisatie verbeteren, het risico op allergieën en astma verminderen en het algemene welzijn van kinderen verbeteren. Dit concept van '*microbiota-rewilding*' is onlangs onderzocht in enkele dierenstudies. Deze studies ondersteunen dat een diverse microbiota afkomstig uit de *wilde* buitenomgeving bijdraagt aan een betere stimulatie van het immuunsysteem. Wellicht moeten we teruggaan naar onze oorsprong, terug naar de natuur en de wildernis in.

Concluderend, we zien steeds meer de bomen door het bos doordat dit proefschrift licht werpt op de relevantie van specifieke gastheer- en omgevingsfactoren en de eerste kolonisten in de ontwikkelende darm van het vroege leven. Door bij te dragen aan de groeiende kennis over de aard en de mogelijkheden van de microbiota in het vroege leven heeft ons onderzoek nieuwe aanknopingspunten opgeleverd om de precieze mechanismen van de microbiota ontwikkeling en de impact ervan op de gastheer verder te exploreren. Aanvullende functionele studies die diermodellen of co-cultuursystemen integreren met bevindingen of monsters van uitgebreide longitudinale menselijke geboortecohorten zullen helpen bij het begrijpen van de inter-individuele microbiële variatie en de invloed hiervan op het ontstaan van ziekten. Het onderzoeksgebied van de microbiologie focust zich al lange tijd op analytische methoden binnen de eigen grenzen, maar dit proefschrift benadrukt dat de microbiële interacties complex zijn en alleen kunnen worden ontrafeld door grensoverschrijdend onderzoek.

## Zusammenfassung

Die Zusammensetzung der intestinalen Mikrobiota ist in der frühen Kindheit sehr dynamisch und hat einen starken Einfluss auf das Immunsystem des Wirtes. Störungen in der Bildung der Mikrobiota während dieser Kindheit sind mit einer unzureichenden Entwicklung des Immunsystems verbunden und können daher lebenslange Folgen für die Anfälligkeit für entzündliche und immunvermittelte Krankheiten haben. Daher ist ein detailliertes Verständnis der Faktoren, die die Mikrobiota während des frühen Lebens prägen, von besonderer Bedeutung.

**Kapitel 2** beinhaltet einen Überblick über die Herkunft und Bedeutung von Pionierbakterien und die bekannten evolutionären Faktoren, die die Entwicklung der Darmmikrobiota bei Kindern beeinflussen. Die Mikrobiota entwickelt sich in den ersten Lebensjahren zu einem komplexen und vielfältigen Ökosystem unter dem Einfluss einer Vielzahl von Umwelt- und mütterlichen Faktoren, z.B. Art der Geburt, Medikamenteneinnahme, Muttermilch *versus* Flaschenernährung und andere Ernährungsfaktoren. Diese mikrobielle Reifung hängt zum Teil von der Verteilung der Mikroben (Dispersion) aus der Umwelt ab und wird darüber hinaus von der Reihenfolge in der verschiedene Bakterienarten den Darm besiedeln (historische Kontingenz) und von Faktoren die bestimmten Bakterien einen selektiven Vorteil bieten (Habitat-Filterung), beeinflusst.

In **Kapitel 3** haben wir die Entwicklung und Sukzession von Bakterien bis zum Erwachsenenalter sowohl im Dün- als auch im Dickdarm von Mäusen genau erfasst und untersucht, ob Stoffwechselfaktoren des Wirts an diesem Reifungsprozess beteiligt sind. Der bakterielle Artenreichtum und die Diversität nahmen zunächst sowohl im Dünndarm als auch im Dickdarm innerhalb der ersten 24 Stunden nach der Geburt ab. Dies verdeutlicht dass es nicht möglich ist den neonatalen Darm zu diesem Zeitpunkt dauerhaft zu besiedeln. Des Weiteren war die gesamte Mikrobiota-Zusammensetzung sich zwischen Dick- und Dünndarm kurz nach der Geburt sehr ähnlich. Die tiefgreifendsten Veränderungen während der Bildung einer organspezifischen Mikrobiota, wurden während des Stillens beobachtet. Diese Sukzession der Bakterien und die Gesamtreife des mikrobiellen Ökosystems hängen auch von Schwankungen der verfügbaren Substrate im Darm aufgrund von Stoffwechseleränderungen ab. Deshalb haben wir die Konzentrationen von Gallensäuren, Aminosäuren, biogenen Aminen,

Acylcarnitinen, Glycerophospholipiden, Sphingolipiden und Zuckern in verschiedenen Altersstufen gemessen. Wir identifizierten die Gallensäuren als potenzielle Treiber für die Entwicklung der Mikrobiota-Zusammensetzung. Die Kausalität dieses Befundes wurde durch orale Verabreichung von Gallensäuren bestätigt. Die ausgewählten Gallensäuren (z. B. Tauro- $\beta$ -Muricholsäure und Ursodeoxycholsäure) förderten insbesondere das Wachstum von Laktobazillen im Dünndarm. Diese schnelle Vermehrung der Laktobazillen nach der Verabreichung von Gallensäuren könnte auch den Schutz vor Infektionen, die sogenannte Kolonisationsresistenz, fördern.

In **Kapitel 4** haben wir gezeigt, dass unter Umweltfaktoren wie Hygiene, Medikamenteneinnahme, Lebensstil und Ernährung, das Vorhandensein von methanogenen Archaeen (*Methanobrevibacter smithii*) im Stuhl von Kindern im Alter von 6-10 Jahren positiv mit dem Konsum von Bio-Joghurt und -Milch korreliert ist. Anschließend bestätigten wir das Vorhandensein dieser Archaeen in Milchprodukten, was darauf hindeutet, dass der Verzehr von Milchprodukten die Darmbesiedlung mit Archaeen bei Kindern fördern kann. Die Besiedlung mit methanogenen Archaeen ist wichtig wegen ihrer vielen vorteilhaften (immunmodulierenden) Wirkungen und wegen ihrer potenziellen Schutzwirkung gegen Asthma, atopische Dermatitis und Sensibilisierung auf Allergene, wie in der gleichen Kohorte von Kindern gezeigt wurde.

In **Kapitel 5** haben wir die Entwicklung der intestinalen Mikrobiota in einer Kohorte von Neugeborenen zu verschiedenen Zeitpunkten im ersten Lebensjahr und einmalig im Schulalter untersucht. Wir haben gezeigt, dass das Beenden des Stillens stärker mit Veränderungen in der mikrobiellen Zusammensetzung zusammenhängt als die Einführung von fester Nahrung. Wir bestätigten des Weiteren, dass die *Bacteroide* am stärksten von der Art der Geburt beeinflusst werden. Die geringeren *Bakteroidzahlen* bei den per Kaiserschnitt geborenen Säuglingen blieben in unserer Studie bis zum Alter von 31 Wochen postpartal statistisch signifikant, auch nach sorgfältiger Adjustierung für andere Confounder, wie z.B. die Art der Ernährung. Aus unserer [**Kapitel 5**] und früheren Arbeiten geht hervor, dass es hauptsächlich diese *Bacteroid-Stämme* sind, die während der vaginalen Geburt von der Mutter auf das Kind übertragen werden. Die Störung der Mikrobiota im frühen Leben kann eine kausale Rolle bei der Entwicklung von Erkrankungen wie chronischen Darmentzündungen, Adipositas und Allergien spielen. Die Art der Entbindung hat nicht nur einen wichtigen Einfluss auf die

Darmmikrobiota von Kindern, sondern erhöht auch das Risiko auf Allergien und Asthma, die am häufigsten vorkommenden chronischen Kinderkrankheiten. Allerdings beruhen diese Erkenntnisse meist auf Querschnittsuntersuchungen, nur wenige Studien haben die Entwicklung der Darmmikrobiota vor dem Auftreten allergischer Symptome untersucht. In **Kapitel 5** haben wir darum die Veränderungen der Mikrobiota mit der Entwicklung von Allergien und Asthma in Verbindung gebracht. Wir haben gezeigt, dass die Gesamtzusammensetzung der Mikrobiota, der Entwicklungsstand und die Menge spezifischer Bakterienspezies mit dem Risiko auf allergische Erkrankungen im späteren Leben assoziiert sind. Zum Beispiel fanden wir mit Hilfe von Längsschnittdaten und sorgfältiger Korrektur potenzieller Störfaktoren heraus, dass die Menge an *Lachnobacterium* während der gesamten Kindheit bei Kindern, die atopische Dermatitis oder Asthma entwickelten, signifikant niedriger war. Im Gegensatz dazu war eine geringere Menge an *Faecalibacterium* ein Hinweis auf ein erhöhtes Risiko, ein atopisches Ekzem zu entwickeln.

Obwohl Asthma bereits in den ersten Lebensjahren beginnen kann [**Kapitel 5**], kann eine klinische Diagnose bei der überwiegenden Mehrheit der Kinder erst ab dem 6. Lebensjahr gestellt werden. Das sogenannte „Wheezing“ kann schon in jüngeren Jahren erkannt werden, aber nur ein Drittel der keuchenden Kleinkinder entwickelt schließlich Asthma. Ein zuverlässiger Biomarker zur Vorhersage, welche keuchenden Kinder Asthma entwickeln werden, fehlt derzeit. Daher untersuchten wir in **Kapitel 6** die intestinale Mikrobiota von 202 keuchenden Kindern und 50 gesunden Kontrollkindern im Alter von 2-4 Jahren und den Zusammenhang mit Asthma im Alter von 6 Jahren. Die mikrobielle Diversität und die gesamte mikrobielle Zusammensetzung in diesem Alter waren weder mit dem Keuchen assoziiert noch prädiktiv für die Entwicklung von Asthma. Es wurde jedoch festgestellt, dass bestimmte Bakterien, wie z. B. *Escherichia*, mit einem erhöhten Risiko für die Entwicklung von Asthma verbunden sind.

Sowohl in **Kapitel 5** als auch in **Kapitel 6** haben wir versucht, Veränderungen in der Zusammensetzung der Mikrobiota mit der Entwicklung von Asthma in Verbindung zu bringen. In beiden Studien wurden die Kinder im Schulalter (ab dem 6. Lebensjahr) klinisch untersucht. Der Hauptunterschied zwischen diesen Studien war das Alter, in dem die Mikrobiota charakterisiert wurde. In **Kapitel 5** wurde die intestinale Mikrobiota im ersten Lebensjahr untersucht, während in **Kapitel 6** die Mikrobiota im Vorschulalter untersucht wurde. Die in **Kapitel 6** beschriebenen mäßigen mikrobiellen Störungen lassen vermuten, dass die frühe

mikrobielle Stimulation in diesem Alter abgeschlossen sein könnte. Diese Befunde stehen im Einklang mit der sogenannten „*Window of Opportunity*“-Hypothese, der kritischen Zeitraum in den ersten Lebensmonaten, in dem die Mikrobiota ihren stärksten Einfluss auf die Entwicklung des Immunsystems hat. Dies unterstreicht einmal mehr die Bedeutung eines guten Timings der Probenentnahme und fragt nach Studien mit regelmäßigen klinischen Untersuchungen im Kindesalter in Kombination mit einer detaillierten immunologischen Phänotypisierung.

Die orale Verabreichung spezifischer nützlicher Bakterien, so genannter Probiotika, kann eine Möglichkeit sein, die intestinale Mikrobiota gezielt zu manipulieren, um klinische Vorteile zu erzielen. In **Kapitel 7** beschreiben wir, wie die Verabreichung von verschiedenen Stämmen von Laktobazillen und Bifidobakterien die Mikrobiota von Frühgeborenen beeinflusst. Wir konnten zeigen, dass die Verabreichung von Probiotika mit Veränderungen des mikrobiellen Ökosystems verbunden war und dass die probiotischen Stämme teilweise auch nach Beendigung der Verabreichung weiter kolonisieren. Die Verabreichung von probiotischen Bakterienstämmen, insbesondere von Bifidobakterien, wurde mit einer Reduktion der mit nekrotisierender Enterokolitis (NEC) assoziierten Bakterien sowie der Inzidenz dieser tödlichen Erkrankung in Verbindung gebracht. Unsere Studie unterstreicht, dass die frühe Kindheit ein kritischer Zeitraum ist, in dem eine Manipulation der Mikrobiota möglich ist und die Gesundheit gefördert werden kann.

Die Studien in dieser Dissertation, in denen die Beteiligung der Mikrobiota an verschiedenen Krankheiten beschrieben wird [**Kapitel 5, Kapitel 6**] und in denen die wertvollen Eigenschaften bestimmter Pionierbakterien hervorgehoben werden [**Kapitel 2, Kapitel 3, Kapitel 4, Kapitel 7**], können das Bewusstsein dafür schärfen, Bakterien gezielte als Medikamente einzusetzen. Spezifische Stämme von z. B. *Lachnobacterium* oder *Bacteroides* können möglicherweise nützlich sein, um allergische Erkrankungen oder eine verzögerte Entwicklung der Mikrobiota bei nicht vaginal geborenen Kindern zu verhindern [**Kapitel 5**].

In uns lebt ein Regenwald, der aus verschiedenen Mikroben Arten besteht, die im Darm ein sehr komplexes Ökosystem bilden. Die Störung solcher Ökosysteme kann dramatische Folgen haben, und was die menschliche Darmmikrobiota betrifft, kann dies einen großen Einfluss auf unsere Gesundheit haben. Seit der Entdeckung des mikrobiellen Lebens durch van Leeuwenhoek vor 300 Jahren haben wir gerade erst angefangen, die Auswirkungen und die Rolle unserer kleinen

Freunde in diesem Wald zu verstehen. Die intestinale Mikrobiota ist ein fester Bestandteil unserer Physiologie, die von Geburt an Wurzeln schlägt und wächst. Die Mikroben Arten, die sich zuerst ansiedeln, spielen eine entscheidende Rolle bei der Entwicklung dieses Ökosystems und können die endgültige Zusammensetzung und Funktionalität der Mikrobiota für den Rest unseres Lebens beeinflussen. Die Etablierung dieser Pioniere im Darmtrakt von Säuglingen wird teilweise durch mütterliche Faktoren, Medikamente, Wirtseigenschaften, Ernährung und Umwelt beeinflusst. Ein besseres Verständnis des Zusammenspiels zwischen der intestinalen Mikrobiota, der Physiologie unseres Körpers und den Faktoren, die die mikrobielle Entwicklung beeinflussen, ist wichtig, da die Mikrobiota eine wesentliche Rolle in unseren Stoffwechsel, dem Immunsystem und anderen Prozessen spielt. Diese Dissertation gibt innovative Einblicke in spezifische Faktoren, die die Zusammensetzung der Mikrobiota im Darm beeinflussen und zur Variation zwischen Individuen im frühen Leben beitragen. Sie zeigt auch, wie Veränderungen in der mikrobiellen Zusammensetzung und deren Manipulation mit der Entstehung von entzündlichen Erkrankungen wie Allergien zusammenhängen.

Ein Mangel an bestimmten Mikroben könnte teilweise auf die mangelnde Umweltexposition von Kindern zurückzuführen sein, die immer mehr Zeit in der städtischen Umgebung verbringen. Aktivitäten im Freien in einer natürlichen Umgebung mit hoher Biodiversität können die mikrobielle Besiedlung verbessern, das Risiko von Allergien und Asthma verringern und das allgemeine Wohlbefinden von Kindern verbessern. Dieses Konzept des „*microbiota-rewilding*“ wurde kürzlich in einigen Tierstudien untersucht. Diese Studien unterstützen, dass eine vielfältige Mikrobiota, die aus der *freien* Natur stammt, zu einer besseren Stimulation des Immunsystems beiträgt. Vielleicht sollten wir zu unseren Ursprüngen zurückkehren, zurück zur Natur und in die Wildnis.

Abschließend sehen wir immer mehr Bäume im Wald, denn diese Dissertation beleuchtet die Relevanz von spezifischen Wirts- und Umweltfaktoren und den ersten Siedlern im sich entwickelnden Darm des frühen Lebens. Indem wir zum wachsenden Wissen über die Natur und die Möglichkeiten der Mikrobiota im frühen Leben beitragen, haben unsere Forschungsbemühungen neue Anhaltspunkte für die weitere Erforschung der genauen Mechanismen der Mikrobiota-Entwicklung und ihrer Auswirkungen auf den Wirt geliefert. Zusätzliche funktionelle Studien, die Tiermodelle oder Co-Kultursysteme mit Befunden oder Proben erweiterter longitudinaler menschlicher Geburtskohorten integrieren, werden helfen, die interindividuelle mikrobielle Variation und ihren Einfluss auf das Auftreten von Krankheiten zu verstehen. Das Forschungsgebiet der Mikrobiologie hat sich lange auf

Analysemethoden innerhalb der eigenen Grenzen konzentriert, aber diese Dissertation betont, dass die mikrobiellen Wechselwirkungen komplex sind und nur durch grenzüberschreitende Forschung entschlüsselt werden können.

## Impact

The dramatically increasing prevalence of non-communicable diseases (NCDs) such as diabetes, obesity and allergies and related economic and societal impact is a worldwide public health burden [1]. NCDs are globally the primary cause of mortality with about 60% of deaths and still numbers are steadily rising [2]. Notably, atopic diseases including allergies and asthma are the most prevalent NCDs that arise in early life [3]. Atopic diseases often have relatively moderate disease courses, but the chronicity and high prevalence pursue in a large societal and economic impact. In order to support a sustainable health care system, the primary, secondary and tertiary prevention of chronic NCD's is pivotal. These diseases are influenced by multiple environmental and genetical factors [3], and increasing evidence suggests that also the infant gut microbiota composition profoundly contribute to their etiology [5]. Although the genetic aspects of these conditions are unchangeable, the gut microbiota allows manipulation which makes it an intriguing target for new therapies to cure and prevent NCDs. Moreover, a better understanding of the natural dynamics of intestinal microbial maturation and subsequent insight into effective strategies to manipulate the microbiota could also be important for other diseases with an infectious component, i.e. necrotizing enterocolitis (NEC).

Firstly, this thesis studied the influence of host and environmental factors on early-life microbial colonization apart from the focus on a specific disease. These studies identified the neonatal subgroups with a disturbed microbiota, i.e. via caesarean birth, that might benefit most from a possible microbial treatment and could serve as a basis for personalized medicine. We further showed that the presence of methane-producing archaea in children was positively correlated with the consumption of organic yogurt and milk. Subsequently, it was confirmed that these archaea are indeed present in milk products and that dairy consumption can therefore promote the colonization of these microorganisms. Moreover, in addition to these dietary factors, bile metabolites appeared to have an important influence on the development of the small intestine microbiota of newborn mice, which was confirmed by oral administration of bile acids TCA and  $\beta$ TMCA. Importantly, this study was the first to show a causal role of host-derived metabolites in neonatal murine microbiota development and might have accelerated research into this area. Insight into these dietary and host factors that influence the colonization process could act as leads for future (nutritional) interventions.



Secondly, this thesis described if manipulation of the neonatal microbiota indeed can be achieved in order to improve health. We showed that the administration of probiotics, especially bifidobacteria, to preterm infants was associated with a lower risk of developing NEC, a life-threatening intestinal inflammation. Therefore, this thesis confirmed and emphasized that early childhood is a critical period in which targeted manipulation of the intestinal microbiota is possible to promote a healthy future.

Furthermore, the studies in this thesis also identified which factors influence the early development of the microbiota, and how changes in the microbiota during this life-phase relate to the onset of allergic diseases. For instance, the microbiota composition of children in two birth cohorts was investigated in relation to the development of asthma and allergies. Specific bacteria such as *Lachnobacterium* were reduced in children who subsequently developed atopic dermatitis and asthma, while a reduction in *Faecalibacterium* was indicative of atopic dermatitis. An increase in *Escherichia* was associated with an increased risk of asthma among children with pre-school wheeze. The examination of these specific bacteria is pivotal in order to both better understand the onset of these diseases and support the development of more targeted interventions. These commensal bacteria could serve as potential next-generation probiotics in the prevention of allergies or to protect against NEC, and the identified microbiota profiles of infants that did not develop such disease could be part of synthetic fecal transplants. Moreover, it also provides ways to explore new functional mechanisms in the pathogenesis of these diseases. The forthcoming interfering mechanisms, i.e. postbiotics, or the stimulation via probiotics may facilitate the development of low-cost and easy to handle treatments for affordable microbial therapies.

These non-invasive supplements to promote intestinal immune-homeostasis would contribute to improved emotional impact as well, in particular in young vulnerable infants and their parents. In addition, the identified bacteria in this thesis could also be utilized for diagnoses in the early detection of these diseases such as asthma among wheezing pre-school children. In line with this, electronic home monitoring of these bacteria might perhaps be a revolutionary scope in the future.

Alltogether, the observation that certain conditions are associated with specific changes in the intestinal microbiota composition makes it an intriguing target for new treatments, diagnoses

and personalized medicine. These findings will certainly increase the understanding of the establishing beneficial microbiota and will generate new leads for follow-up intervention studies to identify the infants that need it most and improve their gut health. Oral administration of the identified new leads (i.e. organic milk, *Faecalibacterium*, bile acids, etc.) might be a way for targeted manipulation of the gut microbiota to achieve clinical benefits, i.e. to decrease the incidence of infections by pathogens and prevent the development of immune-mediated diseases. The neonatal window may provide the best opportunity for targeted manipulation of the microbiota, potentially even with long-lasting positive health consequences which time will tell.

*Importantly, the impact of this thesis on the environment is remarkable. The thesis has only been printed for specific delegates and thereby about 50% less paper is used compared to the average common orders. Furthermore, it has been printed on 'paper' consisting of 85% to 100% of agricultural waste. The latter resulted in a reduction of 20% CO<sub>2</sub> and 47% lower eco-footprint (general environmental impact) compared to common FSC-certified paper. The improved environmental impact comprises eco-toxicity to soil and water, particulate matter, human toxicity, influence of land use, ozone layer degradation, smog formation, depletion of fossil raw materials, and acidification of soil.*

## References

1. Gouda, H.N., et al., *Burden of non-communicable diseases in sub-Saharan Africa, 1990-2017: results from the Global Burden of Disease Study 2017*. Lancet Glob Health, 2019. 7(10): p. e1375-e1387.
2. World Health Organization. *Global action plan for the prevention and control of noncommunicable diseases 2013-2020*. World Health Organization, 2013. <https://apps.who.int/iris/handle/10665/94384>
3. Ogoina, D. and G.C. Onyemelukwe, *The role of infections in the emergence of non-communicable diseases (NCDs): Compelling needs for novel strategies in the developing world*. J Infect Public Health, 2009. 2(1): p. 14-29.
4. Ober, C. and T.C. Yao, *The genetics of asthma and allergic disease: a 21st century perspective*. Immunol Rev, 2011. 242(1): p. 10-30.
5. Renz, H. and C. Skevaki, *Early life microbial exposures and allergy risks: opportunities for prevention*. Nat Rev Immunol, 2020.

## Acknowledgements



*This is it! The official end of my student-era. A colorful chapter in life, from which I gained a lot. An international journey full of stunning experiences, challenges, but also some heavy disappointments. I learned to deal with these aspects and to keep believing in my opportunities and the path ahead.*

*This journey, together with all people mentioned below that took part in it, developed me into becoming the researcher I'm today and I'm very thankful for that.*

Als eerste dank aan alle kinderen die een prachtige ingekleurde tekening hebben bijgedragen voor dit proefschrift: **Elin, Raf, Elif, Emre, Paul, Nicolas, Julia, Noa** en mijn eigen dochter **Maya**. Jullie vrolijke en bonte illustraties maken het boekje helemaal af!

Secondly, I would like to thank my (co)promotors:

**John**, tie your ears (not ‘nuts’): je was letterlijk de beste supervisor die ik me ooit had kunnen voorstellen. Je motiveerde en inspireerde me met je passie voor goed wetenschappelijk onderzoek en je haalde echt het beste uit mezelf. Zonder je betrokkenheid, oog voor detail (mits de juiste cameralens mee) en het altijd ontvankelijk zijn voor vragen en kritiek was mijn promotietraject niet ontplooid zoals het nu is. Het zette ‘the points on i’. Je wist de perfecte balans te vinden - behalve op skilatten - tussen professionele aansturing en slechte humor. Dank voor de slapeloze nachten, zowel tijdens alle roadtrips door je gesnurk, als op het lab tussen de muizen (**figuur** - credits: **John**).

**Mathias**, I really appreciated you as my doctor-father, although you sometimes didn’t make my PhD easier when you suggested a new project topic almost every week. I enjoyed our lengthy intellectual discussions, the freedom you gave me and the way you tried to push everything to the highest level.

**Paul**, thanks for guiding me all these years, for your fresh input and thinking along with me.

Furthermore, I would like to thank all **co-authors** for their contribution to the studies included in this thesis. I learned a lot from doing research together in a collaborative manner, with colleagues from different backgrounds, as well as the writing process.

To my colleagues in Aachen: **Shahed, Alina, Martina, Nour, Aline, Ambre, Regina, Maike, Stefan, Johannes, Annika, Natalia, Fine, Kaiyi** and all others, thank you for the nice work environment (despite the green), the endless journal clubs, all the support and the weekly cakes. I really enjoyed being there with you right from the start, at a time we all could still fit in one office, and being involved in all developments.

**Martina**, thanks for your technical support and especially all the quality time and conversations we had about our wonderful lives. You understood me completely and always thought further ahead despite me being a bit chaotic at times. I feel privileged to be able to work close together with you and will definitely miss this fun.

Thanks to all colleagues in Maastricht, including **Erik, Christel, Charlotte, Carla, Chris, Mayk, Kevin, Gianluca, Brian, Liene, Melissa, Nader, Birke, Casper, Julius, David, Heike, Judith, Lars,**

**Matthew, Tessa, Wesley, Giang**, and the rest of the MMB-department for all cosy (coffee-) moments and keeping up my cafeïne-level despite the terrible coffeemachines. **Mayk**, dank voor je strakke illustraties voor de publicaties behorende bij dit proefschrift (chapter 2 & chapter 5) waarmee ik op veel congressen kon pronken.

A special thanks to my(former) roomie's **Brian, Gianluca and Liene** who could handle my distracting attitude. **Gianluca**, thanks for letting me steal your awfull coffee when needed and working together with you on two shared first-author master-pieces. **Brian**, mijn 'back2back'-kamergenoot, fijn dat ik altijd met je kon sparren over van alles. **Liene**, het was enorm fijn om met je samen te werken, te reizen en te bouwen aan het Lucki-project. Ik heb genoten van onze toffe tijd in New York, Canada en Ierland, maar ook tot laat op kantoor met warme persoonlijke gesprekken en toekomstvisies. Hoewel mijn stinktrui jou prikkelde, hield je mij juist scherp door vooral je structurele manier van werken (waar ik nog steeds jaloers op ben). Dank voor al het plezier dat ik met je heb gehad door o.a. je vrolijke verschijning, uitlachen om John's uitglijders en de slappe lach tijdens keynote-sessies op conferenties (InVivo). En ik blijf dromen over je zwembad!

Aan mijn lieve paranimfen: **Heike**, dankje dat ik altijd bij je terecht kon, zowel op UM als bij je thuis in Aachen. Van gezellig kletsen over je nieuwste tatoage's en geklaag over lijn 350 tot complete brainstorm-sessies. Ik waardeer dat je altijd stond te springen (behalve op de trampoline op onze bruiloft) om nieuwe projecten en ideeën te bedenken en hoe je dit stiekem ijverig op tientallen memo's schreef. En ook dank dat ik je 'bus-office' excuus mag gebruiken! **Casper**, dank dat je telkens voor me klaarstond. Althans, voor de koffie-pauze's, Tripel Karmeliet, organiseren van (kerst)evenementjes en de ontelbare vrijmibo's - onze gedeelde passie. Thembi was vaak nog niet eens open of we zaten er al. Ik prijs je attente karakter: geen proost-gelegenheid - van carnaval tot zakelijke plenaire sessies - ging aan je voorbij. We hebben meer bitterballen op dan dagen dat ik aan mijn PhD heb besteed, en ik appreciëer je lef om toch altijd de laatste te pakken. Dank ook voor je samenwerkingen met muffin-onderhandelingen en het niet ondersteunen (mislukken) van mijn Dry January.

**Max, Claudia, Alex, Svenja, Leslie, Thomas, Carmen, Camiel, Sanne, Joël, Mirella, Mick, Loes, Mandy, Roger**; vrienden voor het (MLW)leven. Jullie bedankt voor de (studie)basis, alle afleiding en het serieus nemen van mijn poeponderzoek. Zonder de ontelbare feestjes,

festivals en ons eigen Postival/Maria-Hopeloos geen overlevingskans voor mijn PhD. **Max & Carmen**, dank voor alle mooie koffie-momenten op de uni en vooral het delen van elkaar's PhD-belevenissen de afgelopen jaren. **Mick & Joël**, onze gezamenlijke Duitse avonturen waren prachtig; van torens bouwen tijdens carnaval tot alle minibars leegdrinken tijdens de retreat. **Joël**, wat mis ik onze ooit dagelijkse gesprekken nu al; je ongezouten (sterke) mening, trance-updates en gestoei met het Duitse systeem. **Mick**, eeuwig bedankt dat je me hebt overtuigd over de grens te gaan werken!

**Melih**, vielen Dank voor de mega-gezellige autoritten. Het was me een waar genoegen om telkens met je op en neer naar Aachen de pendelen en anders niet gelukt zo vaak de grens te passeren! Ik waardeer ook je enthousiaste deelname samen met **Marlies, Emre & Elif** aan de studies, maar natuurlijk vooral de vriendschap die daaruit is ontstaan.

**Lieber Reiner & Monika**, danke auch an euch für die Unterstützung und alle Ratschläge. Ich danke es zum Teil auch euch, dass ich mich für diese Aachener Route überhaupt erst entschieden habe.

Lieve **mama en papa**, dank dat jullie mij de mogelijkheid hebben gegeven deze 'hobby' te laten uitgroeien tot deze mijlpaal. Alleen met jullie steun en gegeven vrijheid heb ik het zo ver kunnen schoppen! **Sven**, broertje, je was de beste uitlaatklep tijdens mijn promotie-traject, inclusief de (de)motiverende peptalks en Duvelse dilemma's. En dankzij je groene invloed en kijk op de wereld is de 'nature' uit de titel mede aan jou te danken.

**Johanna**, dank voor álle hulp bij vooral de vele keuzes en dat je altijd in me bent blijven geloven (bijv. dat ik het levende bewijs ben dat domme vragen wél bestaan). De gedeelde liefde voor ontwikkeling van natuur en kinderen hoop ik nog lang met je te exploreren.

Lastly, I would like to thank '**Thesaurus**' ! I could never have made it till the end without you. I can't imagine how the language in my thesis would look otherwise. You made this thesis to be the crown of my work; the cherry on the cake.

## About the author

Niels van Best was born on the 6<sup>th</sup> of October 1987 in Eindhoven, the Netherlands. After finishing high school (VWO) at the ROC/Van Maerlantlyceum in 2007, he started his studies Biomedical Sciences (Molecular Life Sciences) at Maastricht University. He continued with the master Biomedical Sciences with a specialization in Nutrition and Translational Research In Metabolism (NUTRIM). During this master, he investigated the impact of obesity-associated dysfunction of intestinal epithelial cells on the secretion of antimicrobial peptides and microbiota composition in humans. For the final Master thesis, he studied the interaction of complement factor three and diet on the gut microbiota composition in mice at both the department of General Surgery and Medical Microbiology, Maastricht University. After obtaining his master's degree in 2014, he started as a research-assistant at the department of General Surgery, Maastricht University. In 2015, he started his PhD research at the institute of Medical Microbiology at the RWTH Uniklinik Aachen in collaboration with the department of Medical Microbiology of Maastricht University/MUMC+. During his Phd-trajectory, he worked on factors that influence gut microbiota development in early life and associations with the onset of allergic diseases under supervision of Prof. Mathias Hornef, Dr. John Penders, and Prof. Paul Savelkoul. Moreover, he gave multiple oral presentations at national and international conferences and was awarded several grants and prizes, including the NVGE student award, the Gut Day 2019 presentation award, an ICMI travel award and a WUN-InFlame early career researcher scholarship. Furthermore, he twice received a seeding grant from the NUTRIM Liver & Digestive Health division. After finishing his PhD(-thesis) in 2020, he was awarded with the Kootstra Talent Fellowship of the Faculty of Health, Medicine and Life Sciences of Maastricht UMC+, which allows him to start as a Postdoctoral Fellow to further study host-microbial interactions in early life maintaining the strong collaboration with RWTH Uniklink Aachen.

## List of publications

**N. van Best**, Trepels-Kottek, P. Savelkoul, T. Orlikowsky, M.W. Hornef, J. Penders. Probiotic-mediated changes in the developing neonatal microbiota. *Gut Microbes*. 2020 Nov 9;12(1):1-16.

R. Pieper, **N. van Best**, K. van Vorst, F. Ebner, M. Reissmann, M. W. Hornef, M. Fulde. Towards a porcine in vivo model to analyze the pathogenesis of TLR5-dependent enteropathies. *Gut Microbes*. 2020 Nov 9;12(1):1782163

A. Riba, K. Hassani, A. Walker, **N. van Best**, D. Zeschwitz, T. Anslinger, N. Sillner, S. Rosenhain, D. Eibach, O. Maiga-Ascofaré, U. RolleKampczyk, M. Basic, A. Binz, S. Mocek, B. Sodeik, R. Bauerfeind, A. Mohs, C. Trautwein, F. Kießling, J. May, M. Klingenspor, F. Gremse, P. Schmitt-Kopplin, A. Bleich, N. Torow, M. von Bergen, M. W. Hornef. Disturbed bile homeostasis and microbiota composition by intestinal protozoan infection causing metabolic dysregulation and growth impairment. *Sci Transl Med*. 2020 Oct 14;12(565):eaay7019.

**N. van Best\***, G. Galazzo\*, B. J. Benedikter, K. Janssen, L. Bervoets, C. Driessen, M. Oomen, M. Lucchesi, P. H. Van Eijck, H. Becker, M. W. Hornef, P. H Savelkoul, F. Stassen, P. Wolffs, J. Penders. How to count our microbes? The effect of different Quantitative Microbiome Profiling approaches. *Front. Cell. Infect. Microbiol*. 2020 Aug 7.

**N. van Best**, U. Rolle-Kampczyk, F. G. Schaap, M. Basic, S. Olde Damink, A. Bleich, P. Savelkoul, M. von Bergen, J. Penders, M. W. Hornef. Bile acids drive the newborn's gut microbiota maturation. *Nature Communications*. 2020 July 23. doi; 10.1038/s41467-020-17183-8.

**N. van Best\***, G. Galazzo\*, L. Bervoets, I. O. Dapaah, P. H. Savelkoul, M. W. Hornef, GI-MDH consortium; S. Lau, E. Hamelmann, J. Penders. 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 May;158(6):1584-1596.



**N. van Best\***, M. A. G. E. Bannier\*, L. Bervoets, P. H. M. Savelkoul, M. W. Hornef, K. D. G. van de Kant, Q. Jöbssis, E. Dompeling, J. Penders. Gut Microbiota in Wheezing Preschool Children and the Association With Childhood Asthma. *Allergy*. 2019 Dec 15. doi: 10.1111/all.14156.

Bennek E, Mandić AD, Verdier J, Roubrocks S, Pabst O, **N. van Best**, Benz I, Kufer T, Trautwein C, Sellge G. Subcellular antigen localization in commensal *E. coli* is critical for T cell activation and induction of specific tolerance. *Mucosal Immunol*. 2018 Oct 16

**N. van Best\***, JA van de Pol\*, CA Mbakwa, C Thijs, PH Savelkoul, IC Arts, MW Hornef, M Mommers, J Penders. Gut Colonization by Methanogenic Archaea Is Associated with Organic Dairy Consumption in Children. *Front Microbiol*. 2017 Mar 10;8:355

Gautheron J, Vucur M, Schneider AT, Severi I, Roderburg C, Roy S, Bartneck M, Schrammen P, Diaz MB, Ehling J, Gremse F, Heymann F, Koppe C, Lammers T, Kiessling F, **N. van Best**, Pabst O, Courtois G, Linkermann A, Krautwald S, Neumann UP, Tacke F, Trautwein C, Green DR, Longerich T, Frey N, Luedde M, Bluher M, Herzig S, Heikenwalder M, Luedde T. The necroptosis-inducing kinase RIPK3 dampens adipose tissue inflammation and glucose intolerance. *Nat Commun*. 2016 Jun 21;7:11869

**N. van Best**, M.W. Hornef, P. Savelkoul, J. Penders. On the origin of species: factors shaping the establishment of infant's gut microbiota. *Birth Defects Res C Embryo Today*. 2015 Dec;105(4):240-51

**N. van Best**, P. Jansen, S. Rensen. The gut microbiota of nonalcoholic fatty liver disease: current methods and their interpretation. *Hepatol Int*. 2015 Jul;9(3):406-15

D. Tedjo, D. Jonkers, P. Savelkoul, A. Masclee, **N. van Best**, M. Pierik, J. Penders. The effect of sampling and storage on the fecal microbiota composition in healthy and diseased subjects. *PLoS One*. 2015 May 29;10(5):e0126685

*\*shared first authorship*