The infant gut microbiome and resistome in health and disease

Marta Reyman

The infant gut microbiome and resistome in health and disease

Marta Reyman

The infant gut microbiome and resistome in health and disease

ISBN/EAN 978-94-6423-212-7

Cover design by Charlotte Margo | @charlypencil
Lay-out and printing by ProefschriftMaken || www.proefschriftmaken.nl

The printing of this thesis was financially supported by Spaarne Gasthuis Academy, Diakademie, Infection & Immunity Utrecht, La Roche-Posay, Olmed and Yakult Nederland B.V.

The infant gut microbiome and resistome in health and disease

Het darmmicrobioom en -resistoom van de zuigeling in gezondheid en ziekte

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op donderdag 1 juli 2021 des middags te 2.15 uur

door

Marta Reyman

geboren op 5 maart 1988 te Vlaardingen

Promotoren:
Prof. dr. E.A.M. Sanders
Prof. dr. D. Bogaert

Copromotor:
Dr. M.A. van Houten

Dla Mamusi i Tatusia

Table of Contents

CHAPTER 1

General introduction

In the past decade, it has become apparent that the human body consists of more than solely human genetic material. Our human cells are greatly outnumbered by the micro-organisms that live on our bodies' surfaces.[1] These microbes, including bacteria, viruses, fungi and archaea, are collectively referred to as our microbiome. The gut microbiome boasts the highest number and density of bacteria in the human body, making its composition and functional capacity subjects of extensive research.[2,3] This has led to the understanding that our gut bacteria perform various vital functions, ranging from the digestion of nutrients to providing (colonization) resistance against potential pathogens.[4]

Although the co-evolution of microbes and their human hosts spans human existence,[5] the crucial development of one individual's microbiome is believed to occur in a narrow timeframe, which starts at conception and spans the first 1000 days of life.[6] Importantly, early-life microbial colonization coincides with a similarly time-limited period, during which the immune system is highly receptive to microbial training. In the course of this so-called window of opportunity, the interaction between the microbiome and host plays a pivotal role in priming the host's immunity.[7] In this way, a perturbed microbial development could form the basis of increased susceptibility to disease later in life. For instance, in respiratory disease, bacterial components and metabolites in the gut and lungs have been shown to modulate the immune system both locally and systemically, and specific bacteria have been associated with asthma, chronic obstructive pulmonary disease and respiratory tract infections (RTIs).[8]

Multiple lifestyle and environmental factors have the potential to perturb early-life microbiome development, such as birth by caesarean section (CS), formula feeding and antibiotics.[9–16] A large body of evidence suggests that birth by CS affects the gut microbiome composition in early-life, for example by reducing the relative abundance of health-associated and anti-inflammatory genera as *Bifidobacterium* and *Bacteroides*, whilst enriching the gut with potential pathogens such as *Enterococcus* and *Klebsiella,* when compared to vaginal delivery.[17,18] Regrettably, the incidence of CS births is steeply rising, among other reasons due to CS birth being plannable, with the percentage of CS births even exceeding 40% in Latin American countries.[19,20] A recent study, however, has sparked debate by ascribing the microbial effects previously attributed to CS birth to the accompanying maternal intrapartum antibiotic prophylaxis (IAP).[21] As expected based on clinical guidelines,[22–24] IAP was co-linear with CS birth in this study, so these factors could not be studied independently. Consequently, the question remains unanswered whether CS birth in itself, or rather IAP, is responsible for neonatal gut microbiome perturbation.

Nevertheless, antibiotics in general are widely accepted as drivers of microbial change.[25,26] Antibiotic administration in early-life reduces the alpha diversity, or species richness, of the gut microbiome. Also, it leads to an increased relative abundance of, among others, *Klebsiella* and *Enterobacter* genera.[27,28] Additionally, antibiotic pressure promotes the progression of antimicrobial resistance (AMR), which can complicate the successful treatment of infections.[29] While antibiotics have existed in nature for billions of years as compounds produced by microbes to kill or inhibit the growth of other microbes competing for the same scarce resources, the miracle drug penicillin was only discovered by Alexander Fleming in 1928.[30] Though in the following years multiple classes of antibiotics were developed to treat wide-ranging infections, their misuse has rapidly increased the problem of AMR, which could lead to a global health crisis in the near future. In a review commissioned by the Wellcome Trust, it has been estimated that by the year 2050, infections with resistant organisms will lead to losses of 10 million lives a year and a cumulative 100 trillion US dollars of economic output, if solutions are not quickly found to slow down the rise of AMR.[31]

Despite this, broad-spectrum antibiotics are still commonly prescribed in newborns for suspected early onset neonatal sepsis (sEONS). Antibiotics are actually the most common medication prescribed on the neonatal ward.[32] Though potentially life-saving in case of proven sepsis, antibiotics are nonetheless overused for this indication: 4-10% of all neonates receive antibiotics for a suspected infection, whereas only an estimated 1 neonate in 1000 will develop a proven infection. This likely results in unnecessary treatment of >90% of all treated neonates.[33–37] Interestingly, large differences exist between countries in antibiotic prescribing behavior for this indication, with Norway having a relatively low prescription rate of around 2%.[38] A web-based survey amongst clinicians suggests that the prescription rate in Switzerland may be even lower, due to the national guideline stimulating observation over immediate treatment of neonates with low risk of EONS.[39,40] Although antibiotic stewardship programs focus mainly on shortening the duration of antibiotic courses, in clinical practice some infants are treated empirically for suspected sepsis beyond 48 hours, despite negative culture results.[41] It is as of yet uncertain what the influence is of broad-spectrum antibiotics on the assembling infant microbiome and the antibiotic resistance gene pool (resistome), as most studies have been performed in small cohorts of preterm neonates.[27,28]

A link has recently been found, though, between antibiotic-induced gut microbial changes and infant respiratory health. In a longitudinal study, researchers suggest that a reduced incidence of pediatric asthma might be due to restrained antibiotic

use in infancy, through preservation of the gut microbial community.[42] This is in line with the concept of the gut-lung axis, where the microbiome of two distal sites, namely the gut and lungs, may be interconnected, and so play a role in modulating the host's immune system and the susceptibility to respiratory disease. Already, in earlier studies based on the Microbiome Utrecht Infant Study (MUIS) cohort, we have correlated the development of the nasopharyngeal and oropharyngeal microbiome with development of RTIs. Our research group has provided evidence that an accelerated maturation of the nasopharyngeal microbiome is associated with the number of RTIs experienced over the first year of life.[43] Additionally, a loss of topography of the upper respiratory microbiome, with an influx of oral bacteria in the nasopharynx, appears to precede RTI episodes.[44] Currently, research linking the infant respiratory with the gastrointestinal microbiome is lacking. Yet, approaching microbiome data from a more holistic, overall microbial network perspective, by combining the microbial data of multiple anatomical niches, may provide further insight into the pathogenesis of RTIs. By studying network features in relation to outcomes such as respiratory health, we might gain a better understanding of the potential beneficial aspects of the microbiome for the human host.

Despite the fact that an increasing number of studies are being performed to relate the gut microbiome to specific health outcomes or interventions, the problem exists that fecal material is not always readily available for collection.[45] Therefore, it would be extremely useful to establish alternative, reliable sampling methods for gut microbiome research. In adults, rectal swabs have already been confirmed to have a satisfactory concordance with fecal samples when studying the gut microbiota.[46,47] In the pediatric population, however, the performance of rectal swabs in the analysis of the overall gut microbiota composition is unknown.

In this thesis, we aim to address the abovementioned challenges by answering the following questions:

- How does the gut microbiome develop in early-life and what is the influence of mode of delivery, independent of maternal antibiotics?
- Is there an association between gut microbiome development and the susceptibility to respiratory infections?
- Are rectal swabs a reliable proxy for fecal samples to study the gut microbiome in very early infancy?
- What are consequences of early-life broad-spectrum antibiotics on gut microbiome and resistome development in the first year of life?

- Which of the three most commonly prescribed antibiotic combinations in the Netherlands for sEONS causes the least ecological harm?
- Can we identify cross-niche microbial networks spanning the infant upper respiratory and gut microbiome?
- Are cross-niche microbial networks in infants associated with respiratory health?

In **chapter 2** we describe the normal development of the gut microbiome and the influence of environmental factors, such as mode of delivery, by studying 120 infants participating in the MUIS study. In this healthy birth cohort, mothers giving birth via CS received prophylactic antibiotics only after clamping of the umbilical cord, preventing direct transfer of antibiotics from mother to infant. This enables us to assess the effect of delivery mode on the gut microbiome, independent of maternal antibiotics. Furthermore, our study design allows us to study possible associations between the early-life gut microbiome and susceptibility to respiratory infections in the first year of life. To confirm that rectal swabs are a suitable alternative for fecal samples in fecal microbiome studies in infants, we collected both sample types in the ZEBRA trial (Zuigelingen En Bacteriële Resistentie na Antibiotica). The primary aim of the ZEBRA trial was to study the effects of the most commonly prescribed antibiotic combinations in the Netherlands for sEONS on the infant gut microbiome and resistome. We enrolled 147 children with sEONS for whom broad-spectrum antibiotics were indicated, and randomly allocated them to treatment with penicillin + gentamicin, co-amoxiclav + gentamicin or amoxicillin + cefotaxime. We collected rectal swabs and/or fecal samples before the initiation of antibiotic treatment, directly after antibiotic treatment cessation and at 1, 4 and 12 months of life. In **chapter 3** we compare 131 pairs of rectal swabs and fecal samples collected at the first time points. Then, in **chapter 4** we describe the main results of the ZEBRA trial with respect to the effects of early-life broad-spectrum antibiotics on the developing gut microbiome and resistome. Next, in **chapter 5**, we describe infant cross-niche microbial networks, through combining samples collected from the gut, nasopharynx and oral cavity of the MUIS cohort infants, and relate these to the susceptibility to RTIs. Finally, in **chapter 6** we review and discuss the results of the studies presented in this thesis in relation to the current literature and formulate recommendations for future studies.

REFERENCES

1 Sender R, Fuchs S, Milo R. Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* 2016; **164**: 337–40.

2 Ley RE, Peterson DA, Gordon JI. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell. 2006; **124**: 837–48.

3 Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: The unseen majority. *Proc Natl Acad Sci U S A* 1998; **95**: 6578–83.

4 Foster KR, Schluter J, Coyte KZ, Rakoff-Nahoum S. The evolution of the host microbiome as an ecosystem on a leash. *Nature* 2017; **548**: 43–51.

5 Moeller AH, Caro-Quintero A, Mjungu D, *et al.* Cospeciation of gut microbiota with hominids. *Science* 2016; **353**: 380–2.

6 Robertson RC, Manges AR, Finlay BB, Prendergast AJ. The Human Microbiome and Child Growth – First 1000 Days and Beyond. *Trends Microbiol* 2019; **27**: 131–47.

7 Gensollen T, Iyer SS, Kasper DL, Blumberg RS. How colonization by microbiota in early life shapes the immune system. *Science* 2016; **352**: 539–44.

8 Budden KF, Gellatly SL, Wood DLA, *et al.* Emerging pathogenic links between microbiota and the gut–lung axis. *Nat Rev Microbiol* 2016; **15**: 55–63.

9 Penders J, Thijs C, Vink C, *et al.* Factors Influencing the Composition of the Intestinal Microbiota in Early Infancy. *Pediatrics* 2006; **118**: 511–21.

10 Dominguez-Bello MG, Costello EK, Contreras M, *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**: 11971–5.

11 Bäckhed F, Roswall J, Peng Y, *et al.* Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* 2015; **17**: 690–703.

12 Bokulich NA, Chung J, Battaglia T, *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci Transl Med* 2016; **8**: 343ra82-343ra82.

13 Martin R, Makino H, Cetinyurek Yavuz A, *et al.* Early-Life Events, Including Mode of Delivery and Type of Feeding, Siblings and Gender, Shape the Developing Gut Microbiota. *PLoS One* 2016; **11**: e0158498.

14 Stokholm J, Thorsen J, Chawes BL, *et al.* Cesarean section changes neonatal gut colonization. *J Allergy Clin Immunol* 2016; **138**: 881-889.e2.

15 Yasmin F, Tun HM, Konya TB, *et al.* Cesarean Section, Formula Feeding, and Infant Antibiotic Exposure: Separate and Combined Impacts on Gut Microbial Changes in Later Infancy. *Front Pediatr* 2017; **5**: 200.

16 Hill CJ, Lynch DB, Murphy K, *et al.* Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort. *Microbiome* 2017; **5**: 4.

17 Tamburini S, Shen N, Wu HC, Clemente JC. The microbiome in early life: implications for health outcomes. *Nat Med* 2016; **22**: 713–22.

18 Boucher HW, Talbot GH, Bradley JS, *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin Infect Dis* 2009; **48**: 1–12.

19 Betrán AP, Ye J, Moller A-B, Zhang J, Gülmezoglu AM, Torloni MR. The Increasing Trend in Caesarean Section Rates: Global, Regional and National Estimates: 1990-2014. *PLoS One* 2016; **11**: e0148343.

20 Mariani GL, Vain NE. The rising incidence and impact of non-medically indicated pre-labour cesarean section in Latin America. *Semin Fetal Neonatal Med* 2019; **24**: 11–7.

21 Chu DM, Ma J, Prince AL, Antony KM, Seferovic MD, Aagaard KM. Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat Med* 2017; **23**: 314-26.

22 NICE. Caesarean section clinical guideline [CG132]. London, 2011.

23 Mackeen AD, Packard RE, Ota E, Berghella V, Baxter JK. Timing of intravenous prophylactic antibiotics for preventing postpartum infectious morbidity in women undergoing cesarean delivery. *Cochrane Database Syst Rev* 2014; **2014**. DOI:10.1002/14651858.CD009516.pub2.

24 Bailey SR, Field N, Townsend CL, Rodger AJ, Brocklehurst P. Antibiotic prophylaxis for women undergoing caesarean section and infant health. *BJOG* 2016; **123**: 875–6.

25 Zeissig S, Blumberg RS. Life at the beginning: perturbation of the microbiota by antibiotics in early life and its role in health and disease. *Nat Immunol* 2014; **15**: 307–10.

26 Sana TG, Monack DM. Microbiology: The dark side of antibiotics. *Nature* 2016; **534**: 624–5.

27 Greenwood C, Morrow AL, Lagomarcino AJ, *et al.* Early Empiric Antibiotic Use in Preterm Infants Is Associated with Lower Bacterial Diversity and Higher Relative Abundance of Enterobacter. *J Pediatr* 2014; **165**: 23–9.

28 Gibson MK, Wang B, Ahmadi S, *et al.* Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat Microbiol* 2016; **1**: 16024.

29 van Schaik W. The human gut resistome. *Philos Trans R Soc Lond B Biol Sci* 2015; **370**: 20140087.

30 Davies J, Davies D. Davies J, Davies D. Origins and Evolution of Antibiotic Resistance. *Microbiol Mol Biol Rev* 2010; **74**: 417–433.

31 O'Neill J (chair). Review on Antimicrobial Resistance. a Rev. Spons. by Wellcome Trust. 2016. https://amr-review.org/ (accessed Feb 21, 2017).

32 Hsieh EM, Hornik CP, Clark RH, Laughon MM, Benjamin DK, Smith PB. Medication use in the neonatal intensive care unit. *Am J Perinatol* 2014; **31**: 811–21.

33 Mukhopadhyay S, Eichenwald EC, Puopolo KM. Neonatal early-onset sepsis evaluations among well-appearing infants: projected impact of changes in CDC GBS guidelines. *J Perinatol* 2013; **33**: 198–205.

34 Escobar GJ, Puopolo KM, Wi S, *et al.* Stratification of Risk of Early-Onset Sepsis in Newborns ≥ 34 Weeks' Gestation. *Pediatrics* 2014; **133**: 30–6.

35 Weston EJ, Pondo T, Lewis MM, *et al.* The burden of invasive early-onset neonatal sepsis in the united states, 2005-2008. *Pediatr Infect Dis J* 2011; **30**: 937–41.

36 Stocker M, van Herk W, el Helou S, *et al.* Procalcitonin-guided decision making for duration of antibiotic therapy in neonates with suspected early-onset sepsis: a multicentre, randomised controlled trial (NeoPIns). *Lancet* 2017; **390**: 871–81.

37 Vergnano S, Menson E, Kennea N, *et al.* Neonatal infections in England: The neonIN surveillance network. *Arch Dis Child Fetal Neonatal Ed* 2011; **96**. DOI:10.1136/adc.2009.178798.

38 Thaulow CM, Berild D, Blix HS, Brigtsen AK, Myklebust TÅ, Eriksen BH. Can We Optimize Antibiotic Use in Norwegian Neonates? A Prospective Comparison Between a University Hospital and a District Hospital. *Front Pediatr* 2019; **7**. DOI:10.3389/fped.2019.00440.
39 Martin S, Christoph B, Jane M, Eric G. Recommendations for term and late preterm infants at risk for perinatal bacterial infection. *Swiss Med Wkly* 2013; **143**. DOI:10.4414/smw.2013.13873.
40 Van Herk W, El Helou S, Janota J, *et al.* Variation in Current Management of Term and Late-preterm Neonates at Risk for Early-onset Sepsis: An International Survey and Review of Guidelines. *Pediatr Infect Dis J* 2016; **35**: 494–500.
41 Wirtschafter DD, Padilla G, Suh O, Wan K, Trupp D, Simon Fayard EE. Antibiotic use for presumed neonatally acquired infections far exceeds that for central line-associated blood stream infections: An exploratory critique. *J Perinatol* 2011; **31**: 514–8.
42 Patrick DM, Sbihi H, Dai DLY, *et al.* Decreasing antibiotic use, the gut microbiota, and asthma incidence in children: evidence from population-based and prospective cohort studies. *Lancet Respir Med* 2020; **11**: 1094-1105.
43 Bosch AATM, de Steenhuijsen Piters WAA, van Houten MA, *et al.* Maturation of the Infant Respiratory Microbiota, Environmental Drivers, and Health Consequences. A Prospective Cohort Study. *Am J Respir Crit Care Med* 2017; **196**: 1582–90.
44 Man WH, Clerc M, de Steenhuijsen Piters WAA, *et al.* Loss of microbial topography between oral and nasopharyngeal microbiota and development of respiratory infections early in life. *Am J Respir Crit Care Med* 2019; **200**: 760–70.
45 den Hertog J, van Leengoed E, Kolk F, *et al.* The defecation pattern of healthy term infants up to the age of 3 months. *Arch Dis Child Fetal Neonatal Ed* 2012; **97**: F465-70.
46 Budding AE, Grasman ME, Eck A, *et al.* Rectal swabs for analysis of the intestinal microbiota. *PLoS One* 2014; **9**: e101344.
47 Bassis CM, Moore NM, Lolans K, *et al.* Comparison of stool versus rectal swab samples and storage conditions on bacterial community profiles. *BMC Microbiol* 2017; **17**: 78.

CHAPTER 2

Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life

Marta Reyman, Marlies A. van Houten, Debbie van Baarle, Astrid A.T.M. Bosch, Wing Ho Man, Mei Ling J.N. Chu, Kayleigh Arp, Rebecca L. Watson, Elisabeth A.M. Sanders, Susana Fuentes*, Debby Bogaert*

* These authors contributed equally to this work

Nature Communications 2019; **10***: 4997*

ABSTRACT

The early life microbiome appears to be affected by mode of delivery, but this effect may depend on intrapartum antibiotic exposure. We here assess the effect of delivery mode on gut microbiota, independent of intrapartum antibiotics, by postponing routine antibiotic administration to mothers until after cord clamping in 74 vaginally delivered and 46 caesarean section born infants. The microbiota differs between caesarean section born and vaginally delivered infants over the first year of life, showing enrichment of *Bifidobacterium* spp., and reduction of *Enterococcus* and *Klebsiella* spp. in vaginally delivered infants. The microbiota composition at one week of life is associated with the number of respiratory infections over the first year. The taxa driving this association are more abundant in caesarean section born children, providing a possible link between mode of delivery and susceptibility to infectious outcomes.

INTRODUCTION

The impact of the human microbiome on health is becoming increasingly clear, with perturbations being associated with various (immune) disorders ranging from allergies and obesity to inflammatory bowel disease[1]. The gastrointestinal (GI) tract is of particular relevance to human health, as it contains the majority and most diverse set of human commensal bacteria[2]. Early-life gut microbiota development is crucial for a balanced priming of the immune system, which occurs early in life in the so-called window of opportunity[3]. Mode of delivery is considered a critical influential factor on gut microbiota development. Birth by caesarean section (CS) has been associated with adverse effects on immune development, predisposing to infections, allergies and inflammatory disorders[4–6]. The rising incidence of CS births is alarming, reaching up to 40.5% of all births in some countries[7]. Research into the impact of CS birth on microbiota and health argues that results may be largely affected by intrapartum antibiotics[8]. Furthermore, the diminished success of breastfeeding after CS[9] adds to alterations in normal microbiota development[10].

In our study, we investigate a cohort of 120 healthy children not directly exposed to intrapartum antibiotics (Microbiome Utrecht Infant Study [MUIS]). In case of CS delivery, mothers were administered perioperative prophylaxis only after clamping of the umbilical cord, making it possible to focus on the independent effects of delivery mode on the gut microbiota. The infants were intensively sampled directly after birth throughout the first year of life. Fecal samples were collected from mothers, and a broad scale of clinical data and environmental and lifestyle characteristics was obtained for all participants. Primarily, we characterized the infant fecal microbiota composition and dynamics over the first year of life and assess, next to delivery mode, the effect of multiple variables, such as feeding type and early-life antibiotic use. Secondarily, we assess the effect of the observed delivery mode-induced gut microbiota alterations on infant health.

Here, we report on differences in the fecal microbiota between CS and vaginally delivered (VD) infants over the first year of life, independent of maternal antibiotics. In VD infants we find evidence for fecal microbiota seeding from mother to infant and a more stable microbiota development in early life compared with CS infants. Regarding specific taxa, VD infants show, among others, an enrichment of health-associated *Bifidobacterium* spp., and a reduction of potentially pathogenic *Enterococcus* and *Klebsiella* spp. The overall microbiota composition differs most pronouncedly between the delivery mode groups at 1 week of age. At this early timepoint, the microbiota composition is associated with the number of respiratory

infections (RIs) a child will suffer from in the first year of life. Taxa strongly associated with more RIs are more abundant in CS children, providing a possible link between mode of delivery and susceptibility to infectious outcomes.

METHODS

Study population and sample collection

Data and fecal samples were available from 120 children, of which 74 were VD and 46 born by CS, and their mothers, who had participated in the prospective MUIS study consisting of healthy, full term Dutch children[11]. These 120 children had all (1) completed the 1-year follow-up (except for two children, for whom samples were only collected until 6 months, after which they dropped out due to moving out of the area) and (2) had at least five fecal samples available from ten timepoints. Recruitment took place during pregnancy and written informed consent was obtained from both parents. Ethical approval was granted by the national ethics committee in the Netherlands, METC Noord-Holland (committee on research involving human subjects, M012-015, NTR3986). The study was conducted in accordance with the European Statements for Good Clinical Practice. The study had originally been powered based on the abundance and distribution of previously published microbiota data from infants[12], ensuring a power of 0.8 to detect at least significant differences in alpha and beta diversity between groups, as well as differences in abundance of the 25 most important operational taxonomical units (OTUs), taking into consideration OTUs with high and low variability and abundance and varying effect sizes. This power calculation was later verified by an online (Human Microbiome Project based) tool[13]. We initially aimed to enroll 88 infants, 44 infants per delivery mode group, allowing a drop-out of 10%. Due to uneven enrolment in both arms, approval was granted by the ethical committee to prolong enrolment to ensure sufficient CS recruitment, simultaneously continuing the parallel enrolment of VD infants to prevent seasonal/annual differences in microbiota development between the delivery mode groups. Eventually, 78 VD and 52 CS children were recruited; 10 (7.7%) dropped out after an average of 2 weeks of follow-up.

Hospital and home visits took place within 2 h postpartum (pp), 24-36 h after delivery (d1), at 7 (w1) and 14 (w2) days and at 1, 2, 4, 6, 9 and 12 months (m) of age. Fecal samples were collected by a nurse during hospitalization or by the parents prior to the home visits. On a voluntary basis, mothers provided one fecal sample 2 weeks after childbirth. Sterile fecal containers were used for sample

collection and parents were instructed to store the material directly at -20°C in the (home) freezer. Samples were transported on dry ice and transferred for long-term storage at -80°C until further laboratory processing. Extensive questionnaires on the health status of the children were collected at each visit by research personnel and additionally at 3 months. At baseline, information was collected on prenatal and perinatal characteristics.

RI events were defined as occurrence of fever (>38.0°) in combination with any of the following parent-reported symptoms: cough, wheezing, dyspnea, earache or malaise. Gastrointestinal (GI) complaints were categorized into: postpartum nausea, constipation, diarrhea and vomiting and were noted as being either present or absent since the previous home visit. For each symptom, occurrences over the first year were summed up to a cumulative number.

DNA isolation and sequencing

Fecal samples were thawed and homogenized by vortexing. For bacterial DNA extraction ~100 µl of raw feces was added to 300 µl of lysis buffer (Agowa Mag Mini DNA Isolation Kit, LGC ltd, UK), 500 µl of 0,1-mm zirconium beads (BioSpec products, Bartlesville, OK, USA) and 500 µl of phenol saturated with Tris-HCl (pH 8.0; Carl Roth, GMBH, Germany) in a 96-wells plate. The samples were mechanically disrupted using a Mini-BeadBeater-96 (BioSpec products, Bartlesville, OK, USA) for 2 min at 2100 oscillations per minute. DNA purification was performed using the Agowa Mag Mini DNA Isolation Kit according to the manufacturer's recommendations. The extracted DNA was eluted in a final volume of 60 µl of elution buffer (LGC Genomics, Germany). Samples collected directly postpartum and on day 1 were presumed to have low bacterial abundance and diversity. Therefore, adaptations in the standard DNA isolation procedure were applied: 150 µl of raw feces was added to 350 µl of lysis buffer and bead beating was done for 2*3 min. In total, 300 µl of the aqueous layer was used for extraction with the Agowa Mag Mini DNA Isolation Kit, the binding time of DNA to the magnetic beads was prolonged to 30 min, the magnetic beads were washed twice with wash buffer 1 and the extracted DNA was eluted in a final volume of 40 µl of buffer. As the yield of these early samples was low, another round of DNA isolation was performed on the aliquots of this set with an altered protocol: a small loop of feces (~150 µl, or 100 µl when liquid) was added to 650 µl lysis buffer including zirconium beads and 600 µl of phenol, and the whole aqueous layer was used. DNA blanks and feces pools consisting of a mix of up to 3 random samples served as controls. The amount of bacterial DNA was determined by quantitative polymerase chain reaction (qPCR) as described[14], using primers specific for the bacterial 16S rRNA

gene (forward: CGAAAGCGTGGGGAGCAAA; reverse: GTTCGTACTCCCCAGGCGG; Probe: 6FAM-ATTAGATACCCTGGTAGTCCA-MGB) on the 7500 Fast Real Time system (Applied Biosystems, CA, USA).

For the sequencing of the V4 hypervariable region of the 16S rRNA gene, 1 ng of DNA was amplified using F515/R806 primers and 30 amplification cycles[15,16]. After amplification of the V4 hypervariable region of the 16S rRNA, the amount of amplified DNA per sample was quantified with the dsDNA 910 Reagent Kit on the Fragment Analyzer (Advanced Analytical, IA, USA). Samples that yielded insufficient DNA after amplification, defined as <0.5 ng per µl, were repeated with a higher concentration of template DNA. Each PCR plate included a mock control and three PCR blanks. 16S rRNA sequencing was performed on the Illumina MiSeq platform (Illumina, Eindhoven, the Netherlands) on a total of 1139 samples and 85 controls in 17 runs.

Bioinformatic processing

The sequences were processed in our bioinformatics pipeline, where we applied an adaptive, window-based trimming algorithm (Sickle, version 1.33) to filter out low quality reads maintaining a Phred score threshold of 30 and a length threshold of 150 nucleotides[17]. Error correction was conducted with BayesHammer (SPAdes genome assembler toolkit, version 3.5.0)[18]. Each set of paired-end sequence reads was assembled with PANDAseq (version 2.10) and demultiplexed (QIIME, version 1.9.1)[19,20]. Singleton and chimeric reads (UCHIME) were removed. OTU picking was performed with VSEARCH abundance-based greedy clustering with a 97% identity threshold[21]. OTUs were annotated with the Naïve Bayesian RDP classifier (version 2.2) and the SILVA reference database[22,23]. This resulted in an OTU-table containing 6690 taxa. We made an abundance-filtered dataset selecting OTUs present at a confident level of detection (0.1% relative abundance) in at least two samples[24], henceforth referred to as our raw OTU-table. The raw OTU-table consisted of 623 taxa (0.4% sequences excluded with filtering) and was used for the downstream analyses unless otherwise specified.

Whole genome shotgun sequencing and processing

To validate the taxonomical annotation of the 16S rRNA sequences and some of our 16S rRNA-based findings, we performed whole genome shotgun (WGS) sequencing on a randomly selected subset of 20 fecal samples collected from 10 VD and 10 CS born children at the age of 1 week. Samples were prepared using the Truseq Nano gel free library preparation kit. Using a NovaSeq instrument, 150 base paired-end sequence data was generated from the libraries to yield 750M+750M reads

(two runs). Reads were trimmed using Cutadapt[25] (version cutadapt-1.9.dev2) of amplicon adapter sequences and on quality at the 3' end maintaining a quality threshold of 30 and a minimum read length of 35 base pairs. Per-sample and per-run SAM files were generated with Bowtie2[26] and the MetaPhlAn2[27] database while adhering to recommended parameters and using -no-unal to suppress reporting unaligned reads and the -very-sensitive parameter. Each SAM file was assigned a read group and SAM files from different runs were merged sample-wise using Picard[28]. MetaPhlAn2 was run to identify the bacterial taxa present within each sample. The SAM files generated using Bowtie2 were used as input for MetaPhlAn2, all other parameters were kept as default.

Determination of specific biomarkers by qPCR

To identify the presence of *E. coli, Klebsiella* spp. (including *K. oxytoca* and *K. pneumoniae*) and *Enterococcus* spp. in all the samples collected at 1 week of age (n=119), we used the VetMAX™ MastiType Multi Kit (Applied Biosystems™, CA, USA) according to the manufacturer's instructions. The qPCR test results were analyzed with the recommended Animal Health VeriVet Software, available on Thermo Fisher Cloud. One sample was discarded from the statistical analyses because its Internal Amplification Control did not pass the Ct-value criteria in three out of the four mixes.

Statistical analyses

A statistical analysis scheme showing the flow in and order of analyses to address the primary, secondary and exploratory research questions can be found in Supplementary Fig. 1. All analyses were performed in R version 3.4.3[29] within RStudio version 1.1.383[30] and figures were made using packages ggplot2[31] and ggpubr[32]. For simple, independent comparisons, we considered p-values <0.05 to be significant. However, for all analyses regarding multiple comparisons, we used the Benjamini-Hochberg method to correct for multiple testing[33].

For comparisons of group differences, a two-sample t test, Wilcoxon rank-sum test or chi-square test was used where appropriate. Survival analysis was executed using the packages *survival* and *survminer*[34,35]. The hazard ratio for antibiotic administration was calculated with a Cox proportional hazard model.

Group differences in Shannon alpha diversity were calculated with t-tests and a linear mixed-effect model with participant as random effect while correcting for age and feeding type. Differences in overall gut bacterial community composition were visualized with nonmetric multidimensional scaling plots (nMDS; *vegan*

package[36]). Ordinations were based on the Bray-Curtis (BC) dissimilarity matrix of relative abundance data with parameter trymax 10,000. The overall gut bacterial community composition of children born by emergency CS (n=10) was, although not fully similar, more similar to that of children born by planned CS (Supplementary Fig. 2; permutational multivariate analysis of variance (PERMANOVA)-test, R^2 0.005, p=0.051) than by vaginal birth (R^2 0.006, p=0.002). Because the number of children born by emergency CS was too small to analyze as an independent group, we decided to group emergency and planned CS children together for all delivery mode comparisons.

To study whether the differences in overall gut microbiota community composition between delivery mode groups were not influenced by antibiotics received indirectly through the breastmilk from mothers treated prophylactically after CS, post hoc analyses were performed on a subset of 11 VD and 11 CS infants that received exclusive formula feeding.

Associations between clinical outcome and microbiota composition were analyzed with the adonis2 function (*vegan* package[36]), based on PERMANOVA-tests per timepoint and across all timepoints using 1999 permutations, including all variables that showed a significant association with microbiota composition when tested individually, namely mode of delivery, duration of hospital stay after delivery, breastfeeding at time of sampling, pacifier use, antibiotics in the 4 weeks prior to sampling, siblings <5 years in the household, pets in the household, high education of parents and daycare attendance. For the temporal analyses, age and subject were added to control for repeated measures.

To test the occurrence of fecal seeding (i.e. whether the microbiota composition in VD children was more similar to that of their own mothers than in the CS group), BC dissimilarities were calculated between children's and their own mother's vs. other mothers' samples, stratified per group. We used a linear mixed model to assess the effect of delivery mode (fixed effect) on BC dissimilarity as a dependent variable, while including subject as a random intercept to adequately control for repeated measures and correcting for timepoint, using the lme function of the *nlme* package[37]. *P*-values of our linear mixed models were extracted using the ANOVA function.

The stability of the gut microbiota composition in the first year of life was visualized by measuring the BC dissimilarities between consecutive samples within each participant over time. To test for group differences, the Mann-Whitney test was used.

Individual bacterial taxa and their succession patterns, and potential differences thereof between groups, were studied at the lowest taxonomic annotated level (OTU). Differential abundance testing was executed with smoothing spline analysis of variance (SS-ANOVA, fitTimeSeries function, *metagenomeseq* package[38,39]) allowing not only to detect biomarker OTUs related to mode of delivery, but also to identify the specific intervals in which significant differences existed. For this analysis, the raw OTU-table was filtered and OTUs with >10 reads in ≥50 samples were included, resulting in 306 OTUs (of the 623)[40]. The SS-ANOVA analysis was adjusted for covariates that both had (1) an effect on overall gut microbiota composition and (2) were unevenly distributed between the two delivery mode groups, namely hospital stay duration after birth, breastfeeding at time of sampling and antibiotic use in the 4 weeks prior to sampling. Although duration of ruptured membranes and gestational age were associated with mode of delivery, we did not find an association with gut microbiota composition, therefore these two variables were excluded from downstream fitTimeSeries analyses. A post hoc SS-ANOVA was performed on the subset of children who were exclusively formula fed, specifically testing for differences in the top five most abundant taxa and focusing on the first 2 months of life, as differences between delivery mode groups overall were most pronounced in this period.

We used linear mixed models to test the importance of delivery mode, duration of hospital stay after birth, feeding type and antibiotic use on the relative abundance of the top 5 most abundant taxa. Using the lme function[37], we set the clinical variables as fixed effects and the arcsine square root transformed relative abundances of each taxon of interest as dependent variable, adding subject as random intercept and correcting for age.

To test if the microbiota could predict for cumulative disease parameters (namely fever episodes, thrush, GI symptoms, RI events, general practitioner and specialist consultation and antibiotic prescription) after 1 week of age, we performed a BC-based PERMANOVA on the overall community composition of the samples obtained at this sampling moment. The number of RI events in the first year of life was tested as both a continuous and categorical variable grouped in 0-2 and 3-7 episodes, based on previous studies[40,41]. The fitZig function of the *metagenomeSeq* package was used to assess the driving OTUs behind significant predictions[39], after removing rare features present in less than 10 samples, resulting in 97 OTUs included in this analysis. The analysis was adjusted for delivery mode. Random forest analysis was used to verify these results, setting the categorized number of RI events as outcome and the OTUs present in the samples at 1 week of age as predictors along with

delivery mode and variables also adjusted for in the fitTimeSeries analysis[42]. We also performed the fitZig analysis in a stratified manner for both delivery mode groups.

Results from 16S rRNA sequencing at 1 week of life were validated by untargeted WGS sequencing (subset of 20 infants) and targeted qPCR (on all 120 infants).

RESULTS

Population characteristics

In our study population, 74 children were VD and 46 children were born by CS. Of those, 36 (78%) were born by planned and 10 (22%) by emergency CS. There were two cases of pre-/intrapartum antibiotics, one in each delivery mode group, indicated for maternal fever, both of which were included in the analyses. All but three children were born in the hospital. Baseline characteristics and cumulative disease parameters over the first year of life of all children, stratified by mode of delivery, are shown in Table 1. Clinical variables were evenly distributed over both groups with the exception of gestational age (two-sample *t* test, $p=0.003$), duration of ruptured membranes ($p=0.019$), hospital stay duration (Wilcoxon test, $p<0.001$) and total duration of breastfeeding in the first year of life ($p=0.014$), all being intrinsically related to delivery mode. The number of children receiving exclusive formula feeding did not differ between the two groups. Only 36 children (30%) received antibiotics during their first year of life, some receiving multiple courses (56 courses in total for all children), mostly (80%) indicated for RIs.

Microbiota composition and mode of delivery

Of the 1243 fecal samples available from our 120 participants and their mothers, 1139 (92%) passed the quality criteria for further analysis following DNA extraction and 16S rRNA-based sequencing of the V4 hypervariable region (Supplementary Fig. 3), representing 70,886,595 high quality reads in total. The Good's coverage of the included children's samples was high, with a minimum of 99.56% (median 99.96%). The raw Operational Taxonomical Unit (OTU)-table contained 623 OTUs distributed over seven bacterial phyla, with the Firmicutes generally being the most prominent phylum.

We observed that the infants' overall microbial community composition developed slowly towards an adult-like profile (mothers'), though had not yet reached full maturation to an adult-like composition at 12 months of age (Fig. 1a). This was illustrated by a steady decrease in Bray-Curtis (BC) dissimilarity index between

Table 1. Baseline characteristics

	Vaginal birth	C-section birth	p
n (%)	74 (61.7)	46 (38.3)	
Gender, female (%)	39 (52.7)	24 (52.2)	1.000
Gravidity mothers, median (IQR)	2.00 (1.00, 3.00)	2.00 (2.00, 2.75)	0.638
Gestational age in weeks, mean (sd)	39.75 (1.21)	39.12 (0.84)	***0.003***
Birth weight in grams, mean (sd)	3490.41 (485.87)	3618.00 (459.18)	0.156
Ruptured membranes in hours, mean (sd)	7.32 (9.82)	3.05 (8.39)	***0.019***
Apgar score at 5 minutes (%)			0.737
6	1 (1.4)	0 (0.0)	
7	2 (2.7)	1 (2.2)	
8	1 (1.4)	2 (4.3)	
9	10 (13.7)	8 (17.4)	
10	59 (80.8)	35 (76.1)	
Season of birth (%)			0.557
Winter	13 (17.6)	12 (26.1)	
Spring	17 (23.0)	12 (26.1)	
Summer	29 (39.2)	13 (28.3)	
Fall	15 (20.3)	9 (19.6)	
Hospital stay in dayparts, median (IQR)	3.00 (1.00, 4.75)	12.00 (9.50, 14.00)	***<0.001***
Number of siblings, median (IQR)	1 (1.00, 1.00)	1 (0.00, 1.00)	0.264
Presence of siblings < 5 years of age (%)	40 (54.1)	28 (60.9)	0.587
Presence of pets (%)			0.910
None	41 (55.4)	25 (54.3)	
Cat(s)	16 (21.6)	11 (23.9)	
Dog(s)	6 (8.1)	4 (8.7)	
Cat(s) and dog(s)	3 (4.1)	3 (6.5)	
Other	8 (10.8)	3 (6.5)	
Inhouse smoking (%)	1 (1.4)	2 (4.3)	0.674
Parents finished higher education (%)	60 (81.1)	31 (67.4)	0.138
Breastfeeding in days, median (IQR)	132.50 (7.00, 310.25)	25 .00(1.00, 124.00)	***0.014***
Exclusive formula feeding (%)	11 (14.9)	11 (23.9)	0.316
Age start solid food in days, median (IQR)	130.50 (118.25, 165.00)	128.00 (120.00, 163.00)	0.816
Pacifier use at 1 month of age (%)	53 (71.6)	33 (71.7)	1.000
Antibiotic use in 1st year of life (%)	19 (26.0)	17 (37.8)	0.254
Number of antibiotic courses, median (IQR)	0.00 (0.00, 0.75)	0.00 (0.00, 1.00)	0.119

Continue

Continued

	Vaginal birth	C-section birth	p
Daycare since (%)			0.743
2 months	1 (1.4)	0 (0.0)	
3months	18 (24.3)	11 (24.4)	
4 months	14 (18.9)	12 (26.7)	
6 months	8 (10.8)	7 (15.6)	
9 months	9 (12.2)	3 (6.7)	
12 months	1 (1.4)	0 (0.0)	
> 12 months	23 (31.1)	12 (26.7)	
Fever, median (range)	2.00 (0.00, 4.00)	2.00 (0.00, 5.00)	0.448
Nausea postpartum, median (range)	0.00 (0.00, 2.00)	0.00 (0.00, 1.00)	0.541
Constipation, median (range)	0.00 (0.00, 5.00)	0.00 (0.00, 5.00)	0.496
Diarrhea, median (range)	0.00 (0.00, 2.00)	0.00 (0.00, 2.00)	0.108
Vomiting, median (range)	0.00 (0.00, 3.00)	0.00 (0.00, 1.00)	0.505
Thrush, median (range)	0.00 (0.00, 4.00)	0.00 (0.00, 3.00)	0.316
Respiratory tract infections (%)			0.100
0-2	30 (41.1)	11 (24.4)	
3-7	43 (58.9)	34 (75.6)	

Baseline characteristics stratified by mode of delivery. Categorical variables are shown in absolute numbers with percentages (%); continuous, normally distributed variables as means with standard deviations (SD); continuous, non-normally distributed variables as medians with interquartile ranges (IQR) or ranges where specified. Two sample *t* tests were used to compare means of normally distributed continuous variables; Wilcoxon rank-sum tests were applied to compare medians of non-normally distributed continuous variables; significant differences between categorical variables were tested with chi-square tests. The *p*-values of variables that differed significantly between the two groups are in bold and italicized for clarity.

infants' and mothers' samples over time, with a median index of 0.999 directly postpartum, which reached 0.739 at the end of the first year (Supplementary Table 1). We observed clear differences in the early development of the overall community composition between VD and CS children, with a maximum effect of delivery mode at 1 week of life (PERMANOVA-test, R^2 0.142, adjusted *p*-value 0.003, Benjamini-Hochberg method[33]) and significant differences until the age of 2 months (R^2 0.021, adjusted *p*-value 0.055), after which these differences gradually disappeared (Fig. 2). To rule out this finding was due to the indirect exposure of CS children to maternal antibiotics through breastfeeding, we repeated this analysis post hoc on a subset of 11 VD and 11 CS children who received exclusive formula feeding. We observed similar (at some timepoints even bigger) effects (R^2) of delivery mode on overall community composition until 2 months of life within this subset. The association between delivery mode and composition was still significant at 1 week (Fig. 2 and Supplementary Table 2; R^2 0.215, adjusted *p*-value 0.008) and 2 weeks

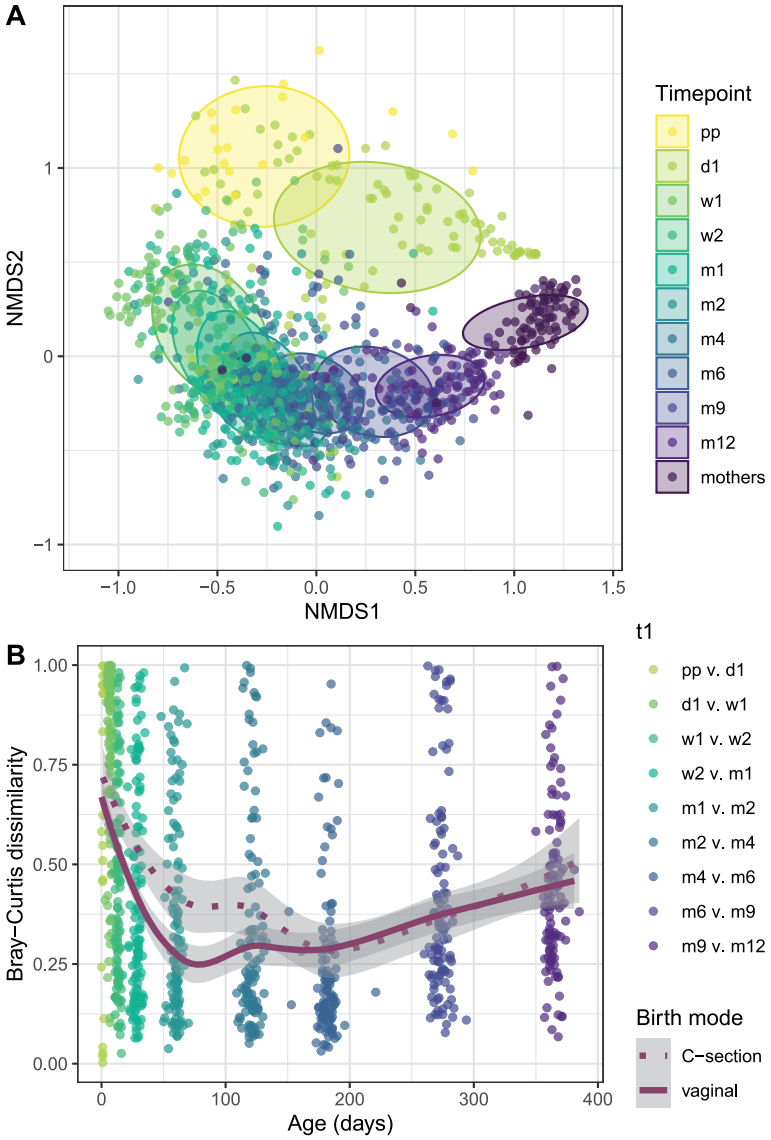


Figure 1. Overall gut microbiota community composition development and stability. a. Non-metric multidimensional scaling (nMDS) plot, based on Bray-Curtis (BC) dissimilarity between samples, with data points and ellipses colored by timepoint. Children's overall gut community composition developed toward a more adult-like pattern in the first year of life, becoming more similar to microbiota of adults (mothers' samples, *n*=87). **b.** As measure of stability, we calculated BC dissimilarities between consecutive sample pairs belonging to an individual per time interval and plotted these at the end of each interval (t+1). Loess lines were fitted over the data points per delivery mode group, and the gray areas represent the 0.95 confidence intervals. Stability was significantly lower in C-section born infants until 2 months of life.


Birth mode
C−section
vaginal
Post partum
n = 6
n = 16
Stress = 0.209
R² 0.114
p.adj = 0.072
Day 1
n = 36
n = 61
R² 0.040
p.adj = 0.012
Week 1
n = 45
n = 74
R² 0.142
p.adj = 0.003
Week 2
n = 46
n = 72
R² 0.100
p.adj = 0.003
Month 1
n = 46
n = 74
R² 0.032
p.adj = 0.020
Month 2
n = 46
n = 73
R² 0.021
p.adj = 0.055
Month 4
n = 40
n = 73
R² 0.012
p.adj = 0.248
Month 6
n = 45
n = 72
R² 0.013
p.adj = 0.236
Month 9
n = 43
n = 72
R² 0.008
p.adj = 0.491
Month 12
n = 43
n = 69
R² 0.008
p.adj = 0.582
Exclusively formula fed subset: week 1
n = 10
n = 11
R² 0.215
p.adj = 0.008
Exclusively formula fed subset: week 2
n = 11
n = 11
R² 0.152
p.adj = 0.044
NMDS1
NMDS2

< Figure 2. NMDS plots of children's samples per timepoint stratified according to mode of delivery. Non-metric multidimensional scaling (nMDS) plots, based on Bray-Curtis (BC) dissimilarity between samples, visualizing the overall gut bacterial community composition stratified for mode of delivery, per timepoint. Each data point represents the microbial community composition of one sample. The ellipses represent the standard deviation of data points belonging to each birth mode group, with the center points of the ellipses calculated using the mean of the coordinates per group. The stress of the ordination, effect sizes (R^2) calculated by multivariate permutational multivariate analysis of variance (PERMANOVA)-tests and corresponding adjusted *p*-values (p.adj) are shown in the plots and *n* represents the biologically independent samples per group.

of life (R^2 0.152, adjusted *p*-value 0.044) in the exclusively formula fed children. For the whole study population, we also found that the microbial community in VD children was more stable when compared with CS children until 2 months of life (Fig. 1b). The BC dissimilarity between consecutive samples until 4 months was higher in CS, compared with VD children (Mann-Whitney test, $p<0.001$, $p=0.001$ and $p=0.042$, for the intervals 1-2 weeks, 2 weeks–1 month and 1-2 months, respectively). In general, alpha diversity increased directly after birth, and again after 4 months of life, coinciding with the age that solid food was introduced to the children's diet (median = 128 days, IQR = 119.8-164.2 days). There were no significant differences in alpha diversity between the two delivery mode groups at any timepoint (Supplementary Fig. 4). When testing the effect of delivery mode on alpha diversity longitudinally with a mixed effect model, no significant effect was found (ANOVA, $p=0.511$), and neither did feeding type have an effect in this model ($p=0.652$).

Covariates that were significantly associated with fecal microbiota composition over time, as tested with the adonis2 function[36] (PERMANOVA-test) were, besides mode of delivery (R^2 0.013, adjusted *p*-value 0.001): age (R^2 0.034, adjusted *p*-value 0.001), breastfeeding (R^2 0.007, adjusted *p*-value 0.001), daycare attendance (R^2 0.006, adjusted *p*-value 0.001), siblings <5 years of age (R^2 0.006, adjusted *p*-value 0.001), pacifier use (R^2 0.005, adjusted *p*-value 0.003) and antibiotics in the 4 weeks prior to sampling (R^2 0.003, adjusted *p*-value 0.03). Pets in the household (R^2 0.006, adjusted *p*-value 0.061) and duration of hospital stay after birth (R^2 0.002, adjusted *p*-value 0.061) showed a trend towards being associated with fecal microbiota composition over time (Supplementary Fig. 5).

Fecal microbiota seeding from mother to infant

To study the existence of direct maternal fecal microbiota seeding during birth, and to assess the role of delivery mode herein, we studied the concordance of the microbiota composition of children's fecal samples and their mother's microbiota

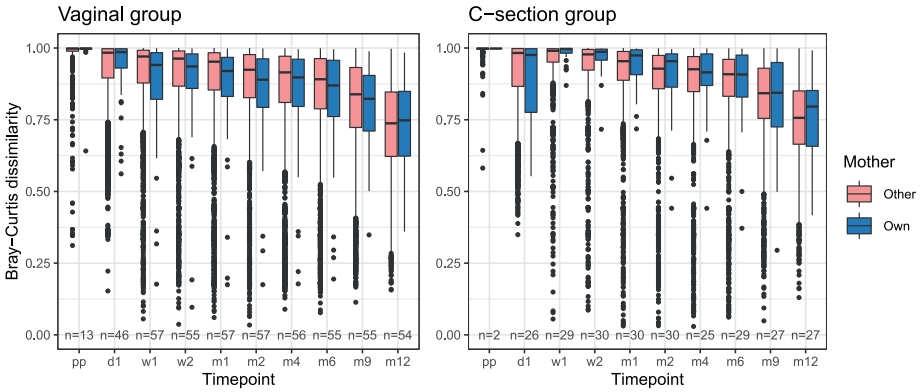


Figure 3. Comparison of overall composition between children and mothers (own vs. other). Children's fecal microbiota were compared with the mothers' fecal microbiota, based on BC dissimilarity and stratified according to mode of delivery. A significantly lower dissimilarity (more comparable microbiota) was observed between a child's microbiota and its own mother vs. other mothers in children born vaginally throughout the first year of life, but not in children born by C-section (linear mixed models, ANOVA; p=0.025 and p=0.271, respectively). Boxplots with medians are shown; the lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles); the upper and lower whiskers extend from the hinge to the largest and smallest value no further than 1.5*IQR from the hinge; outliers are plotted individually. Pp = postpartum, d = day, m = month, n = number of mother-own-infant pair comparison per timepoint.

over time, in relation to the concordance with other mothers' samples. Using a linear mixed model, we found that an infant's fecal microbiota composition was significantly more similar to that of its own mother than to that of other mothers in VD children when studied over the entire first year of life (ANOVA, p=0.025; Fig. 3), but not in CS children (p=0.271). This difference between groups seemed independent of the intravenous antibiotics administered to the mothers in the CS group after cord clamping, as the overall fecal microbiota composition of CS and VD mothers themselves did not differ shortly after birth (PERMANOVA test, R^2 0.013, p=0.351; Supplementary Fig. 6).

Dynamics of microbiota development

The succession pattern of bacterial taxa in the VD children in our study population was consistent with the description of normal early-life gut microbiota development in previous studies (Fig. 4a)[43,44]. Facultative anaerobic genera, such as *Escherichia,* and *Staphylococcus* were highly abundant in the earliest samples, gradually making way for a predominance of the genus *Bifidobacterium*. Using smoothing spline analysis of variance (SS-ANOVA), we observed, among others, that *Bifidobacterium* was more abundant in VD than in CS children from day 1 until day 30, even when

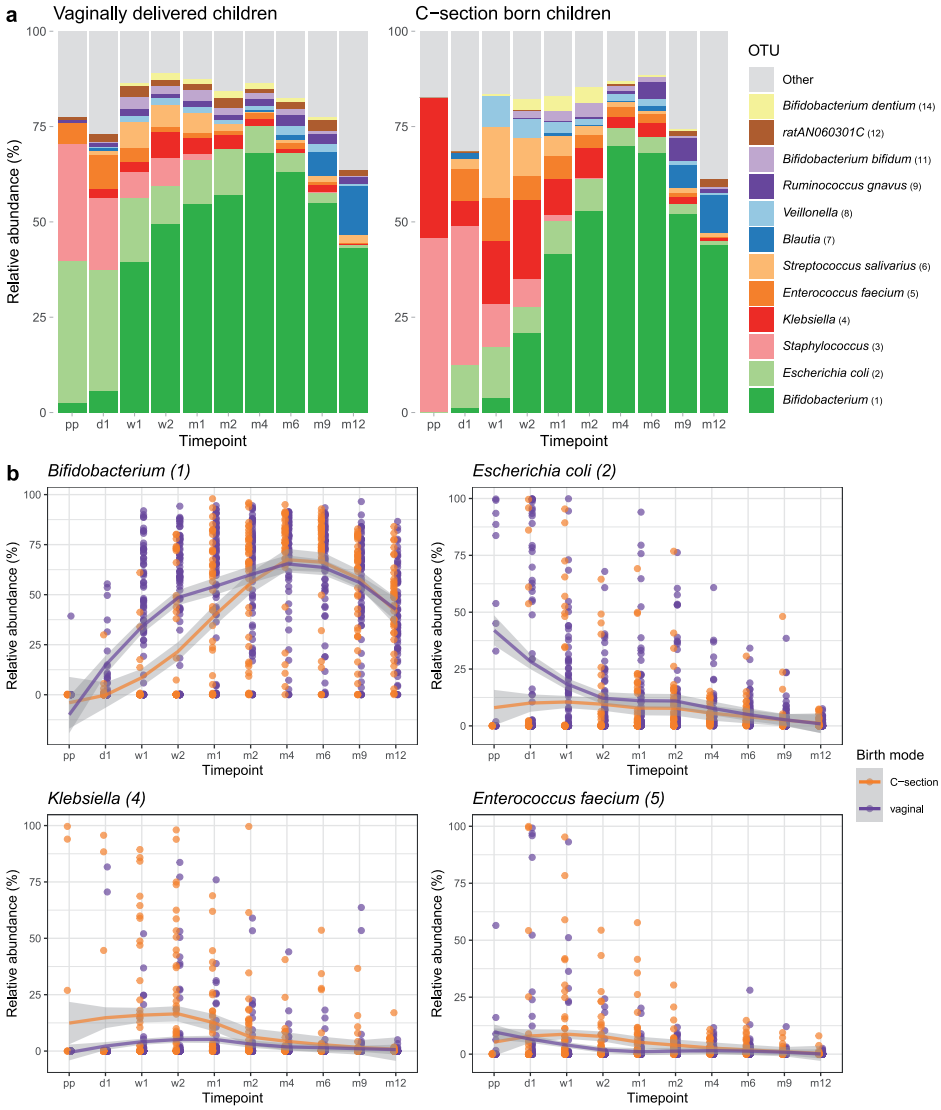


Figure 4. Mean relative abundance of most abundant OTUs. a. Mean relative abundances of the 12 most abundant OTUs are depicted for all samples per timepoint, stratified by birth mode. Pp = postpartum, d = day, m = month. **b.** Mean relative abundances of *Bifidobacterium*, *Escherichia*, *Klebsiella* and *Enterococcus* over time. Loess lines were fitted over the data points per delivery mode group and the gray areas represent the 0.95 confidence intervals.

correcting for breastfeeding (adjusted *p*-value 0.003; Supplementary Table 3). Also, *Escherichia* was more abundant in VD compared with CS born children in the first 85 days of life. In contrast, in CS children we found, among others, higher abundances of *Klebsiella* from birth to day 139 and *Enterococcus* between 7 and 35 days (both adjusted *p*-values 0.003). The dynamics of differences in *Bifidobacterium*, *Escherichia*, *Klebsiella* and *Enterococcus* over time are visualized in Fig. 4b, underlining the effect sizes and duration of differences.

To confirm that these results were not the consequence of indirect antibiotic exposure of infants through breastmilk, we executed a post hoc SS-ANOVA analysis for the subset of exclusively formula fed children. We analyzed the top five most abundant taxa over the first 2 months of life and again found an increased abundance of *Bifidobacterium* (day 5-44, adjusted *p*-value 0.003) in VD infants, and an increased abundance of *Klebsiella* in the CS born infants (day 10-20, adjusted *p*-value 0.020). Also, *Staphylococcus* was found to be more abundant in the CS children from 0 to 6 days (adjusted *p*-value 0.020).

We used mixed effect models to study the potential associations between delivery mode, age, feeding type, antibiotic use and hospital stay duration and the five most abundant taxa. We found that the abundance of *Bifidobacterium* was associated with mode of delivery (ANOVA, $p=0.004$), age ($p<0.001$) and breastfeeding ($p<0.001$). Surprisingly, breastfeeding did not compensate for the lack of *Bifidobacterium* in children born by CS: children born by CS and receiving breastfeeding had less *Bifidobacterium* present in their fecal samples than formula fed VD children (at 1 week of life, Wilcoxon test, $p<0.001$, median relative abundance 0.016% and 45.1%, respectively). The triad CS birth, age and formula feeding was also positively associated with the abundance of *Enterococcus* and *Klebsiella*. In addition, *Klebsiella* abundance was positively associated with having received antibiotics in the previous 4 weeks of life ($p=0.001$). *Escherichia* abundance was only associated with vaginal delivery ($p=0.003$) and age ($p<0.001$). *Staphylococcus* abundance was associated with age and breastfeeding (both $p<0.001$), but not with delivery mode. We did not find an association between duration of hospital stay after birth and any of these taxa.

Among the remaining mode of delivery-associated taxa observed (see Supplementary Table 3), we found to be of particular interest that *Bacteroides* spp., which are considered to be important regulators of intestinal immunity[45], were more abundant in the VD compared with CS children in the first months of life.

Delivery mode-induced microbiota changes and infant health

Since delivery mode is reported to be associated with infant and childhood health, especially regarding respiratory illness[46], we defined a secondary research question, namely whether gut microbiota development is associated with health outcome. Although it was not our aim to study differences in health outcomes between the delivery mode groups in our cohort, we did find a trend toward differences in infectious disease and treatment parameters, specifically parent-reported RI events and antibiotic courses over the first year of life (Table 1, chi-square test, p=0.119 and p=0.100, respectively). Exploring this further with a temporal post hoc analysis, we additionally found a trend towards a lower hazard ratio for antibiotic prescriptions in VD children in the first year of life (Cox proportional hazard model, HR = 0.606, p=0.134).

Altogether, these results supported the validity of our secondary aim to investigate the potential role of delivery mode-induced gut microbiota changes on health. To test this, we studied the association between fecal microbiota composition at 1 week of life (where the maximum effect of mode of delivery on microbiota composition was observed; Fig. 2), and all commonly observed health parameters in the first year of life. We categorized the number of RI events into 0-2 vs. 3-7 RIs, based on previous studies of the respiratory microbiome within this same cohort[40,41]. In these studies, RIs were initially categorized into three groups based on the normal distribution of this variable. The 0-2 RIs group was found to have the most stable development of the nasopharyngeal microbiota when compared with children suffering from >2 RIs in the first year of life, and was defined as the healthy reference group. While there were no correlations between microbiota composition and GI complaints, we observed an association between microbiota composition at 1 week of life and the categorized number of RI events (0-2 vs. 3-7 RI events: PERMANOVA-test, R^2 0.033, adjusted p-value 0.028) as well as number of antibiotic courses prescribed over the first year (R^2 0.024, adjusted p-value 0.055). These two outcomes were related, as the antibiotics prescribed were mostly indicated for RIs. We next aimed to identify the taxa explaining this association between microbiota composition at 1 week and fewer RI events later in life by cross-sectional differential abundance analysis, while adjusting for mode of delivery. We observed, among others, *Bifidobacterium* to be associated with fewer RI events (0-2 vs. 3-7 RI events, zero-inflated Gaussian mixture model, log2 fold change (log2FC) 2.118, adjusted p-value 0.049, Fig. 5), whereas *Klebsiella* and *Enterococcus* were negatively associated with fewer RI events (log2FC -3.242, adjusted p-value 0.007 and log2FC -2.838, adjusted p-value 0.009, respectively). Other taxa found to be negatively associated with fewer RI events encompassed genera such as *Veillonella*

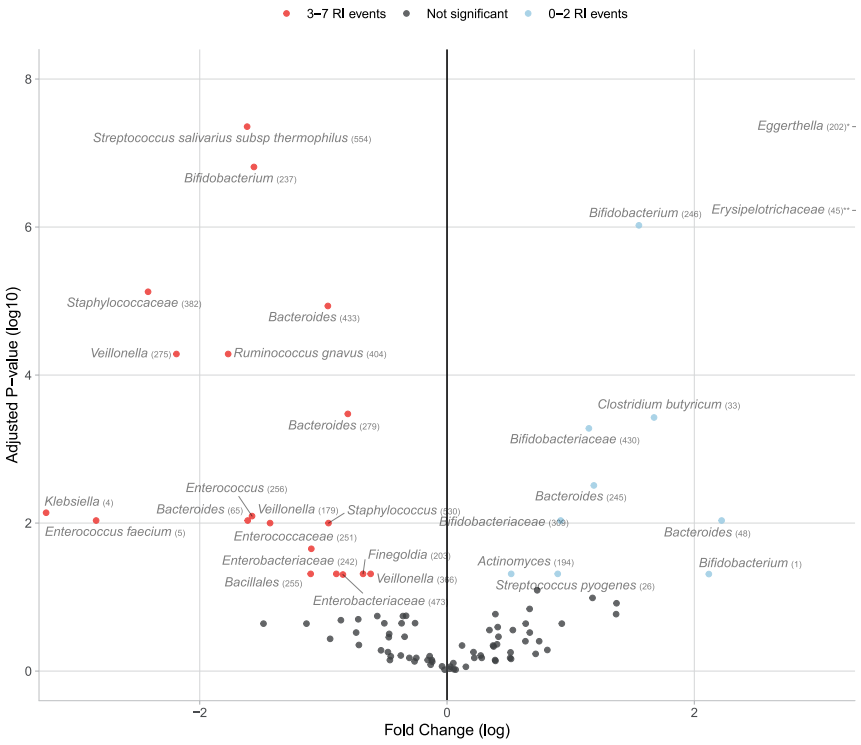


Figure 5. Differentially abundant taxa between 0-2 vs. 3-7 RI events in first year of life. To identify taxa that were differentially abundant between children experiencing more vs. limited respiratory infection (RI) events over the first year of life, fitZig analysis was performed on the 119 samples obtained at week 1 with the rare-features-filtered OTU table containing 97 taxa and contrasts set to 0-2 vs. 3-7 RI events in the first year of life. The blue data points indicate taxa that were significantly more abundant in children having 0-2 RI events, while red points represent taxa that were significantly more abundant in children with 3-7 RI events in the first year of life. The results of two datapoints falling beyond the limits of the plot: **Eggerthella* log2FC 3.512, adjusted *p*-value (log10) 7.357, ***Erysipelotrichaceae* log2FC 6.912, adjusted *p*-value (log10) 6.222, calculated using a zero-inflated Gaussian mixture model.

and *Staphylococcus*. Random forest analysis was used to verify these results and identified once again *Enterococcus, Bifidobacterium* and *Klebsiella* as the most important taxa driving the prediction of the categorized RI events in the first year of life (Supplementary Table 4). Furthermore, a stratified analysis for the VD and CS groups separately, showed similar associations between gut microbiota composition at 1 week of life and number of RI events for the two groups (PERMANOVA, R^2 0.003 and 0.005, *p*-value 0.040 and 0.068, respectively). Taxa associated with the

number of RI events were comparable between the overall and stratified analyses, although for VD children we now only found significance for *Enterococcus* (zero-inflated Gaussian mixture model, log2FC -2.525, adjusted *p*-value 0.074), whereas for the CS children, *Bifidobacterium* (log2FC 2.805), *Klebsiella* (log2FC -6.991) and *Enterococcus* (log2FC; -4.283) were all significantly associated with number of RIs (adjusted *p*-values 0.055, 0.010 and 0.055, respectively).

Validation of the results with metagenomics and targeted qPCR

To validate our primary findings independently, we executed WGS sequencing on a subset of 20 randomly selected samples collected at 1 week of life from 10 VD and 10 CS born children. WGS sequencing yielded a total of 137 unique bacterial taxa. The relative abundances of the top 12 OTUs and species of both sequencing methods are represented in Supplementary Fig. 7, and show highly comparable profiles. The most abundant *Bifidobacterium* species in the WGS dataset were *B. longum*, *B. breve* and *B. adolescentis*. The combined relative abundances of these three species strongly correlated with the most abundant *Bifidobacterium* of the 16S rRNA dataset (Pearson's r 0.95, adjusted *p*-value <0.001). In this way, we could also correlate the *E. coli, Staphylococcus, Klebsiella* and *E. faecium* OTU abundance of the 16S rRNA dataset with high certainty to the *E. coli, S. epidermidis, K. oxytoca* and *E. faecium* species abundance in the WGS dataset (Supplementary Table 5).

We also used the WGS sequencing data to validate the differences found by 16S rRNA sequencing in overall gut microbiota composition between VD and CS born children at 1 week of life. The results from the ordination using the WGS sequencing data are shown in Supplementary Fig. 8. Again, we found a significant effect of delivery mode on the overall gut microbiota composition at 1 week of life using WGS sequencing data (PERMANOVA test, R^2 0.125, p=0.01).

Next, we compared the microbiota profiles obtained by WGS sequencing between VD and CS children and observed that the combination of *B. breve*, *B. longum* and *B. adolescentis* (median relative abundance 72.2% in the VD vs. 0.074% in the CS born children, Wilcoxon test, *p*=0.002), *K. oxytoca* (<0.001% vs. 0.006%, *p*=0.153) and *E. faecium* (0.014% vs. 0.035%, *p*=0.023) were differentially abundant between groups. In this dataset, *E. coli* and *S. epidermidis* did not differ significantly between the delivery mode groups.

To alternatively validate our results on the overall cohort and in a targeted manner, we performed qPCR analyses for *E. coli*, *Klebsiella* spp. and *Enterococcus* spp. on all week 1 samples (*n*=119). The qPCR results confirmed that *E. coli* is more

commonly present in VD children compared with CS children (chi-square test, $p<0.001$), whereas CS children are more often colonized with *Klebsiella* spp. ($p=0.011$) and *Enterococcus* spp. ($p=0.004$) than VD children, corroborating the 16S rRNA and WGS sequencing results (Supplementary Table 6). Finally, qPCR also confirmed that colonization with *Enterococcus* spp. and *Klebsiella* spp. at 1 week of life was positively associated with more RI events in the first year of life, though this difference was only significant for *Enterococcus* spp. ($p=0.015$, Supplementary Table 7).

DISCUSSION

In this study, we were able to investigate the effect of mode of delivery on fecal microbiota development in healthy children, independent of maternal antibiotic exposure, as antibiotics given perioperatively for CS were postponed to after cord clamping. We here describe the dynamics of the fecal microbiota in the first year of life in relation to mode of delivery and assess how early-life mode of delivery-induced microbiota alterations might affect susceptibility to RIs in the first year of life.

We found substantial differences in the gut microbiota composition and stability between VD and CS children, especially in the first months of life, with notably *Bifidobacterium* being more abundant in VD children, consistent with literature[47–51]. Bifidobacteria are health-associated microbes well-known for their use as probiotics[52]. They promote gut health and provide defense against pathogens[1]. We found that in children born by CS, the colonization with *Bifidobacterium* was significantly delayed, which was not affected by feeding type. This suggests that maternal transmission during vaginal delivery is essential in acquiring these bacterial species in early life[4], which was supported by the evidence we found for fecal seeding from mother to child in the VD, but not in CS children. Perhaps therefore not solely vaginal microbiota seeding[53], but also fecal microbiota seeding during vaginal delivery is instrumental in shaping the newborn's gut microbial environment. These data suggest that only after proper initial (vaginal-)fecal seeding takes place, the growth of beneficial groups such as *Bifidobacterium* can be promoted, which can be further enhanced through the prebiotic oligosaccharides present in breast milk[54]. Hence, the stimulation of breastfeeding in women that have delivered by CS, or the advances in prebiotic formulations of modern formula milk, might not be able to correct the lack of *Bifidobacterium* seeding during delivery, as we saw that breastfed CS infants carried these bacteria in lower abundance than

formula fed VD infants. Our data also suggest that this lack of *Bifidobacterium* is not the consequence of antibiotic exposure intrapartum, but merely a consequence of delivery mode itself, which was further supported by the sub-analysis on formula fed infants, limiting the likelihood of exposure to maternal antibiotics through breastmilk. Importantly, the delay in *Bifidobacterium* establishment may have a major impact on the infants' early life and future health, as the window of opportunity for immune priming occurs within the first 1000 days of life[55].

While *Bifidobacterium* was abundantly present in the fecal samples of VD children, the potential pathogenic and proinflammatory *Klebsiella* and *Enterococcus* were more abundant in children born by CS, which is in accordance with previous studies[50,56]. For *Klebsiella,* the difference in abundance lasted for more than 4 months, long after the initial neonatal period. Recently, an increased *Klebsiella*/*Bifidobacterium* ratio in early life was correlated with later development of pediatric allergy[57]. This could be an important link between mode of delivery and increased prevalence of pediatric allergies following CS birth. In addition, bacteria from the *Klebsiella* genus are a common cause of nosocomial infections and act as a reservoir for a diverse scale of antimicrobial resistance genes[58,59]. It is highly likely that the *Klebsiella* bacteria found in our study population were acquired from the hospital (operating room) environment, and in the absence of a stable environment, thrived in the gut of CS infants. These findings might thus imply that mode of delivery not only increases the risk for immunological disorders, but also the risk for (antibiotic resistant) infections.

Although our study was not powered to investigate differences in overall health outcomes between VD and CS children, we still found a trend towards more RI events and a higher need for antibiotics in the first year of life in CS compared with VD children. This early timeframe is clinically important, as early onset of RI is considered a risk factor for recurrent infections[60,61]. As CS incidence is specifically rising in Latin American countries[7], in future studies it would be of interest to assess socio-economic factors as a link between mode of delivery, maternal microbiota characteristics, choice of feeding type and early-life risk of infection.

Uniquely to our study, we were able to relate the fecal microbiota composition at a very early age (one week), where the microbiota differences between mode of delivery groups were largest, to RI events occurring in the first year of life. Our stratified analyses per delivery mode group showed similar results to the overall cohort analysis, with especially *Bifidobacterium, Klebsiella* and *Enterococcus* being associated with RI events independent of mode of delivery, though effect

sizes were larger in the CS group. One reason for this could be a mediating effect of breastfeeding with mode of delivery, since breastfeeding is less common in CS children, and has a known protective and independent effect against infectious diseases[62].

Although these results require validation in preferentially larger cohorts, our findings do suggest that early-life gut microbiota composition might play a role in systemic (immune-mediated) resistance against infectious diseases. We found that especially early-life presence and abundance of *Klebsiella* and *Enterococcus* species, belonging to the ESKAPE pathogen family[63], show a relation with the development of a higher incidence of RI events later in life, whereas *Bifidobacterium* and certain *Bacteroides* species might play a protective role. Potential mechanisms by which bifidobacteria may protect against pathogens and resulting infections is through increasing the local pH and indirectly through increasing the short-chain fatty acid abundance in the gut, promoting gut health[1].

Strengths of our study include high sampling frequency, the broad scale of clinical and epidemiological data collected consistently by trained research nurses, the quality of the sequencing data enabling us to maintain a strict filtering threshold, and the longitudinal character of our study combined with the use of differential abundance testing, providing us with enough statistical power to discern differences in both microbial succession patterns and disease parameters between delivery mode groups.

The most important limitation of our study is the limited number of samples from the postpartum timepoint that elicited sufficient amounts of DNA for characterization of the microbiota, though, inherent to the low dense colonization in neonates in general, this is unlikely to have introduced confounding. Also, sequencing of the V4 hypervariable region of the 16S rRNA gene does not allow for confident reporting of results on a lower taxonomical level than genus level, though this was in part resolved by WGS sequencing validation of a subset of 20 samples. Finally, our observational study was not primarily designed to investigate health differences between delivery mode groups, therefore the power to research correlations between drivers, biomarkers and health consequences was limited.

In conclusion, we here report on modest differences in health characteristics between delivery mode groups, where children born by CS show a tendency toward higher incidence of RI events in early life, as well as a trend in higher need for antibiotics than VD children, with the former being linked to differences in abundance of

several biomarker bacteria. In CS delivered children, the gut microbiota appears less stable with the acquisition of *Bifidobacterium* being delayed when compared with VD children. This delay is independent of feeding type, suggesting that maternal transmission during vaginal delivery is essential in acquiring these bacterial species in early life, which is further supported by the evidence for fecal seeding in the VD children, but not in CS children in our study. The abundance of potential pathogens from the genera *Klebsiella* and *Enterococcus* is higher in children born by CS, and independent of prenatal antibiotic exposure, duration of hospitalization and feeding type. These taxa are also associated with a higher incidence of RI events in the first year of life. These findings provide evidence for a possible link between mode of delivery-induced alterations in the infant gut microbiota and susceptibility to (infectious) diseases.

ACKNOWLEDGEMENTS

The authors are most grateful for the participation of all the children and their families. We would like to acknowledge all the members of the research team of the Spaarne Gasthuis Academy, the Department of Obstetrics and Gynecology of the Spaarne Gasthuis (Hoofddorp and Haarlem, the Netherlands) and participating midwifery clinics. We would also like to acknowledge Edinburgh Genomics for executing the WGS sequencing and Katherine Emelianova for the bioinformatic analysis of the shotgun sequences.

REFERENCES

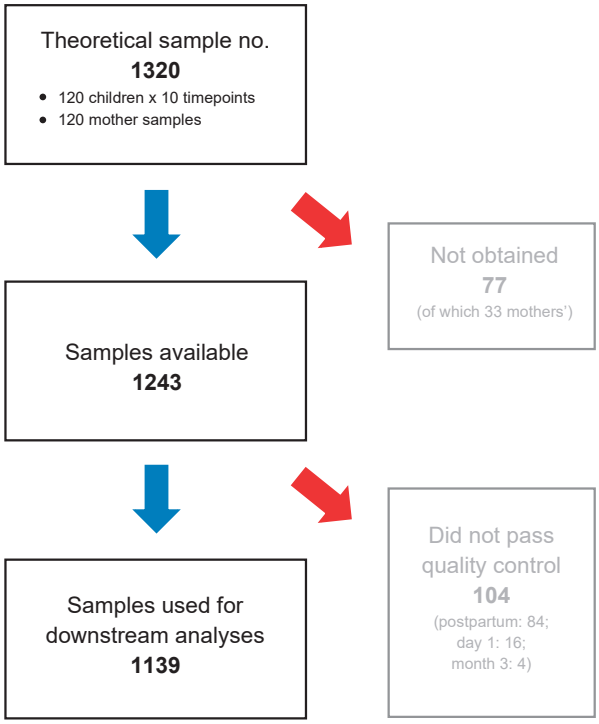
1. Tamburini, S., Shen, N., Wu, H. C. & Clemente, J. C. The microbiome in early life: implications for health outcomes. *Nat. Med.* **22**, 713–722 (2016).
2. Savage, D. C. Microbial ecology of the gastrointestinal tract. *Annu. Rev. Microbiol.* **31**, 107–33 (1977).
3. Gensollen, T., Iyer, S. S., Kasper, D. L. & Blumberg, R. S. How colonization by microbiota in early life shapes the immune system. *Science.* **352**, 539–544 (2016).
4. 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.* **107**, 11971–5 (2010).
5. Sevelsted, A., Stokholm, J., Bønnelykke, K. & Bisgaard, H. Cesarean section and chronic immune disorders. *Pediatrics* **135**, e92-8 (2015).
6. Mueller, N. T. *et al.* Prenatal exposure to antibiotics, cesarean section and risk of childhood obesity. *Int. J. Obes. (Lond).* **39**, 665–70 (2015).
7. Betrán, A. P. *et al.* The Increasing Trend in Caesarean Section Rates: Global, Regional and National Estimates: 1990-2014. *PLoS One* **11**, e0148343 (2016).
8. 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**: 314-26.
9. Dewey, K. G., Nommsen-Rivers, L. A., Heinig, M. J. & Cohen, R. J. Risk factors for suboptimal infant breastfeeding behavior, delayed onset of lactation, and excess neonatal weight loss. *Pediatrics* **112**, 607–19 (2003).
10. Azad, M. B. *et al.* Impact of maternal intrapartum antibiotics, method of birth and breastfeeding on gut microbiota during the first year of life: a prospective cohort study. *BJOG* **123**, 983–93 (2016).
11. Bosch, A. A. T. M. *et al.* Development of Upper Respiratory Tract Microbiota in Infancy is Affected by Mode of Delivery. *EBioMedicine* **9**, 336–345 (2016).
12. Bogaert, D. *et al.* Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. *PLoS One* **6**, e17035 (2011).
13. Mattiello, F. *et al.* A web application for sample size and power calculation in case-control microbiome studies. *Bioinformatics* **32**, 2038–2040 (2016).
14. Biesbroek, G. *et al.* Deep sequencing analyses of low density microbial communities: working at the boundary of accurate microbiota detection. *PLoS One* **7**, e32942 (2012).
15. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–20 (2013).
16. 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.* 4516–22 (2011). doi:10.1073/pnas.1000080107
17. Joshi, N. A. & Fass, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). (2011). Available at: https://github.com/najoshi/sickle.

18. Nikolenko, S. I., Korobeynikov, A. I. & Alekseyev, M. A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* **14**, S7 (2013).
19. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* **13**, 31 (2012).
20. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–6 (2010).
21. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
22. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
23. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
24. Subramanian, S. *et al.* Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**, 417–21 (2014).
25. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
26. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
27. Segata, N. *et al.* Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* **9**, 811–814 (2012).
28. Broad Institute. Picard Tools. *Broad Institute, GitHub Repos.* https://broadinstitute.github.io/picard/ (2018)
29. R Core Team. R: A language and environment for statistical computing. https://www.r-project.org/. (2017).
30. RStudio Team. RStudio: Integrated Development for R. http://www.rstudio.com/. (2016).
31. Wickham, B. *ggplot2: Elegant Graphics for Data Analysis.* (Springer-Verlag New York, 2009).
32. Kassambara, A. ggpubr: 'ggplot2' Based Publication Ready Plots. https://CRAN.R-project.org/package=ggpubr. (2018).
33. Benjamini, Y. & Hochberg, Y. Controlling The False Discovery Rate–A Practical And Powerful Approach To Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
34. Terry, M. & Therneau, M. Title Survival Analysis. https://CRAN.R-project.org/package=survival (2017).
35. Kassambara, A. & Kosinski, M. survminer: Drawing Survival Curves using 'ggplot2'. https://CRAN.R-project.org/package=survminer (2018).
36. Oksanen, J. *et al.* Community Ecology Package. https://CRAN.R-project.org/package=vegan. (2017).
37. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team. {nlme}: Linear and Nonlinear Mixed Effects Models. https://CRAN.R-project.org/package=nlme (2017).

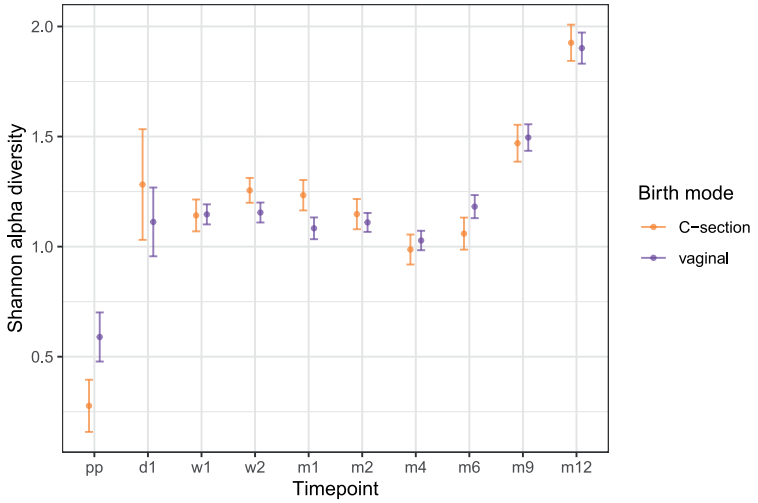
38. Paulson, J. N., Stine, O. C., Bravo, H. C. & Pop, M. metagenomeSeq: Statistical analysis for sparse high-throughput sequencing. http://cbcb.umd.edu/software/metagenomeSeq. (2016).
39. Paulson, J. N., Stine, O. C., Bravo, H. C. & Pop, M. Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* **10**, 1200–2 (2013).
40. Bosch, A. A. T. M. *et al.* Maturation of the Infant Respiratory Microbiota, Environmental Drivers, and Health Consequences. A Prospective Cohort Study. *Am. J. Respir. Crit. Care Med.* **196**, 1582–1590 (2017).
41. Man, W. H. *et al.* Loss of microbial topography between oral and nasopharyngeal microbiota and development of respiratory infections early in life. *Am. J. Respir. Crit. Care Med.* **200**, 760–770 (2019).
42. Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R News* **2**, 18–22 (2002).
43. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–7 (2012).
44. Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **8**, 343ra82-343ra82 (2016).
45. Macfarlane, S. & Macfarlane, G. T. Regulation of short-chain fatty acid production. *Proc. Nutr. Soc.* **62**, 67–72 (2003).
46. Baumfeld, Y. *et al.* Elective cesarean delivery at term and the long-term risk for respiratory morbidity of the offspring. *Eur. J. Pediatr.* **177**, 1653–1659 (2018).
47. Bäckhed, F. *et al.* Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**, 690–703 (2015).
48. Penders, J. *et al.* Factors Influencing the Composition of the Intestinal Microbiota in Early Infancy. *Pediatrics* **118**, 511–521 (2006).
49. Makino, H. *et al.* Mother-to-Infant Transmission of Intestinal Bifidobacterial Strains Has an Impact on the Early Development of Vaginally Delivered Infant's Microbiota. *PLoS One* **8**, e78331 (2013).
50. Adlerberth, I. *et al.* Reduced enterobacterial and increased staphylococcal colonization of the infantile bowel: an effect of hygienic lifestyle? *Pediatr. Res.* **59**, 96–101 (2006).
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.* **8**, 738 (2017).
52. Allen, S. J., Martinez, E. G., Gregorio, G. V & Dans, L. F. Probiotics for treating acute infectious diarrhoea. *Cochrane database Syst. Rev.* CD003048 (2010). doi:10.1002/14651858.CD003048.pub3
53. Dominguez-Bello, M. G. *et al.* Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. *Nat. Med.* **22**, 250–3 (2016).
54. Zivkovic, A. M., German, J. B., Lebrilla, C. B. & Mills, D. A. Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl 1**, 4653–8 (2011).
55. Arrieta, M.-C. *et al.* Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci. Transl. Med.* **7**, 307ra152-307ra152 (2015).
56. Stokholm, J. *et al.* Cesarean section changes neonatal gut colonization. *J. Allergy Clin. Immunol.* **138**, 881-889.e2 (2016).

57. Low, J. S. Y. *et al.* Ratio of Klebsiella/Bifidobacterium in early life correlates with later development of paediatric allergy. *Benef. Microbes* **8**, 681–695 (2017).
58. Rosenthal, V. D. *et al.* International Nosocomial Infection Control Consortium (INICC) report, data summary of 36 countries, for 2004-2009. *Am. J. Infect. Control* **40**, 396–407 (2012).
59. Gibson, M. K. *et al.* Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat. Microbiol.* **1**, 16024 (2016).
60. Kvaerner, K. J., Nafstad, P., Hagen, J. A., Mair, I. W. & Jaakkola, J. J. Recurrent acute otitis media: the significance of age at onset. *Acta Otolaryngol.* **117**, 578–84 (1997).
61. Labout, J. A. M. *et al.* Risk factors for otitis media in children with special emphasis on the role of colonization with bacterial airway pathogens: the Generation R study. *Eur. J. Epidemiol.* **26**, 61–6 (2011).
62. Duijts, L., Jaddoe, V. W. V., Hofman, A. & Moll, H. A. Prolonged and Exclusive Breastfeeding Reduces the Risk of Infectious Diseases in Infancy. *Pediatrics* **126**, e18–e25 (2010).
63. Boucher, H. W. *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **48**, 1–12 (2009).

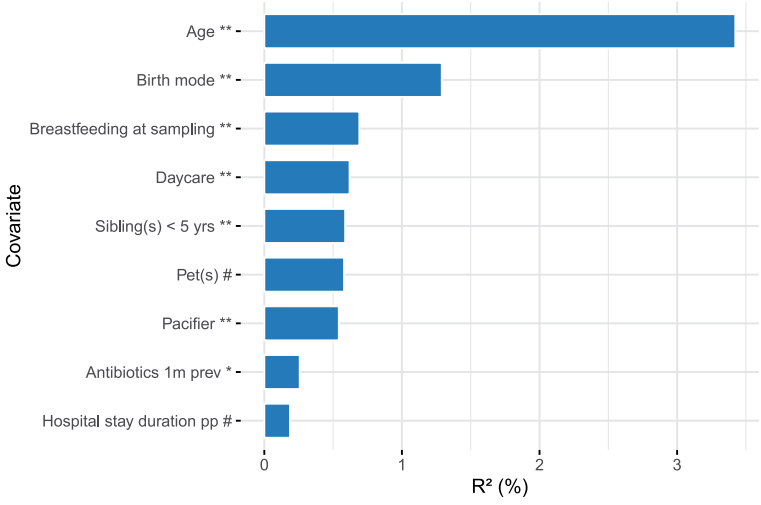
SUPPLEMENTARY INFORMATION



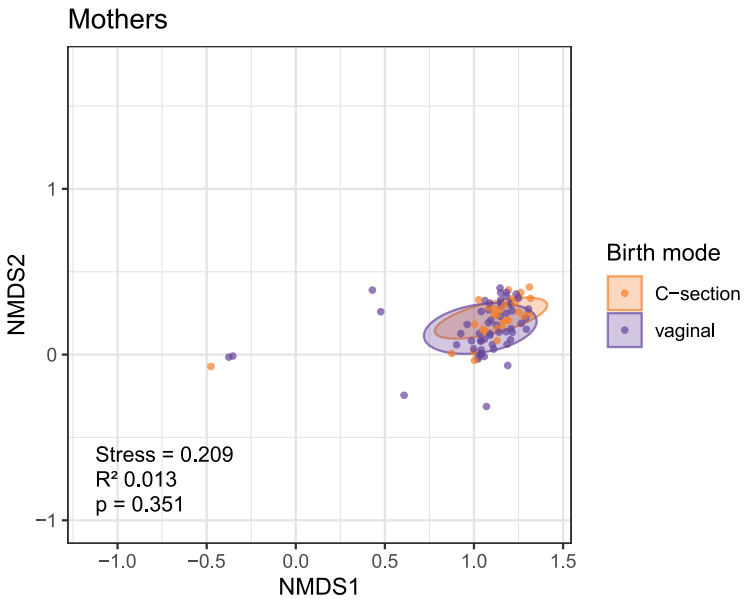
Supplementary Figure 1. Flowchart showing number of samples aimed for, obtained, and of sufficient quality for analyses



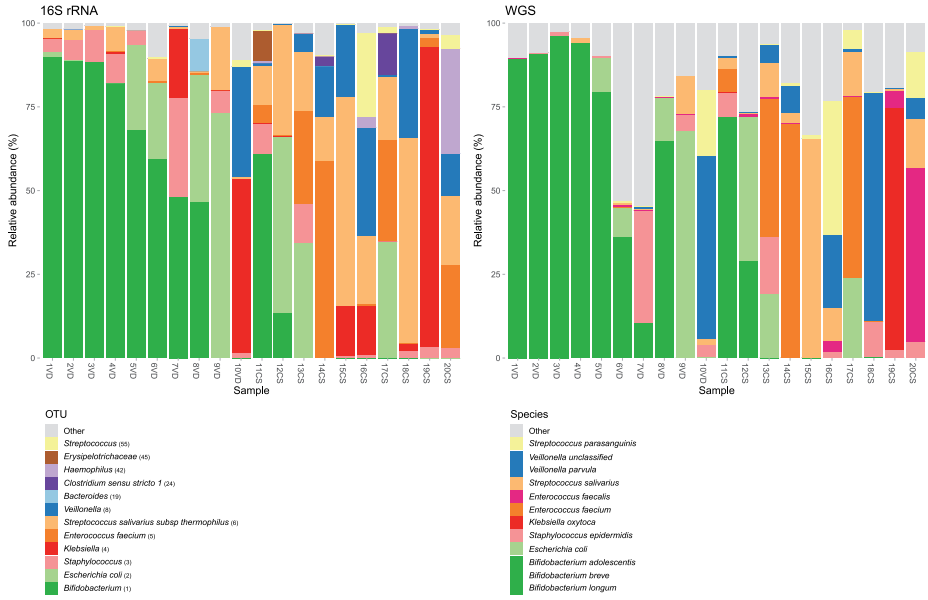
Supplementary Figure 2. Diversity indices. Differences in Shannon alpha diversity of the fecal microbiota between children born vaginally or by C-section plotted per timepoint and of the mother samples obtained two weeks after delivery. The data points represent the means per group per timepoint and the error bars represent the standard errors of the means (+/- 1SD). There were no significant differences found in alpha diversity between the two delivery mode groups at any timepoint.



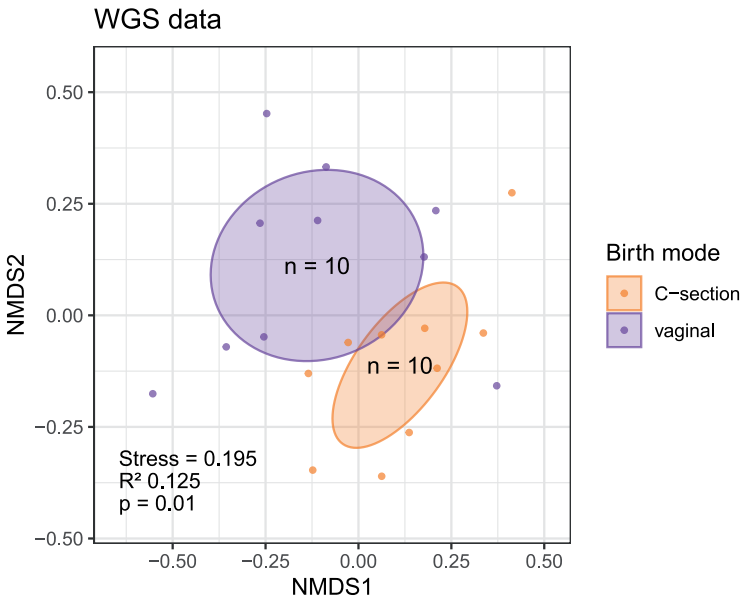
Supplementary Figure 3. Associations between baseline characteristics and fecal microbiota composition. Covariates are shown that were significantly correlated with fecal microbiota at any timepoint in univariate analyses were included in a multivariate permutational multivariate analysis of variance (PERMANOVA) test over all timepoints using 1999 permutations with the strata parameter set to participant. The percentage of variance explained (R^2 (%)) is plotted on the x-axis and # = adjusted p-value <0.1 (in both cases 0.061), * = adjusted p-value <0.05 and ** = adjusted p-value <0.01.



Supplementary Figure 4. NMDS plot of mothers' samples stratified by delivery mode. Non-metric multidimensional scaling (nMDS) plot visualizing the overall gut bacterial community composition of the mothers' samples collected two weeks after delivery. Each data point represents the microbial community composition of one sample. The ellipses represent the standard deviation of data points belonging to one of two groups: mothers that gave birth vaginally (purple; n = 57 biologically independent samples) or by C-section (orange; n = 30 biologically independent samples). The center points of the ellipses were calculated using the mean of the coordinates of each group. The stress, effect size (R^2) calculated by multivariate permutational multivariate analysis of variance (PERMANOVA) test and corresponding p-value are shown in the plot. Even though all mothers that delivered by C-section received antibiotics after the clamping of the umbilical cord, their overall but bacterial community composition 2 weeks after delivery does not differ from mothers that gave birth vaginally and did not receive post-partum antibiotics.



Supplementary Figure 5. Comparison of 16S rRNA OTUs with species found by WGS sequencing. A random subset of 20 samples were analyzed by whole genome shotgun (WGS) sequencing. The mean relative abundances of the 12 most abundant 16S rRNA OTUs were compared to the species found by WGS sequencing. Taxa approaching a similar annotation in both methods are colored similarly to allow for easier visual comparison. VD = vaginal delivery, CS = caesarean section.



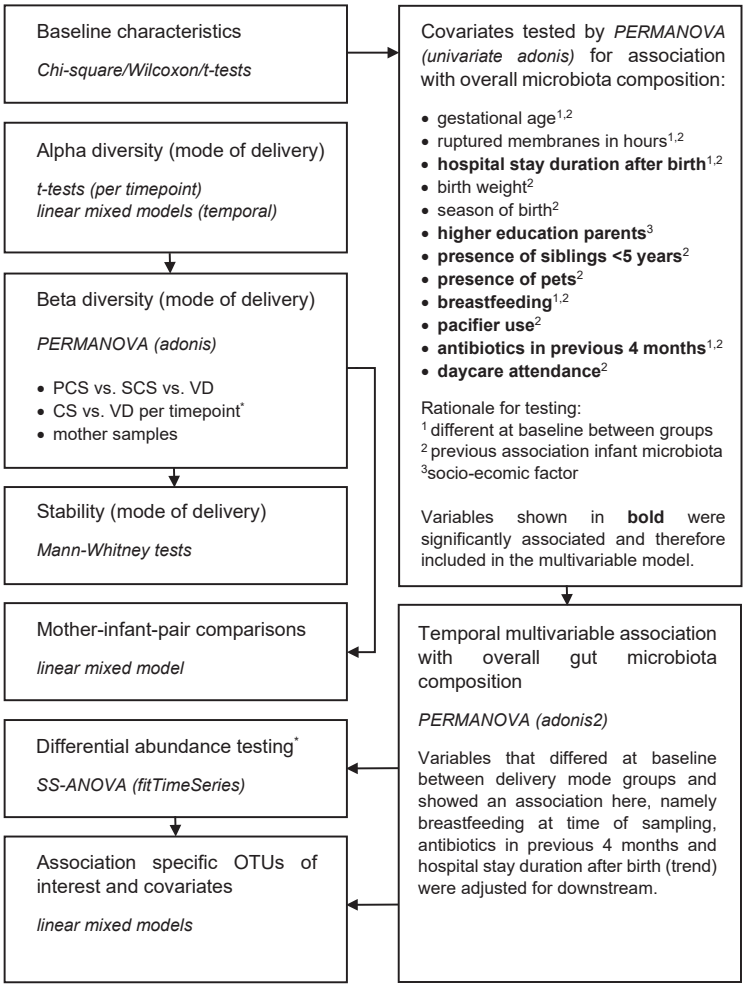
Supplementary Figure 6. NMDS plot of samples analyzed by WGS sequencing. Non-metric multidimensional scaling (nMDS) plot visualizing the differences in overall gut microbial community composition between the delivery mode groups at 1 week of age using the whole genome shotgun (WGS) sequencing data of a randomly selected subset of 20 samples (*n* represents the number of biologically independent samples per group). Each data point represents the microbial community composition of one sample. The ellipses represent the standard deviation of data points belonging to each delivery mode group, with the center points of the ellipses calculated using the mean of the coordinates per group. The stress of the ordination, number of children per group, effect size (R^2) calculated by multivariate permutational multivariate analysis of variance (PERMANOVA) test and corresponding *p*-value are printed in the plot.

Supplementary Figure 7. Statistical analysis scheme

Flow in data analyses (with type of test(s) in italic) to address the following research questions:

Primary research question

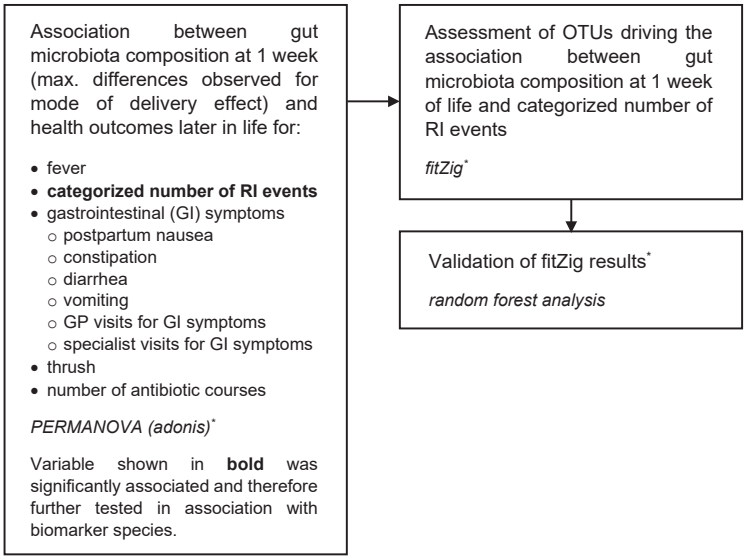
Are there differences in gut microbiota development between delivery mode groups?



* Post hoc testing of differences in beta diversity and differential abundance testing of taxa for subset of children receiving exclusive formula feeding to exclude potential confounding of antibiotic transmission through breastfeeding on delivery-mode related microbiota findings. PCS = primary caesarean section (CS), SCS = secondary CS, VD = vaginally delivered, OTU = operational taxonomical unit

Secondary research question

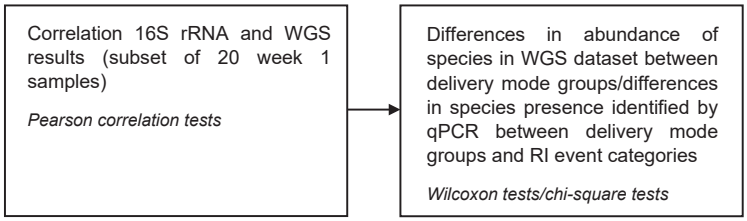
Is gut microbiota composition in early life associated with health?



* Post-hoc analyses on stratified data per delivery mode group to exclude potential confounding effect of delivery mode group on outcome. RI = respiratory infection

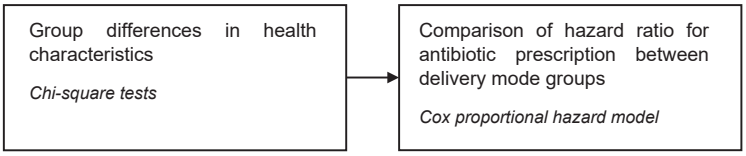
Exploratory research questions

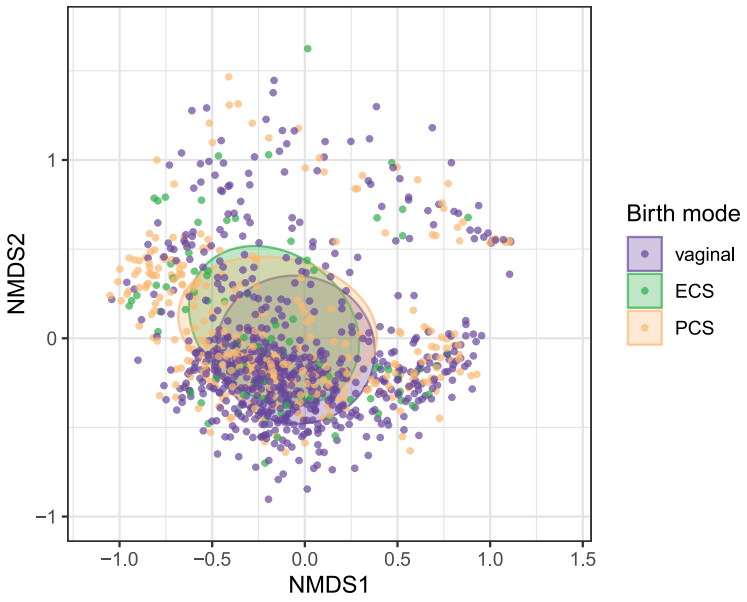
Can 16S rRNA sequencing taxonomical annotation be validated with whole genome shotgun (WGS) sequencing and can the differences in biomarkers between delivery mode groups found by 16S rRNA be confirmed by WGS and qPCR?



Exploratory (post hoc) analyses

Is delivery mode associated with health characteristics?





Supplementary Figure 8. NMDS plot of all samples colored by specific delivery mode. Non-metric multidimensional scaling (nMDS) plot visualizing the overall gut bacterial community composition of all participants' samples colored by delivery mode, specified as vaginal, emergency CS (ECS) or planned CS (PCS) delivery. Each data point represents the microbial community composition of one sample. The ellipses represent the standard deviation of data points belonging to each group. The overall gut bacterial community composition of children born by ECS (n = 10) was more similar to that of children born by PCS (permutational multivariate analysis of variance [PERMANOVA] test, R^2 = 0.005, p = 0.051) than by VD (R^2 = 0.006, p = 0.002), hence the ECS and PCS children were grouped together for delivery mode comparisons.

Supplementary Table 1. Decrease in BC dissimilarity between children's and mothers' samples over time

Timepoint	median BC dissimilarity	IQR25	IQR75
Postpartum	0.999	0.996	1.000
Day 1	0.986	0.903	0.998
Week 1	0.982	0.922	0.996
Week 2	0.971	0.895	0.993
Month 1	0.954	0.864	0.985
Month 2	0.935	0.845	0.978
Month 4	0.923	0.837	0.971
Month 6	0.904	0.807	0.966
Month 9	0.834	0.725	0.924
Month 12	0.739	0.633	0.839

Overview of groupwise Bray-Curtis (BC) dissimilarity indices between children's and mothers' samples per timepoint. IQR = interquartile range.

Supplementary Table 2. Effect of delivery mode on overall gut microbiota composition in exclusively formula fed children

Timepoint	R^2	Adjusted *p*-value
Day 1	0.082	0.441
Week 1	0.215	0.008
Week 2	0.152	0.044
Month 1	0.038	1.000
Month 2	0.029	1.000

Effect sizes (R^2) calculated by multivariate permutational multivariate analysis of variance (PERMANOVA)-tests and corresponding adjusted *p*-values are shown for the effect of delivery mode on overall gut microbiota composition of the exclusively formula fed subset of participants. The results for the post-partum timepoint are not shown because there was only one sample available per group. The period until 2 months was studied as this was where the biggest differences in microbiota were found for the overall cohort.

Supplementary Table 3. FitTimeSeries results of differentially abundant taxa between delivery mode groups

	OTU	Interval no.	Interval start	Interval end	Area	Adjusted *p*-value
1	*Bifidobacterium* (1)	interval:1	1	30	112,712	0,003287
2	*Escherichia coli* (2)	interval:1	0	85	167,6269	0,003287
3	*Klebsiella* (4)	interval:1	0	139	-276,087	0,003287
4	*Enterococcus faecium* (5)	interval:1	7	35	-65,1366	0,003287
5	*Enterococcus faecium* (5)	interval:2	269	361	133,5926	0,016636
6	*Enterococcus faecium* (5)	interval:3	362	368	10,49423	0,003287
7	*Streptococcus salivarius subsp thermophilus* (6)	interval:2	0	25	-21,0815	0,011449
8	*Veillonella* (8)	interval:1	0	73	-100,675	0,003287
9	*Bifidobacterium bifidum* (11)	interval:1	0	68	79,26634	0,003287
10	*ratAN060301C* (12)	interval:1	0	180	191,876	0,003287
11	*Pseudobutyrivibrio* (13)	interval:1	360	379	23,58741	0,045462
12	*Bifidobacterium dentium* (14)	interval:1	0	1	0,81083	0,003287
13	*Bifidobacterium dentium* (14)	interval:2	43	74	-28,6032	0,009705
14	*Bifidobacterium dentium* (14)	interval:3	247	282	40,26375	0,003287
15	*Collinsella* (16)	interval:1	0	170	125,4543	0,00794
16	*Faecalibacterium* (17)	interval:1	0	10	-8,48254	0,04458
17	*Faecalibacterium* (17)	interval:2	203	354	-252,214	0,039765
18	*Faecalibacterium* (17)	interval:3	355	378	-34,3762	0,03537
19	*Streptococcus gallolyticus subsp macedonicus* (18)	interval:1	0	353	520,936	0,003287
20	*Bacteroides* (19)	interval:1	0	303	683,5567	0,003287
21	*Clostridium sensu stricto 1* (21)	interval:1	0	191	-204,811	0,003287
22	*Lactobacillus* (23)	interval:1	0	184	-106,637	0,045462
23	*Clostridium sensu stricto 1* (24)	interval:1	0	146	-107,426	0,003287
24	*Veillonella* (27)	interval:1	0	385	458,6269	0,003287
25	*Clostridium butyricum* (29)	interval:1	231	385	92,90985	0,003287
26	*Lachnospiraceae* (30)	interval:1	0	103	-43,8028	0,009705
27	*Bacteroides* (35)	interval:1	0	374	513,9418	0,003287
28	*Peptostreptococcaceae* (36)	interval:1	0	308	-208,9	0,003287
29	*Peptostreptococcaceae* (37)	interval:1	17	211	-62,1764	0,02857
30	*Eubacterium hallii* (39)	interval:1	365	385	50,32886	0,047501
31	*Lactobacillus acidophilus* (40)	interval:1	0	111	40,92728	0,003287
32	*Bacteroides* (48)	interval:1	29	272	179,8846	0,003287
33	*Bacteroides* (53)	interval:1	0	241	97,51382	0,003287
34	*Dorea* (54)	interval:1	42	263	-144,523	0,003287

Continue

Continued

	OTU	Interval no.	Interval start	Interval end	Area	Adjusted *p*-value
35	*Streptococcus* (55)	interval:1	0	57	-43,929	0,003287
36	*Streptococcus* (55)	interval:2	169	259	50,76685	0,005661
37	*Proteus mirabilis* (63)	interval:1	0	59	10,77698	0,003287
38	*Bacteroides* (65)	interval:1	8	317	184,1286	0,003287
39	*Lachnospiraceae* (68)	interval:1	248	385	35,62635	0,016636
40	*Streptococcus* (69)	interval:1	0	77	-35,2993	0,021978
41	*Lactobacillus fermentum* (75)	interval:1	68	385	171,5	0,003287
42	*Ruminococcaceae* (76)	interval:1	268	324	-60,6412	0,015506
43	*Ruminococcaceae* (76)	interval:2	362	385	39,14089	0,003287
44	*Roseburia* (77)	interval:1	253	280	-27,9272	0,003287
45	*Roseburia* (77)	interval:2	325	363	53,83287	0,031899
46	*Coprococcus* (92)	interval:1	365	385	19,34755	0,029643
47	*Clostridium sensu stricto* (95)	interval:1	58	385	-75,2924	0,02857
48	*Lachnospiraceae* (101)	interval:1	328	385	-25,165	0,046671
49	*Eubacterium desmolans* (106)	interval:1	214	385	-121,624	0,003287
50	*Rothia* (113)	interval:1	0	59	-40,3232	0,003287
51	*Rothia* (113)	interval:2	165	269	48,55794	0,003287
52	*Megasphaera* (122)	interval:1	14	42	-7,17491	0,00794
53	*Megasphaera* (122)	interval:2	223	360	75,28518	0,03675
54	*bacterium mpn isolate group 25* (124)	interval:1	370	385	15,47146	0,032815
55	*Lachnospiraceae* (127)	interval:1	29	274	-50,5619	0,03537
56	*Roseburia* (133)	interval:1	276	358	-98,0038	0,023743
57	*Roseburia* (133)	interval:2	367	385	25,386	0,011449
58	*Marvinbryantia* (137)	interval:1	18	26	1,600834	0,003287
59	*Ruminococcaceae* (138)	interval:1	15	24	2,685455	0,003287
60	*Bifidobacterium* (147)	interval:1	0	44	14,10129	0,003287
61	*Bifidobacterium* (147)	interval:2	282	385	-40,6182	0,011449
62	*Subdoligranulum* (148)	interval:1	177	385	-83,4036	0,03066
63	*Lachnospiraceae* (153)	interval:1	139	385	86,92777	0,021978
64	*Lachnospiraceae* (155)	interval:3	288	356	-74,418	0,042053
65	*Varibaculum* (161)	interval:1	10	59	-17,5114	0,005661
66	*Varibaculum* (161)	interval:2	118	154	13,81666	0,029643
67	*Blautia* (169)	interval:1	14	25	4,242241	0,003287
68	*Parasutterella* (171)	interval:1	75	385	102,9637	0,00794
69	*Bifidobacterium animalis* (175)	interval:1	0	153	-27,2805	0,00794
70	*Anaerosporobacter* (181)	interval:1	368	385	15,52433	0,03066

Continue

Continued

	OTU	Interval no.	Interval start	Interval end	Area	Adjusted *p*-value
71	*Bifidobacterium* (182)	interval:1	42	195	-36,15	0,009705
72	*Escherichia Shigella* (185)	interval:1	0	41	13,9863	0,003287
73	*Sutterella* (187)	interval:1	49	259	37,64495	0,025232
74	*Streptococcus* (189)	interval:1	0	33	-8,99449	0,020585
75	*Streptococcus* (189)	interval:2	181	277	25,44483	0,003287
76	*Bifidobacteriaceae* (198)	interval:1	119	385	-90,3912	0,009705
77	*Desulfovibrio* (201)	interval:1	128	385	-48,1528	0,032815
78	*Finegoldia* (203)	interval:1	0	159	-32,6427	0,003287
79	*Bifidobacterium dentium* (205)	interval:1	120	385	-66,4652	0,016636
80	*Collinsella* (211)	interval:1	18	247	72,1091	0,049817
81	*Collinsella* (215)	interval:1	0	385	138,1499	0,003287
82	*Lachnospiraceae* (221)	interval:1	0	5	-1,23682	0,04458
83	*Lachnospiraceae* (221)	interval:2	268	271	0,732577	0,005661
84	*Lachnospiraceae* (221)	interval:3	323	361	-23,455	0,040759
85	*Lachnospiraceae* (221)	interval:4	376	385	6,209622	0,031899
86	*Bacteroides* (245)	interval:1	183	385	-62,2234	0,009705
87	*Bifidobacterium* (246)	interval:1	0	173	45,25739	0,003287
88	*Enterococcaceae* (251)	interval:1	0	41	-10,8802	0,013437
89	*Klebsiella* (252)	interval:1	0	113	-44,4969	0,003287
90	*Enterococcus* (256)	interval:1	0	70	-22,7148	0,003287
91	*Bifidobacterium breve* (257)	interval:1	0	369	205,9727	0,003287
92	*Bifidobacterium breve* (261)	interval:1	0	234	46,40682	0,011449
93	*Negativicoccus* (263)	interval:1	0	175	-32,338	0,003287
94	*Streptococcus* (268)	interval:1	0	385	143,3426	0,003287
95	*Anaerostipes* (271)	interval:1	0	2	0,232903	0,003287
96	*Veillonella* (275)	interval:1	0	46	-5,37729	0,047501
97	*Eubacterium hallii* (278)	interval:1	271	348	-39,4095	0,005661
98	*Eubacterium hallii* (278)	interval:2	365	385	14,2956	0,003287
99	*Bacteroides* (279)	interval:1	17	199	46,71601	0,003287
100	*Citrobacter sedlakii* (288)	interval:1	0	141	-18,3047	0,005661
101	*bacterium NLAE zl C423* (298)	interval:1	0	183	21,71035	0,003287
102	*Bifidobacteriaceae* (299)	interval:1	0	49	6,191564	0,003287
103	*bacterium NLAE zl C350* (303)	interval:1	0	39	5,321786	0,003287
104	*Bifidobacteriaceae* (309)	interval:1	346	385	9,008275	0,003287
105	*Bacteroides* (310)	interval:1	49	385	80,96046	0,016636
106	*Bifidobacterium* (320)	interval:1	0	95	10,17972	0,003287
107	*Bacteroides* (326)	interval:1	143	239	14,71285	0,016636

Continue

Continued

	OTU	Interval no.	Interval start	Interval end	Area	Adjusted *p*-value
108	*Bacteroides fragilis* (327)	interval:1	0	374	149,0278	0,003287
109	*Coprobacillus* (331)	interval:1	43	167	-21,4177	0,005661
110	*Gardnerella* (333)	interval:1	105	385	52,21995	0,009705
111	*Peptostreptococcaceae* (335)	interval:1	65	223	-17,8886	0,040759
112	*Leuconostoc* (341)	interval:1	29	287	36,92326	0,003287
113	*Pseudobutyrivibrio* (349)	interval:1	91	385	45,94751	0,005661
114	*Veillonella* (366)	interval:1	142	168	-6,0963	0,034251
115	*Pasteurellaceae* (385)	interval:1	0	50	-5,49709	0,021978
116	*Alistipes* (388)	interval:1	120	385	-51,374	0,011449
117	*Aggregatibacter* (390)	interval:1	0	127	-15,8943	0,005661
118	*Odoribacter* (398)	interval:1	323	360	-21,2673	0,005661
119	*Odoribacter* (398)	interval:3	371	385	8,259985	0,003287
120	*Streptococcus gallolyticus subsp macedonicus* (399)	interval:1	0	310	71,0579	0,003287
121	*Bifidobacteriaceae* (430)	interval:1	34	152	-27,4874	0,009705
122	*Bacteroides* (433)	interval:1	0	129	12,50547	0,013437
123	*Bifidobacterium* (435)	interval:1	0	84	14,61852	0,003287
124	*bacterium NLAE zl G195* (451)	interval:1	0	192	23,93176	0,003287
125	*Actinomyces sp oral clone DR002* (463)	interval:1	22	188	-28,9075	0,016636
126	*Lactobacillus* (478)	interval:1	19	168	-30,5026	0,029643
127	*Bacteroides* (480)	interval:1	39	234	22,8699	0,00794
128	*Propionibacterium* (483)	interval:1	54	148	-8,69648	0,03066
129	*Anaerostipes* (488)	interval:1	238	301	-23,7484	0,025232
130	*Bacteroides* (491)	interval:1	0	287	38,98332	0,003287
131	*Lachnospiraceae* (496)	interval:1	116	202	-14,4411	0,032815
132	*Peptostreptococcaceae* (499)	interval:1	49	295	19,37888	0,005661
133	*Streptococcus* (502)	interval:1	0	85	-11,6154	0,003287
134	*Veillonella* (518)	interval:2	1	133	-18,0585	0,003287
135	*Staphylococcus* (530)	interval:1	0	24	-2,91257	0,048668

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in OTUs existed between the delivery mode groups, adjusted for duration of hospital stay after birth, breastfeeding at time of sampling and antibiotic use in the 4 weeks prior to sampling. A positive area value indicates that the abundance of a specific OTU is higher in the VD group, while a negative area value indicates that the abundance of an OTU is higher in the CS group. To correct for multiple testing, the Benjamini-Hochberg method was applied and only results with an adjusted *p*-values of < 0.05 are shown.

Supplementary Table 4. Random forest validation of biomarker species found at 1 week of life associated with RI events

OTU	Mean decrease in Gini coefficient
Bifidobacterium (1)	1.372
Enterococcus faecium (5)	1.300
Klebsiella (4)	1.190
Staphylococcus (3)	1.143
Streptococcus (55)	1.014
Haemophilus (42)	1.012
Streptococcus salivarius subsp thermophilus (6)	0.980
Streptococcus (189)	0.890
Bifidobacterium (435)	0.866
Rothia (113)	0.810

Random forest analysis was performed on the samples collected at 1 week of life to validate the biomarker species found with fitZig analysis to be associated with respiratory infection (RI) events later in life. The categorized RI events (0-2 or 3-7) were set as outcome and the relative abundance of OTUs as predictors, together with delivery mode, duration of hospital stay after birth, feeding type and antibiotics in the 4 weeks prior to sampling. Biomarkers most discriminative of outcome are shown. Again, *Bifidobacterium*, *Enterococcus* and *Klebsiella* were found to be important taxa in the association between microbiota composition at 1 week of life and the number of RI events in the first year.

Supplementary Table 5. Correlation between top five most abundant 16S rRNA OTUs and WGS species

16S rRNA OTUs	WGS species	Pearson's r	Adjusted *p*-value
Bifidobacterium (1)	*Bifidobacterium longum* *Bifidobacterium breve* *Bifidobacterium adolescentis*	0.95	<0.001
Escherichia coli (2)	*Escherichia coli*	0.95	<0.001
Staphylococcus (3)	*Staphylococcus epidermidis*	0.86	<0.001
Klebsiella (4)	*Klebsiella oxytoca*	0.83	<0.001
Enterococcus faecium (5)	*Enterococcus faecium*	0.92	<0.001

The top five most abundant OTUs of the 16S rRNA dataset were correlated with the corresponding whole genome shotgun (WGS) sequencing species using Pearson correlations. The relative abundances of the three most abundant *Bifidobacterium* species of the WGS dataset were combined. An adjusted *p*-value of < 0.05 stands for a significant correlation (not a significant difference).

Supplementary Table 6. Confirmation of significant differences in *E. coli*, *Klebsiella* spp. and *Enterococcus* spp. between delivery mode groups by qPCR

qPCR positive for	Vaginal birth	C-section birth	*p*
n (%)	74	46	
E. coli (%)	53 (71.6)	16 (36.4)	*< 0.001*
Klebsiella spp. (%)	7 (9.5)	13 (29.5)	*0.011*
Enterococcus spp. (%)	40 (54.1)	36 (81.8)	*0.004*

Quantitative polymerase chain reaction (qPCR) analysis was performed for *E. coli*, *Klebsiella* spp. and *Enterococcus* spp. on all samples of infants collected at 1 week of life (*n* = 119). One sample was discarded from analysis due to an abnormal signal of its Internal Amplification Control. The qPCR results corroborated the 16S rRNA and whole genome sequencing findings: VD children more often have *E. coli* in their samples compared to CS children, whereas CS children more often have *Klebsiella* spp. and *Enterococcus* spp. present in their samples than VD children (chi-square tests).

Supplementary Table 7. Confirmation of association between *Klebsiella* spp. and *Enterococcus* spp. colonization at 1 week of life and more RI events later in life

qPCR positive for	0-2 RI events	3-7 RI events	*p*
n (%)	41	75	
Klebsiella spp. (%)	4 (9.8)	16 (21.3)	0.187
Enterococcus spp. (%)	20 (48.8)	55 (73.3)	*0.015*

Quantitative polymerase chain reaction (qPCR) analysis was performed for, amongst others, *Klebsiella* spp. and *Enterococcus* spp. on all samples of infants collected at 1 week of life (n = 119). One sample was discarded from analysis due to an abnormal signal of its Internal Amplification Control. Information on RI events in the first year of life was unavailable for two participants. Presence of *Enterococcus* spp. was associated with more respiratory infection (RI) events in the first year of life (chi-square test).

CHAPTER 3

Rectal swabs are a reliable proxy for faecal samples in infant gut microbiota research based on 16S-rRNA sequencing

Marta Reyman, Marlies A. van Houten, Kayleigh Arp, Elisabeth A.M. Sanders, Debby Bogaert

Scientific Reports 2019; 9: 16072

ABSTRACT

Rectal swabs are potentially a valuable method for monitoring the gut microbiome in research and clinical settings, where it is important to adhere to strict timing, or where acute sampling is needed. It is currently unknown whether rectal swabs give comparable results to faecal samples regarding microbiota community composition in neonates and infants. To study how well the two sampling methods correlate in infants, we compared the 16S-rRNA-based sequencing results of 131 paired rectal swabs and faecal samples collected from 116 infants at two timepoints in early life. The paired samples were highly comparable regarding both diversity and overall community composition, and strongly correlated on taxonomical level. We observed no significant nor relevant contribution of sampling method to the variation in overall gut microbiota community composition in a multivariable model. Our study provides evidence supporting the use of rectal swabs as a reliable proxy for faecal samples in infant gut microbiota research.

INTRODUCTION

The interest in studying the gut microbiome in relation to health and disease is rapidly growing[1]. With continuing advances in rapid sequencing technology, monitoring of the gut microbiome in a clinical setting and conducting longitudinal microbiota studies into cause-effect relationships are becoming more feasible[2]. Currently, the routine sampling method for gut microbiota analysis is the collection of faeces[3]. However, it can be problematic to collect faeces within a narrow timeframe, as stool is not always readily available, especially in early life when the frequency of defaecation varies greatly[4]. Rectal swabs, on the other hand, can be collected easily and at any time, allowing flexible and consistent sampling between individuals and in relation to interventions. The collection of rectal swabs is already being applied for the screening of specific pathogens and multi-resistant organisms in the clinical setting[5]. Previous studies have shown that faecal samples and rectal swabs show satisfactory concordance when studying the gut microbiota in adults[6,7]. In the paediatric population, rectal swabs have been compared to faecal samples with respect to the detection of specific pathogens, such as norovirus[8], but not yet on their performance in analysing the overall microbiota composition. The infant gut microbiota is very dynamic in the first weeks of life[9,10], so it would be valuable to verify whether rectal swabs give a reliable representation of its composition in this period. If so, rectal swabs could be an ideal sampling method for the monitoring of treatment or study interventions on the neonatal ward or paediatric intensive care unit or in (longitudinal) population-based studies.

In our study, we compared the alpha and beta diversity between rectal swabs and faecal samples collected at the same time from the same individual at two sampling moments, and studied how well the sampling methods correlate on taxonomical level. The objective of our study was to determine whether rectal swabs are a good proxy for faecal samples in infant gut microbiota research.

METHODS

Study population and sample collection

To study whether rectal swabs are a good proxy for faecal samples in infant gut microbiota studies, we used a subset of 131 paired faecal and rectal swab samples from 116 children participating in the Dutch randomised controlled ZEBRA study. The ZEBRA study aims to evaluate the effects of antibiotic treatment indicated for (suspected) neonatal sepsis in the first week of life on the developing gut microbiota.

Written informed consent was obtained from both parents. Ethical approval was granted by the national ethics committee in the Netherlands, METC Noord-Holland (committee on research involving human subjects, M014-024, NTR5119). The study was conducted in accordance with the European Statements for Good Clinical Practice.

Rectal swabs were collected using FaecalSwab™ kits (Copan Diagnostics, CA, USA) by trained physicians or research personnel before the start of antibiotic treatment (timepoint 1) and 24-48 hours after cessation of antibiotic therapy (timepoint 2). Faecal samples were obtained at the same timepoints, usually directly after the rectal swab, this being a stimulatory trigger for defecation, and stored in sterile faecal containers by a nurse during hospital stay, or by the parents if the participant was already discharged at the later timepoint. All material was directly stored at -20°C before being transferred (<2 weeks) to a -80°C freezer until further laboratory processing. We only analysed paired samples that were obtained within 24 hours of one another, and in the case of timepoint 1 were also both obtained strictly before the start of antibiotic treatment.

DNA isolation and sequencing

Bacterial DNA was isolated from faecal samples as previously described[11]. We used approximately 100 µl of faeces, 300 µl of lysis buffer, 500 µl zirconium beads and 500 µl of phenol, and performed an extra phenol/chloroform step. Samples collected on day 1 were presumed to have low bacterial abundance. Therefore, further adaptations were applied as described previously[12], with the additional changes of using 150 µl instead of 100 µl of faeces (or 100 µl of material in the case of rectal swabs) and implementing an extra step with wash buffer 1. DNA blanks and a positive control consisting of a mix of up to three random faecal samples were used for quality control. The amount of bacterial DNA was determined by quantitative polymerase chain reaction (qPCR) as previously described[12].

After amplifying the V4 hypervariable region of the 16S rRNA, quantification of the amount of amplified DNA per sample was executed with the dsDNA 910 Reagent Kit on the Fragment Analyzer (Advanced Analytical, IA, USA). Samples yielding insufficient DNA after amplification, defined as <0.5 ng/µl, were repeated with a higher concentration of template DNA. A mock control and three PCR blanks were included in each PCR plate. 16S rRNA sequencing was performed on the Illumina MiSeq platform (Illumina, Eindhoven, the Netherlands).

Bioinformatic processing

The samples and their sequences described in this manuscript are part of a larger dataset existing of 2176 samples and controls, and together were processed using our in-house bioinformatics pipeline[13]. In short, we applied an adaptive, window-based trimming algorithm (Sickle, version 1.33) to filter out low quality reads, maintaining a Phred score threshold of 30 and a length threshold of 150 nucleotides[14]. Error correction was performed with BayesHammer (SPAdes genome assembler toolkit, version 3.5.0)[15]. Each set of paired-end sequence reads was assembled using PANDAseq (version 2.10) and demultiplexed (QIIME, version 1.9.1)[16,17]. Singleton and chimeric reads (UCHIME) were removed. Operational Taxonomical Unit (OTU) picking was conducted with VSEARCH abundance-based greedy clustering with a 97% identity threshold[18]. OTUs were annotated using the Naïve Bayesian RDP classifier (version 2.2) and the SILVA reference database[19,20]. This resulted in an OTU-table containing 18,951 taxa in total. We created an abundance-filtered dataset selecting OTUs present at a confident level of detection (0.1% relative abundance) in at least two samples[21], hereafter referred to as our raw OTU-table. The raw OTU-table consisted of in total 730 taxa (0.49% sequences excluded with filtering). Next, we used both the prevalence and frequency methods of the *decontam* package[22] to exclude possible contaminants, discarding 35 taxa, and thus retaining 695 taxa in total. The subset of paired samples studied here contained only 372 of these taxa.

Statistical analyses

All analyses were performed in R version 3.4.3[23] within RStudio version 1.1.383[24] and figures were made using packages ggplot2[25] and ggpubr[26]. The alpha diversity of the two sampling methods was compared using the observed species richness and Shannon diversity. When rarefying to a sequencing depth of 25,000 reads after filtering and decontamination (lowest quartile), or even 15,000 reads, the differences found using the raw data remained, so from here the raw (unrarefied) data was used for the comparisons and correlations in alpha diversity. Group differences were tested for using Wilcoxon tests. The correlation in alpha diversity between paired samples was calculated with Pearson.

The effect of sampling method on composition was analysed univariately with permutational analysis of variance (PERMANOVA)-tests with 1999 permutations (adonis function; *vegan* package[27]) for all samples, and also stratified per timepoint, to prevent confounding by repeated measures. To visualise differences in composition we generated non-metric multidimensional scaling plots (nMDS; vegan package[27]). Ordinations were based on the Bray-Curtis (BC) dissimilarity matrix of relative abundance data with parameter trymax 10,000. To test whether

paired faecal samples and rectal swabs (within child comparison) were more similar in microbiota composition than unpaired samples (between children comparison), we calculated the BC similarities (1 – BC dissimilarity) between all samples and compared the level of similarity between the paired and unpaired samples using Wilcoxon. For the paired samples, we also tested the correlation between composition and the difference in collection time and reads between the two sampling methods.

We performed a temporal, multivariable PERMANOVA-test (adonis2 function, *vegan* package[27], 1999 permutations) to test whether sampling method contributes to the variation in overall gut microbiota community composition and how this relates to other known drivers of community composition. First, we tested covariates known to be associated with gut microbiota composition (age, delivery mode, feeding type) univariately with PERMANOVA-tests as described above. Only covariates that showed a significant association in the univariate analysis (age, delivery mode), were included in a multivariable model along with sampling method, whilst setting the strata parameter to individual.

Finally, we evaluated whether the sampling methods correlated on taxonomical level by calculating the Pearson correlation coefficient for all individual taxa based on their relative abundance. We also calculated the delta mean relative abundance between the two sampling methods to show the differences found.

P-values or, where applicable, adjusted p-values calculated using the Benjamini-Hochberg method[28], <0.05 were deemed significant.

RESULTS

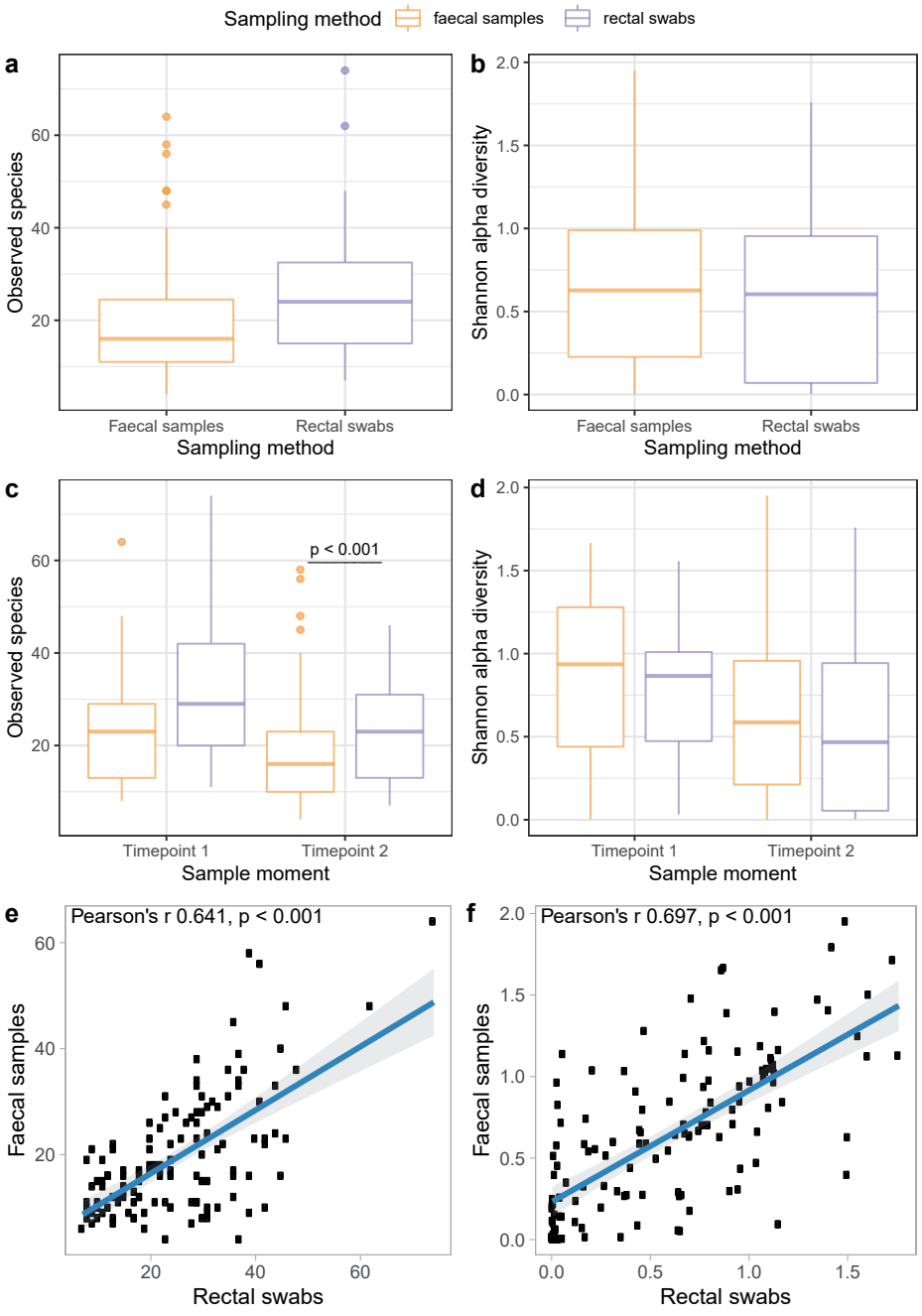
After quality filtering, there were 131 closely paired faecal and rectal swab samples available from a subset of 116 neonates with suspected early onset neonatal sepsis participating in the ZEBRA study. A closely paired sample is defined as a matched faecal sample and rectal swab collected within a timeframe of 24 hours from an individual participant. Paired samples were collected at two timepoints: before start (timepoint 1) and after cessation (timepoint 2) of antibiotic treatment. For ethical reasons the start of antibiotic treatment could not be delayed whilst waiting for a faecal sample, so per our study protocol we collected rectal swabs from all neonates before the start of antibiotic treatment, and additionally, a faecal sample if the neonate had defecated before start of treatment as well, to avoid treatment bias. The

number of times that both the faecal sample and rectal swab were collected before antibiotic administration had started was therefore restricted to 21/116 infants. At the second timepoint, we managed to obtain closely paired faecal samples and rectal swabs in 110/116 infants (median time difference between paired samples: 0 hours, range 0-24 hours). In a total of 23 cases the rectal swab was collected before the faecal sample (range 5 minutes to 24 hours), in 20 cases the rectal swab was collected after the faecal sample (range 15 minutes to 22,5 hours) and in 88 cases the rectal swab and faecal sample were collected at almost the same time, where in practice a trained physician collected the rectal swab first and a nurse collected the faeces shortly hereafter, the rectal swab being a stimulatory trigger for defaecation. The median age at sampling was 1 day for timepoint 1 and 6 days for timepoint 2. Further sample characteristics are detailed in Supplementary Table S1. Of the 116 infants included in this study, 15 had a paired sample available for both timepoint 1 and 2. Therefore, for analyses comparing sampling methods overall, we stratified per timepoint to take repeated measures into account. In order to answer our primary research question, namely if a rectal swab reliably reflects the fecal composition of an individual, we included all 131 paired samples, regardless of timepoint, for correlation analyses.

Paired faecal samples and rectal swabs are comparable in alpha and beta diversity

The 262 samples analysed in this study represented 13,026,207 high quality Illumina Miseq sequences with a median Good's coverage of 99.99% (range 99.72-100%). These sequences were annotated to 372 unique taxa. 270 taxa, representing 99.99% of all sequences, were present in both the faecal samples and rectal swabs. The combined relative abundance of the 52 taxa found only in faeces, and the 50 taxa found only in rectal swabs, was only 0.01%.

Overall, we found no differences in alpha diversity between the faecal samples and rectal swabs as measured by using observed species richness and Shannon diversity (Figure 1a and b). Since we found a correlation between alpha diversity and timepoint (Wilcoxon test $p=0.004$), we also performed the analyses stratified per timepoint (Figure 1c). At timepoint 1, we did not find a significant difference in observed species richness and Shannon diversity (Figure 1d) between the sampling methods, though at timepoint 2 species richness was significantly higher in the rectal swabs compared to the faecal samples (Wilcoxon test, median 23 [range 7-42] versus 16 [range 4-58] OTUs, $p<0.001$). When analysing the paired data, however, we found a strong correlation between paired faecal samples and rectal swabs for


Sampling method
faecal samples
rectal swabs
a
Observed species
60
40
20
Faecal samples
Rectal swabs
Sampling method
b
Shannon alpha diversity
2.0
1.5
1.0
0.5
0.0
Faecal samples
Rectal swabs
Sampling method
c
Observed species
p < 0.001
Timepoint 1
Timepoint 2
Sample moment
d
Shannon alpha diversity
Timepoint 1
Timepoint 2
Sample moment
e
Pearson's r 0.641, p < 0.001
Faecal samples
Rectal swabs
f
Pearson's r 0.697, p < 0.001
Faecal samples
Rectal swabs

< Figure 1. Alpha diversity measures. Differences are shown in observed species richness and Shannon diversity, respectively, between faecal samples and rectal swabs overall (**a**, **b**) and stratified per timepoint (**c**, **d**), to account for repeated measures of 15 participants from whom a paired sample was obtained at both timepoint 1 and 2. At timepoint 2 there was a significant difference in observed species richness between the faecal samples and rectal swabs (Wilcoxon test, median 16 versus 24, $p<0.001$). No significant difference in Shannon diversity was found between the sampling methods. Boxplots with medians are shown; the lower and upper hinges correspond to the first and third quartiles (the 25^{th} and 75^{th} percentiles); the upper and lower whiskers extend from the hinge to the largest and smallest value no further than 1.5*IQR from the hinge; outliers are plotted individually. The observed species richness (**e**) and Shannon diversity (**f**) of the faecal samples are plotted against the observed species richness and Shannon diversity of their paired rectal swabs. A strong and significant correlation was found for both alpha diversity measures between the two sampling methods (Pearson's r 0.641 and 0.697, respectively. Both $p<0.001$).

both species richness and Shannon diversity indices (Figure 1e and 1f; Pearson's r=0.641, and 0.697, respectively: both $p<0.001$).

Regarding overall community composition, the two sampling methods did not differ at either timepoint, neither in an overall analysis, nor after stratification per timepoint (PERMANOVA-test), R^2 0.006 with p=0.994 at timepoint 1 and R^2 0.002 with p=0.897 at timepoint 2; Supplementary Figure 1). Furthermore, the composition of paired samples was significantly more similar than that of unpaired samples (Figure 2; median Bray-Curtis [BC] similarity 0.866; versus median inter-individual similarity 0.007, $p<0.001$). Neither the time between the collection of a paired faecal sample and rectal swab, nor the difference in reads between the sample types, was correlated with community composition similarity as measured by BC.

Effect of sampling method compared to other clinical variables on microbial community composition

To evaluate the importance of sampling method for the overall observed variation in microbial community composition, we performed a multivariable PERMANOVA-test. First, we tested variables known to be associated with microbiota composition (age, delivery mode, feeding type) univariately and only included variables that showed a significant association in our study (age, delivery mode) in the multivariable model, along with sampling method. The contribution of sampling method to the variation in community composition was minimal and not significant (Figure 3; R^2 0.002 p=0.835), while participant's age was the most important explanatory variable (R^2 0.016, p=0.002).

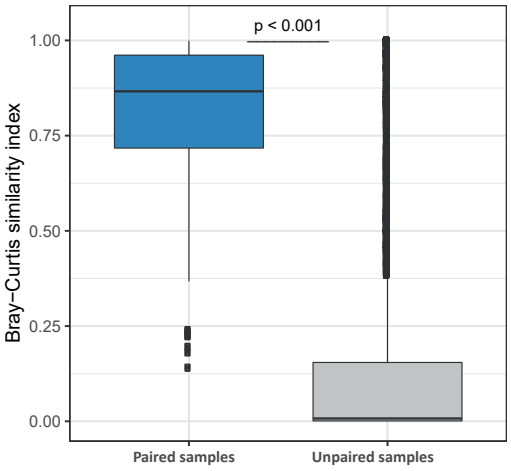


Figure 2. Concordance in microbiota community composition of paired rectal swabs and faecal samples. The concordance in microbiota community composition of paired faecal samples and rectal swabs was compared to the concordance of unpaired samples, using BC similarity (1-BC dissimilarity). A BC similarity index of 1 indicates an identical composition, while an index of 0 indicates the opposite. A significantly higher similarity was observed between paired samples (median BC similarity 0.866) than between unpaired samples (median BC similarity 0.007), as calculated by Wilcoxon. Boxplots with medians are shown; the lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles); the upper and lower whiskers extend from the hinge to the largest and smallest value no further than 1.5*IQR from the hinge; outliers are plotted individually.

Paired faecal samples and rectal swabs correlate strongly on taxonomical level

The paired faecal samples and rectal swabs had a similar taxonomical composition with respect to the most abundant OTUs (Figure 4). We performed Pearson correlation tests between the faecal samples and rectal swabs for all 198 testable OTUs, meaning that the 102 OTUs that were uniquely present in either the faecal samples or rectal swabs were excluded. The correlations of the top 15 most abundant taxa are shown in Table 1, all showing strong (Pearson's r>0.60) to very strong (Pearson's r>0.80) correlations. A total of 150 out of 198 testable OTUs correlated strongly and significantly in relative abundance between the paired samples with a median Pearson's r of 0.92 (IQR 0.90-1.00, adjusted p-value < 0.05). Together, these 150 taxa had a combined relative abundance of 98.5% of all sequences observed in our dataset. The comprehensive results of the correlations of all OTUs can be found in Supplementary Table 2.

Table 1. Correlation of the top 15 most abundant OTUs between the two sampling methods.

OTU	mean RA faeces	mean RA rs	Δ mean RA	Pearson's r
Escherichia coli (2)	23.39%	24.19%	-0.80%	0.95
Enterococcus faecium (5)	18.46%	15.61%	2.85%	0.96
Staphylococcus epidermidis (3)	13.38%	16.70%	-3.32%	0.76
Klebsiella (4)	11.17%	11.82%	-0.65%	0.99
Bacteroidales *ratAN060301*C (7)	3.84%	4.60%	-0.76%	0.94
Streptococcus salivarius subsp thermophilus (6)	5.00%	2.45%	2.55%	0.83
Streptococcus (11)	4.52%	2.66%	1.86%	0.68
Bacteroides (14)	2.79%	3.04%	-0.24%	0.98
Bifidobacterium (1)	2.79%	2.50%	0.29%	0.97
Streptococcus (30)	2.26%	1.15%	1.11%	0.70
Parabacteroides (37)	1.41%	1.22%	0.19%	0.99
Clostridium butyricum (26)	0.92%	1.71%	-0.79%	0.95
Veillonella (9)	0.81%	1.30%	-0.49%	0.68
Rothia (43)	0.65%	0.91%	-0.26%	0.98
Clostridium sensu stricto 1 (15)	0.67%	0.77%	-0.10%	0.90

We studied, among others, the correlation of the relative abundance of the 15 most abundant Operational Taxonomical Units (OTUs) between the paired faecal samples and rectal swabs, using the Pearson's correlation coefficient. For all 15 OTUs (with a combined relative abundance of 92.5%) we found a strong (Pearson's r 0.60-0.79) to very strong (Pearson's r 0.80-1) correlation between the paired faecal samples and rectal swabs. RA = relative abundance; rs = rectal swabs; Δ mean RA = difference in mean RA between faecal samples and rectal swabs calculated for each OTU. A comprehensive list of results of the correlations of all OTUs can be found in Supplementary Table 2.

DISCUSSION

Rectal swabs are more flexible to collect than faecal samples for research and clinical purposes, and could therefore be a valuable tool in cases where it is important to adhere to strict sampling timeframes, as when sampling around time sensitive interventions. In this study, we showed that paired infants' rectal swabs and faecal samples correlate well on alpha diversity, are comparable in overall community composition and correlate strongly on taxonomical level.

The observed species richness did not differ significantly between the two sampling methods at the earliest timepoint and correlated strongly between the paired faecal samples and rectal swabs. At the second timepoint, the observed species richness

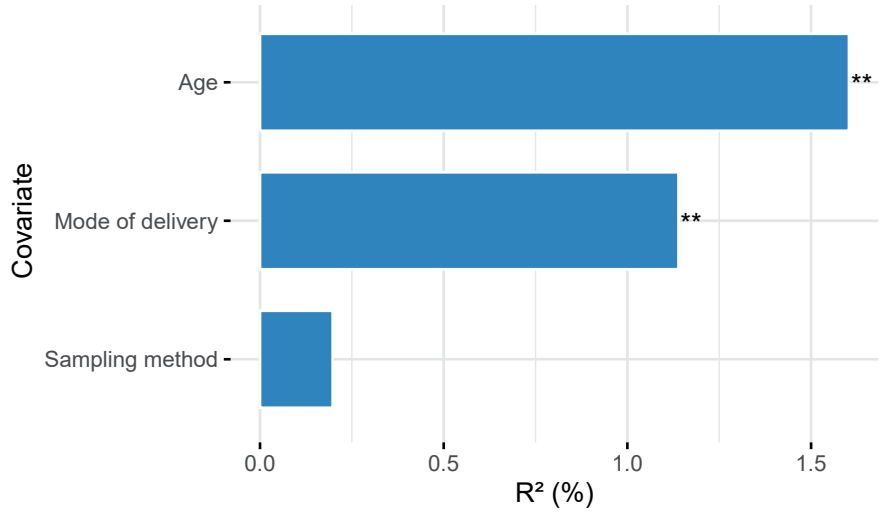


Figure 3. Variance explained in microbiota community composition by clinical variables. Covariates known to be associated with gut microbiota composition (age, delivery mode, feeding type) were first tested univariately with permutational multivariate analysis of variance (PERMANOVA)-tests using 1999 permutations, followed by a multivariable, temporal PERMANOVA, including sampling method to test what effect sampling method has on community composition in a clinical context. All samples were used with 1999 permutations and the strata parameter set to individual. The percentage of variance explained (R^2 (%)) is plotted on the x-axis. Significant associations with microbiota composition (p=0.002 for age and p=0.005 for mode of delivery) are visualised in order of importance. Sampling method explained only a minor and non-significant component of the variance in community composition compared to all other parameters in the model (p=0.835).

differed significantly between faecal samples and rectal swabs, even after filtering our OTU table of possible contaminants and also when repeating this analysis with rarefied data. However, this difference in diversity was contrary to what we expected: if at all different, we expected the rectal swabs, due to the smaller amount of material collected with this method, to pick up fewer species, but the opposite was the case, showing at least no taxa were missed with this type of sampling. Importantly, no differences in Shannon diversity were found between the sampling methods overall, or at either timepoint, and a clear correlation between paired samples was found. We also found an equal number of taxa that were unique to either sampling method (52 in faecal samples and 50 in rectal swabs), though the combined relative abundance of these was extremely low (<0.01%) and therefore less relevant for overall community structure. As a result, we found no difference in overall microbial community composition between the two sampling methods. Also, the concordance between the microbiota of a paired faecal sample and rectal

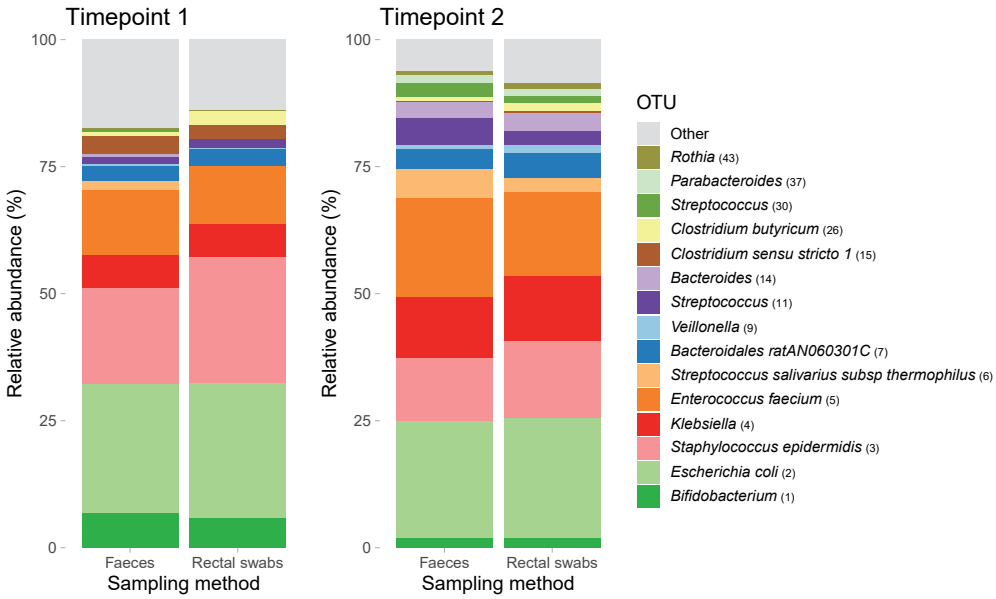


Figure 4. Relative abundance of the top 15 most abundant OTUs per timepoint and sampling method. Mean relative abundance of the top 15 most abundant Operational Taxonomical Units (OTUs) in this dataset, stratified per timepoint and sampling method.

swab collected from the same participant was high, confirming previous findings in adults[7].

In a broader clinical context, we found it interesting to establish that sampling method did not significantly explain variation in microbiota composition, as opposed to known drivers such as age[29], further supporting that rectal swabs are an appropriate proxy for faecal samples in infant gut microbiota studies. Unfortunately, we could not study the effect of antibiotics on composition in the multivariable model, because antibiotic treatment was colinear with timepoint in our study (sampling moments were before the start and after cessation of antibiotic treatment), and therefore age.

With respect to individual taxa, we found a high abundance of facultative anaerobic genera, such as *Escherichia coli,* and *Staphylococcus epidermidis* in the earliest samples, consistent with the description of normal early life gut microbiota development in previous studies[30,31]. The paired faecal samples and rectal swabs showed a strong correlation for most bacteria abundantly present in the infant microbiota. The combined relative abundance of the taxa with a strong to very

strong correlation (Pearson's r>0.60) between the sampling methods was above 98%, including predominant and clinically relevant taxa such as *Klebsiella* and *Enterococcus faecium,* which are known reservoirs for antibiotic resistance genes[32–34], as well as taxa like *Bifidobacterium* which are associated with various beneficial functions[35]. Altogether, this underlines that rectal swabs are a very good proxy for faecal samples in microbiota analyses.

ACKNOWLEDGEMENTS

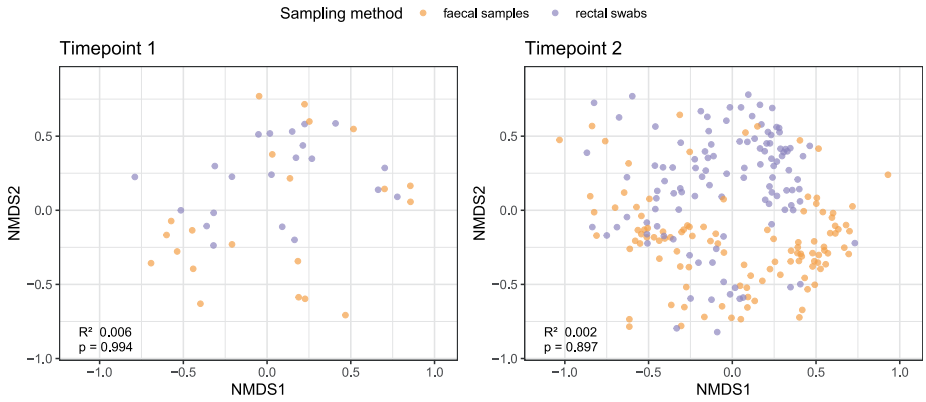
The authors are most grateful for the participation of all the children and their families. We would like to acknowledge all the members of the research team of the Spaarne Gasthuis Academy, the laboratory staff of the University Medical Center Utrecht and the Netherlands Organisation for Applied Scientific Research, and the Neonatology Departments of the Spaarne Gasthuis Hoofddorp and Haarlem, Diakonessenhuis Utrecht and Tergooiziekenhuis Blaricum for the help in participant recruitment. The ZEBRA research was financed by ZonMw Priority Medicines grant 205300001 and CSO grant SCAF/16/03.

REFERENCES

1. Kundu, P., Blacher, E., Elinav, E. & Pettersson, S. Our Gut Microbiome: The Evolving Inner Self. *Cell* **171**, 1481–1493 (2017).
2. Leggett, R. M. *et al.* Rapid MinION metagenomic profiling of the preterm infant gut microbiota to aid in pathogen diagnostics. *bioRxiv* 180406 (2017). doi:10.1101/180406
3. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
4. den Hertog, J. *et al.* The defecation pattern of healthy term infants up to the age of 3 months. *Arch. Dis. Child. Fetal Neonatal Ed.* **97**, F465-70 (2012).
5. Siegel, J. D., Rhinehart, E., Jackson, M., Chiarello, L. & Healthcare Infection Control Practices Advisory Committee. Management of multidrug-resistant organisms in health care settings, 2006. *Am. J. Infect. Control* **35**, S165-93 (2007).
6. Budding, A. E. *et al.* Rectal swabs for analysis of the intestinal microbiota. *PLoS One* **9**, e101344 (2014).
7. Bassis, C. M. *et al.* Comparison of stool versus rectal swab samples and storage conditions on bacterial community profiles. *BMC Microbiol.* **17**, 78 (2017).
8. Gibory, M. *et al.* Rotavirus detection in bulk stool and rectal swab specimens in children with acute gastroenteritis in Norway. *J. Clin. Virol.* **97**, 50–53 (2017).
9. 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.* **8**, 738 (2017).
10. Arrieta, M.-C., Stiemsma, L. T., Amenyogbe, N., Brown, E. M. & Finlay, B. The intestinal microbiome in early life: health and disease. *Front. Immunol.* **5**, 427 (2014).
11. Zaura, E., Keijser, B. J. F., Huse, S. M. & Crielaard, W. Defining the healthy core microbiome of oral microbial communities. *BMC Microbiol.* **9**, 259 (2009).
12. Biesbroek, G. *et al.* Deep sequencing analyses of low density microbial communities: working at the boundary of accurate microbiota detection. *PLoS One* **7**, e32942 (2012).
13. Bosch, A. A. T. M. *et al.* Maturation of the Infant Respiratory Microbiota, Environmental Drivers, and Health Consequences. A Prospective Cohort Study. *Am. J. Respir. Crit. Care Med.* **196**, 1582–1590 (2017).
14. Joshi, N. A. & Fass, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). (2011). Available at: https://github.com/najoshi/sickle.
15. Nikolenko, S. I., Korobeynikov, A. I. & Alekseyev, M. A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* **14**, S7 (2013).
16. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* **13**, 31 (2012).
17. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–6 (2010).

18. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
19. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
20. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
21. Subramanian, S. *et al.* Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**, 417–21 (2014).
22. Davis, N. M., Proctor, D., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *bioRxiv* 221499 (2017). doi:10.1101/221499
23. R Core Team. R: A language and environment for statistical computing. https://www.r-project.org/. (2017).
24. RStudio Team. RStudio: Integrated Development for R. http://www.rstudio.com/. (2016).
25. Wickham, B. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2009).
26. Kassambara, A. ggpubr: 'ggplot2' Based Publication Ready Plots. https://CRAN.R-project.org/package=ggpubr. (2018).
27. Oksanen, J. *et al.* Community Ecology Package. https://CRAN.R-project.org/package=vegan. (2017).
28. Benjamini, Y. & Hochberg, Y. Controlling The False Discovery Rate–A Practical And Powerful Approach To Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
29. Lozupone, C. A. *et al.* Meta-analyses of studies of the human microbiota. *Genome Res.* **23**, 1704–14 (2013).
30. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–7 (2012).
31. Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **8**, 343ra82-343ra82 (2016).
32. Gibson, M. K. *et al.* Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat. Microbiol.* **1**, 16024 (2016).
33. Boucher, H. W. *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **48**, 1–12 (2009).
34. Peterson, L. R. Bad bugs, no drugs: no ESCAPE revisited. *Clin. Infect. Dis.* **49**, 992–3 (2009).
35. Kwak, M.-J. *et al.* Evolutionary architecture of the infant-adapted group of Bifidobacterium species associated with the probiotic function. *Syst. Appl. Microbiol.* **39**, 429–439 (2016).

SUPPLEMENTARY INFORMATION



Supplementary Figure S1. NMDS plots. Non-metric multidimensional scaling (nMDS) plot based on Bray-Curtis (BC) dissimilarity between samples, visualizing the differences in overall gut microbiota community composition between sampling methods, per timepoint. Each data point represents the microbial community composition of one sample and is coloured by sampling method. The stress of the ordination was 0.280 (arbitrary). The effect sizes (R^2) and p-values calculated by permutational multivariate analysis of variance (PERMANOVA)-tests are shown in the plots. No significant differences in overall gut microbiota community composition between sampling methods were found.

Supplementary Table S1. Sample characteristics

	Timepoint 1	**Timepoint 2**
Total number of samples	**42**	**220**
Number of sample pairs*	21	110
Age of participants at sampling in days, median (IQR)	1 (1-2)	6 (4-8.8)
faecal samples	1 (1-2)	6 (4-9)
rectal swabs		
Time difference between paired sample collection in hours, median (IQR)	0 (0-0)	0 (0-2.5)
Reads per sample, median (IQR)		
faecal samples	34,142 (27,433- 40,674)	42,146 (29,216-55,482)
rectal swabs	41,252 (26,744- 55,923)	50,519 (42,058-63,011)

* Sample pair is defined as a matched faecal sample and rectal swab collected within 24 hours from an individual participant.

An overview of the sample numbers and characteristics is shown per sample moment. IQR = interquartile range.

Supplementary Table S2. Correlation of all OTUs between the two sampling methods

OTU	mean RA faeces (%)	mean RA rs (%)	delta mean RA (%)	Pearson's r	adjusted p-value
Escherichia coli (2)	23.39	24.19	-0.80	0.95	0.00
Enterococcus faecium (5)	18.46	15.61	2.85	0.96	0.00
Staphylococcus epidermidis (3)	13.38	16.70	-3.32	0.76	0.00
Klebsiella (4)	11.17	11.82	-0.65	0.99	0.00
Bacteroidales *ratAN060301C* (7)	3.84	4.60	-0.76	0.94	0.00
Streptococcus salivarius subsp thermophilus (6)	5.00	2.45	2.55	0.83	0.00
Streptococcus (11)	4.52	2.66	1.86	0.68	0.00
Bacteroides (14)	2.79	3.04	-0.24	0.98	0.00
Bifidobacterium (1)	2.79	2.50	0.29	0.97	0.00
Streptococcus (30)	2.26	1.15	1.11	0.70	0.00
Parabacteroides (37)	1.41	1.22	0.19	0.99	0.00
Clostridium butyricum (26)	0.92	1.71	-0.79	0.95	0.00
Veillonella (9)	0.81	1.30	-0.49	0.68	0.00
Rothia (43)	0.65	0.91	-0.26	0.98	0.00
Clostridium sensu stricto 1 (15)	0.67	0.77	-0.10	0.90	0.00
Bacteroides (31)	0.74	0.65	0.09	0.96	0.00
Lachnospiraceae (44)	0.65	0.68	-0.02	1.00	0.00
Streptococcus pyogenes (17)	0.46	0.76	-0.30	0.99	0.00
Peptostreptococcaceae (32)	0.41	0.80	-0.39	0.10	0.33
Streptococcus gallolyticus subsp macedonicus (16)	0.72	0.34	0.38	0.72	0.00
Corynebacterium (47)	0.45	0.57	-0.12	0.33	0.00
Haemophilus (38)	0.48	0.46	0.02	0.75	0.00
Prevotella (27)	0.22	0.59	-0.37	0.97	0.00
Bacteroides (33)	0.15	0.65	-0.50	0.78	0.00
Bifidobacterium animalis (36)	0.30	0.40	-0.11	1.00	0.00
Erysipelotrichaceae (34)	0.27	0.43	-0.16	1.00	0.00
Megamonas (46)	0.18	0.30	-0.12	0.99	0.00
Lactobacillus (40)	0.30	0.15	0.15	1.00	0.00
Clostridium sensu stricto 1 (23)	0.31	0.13	0.18	0.33	0.00
Streptococcus anginosus subsp whileyi (55)	0.32	0.07	0.25	0.69	0.00
Stenotrophomonas maltophilia (91)	0.23	0.12	0.11	1.00	0.00
Bacteroides (94)	0.21	0.12	0.09	1.00	0.00
Parasutterella (88)	0.06	0.26	-0.20	1.00	0.00

Continue

Continued

OTU	mean RA faeces (%)	mean RA rs (%)	delta mean RA (%)	Pearson's r	adjusted p-value
Veillonella (29)	0.16	0.15	0.01	1.00	0.00
Corynebacterium striatum (77)	0.02	0.28	-0.26	0.52	0.00
Ruminococcus gnavus CC55 001C (10)	0.10	0.12	-0.02	1.00	0.00
Gemella (120)	0.01	0.16	-0.14	0.99	0.00
Bifidobacterium (63)	0.13	0.03	0.10	0.87	0.00
Bacillales (147)	0.08	0.01	0.07	0.60	0.00
Parabacteroides goldsteinii dnLKV18 (133)	0.00	0.09	-0.08	1.00	0.00
Lactococcus lactis (78)	0.06	0.02	0.03	0.46	0.00
Phascolarctobacterium (219)	0.01	0.05	-0.04	1.00	0.00
Sutterella (134)	0.02	0.05	-0.03	0.72	0.00
Bacillus cereus (68)	0.01	0.06	-0.05	1.00	0.00
Lactobacillales (142)	0.00	0.06	-0.06	0.84	0.00
Tepidimonas (95)	0.02	0.05	-0.03	0.20	0.03
Bacteroides (58)	0.02	0.05	-0.03	0.99	0.00
Eubacterium sp CS1 Van (96)	0.05	0.00	0.05	0.81	0.00
Bacteroides (194)	0.03	0.01	0.02	0.98	0.00
Neisseria (121)	0.01	0.04	-0.03	0.10	0.31
Parabacteroides distasonis (207)	0.02	0.02	0.01	0.98	0.00
Aquabacterium (152)	0.01	0.03	-0.02	0.15	0.12
Bacteroides (249)	0.02	0.01	0.01	0.95	0.00
Acinetobacter (80)	0.03	0.01	0.02	1.00	0.00
Peptostreptococcaceae (50)	0.01	0.02	0.00	1.00	0.00
Nitriliruptor (126)	0.01	0.02	-0.01	0.24	0.01
Acinetobacter (143)	0.01	0.02	-0.01	0.35	0.00
Streptococcus (213)	0.03	0.00	0.02	0.14	0.15
Bifidobacterium animalis (311)	0.00	0.03	-0.03	1.00	0.00
Comamonadaceae (130)	0.01	0.02	-0.01	0.11	0.25
Prevotella bivia (102)	0.00	0.02	-0.02	0.99	0.00
Enterococcus (186)	0.02	0.00	0.02	0.46	0.00
Staphylococcaceae (246)	0.02	0.00	0.01	0.75	0.00
Bacteroides (236)	0.01	0.01	0.00	0.33	0.00
Burkholderia (156)	0.01	0.01	-0.01	0.29	0.00
Bilophila wadsworthia 3 1 6 (123)	0.02	0.00	0.01	0.02	0.93
Lactococcus (189)	0.02	0.00	0.01	0.19	0.04
Veillonella (76)	0.01	0.01	0.00	0.19	0.04

Continue

Continued

OTU	mean RA faeces (%)	mean RA rs (%)	delta mean RA (%)	Pearson's r	adjusted p-value
Acinetobacter (157)	0.01	0.01	0.00	0.05	0.73
Enhydrobacter (119)	0.01	0.01	0.00	0.06	0.59
Veillonella (87)	0.01	0.01	0.00	0.75	0.00
Lactobacillus reuteri (106)	0.02	0.00	0.02	0.10	0.34
Actinomyces (127)	0.01	0.01	0.00	0.95	0.00
Streptococcus (303)	0.02	0.00	0.01	0.43	0.00
Lactobacillus jensenii (271)	0.01	0.01	0.00	1.00	0.00
Actinomyces (292)	0.01	0.00	0.01	0.86	0.00
Bifidobacterium animalis (202)	0.01	0.01	0.00	1.00	0.00
Bacteroides (214)	0.01	0.01	0.00	1.00	0.00
Pseudomonas fluorescens (144)	0.00	0.02	-0.01	0.04	0.80
Bacteroides (423)	0.01	0.01	0.00	0.99	0.00
Bacillales (360)	0.01	0.00	0.01	0.62	0.00
Streptococcus (192)	0.01	0.00	0.01	0.31	0.00
Bacteroides (222)	0.01	0.01	0.00	0.92	0.00
Finegoldia (164)	0.00	0.01	-0.01	0.07	0.55
Collinsella (24)	0.00	0.01	-0.01	0.06	0.65
Escherichia Shigella (154)	0.01	0.01	0.00	0.63	0.00
Bacteroides (426)	0.01	0.00	0.01	0.99	0.00
Moraxella atlantae (514)	0.01	0.01	0.00	1.00	0.00
Streptococcus (338)	0.01	0.00	0.01	0.97	0.00
Lactobacillus salivarius (85)	0.00	0.01	0.00	1.00	0.00
Enterococcaceae (191)	0.01	0.00	0.01	0.92	0.00
Sphingobium (257)	0.01	0.00	0.01	-0.01	0.94
Bacteroides plebeius DSM 17135 (183)	0.00	0.01	-0.01	1.00	0.00
Bacteroides (70)	0.01	0.01	0.00	0.40	0.00
Actinomyces (73)	0.01	0.00	0.01	1.00	0.00
Erysipelotrichaceae (53)	0.00	0.01	0.00	0.92	0.00
Enterococcus (376)	0.01	0.00	0.00	0.40	0.00
Bacteroides (210)	0.00	0.01	0.00	0.90	0.00
Clostridium difficile 630 (345)	0.00	0.01	-0.01	0.74	0.00
Paracoccus (266)	0.00	0.01	0.00	-0.02	0.94
Corynebacterium aurimucosum ATCC 700975 (330)	0.01	0.00	0.00	1.00	0.00
Pasteurella pneumotropica (301)	0.01	0.00	0.01	0.07	0.51
Firmicutes (169)	0.01	0.00	0.01	1.00	0.00

Continue

Continued

OTU	mean RA faeces (%)	mean RA rs (%)	delta mean RA (%)	Pearson's r	adjusted p-value
Fusicatenibacter saccharivorans (20)	0.00	0.00	0.00	1.00	0.00
Caulobacteraceae (520)	0.01	0.00	0.01	0.99	0.00
Lachnospiraceae (83)	0.01	0.00	0.00	1.00	0.00
Klebsiella (200)	0.01	0.00	0.00	0.67	0.00
Enterobacteriaceae (205)	0.01	0.00	0.01	1.00	0.00
Bifidobacterium (12)	0.01	0.00	0.01	0.23	0.01
Veillonella sp DNF00869 (201)	0.01	0.00	0.00	0.14	0.15
Leuconostoc (275)	0.01	0.00	0.01	-0.01	0.94
Aggregatibacter (356)	0.00	0.01	0.00	0.20	0.03
Staphylococcaceae (452)	0.01	0.00	0.00	0.74	0.00
Peptostreptococcaceae (396)	0.00	0.01	0.00	0.81	0.00
Bacteroides (324)	0.00	0.00	0.00	0.94	0.00
Peptostreptococcus (98)	0.00	0.01	-0.01	0.14	0.13
Bifidobacterium (161)	0.00	0.01	0.00	1.00	0.00
Corynebacterium freneyi (429)	0.00	0.00	0.00	0.33	0.00
Clostridium difficile 630 (359)	0.00	0.00	0.00	0.57	0.00
Bacteroides (48)	0.00	0.01	-0.01	0.40	0.00
Coriobacteriaceae bacterium WAL 18889 (151)	0.00	0.00	0.00	0.01	0.94
Staphylococcus (507)	0.00	0.00	0.00	0.43	0.00
Streptococcus gallolyticus subsp macedonicus (464)	0.00	0.00	0.00	0.98	0.00
Staphylococcus (427)	0.00	0.00	0.00	0.21	0.02
Chryseobacterium (448)	0.00	0.00	0.00	-0.01	0.94
Lactobacillus (371)	0.00	0.00	0.00	0.79	0.00
Staphylococcus (489)	0.00	0.00	0.00	0.17	0.06
Staphylococcaceae (529)	0.00	0.00	0.00	0.77	0.00
Streptococcus (537)	0.00	0.00	0.00	0.06	0.63
Bacteroides (645)	0.00	0.00	0.00	0.98	0.00
Actinomyces sp 2002 2301122 (203)	0.00	0.00	0.00	0.50	0.00
Bifidobacterium (272)	0.00	0.00	0.00	1.00	0.00
Bifidobacterium bifidum NCIMB 41171 (13)	0.00	0.00	0.00	0.94	0.00
Bifidobacterium (336)	0.00	0.00	0.00	0.92	0.00
Bacteroides (248)	0.00	0.00	0.00	1.00	0.00
Gardnerella vaginalis 0288E (71)	0.00	0.00	0.00	0.31	0.00

Continue

Continued

OTU	mean RA faeces (%)	mean RA rs (%)	delta mean RA (%)	Pearson's r	adjusted p-value
Serratia marcescens (575)	0.00	0.00	0.00	0.14	0.15
Bifidobacterium (146)	0.00	0.00	0.00	0.69	0.00
Veillonella (115)	0.00	0.00	0.00	0.23	0.01
Bacteroides (400)	0.00	0.00	0.00	0.99	0.00
Bacilli (471)	0.00	0.00	0.00	0.71	0.00
Enterobacteriaceae (312)	0.00	0.00	0.00	0.74	0.00
Bacteroides (603)	0.00	0.00	0.00	0.99	0.00
Enterobacteriaceae (632)	0.00	0.00	0.00	0.97	0.00
Corynebacterium (641)	0.00	0.00	0.00	-0.01	0.94
Micrococcus luteus (545)	0.00	0.00	0.00	0.01	0.94
Streptococcus urinalis (389)	0.00	0.00	0.00	1.00	0.00
Flavonifractor (111)	0.00	0.00	0.00	1.00	0.00
Anaerococcus tetradius (329)	0.00	0.00	0.00	0.46	0.00
Janthinobacterium lividum (430)	0.00	0.00	0.00	0.02	0.94
Sutterella (605)	0.00	0.00	0.00	1.00	0.00
Bacillales (615)	0.00	0.00	0.00	0.46	0.00
Corynebacterium (328)	0.00	0.00	0.00	0.58	0.00
Corynebacterium (321)	0.00	0.00	0.00	0.64	0.00
Veillonella (300)	0.00	0.00	0.00	0.94	0.00
Corynebacteriaceae (504)	0.00	0.00	0.00	0.03	0.85
Staphylococcus (588)	0.00	0.00	0.00	0.47	0.00
Enterobacteriaceae (626)	0.00	0.00	0.00	0.98	0.00
Phascolarctobacterium (165)	0.00	0.00	0.00	0.32	0.00
Moryella (60)	0.00	0.00	0.00	1.00	0.00
Leuconostoc mesenteroides (282)	0.00	0.00	0.00	0.92	0.00
Dermacoccus (382)	0.00	0.00	0.00	-0.01	0.94
Acinetobacter ursingii ANC 3649 (388)	0.00	0.00	0.00	-0.01	0.94
Halomonas (494)	0.00	0.00	0.00	0.02	0.94
Bifidobacterium animalis (647)	0.00	0.00	0.00	1.00	0.00
Pseudoxanthomonas (667)	0.00	0.00	0.00	-0.02	0.94
Peptoniphilus (129)	0.00	0.00	0.00	-0.03	0.93
Methylobacterium radiotolerans (437)	0.00	0.00	0.00	-0.01	0.94
Bacteroides (535)	0.00	0.00	0.00	1.00	0.00
Dermabacter (335)	0.00	0.00	0.00	1.00	0.00
Streptococcus (425)	0.00	0.00	0.00	-0.01	0.94

Continue

Continued

OTU	mean RA faeces (%)	mean RA rs (%)	delta mean RA (%)	Pearson's r	adjusted p-value
Corynebacterium sp 1145 (518)	0.00	0.00	0.00	0.98	0.00
Raoultella ornithinolytica (495)	0.00	0.00	0.00	1.00	0.00
Corynebacterium propinquum (408)	0.00	0.00	0.00	-0.01	0.94
Bacteroides (681)	0.00	0.00	0.00	0.99	0.00
Bifidobacterium (233)	0.00	0.00	0.00	1.00	0.00
Rhizobium (561)	0.00	0.00	0.00	-0.02	0.94
Lactobacillus (42)	0.00	0.00	0.00	1.00	0.00
Vogesella (587)	0.00	0.00	0.00	-0.02	0.94
Streptococcus (521)	0.00	0.00	0.00	0.25	0.01
Blautia (8)	0.00	0.00	0.00	0.28	0.00
Plesiomonas (485)	0.00	0.00	0.00	0.94	0.00
Enterobacteriaceae (674)	0.00	0.00	0.00	1.00	0.00
Lactobacillus paracasei (18)	0.00	0.00	0.00	0.22	0.02
Bifidobacterium animalis (591)	0.00	0.00	0.00	1.00	0.00
Roseomonas (684)	0.00	0.00	0.00	-0.01	0.94
Bacteroides (380)	0.00	0.00	0.00	0.49	0.00
Bacteroides (482)	0.00	0.00	0.00	1.00	0.00
Ruminococcaceae (243)	0.00	0.00	0.00	0.05	0.67
Methylophilus (451)	0.00	0.00	0.00	0.11	0.26
Bifidobacterium animalis (293)	0.00	0.00	0.00	1.00	0.00
Bifidobacterium (223)	0.00	0.00	0.00	0.95	0.00
ratAN060301C (268)	0.00	0.00	0.00	0.54	0.00
Bifidobacterium animalis (560)	0.00	0.00	0.00	1.00	0.00
Alistipes (215)	0.00	0.00	0.00	0.99	0.00
Turicibacter (220)	0.00	0.00	0.00	-0.01	0.94
Staphylococcus (565)	0.00	0.00	0.00	0.24	0.01
Acinetobacter calcoaceticus (612)	0.00	0.00	0.00	-0.01	0.94
Bifidobacterium breve (473)	0.00	0.00	0.00	0.46	0.00
Aerococcus christensenii (61)	0.00	0.00	0.00	-0.01	0.94
Bacteroides (635)	0.00	0.00	0.00	0.97	0.00
Bacteroides (663)	0.00	0.00	0.00	1.00	0.00
Bacteroides (658)	0.00	0.00	0.00	1.00	0.00
Veillonella (502)	0.00	0.00	0.00	0.97	0.00
Peptostreptococcaceae (309)	0.00	0.00	0.00	-0.01	0.94
Ruminococcus gnavus CC55 001C (366)	0.00	0.00	0.00	1.00	0.00

Continue

Continued

OTU	mean RA faeces (%)	mean RA rs (%)	delta mean RA (%)	Pearson's r	adjusted p-value
Enterobacteriaceae (199)	0.00	0.00	0.00	0.13	0.19
Blautia (25)	0.00	0.00	0.00	0.02	0.94
Anaerostipes (22)	0.00	0.00	0.00	0.84	0.00
Bacteroides (242)	0.00	0.00	0.00	1.00	0.00
Peptoniphilus (413)	0.00	0.00	0.00	1.00	0.00
Varibaculum (117)	0.00	0.00	0.00	0.59	0.00
Subdoligranulum (136)	0.00	0.00	0.00	1.00	0.00
Anaerococcus (105)	0.00	0.00	0.00	-0.02	0.94
Bifidobacterium animalis (182)	0.00	0.00	0.00	1.00	NA
Lachnospiraceae (670)	0.00	0.00	0.00	1.00	NA
Sutterella (197)	0.00	0.00	0.00	0.28	0.00
Paracoccus (297)	0.00	0.00	0.00	0.46	0.00
Atopobium (174)	0.00	0.00	0.00	-0.01	0.94
Clostridium sensu stricto 1 (124)	0.00	0.00	0.00	-0.01	0.94
Subdoligranulum (45)	0.00	0.00	0.00	0.04	0.77
Bifidobacterium breve (634)	0.00	0.00	0.00	1.00	0.00
Subdoligranulum (81)	0.00	0.00	0.00	0.46	0.00
Anaerococcus (476)	0.00	0.00	0.00	-0.01	0.94
Anaerococcus (418)	0.00	0.00	0.00	-0.01	0.94
Lactobacillus delbrueckii subsp bulgaricus (190)	0.00	0.00	0.00	-0.01	0.94
Clostridium sensu stricto 1 (557)	0.00	0.00	0.00	-0.01	0.94
Clostridium sensu stricto 1 (462)	0.00	0.00	0.00	0.75	0.00
Phascolarctobacterium (317)	0.00	0.00	0.00	0.94	0.00
Eggerthella (178)	0.00	0.00	0.00	0.98	0.00
Bacteroides (247)	0.00	0.00	0.00	0.90	0.00
Xanthomonadales (274)	0.00	0.00	0.00	0.94	0.00
Lachnospiraceae (86)	0.00	0.00	0.00	0.37	0.00
Clostridium sensu stricto 1 (548)	0.00	0.00	0.00	0.74	0.00
Bacteroides (470)	0.00	0.00	0.00	0.77	0.00
Corynebacterium (460)	0.00	0.00	0.00	0.56	0.00
Faecalibacterium (21)	0.00	0.00	0.00	0.63	0.00
Streptococcus gallolyticus subsp macedonicus (350)	0.00	0.00	0.00	0.63	0.00
Clostridium sensu stricto 1 (621)	0.00	0.00	0.00	0.93	0.00
Prevotella melaninogenica (639)	0.00	0.00	0.00	-0.01	0.94
Negativicoccus sp S5 A15 (260)	0.00	0.00	0.00	-0.01	0.94

Continue

Continued

OTU	mean RA faeces (%)	mean RA rs (%)	delta mean RA (%)	Pearson's r	adjusted p-value
Dermabacter (617)	0.00	0.00	0.00	-0.01	0.94
Veillonella (454)	0.00	0.00	0.00	0.73	0.00
Bacteroides fragilis CL03T00C08 (291)	0.00	0.00	0.00	0.18	0.05
Alloscardovia (643)	0.00	0.00	0.00	1.00	0.00
Enterobacteriaceae (264)	0.00	0.00	0.00	0.88	0.00
Thermus (167)	0.00	0.00	0.00	-0.01	0.94
Bifidobacterium (347)	0.00	0.00	0.00	-0.01	0.94
Pantoea agglomerans (230)	0.00	0.00	0.00	-0.02	0.94
Veillonella (465)	0.00	0.00	0.00	0.19	0.04
Ruminococcus bromii L2 63 (62)	0.00	0.00	0.00	0.14	0.14
Candidatus Nitrososphaera (542)	0.00	0.00	0.00	1.00	0.00
Ferriphaselus (481)	0.00	0.00	0.00	-0.01	0.94
Bifidobacterium (228)	0.00	0.00	0.00	1.00	0.00
Peptostreptococcaceae (75)	0.00	0.00	0.00	-0.02	0.94
Eubacterium hallii DSM 3353 (52)	0.00	0.00	0.00	1.00	0.00
Bifidobacteriaceae (344)	0.00	0.00	0.00	1.00	0.00
Veillonella (555)	0.00	0.00	0.00	-0.01	0.94
Butyricimonas (503)	0.00	0.00	0.00	1.00	0.00
Blautia (79)	0.00	0.00	0.00	-0.01	0.94
Bacteroides (493)	0.00	0.00	0.00	1.00	0.00
Escherichia Shigella (469)	0.00	0.00	0.00	0.31	0.00
Gardnerella (332)	0.00	0.00	0.00	0.49	0.00
Bifidobacteriaceae (122)	0.00	0.00	0.00	-0.01	0.94
Bacteroides (166)	0.00	0.00	0.00	0.90	0.00
Lactobacillus fermentum (49)	0.00	0.00	0.00	-0.01	0.94
Bacteroides (445)	0.00	0.00	0.00	-0.01	0.94
Coprococcus (99)	0.00	0.00	0.00	-0.01	0.94
Bifidobacterium animalis (631)	0.00	0.00	0.00	1.00	0.00
Paraprevotella (508)	0.00	0.00	0.00	1.00	0.00

We studied the correlation of the relative abundance of all Operational Taxonomical Units (OTUs) between the paired faecal samples and rectal swabs, using the Pearson's correlation coefficient. RA = relative abundance; rs = rectal swabs; Δ mean RA = difference in mean RA between faecal samples and rectal swabs calculated for each OTU.

CHAPTER 4

A randomized trial studying the effects of early-life antibiotics on the developing infant gut microbiota and resistome: results from the ZEBRA study

Marta Reyman, Marlies A. van Houten, Rebecca L. Watson, Mei Ling J.N. Chu, Kayleigh Arp, Wouter J. de Waal, Irene Schiering, Frans B. Plötz, Rob J.L. Willems, Willem van Schaik, Elisabeth A.M. Sanders, Debby Bogaert

Submitted

ABSTRACT

We performed a randomized trial to study the effects of early-life antibiotics on the developing infant gut microbiome and on the selection of antimicrobial resistance. We included 147 term born infants treated with broad-spectrum antibiotics for (suspected) early-onset neonatal sepsis in their first week of life. Infants were randomly allocated to three commonly prescribed intravenous antibiotic combinations for this indication in the Netherlands. A subset of 80 non-antibiotic treated infants from a healthy birth cohort served as controls. The overall gut microbial community composition and antimicrobial resistance gene profile majorly shifted directly following treatment and normalized over 12 months. Among others, we found a decreased abundance of *Bifidobacterium* spp. and increased abundance of *Klebsiella* and *Enterococcus* spp. in the antibiotic treated infants compared to controls. Amoxicillin + cefotaxime showed the largest effects on both microbial community composition and antimicrobial resistance gene profile, whereas the penicillin + gentamicin regimen showed least effects.

The importance of the human gut microbiome in health and disease is becoming increasingly clear. Disturbances of the gut microbial community composition after birth are associated with a broad scale of health problems in early infancy and later in life, such as infantile colic, wheezing, allergies, functional gastrointestinal disorders, obesity and generally an altered immune development.[1–6] The various causes of disturbances are also becoming evident. Among others, the effects of caesarean section (CS) delivery, formula feeding (as opposed to breastfeeding) and antibiotics on the developing neonatal gut microbiome have been described.[7,8] Antibiotic treatment in particular alters species diversity (α-diversity) and community composition, i.e. the ecology, of the gut microbiome for a prolonged period of time in both children and adults.[9,10] Antibiotic-induced shifts in gut microbiome composition subsequently result in the selection of antimicrobial resistance (AMR).[11] The ecological side effects of antibiotics may be even more pronounced and persistent when administered in the early assembly phase of the neonatal gut microbiome in the first weeks of life. However, in this phase of life, broad-spectrum antibiotics are prescribed in up to 10% of all neonates because of (suspected) early-onset neonatal sepsis (sEONS).[12,13]

The common causative pathogens of EONS are group B streptococcus, *Listeria monocytogenes* and *Escherichia coli*. These organisms are susceptible to antibiotic regimens involving combinations of penicillin with gentamicin or amoxicillin combined with cefotaxime, making these treatments equally effective.[14] However, already two decades ago, a Dutch study performed on a neonatology intensive care unit highlighted the issue of amoxicillin driving the overgrowth of β-lactamase producing bacteria, such as *Klebsiella* species (spp.).[15] Another concern raised was that third-generation cephalosporins, such as cefotaxime, select for resistant *Enterobacter* spp. strains.[15] The full ecological impacts of empiric antibiotic treatment administered in early-life on microbiome and AMR gene composition are still not well-known.

To identify the antibiotic regimen with the least ecological and AMR selection effects, we enrolled 147 infants born at term, either by natural delivery or by secondary CS (SCS), for whom broad-spectrum antibiotics were indicated in the first week of life because of sEONS, and randomized them over three most commonly prescribed intravenous antibiotic combinations for this indication in the Netherlands.[16] On average, the antibiotic treatment duration was 48 hours. Antibiotic treated infants showed temporarily reduced gut microbial diversity, and major and prolonged ecological perturbations and shifts in AMR gene profile, compared with healthy, term born controls. Also, marked differences were found in ecological perturbation

between the three different antimicrobial regimens, suggesting that, next to adequate treatment, the choice of empirical antibiotics is also relevant for adverse ecological side-effects.

METHODS

Study design and participants

We performed a randomized study in 147 infants who required broad-spectrum antibiotics for treatment of sEONS in their first week of life. Infants were recruited at the Spaarne Gasthuis Hoofddorp and Haarlem, Diakonessenhuis Utrecht and Tergooiziekenhuis Blaricum in the Netherlands (Zuigelingen En Bacteriële Resistentie na Antibiotica, ZEBRA trial). Inclusion criteria were indication for broad-spectrum antibiotic treatment due to sEONS in the first seven days of life, birth by vaginal delivery or SCS, gestational age of ≥36 weeks, absence of prenatally established underlying morbidity, parental age of ≥18 years and the ability of parents to understand the Dutch or English language. A subset of healthy, term born infants from the Dutch Microbiome Utrecht Infant Study, served as controls.[17] We included 80 out of 120 infants from this birth cohort that were born vaginally or by SCS, had not received antibiotics in the first week of life and whose samples could be age-matched to those of the sEONS infants. Written informed consent was obtained from both parents. Both studies were approved by the Dutch National Ethics Committee (Netherlands Trial Registry NL4882 and NL3821). The randomized trial conformed to the Consolidated Standards of Reporting Trials (CONSORT) guidelines (Supplementary Information). In four cases a switch to a different antibiotic was made after blood/urine culture results became known and in 13 cases a switch to a different antibiotic was made in order for the infant to finish the treatment orally. These infants were also included in the intention to treat analysis as these practices reflect the clinical setting.

Power calculation

The study was initially powered making use of previously published infant microbiota data, ensuring a power of 0.8 to detect at least significant differences in alpha and beta diversity between groups, and at least two-fold differences in abundance of the 25 most important Operational Taxonomical Units (OTUs).[18] For power calculations, we used data of OTUs with high and low variability and abundance, and varying effect sizes. Our power calculation was verified by the online (HMP-based) tool as soon as this became available.[19] We initially aimed to enrol 132 infants, 44 infants per antibiotic regimen, allowing a drop-out of 10%. Due to

the accidental loss of a set of samples from 11 participants, approval was granted by the ethical committee to prolong enrolment in order to replace the lost samples, to ensure power of the study. Eventually, 147 infants were enrolled, 49 per antibiotic regimen.

Randomization and masking

Infants were randomly allocated 1:1:1 to three most commonly prescribed intravenous antibiotic combinations for sEONS in the Netherlands, namely penicillin + gentamicin, co-amoxiclav + gentamicin or amoxicillin + cefotaxime. Blinding was not performed, as this was not deemed relevant in our observational study with shifts in gut microbiome and AMR gene composition as primary study outcomes. The sequence with which participants were allocated to the groups was generated with the Research Manager software (Cloud9 Software BV) by an unaffiliated research nurse using 11 blocks of 12 (4:4:4) and one block of 15 (5:5:5) to enable a balanced randomization over the three regimens in the four hospitals. Allocation concealment was achieved using sealed, opaque envelopes which were delivered to the hospitals. Initial enrolment was performed by trained physicians who assigned the groups through selecting one of the sealed envelopes. Neither the research nurse who generated the allocation sequence nor the enrolling physicians were involved with the rest of the trial.

Sample collection

The duration of antibiotic treatment was not specified in the protocol, but was prescribed by the responsible physician depending on the clinical signs, reflecting the clinical setting. Rectal swabs and/or faeces were collected strictly before the start of antibiotic treatment (median sample age 1 day, interquartile range [IQR] 0-1 days) and 24-48 hours after treatment cessation (median 32 hours after cessation, IQR 27.6-40.1 hours; median sample age 6 days, IQR 4-8 days) and at 1, 4 and 12 months of age (median 31 days, IQR 30-34 days; median 123 days, IQR 120-127 days; median 367 days, IQR 365-371, respectively). The sample moments before and immediately after antibiotic treatment corresponded with 1 day (median sample age 1 day, IQR 1-1 days) and 1 week of life (median 6 days, IQR 6-7 days) of the controls (eFigure 1). Rectal swabs were collected using FaecalSwab™ kits (Copan Diagnostics, CA, USA) by trained physicians or research personnel before and 24-48 hours after antibiotic treatment. Faecal samples were obtained at the same timepoints and additionally at 1, 4 and 12 months and stored in sterile faecal containers by a nurse during hospital stay or by the parents if the participant was already discharged at the later timepoints. Since we previously showed that rectal swabs can be used as a proxy for faeces when the latter is not available, we used rectal swabs where

needed to ensure proper before-after treatment microbiota comparisons.[20] All materials were directly stored at -20°C before being transferred (<2 weeks) to a -80°C freezer until further laboratory processing. Perinatal characteristics of the participants were obtained at baseline and parents filled in online questionnaires about health characteristics at age 1, 4, 6, 8, 10 and 12 months (for the control infants a questionnaire was filled in at 9 months instead of at 8 and 10 months).

DNA isolation and sequencing

Bacterial DNA was isolated from faecal material using the Mag Mini DNA Isolation Kit (LGC ltd, UK) as previously described.[21] We used approximately 100 µl of faeces or 200 µl of swab material, 300 µl of lysis buffer, 500 µl zirconium beads and 500 µl of phenol, and performed an extra phenol/chloroform step. Further adaptations were applied to the samples collected at the first time point, being expected to be of low bacterial abundance and diversity, with the additional changes of using 150 µl of faeces (or 100 µl of material in the case of rectal swabs) and implementing an extra step with wash buffer 1.[22] DNA blanks and a positive control consisting of a mix of up to six random faecal samples were used for quality control. The amount of bacterial DNA was determined by quantitative polymerase chain reaction (qPCR) as previously described using universal primers specifically designed to amplify the bacterial 16S rRNA genes (Forward: 5'-CGAAAGCGTGGGGAGCAAA-3'; Reverse: 5'-GTTCGTACTCCCCAGGCGG-3'; Probe: 5'-6-FAM-ATTAGATACCCTGGTAGTCCA-MGB-3') on the 7500 Fast Real Time qPCR system (Applied Biosystems).[22]

For the sequencing of the V4 hypervariable region of the 16S rRNA gene, approximately 500 pg of DNA was amplified using F515/R806 primers and 30 amplification cycles.[23,24] Primers included Illumina adapters and a unique 8-nt sample index sequence key.[23] After amplification, quantification of the amount of amplified DNA per sample was executed with the dsDNA 910 Reagent Kit on the Fragment Analyzer (Advanced Analytical, IA, USA). Samples yielding insufficient DNA after amplification, defined as <0.5 ng/µl, were repeated with a higher concentration of template DNA. A mock control, positive control, DNA isolation blank and up to three PCR blanks were included in each PCR plate. Amplicons were pooled equimolarly and purified from 1.2% agarose gel using the Gel Extraction Kit (Qiagen, Hilden, Germany). The library was quantified using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Thermo Fisher Scientific, MA, USA). 16S rRNA sequencing was performed on the Illumina MiSeq platform (Illumina, Eindhoven, the Netherlands).

Bioinformatic processing

The samples and their sequences described in this manuscript are part of a larger dataset existing of 2176 faecal/rectal swab samples and controls, and together were processed using our in-house bioinformatics pipeline.[25] In short, we applied an adaptive, window-based trimming algorithm (Sickle, version 1.33) setting the quality threshold to q=30 and length threshold to 150 nucleotides.[26] Error correction was conducted with BayesHammer (SPAdes genome assembler toolkit, version 3.5.0).[27] Each set of paired-end sequence reads was assembled using PANDAseq (version 2.10) and demultiplexed with QIIME (version 1.9.1).[28,29] Singleton and chimeric reads (UCHIME) were removed. OTU picking was performed using VSEARCH abundance-based greedy clustering with a 97% identity threshold.[30] OTU annotation was established using the Naïve Bayesian RDP classifier (version 2.2) and the SILVA reference database.[31,32] This resulted in an OTU-table containing 18,951 taxa in total. We created an abundance-filtered dataset selecting OTUs present at a confident level of detection (0.1% relative abundance) in at least two samples, hereafter referred to as our raw OTU-table.[33] The raw OTU-table consisted of in total 730 taxa (0.49% sequences excluded with filtering). Next, we used both the prevalence and frequency methods of the decontam package with the default threshold of 0.1 to exclude possible contaminants, discarding 35 taxa, and thus retaining 695 taxa in total.[34]

For the purpose of this manuscript, 989 samples from 225 infants (90.7% of obtained samples from 145 antibiotic treated and 80 control infants) were eligible for microbiota analysis (eFigure 2). 16S-rRNA-based sequencing of the V4 hypervariable region resulted in 79,730,049 high quality reads with a minimum Good's coverage of 98.97% (median 99.99%). The raw OTU-table of the 989 samples contained 692 bacterial OTUs distributed over 18 bacterial phyla, with Actinobacteria being the most abundant phylum and Bifidobacterium the most abundant genus.

Metagenomic shotgun sequencing and processing

To validate our 16S rRNA sequencing and qPCR data, we performed metagenomic shotgun sequencing (MGS) on a randomly selected subset of 32 faecal samples collected from 20 antibiotic treated infants post antibiotics and 12 controls at 1 week of age. Sample libraries were prepared using the Truseq Nano gel free library preparation kit. Using a NovaSeq instrument, 150 base paired-end sequence data was generated yielding 1500 M reads (two runs). Reads were trimmed using Cutadapt[35] (version cutadapt-1.9.dev2) maintaining a quality threshold of 30 and a minimum read length of 35 base pairs. Per-sample and per-run SAM files were

generated using Bowtie2 and the MetaPhlAn2 and MEGARes databases,[36–38] while adhering to recommended parameters, including suppressing unaligned reads. Each SAM file was assigned a read group and SAM files from different runs were merged sample-wise using Picard.[39] Merged SAM files for each sample were sorted and indexed using SAMtools.[40] Raw reads and counts of reads mapped to each AMR gene were generated using SAMtools idxstats. To calculate counts per million in order to normalize read counts for library size, counts were divided by a normalisation factor of the library size divided by 1,000,000.

Identification of AMR genes by qPCR and processing

The primer set used in the qPCR assays covered 31 AMR genes, including genes encoding extended spectrum β-lactamases, carbapenemases and proteins involved in vancomycin resistance (eTable 1). The selection of AMR genes was based on two previous studies, aimed to detect the most common AMR genes in the gut microbiota of healthy individuals and clinically relevant AMR genes.[10,41] Of the 81 genes covered in these studies, we selected a total of 31 that were detected in ≥1 sample in either study and were deemed clinically relevant for our pediatric population and the antibiotic regimens being compared.

For the identification of AMR genes through qPCR we used the 96.96 BioMark™ Dynamic Array for Real-Time PCR (Fluidigm Corporation, San Francisco, CA). The concentration of each sample was measured using a specific 16S qPCR and all samples were normalized to 0.1 ng per ul. Any sample below this threshold was included in an undiluted state. Samples were pre-amplified with all primers for a total of 14 cycles, according to the manufacturer's instructions, excluding the 16S rRNA primer, due to the high abundance of 16S rRNA present in the samples, compared to the low abundance of the AMR targets. Primers were used at a concentration of 500 nM. The multiplex qPCR was run using the manufacturer's recommended settings, with the exception of the melting temperature being held at 56°C and the total number of cycles set to 35, which were adjusted according to the melting temperatures of the primers used, previous optimisation and our own internal validation.[10,41] All samples were run in triplicate. On all chips we included the 16S rRNA gene as the reference gene, two positive controls (a sample pool existing of all samples from chip plates 1-5, and a hospital sewage sample) and Non-Template Controls (NTCs). We standardized the within-individual variation across timepoints by including all timepoints from a given individual on a single chip. Individuals were then randomized across chips. Ct values were extracted using the BioMark Real-Time PCR analysis software. NTCs and positive controls were checked within and between chips for errors and consistency. The detection limit on the

Biomark system was set to a Ct value of 20 for AMR targets and 34 (because of lack of pre-amplification) for the 16S RNA reference gene. Sample-primer combinations that were failed by the software were checked manually and either passed or failed according to the suitability of the melting and amplification curves. Melting curve analysis was performed for all reactions within the Biomark Real-Time PCR analysis software to ensure the validity of the measured biological signals. Melting curves of the reactions were compared to those of the positive controls, and reactions showing unreliable S-curves were discarded. A sample was only included when a minimum of two of its triplicate reactions resulted in a Ct value below the detection limit, and for each included sample we calculated an average Ct value. Delta Ct values were calculated using the formula CtAMR – Ct16S rRNA. To account for the preamplification step of the AMR primers and to avoid negative delta Ct values, we added 20 to this value. Depending on the downstream statistical analysis, we transformed the dataset to include binary information or continuous Ct values with an arbitrary value of 32 appointed to negative sample-primer reactions.

Of the 989 samples that were included in the microbiota analyses, 939 samples passed the Fluidigm BioMark™ (Dynamic Array for Real-Time PCR) quality check and were included in the resistome analyses (eFigure 2).

Statistical analysis

All analyses were performed in R version 3.4.3 within RStudio version 1.1.383 and figures were made using package ggplot2 and ggpubr.[42–45] A statistical analysis scheme showing the flow in and order of analyses to address the primary research questions can be found in eFigure 3.

For the microbiome analyses, the numbers of samples analysed of the control group at time point 1 day, 1 week, 1 month, 4 months and 12 months, were: 66, 80, 80, 78 and 74, respectively. For the amoxicillin + cefotaxime treated infants the numbers of samples analysed at time point before antibiotics, after antibiotics, 1 month, 4 months and 12 months, were: 17, 46, 46, 45 and 44, respectively; for the co-amoxiclav + gentamicin treated infants this was: 20, 45, 47, 46 and 44, respectively; and for the penicillin + gentamicin treated infants this was: 23, 48, 47, 46 and 45, respectively. For the resistome analysis, the numbers of samples analysed of the control group at abovementioned time points, were: 55, 76, 78, 77 and 72, respectively; for the amoxicillin + cefotaxime treated infants this was: 11, 40, 46, 45 and 43, respectively; for the co-amoxiclav + gentamicin treated infants this was: 13, 41, 46, 43 and 42, respectively; and for the penicillin + gentamicin treated infants this was: 15, 46, 46, 45 and 45, respectively.

For simple, independent comparisons, we considered p-values <0.05 to be significant. However, for all analyses regarding multiple comparisons, we applied the Benjamini-Hochberg method to correct for multiple testing.[46] For comparisons of group differences, a one-way analysis of variance test, Wilcoxon rank-sum test, Kruskal-Wallis test, or chi-square test was used, where appropriate.

Breastfeeding at time of sampling was defined as an infant receiving any breastfeeding at that time point (albeit exclusive breastfeeding or mixed). Exclusive formula feeding was defined as an infant not having received any breastfeeding at all.

Group differences in Shannon alpha diversity and observed number of genes were calculated using Wilcoxon tests and linear mixed-effect models with participant set as random effect while correcting for age.

The overall gut microbial community composition of faecal samples and rectal swabs was visualized using nonmetric multidimensional scaling plots (nMDS; vegan package[47]). Ordinations were based on the Bray-Curtis (BC) dissimilarity matrix of relative abundance data with parameter trymax 10,000.

Associations between clinical outcome and microbiota composition were analysed with the adonis2 function (vegan package[47]), based on permutational multivariate analysis of variance (PERMANOVA)-tests per timepoint and across all timepoints using 1999 permutations using relative abundance data. To test for multivariate spread, we used the function betadisper (vegan package[47]). The dispersion of the data of the compared groups was significantly heterogenous immediately after antibiotic treatment onwards for the antibiotic treated versus non-treated comparisons ($p<0.001$, $p<0.001$, $p=0.02$ and $p=0.02$, respectively), while this was only the case immediately after antibiotics and at 1 month of age for the comparisons between the separate regimens and controls. In the inter-regimen comparisons, the dispersion of the data was homogenous at all time points. In the multivariable, temporal analyses all variables were included that showed a significant association with microbiota composition when tested individually, and age and subject were added to control for repeated measures.

The stability of the gut microbiota composition in the first year of life was visualized by measuring the BC dissimilarities between consecutive samples within each participant over time. To test for group differences, the Wilcoxon test was used.

Individual bacterial taxa and their succession patterns, and potential differences thereof between groups, were studied at the lowest taxonomic annotated level (OTU). Differential abundance testing was executed with smoothing spline analysis of variance (SS-ANOVA, fitTimeSeries function, metagenomeseq package[48,49]) allowing not only to detect biomarker OTUs related to antibiotic treatment, but also to identify the specific time intervals in which significant differences existed. For this analysis, the raw OTU-table was filtered and OTUs with >10 reads in ≥50 samples were included, resulting in 512 OTUs (of the 692).[25] This filtering step was performed separately for each subset of data used for the different comparisons. The SS-ANOVA analysis was adjusted for covariates that both had (1) an effect on overall gut microbiota composition and (2) were unevenly distributed between the antibiotic treated infants versus controls, namely mode of delivery, and the inter-regimen comparisons were adjusted for prepartum maternal antibiotics. Although gravidity of mothers, gestational age, duration of ruptured membranes, hospital stay duration after birth, presence of siblings <5 years of age, and inhouse smoking were all associated with antibiotic treated infants, we did not find an association between these clinical variables and gut microbiota composition, therefore these six variables were excluded as covariates from downstream fitTimeSeries analyses. While day care attendance and breastfeeding at time of sampling were associated with gut microbiota composition, these variables did not differ at baseline, and therefore were also excluded as covariates from the downstream analyses.

Results from 16S rRNA sequencing at 1 week of life were validated by untargeted MGS (subset of 32 samples). The diversity of AMR genes was defined as the observed number of different AMR genes present in a sample. Comparisons in AMR diversity were calculated using the Wilcoxon test. The overall AMR gene composition, or profiles, of faecal samples and rectal swabs were visualized as described above, with the exception that the ordination was based on the Jaccard index of the binary data. Associations between clinical outcome and AMR gene profiles were analysed similarly to the 16S data, but again using binary data and the Jaccard index. Dispersion of data was only significantly heterogenous in one case, namely at the time point immediately after antibiotic treatment in the comparison of the amoxicillin + cefotaxime regimen versus controls.

Differential abundance testing of the AMR genes was also executed with SS-ANOVA (fitTimeSeries function, metagenomeseq package[48,49]), though no filtering was performed to allow for testing of all 31 AMR genes and no normalisation was applied in the function. The SS-ANOVA analysis was adjusted for covariates that both had (1) an effect on overall AMR gene composition and (2) were unevenly distributed

between the groups compared, so for siblings <5 years of age and mode of delivery in the antibiotic treated infants versus controls and for prepartum maternal antibiotics in the inter-regimen comparisons.

We evaluated the correlation between 16S OTUs and AMR genes by calculating the Pearson correlation coefficient for all individual taxa based on their relative abundance. Positive coefficients indicated a negative correlation between OTU and gene abundance, as a high Ct value indicates low abundance of an AMR gene.

The fitZig function of the metagenomeSeq package[49] was used to assess the group differences in AMR genes as identified by MGS in a subset of 32 samples, after removing rare features present in less than 10 samples, resulting in 343 out of 1504 AMR genes included in this analysis.

RESULTS

Population characteristics

The development of microbiome and resistome development was studied in 147 neonates born ≥36 weeks of gestational age, who were recruited in the period from 16 January 2015 to 13 September 2016 in three different Dutch hospitals, and were followed until their first birthday. Neonates were randomized over the following antibiotic regimens (49 per group): penicillin + gentamicin, co-amoxiclav + gentamicin or amoxicillin + cefotaxime. As controls, 80 age-matched term born infants were included from a previous healthy birth cohort study conducted in two of the same hospitals.[17] All infants were born in hospital except for three spontaneous home births in the antibiotic treated group (one in each regimen) and five in the control group. From two infants in the amoxicillin + cefotaxime group, no samples were available for analysis, so these infants were excluded from further analyses (eFigure 4). Eight (5.4%) children were lost to follow-up due to parents experiencing the collection of samples or completion of questionnaires as too burdensome, or moving abroad.

In line with sEONS characteristics, neonates treated with antibiotics were more often born with SCS, had a slightly shorter gestational age, a lower Apgar score at 5 minutes, and their mothers had on average a longer period of ruptured membranes before delivery (Table 1). Antibiotic treated neonates were more often the first child of parents. As a consequence of the higher rate of SCS and sEONS, the antibiotic treated neonates were hospitalized for a longer period of time. Antibiotic

Table 1. Baseline characteristics

	Amoxicillin + cefotaxime	Co-amoxiclav + gentamicin	Penicillin + gentamicin	Controls	*p*
n	47	49	49	80	
Gender, female (%)	22 (46.8)	26 (53.1)	19 (38.8)	43 (53.8)	0.36
Mode of delivery, secondary C-section (%)	6 (12.8)	10 (20.4)	15 (30.6)	8 (10.0)	*0.02*
Antepartum maternal antibiotics (%)	16 (34.0)	19 (38.8)	29 (59.2)	NA	*0.03*
Season of birth (%)					0.68
Winter	13 (27.7)	14 (28.6)	10 (20.4)	14 (17.5)	
Spring	8 (17.0)	6 (12.2)	8 (16.3)	21 (26.2)	
Summer	18 (38.3)	19 (38.8)	20 (40.8)	31 (38.8)	
Autumn	8 (17.0)	10 (20.4)	11 (22.4)	14 (17.5)	
Gravidity mothers, median (IQR)	1.00 (1.00, 2.00)	1.00 (1.00, 2.00)	1.00 (1.00, 2.00)	2.00 (1.00, 3.00)	*<0.001*
Gestational age in weeks, median (IQR)	40.29 (39.07, 40.93)	40.57 (39.86, 41.14)	40.43 (39.71, 41.29)	39.86 (39.11, 40.71)	*0.01*
Birth weight in grams, mean (SD)	3673.79 (623.99)	3632.41 (506.26)	3633.10 (628.74)	3494.91 (494.13)	0.27
Ruptured membranes in hours, median (IQR)	8.50 (4.00, 21.00)	12.00 (6.00, 20.00)	11.00 (7.00, 17.50)	5.00 (2.50, 8.00)	*<0.001*
Apgar score at 5 minutes (%)					*0.008*
4	0 (0.0)	0 (0.0)	1 (2.0)	0 (0.0)	
5	0 (0.0)	0 (0.0)	2 (4.1)	0 (0.0)	
6	2 (4.3)	3 (6.2)	2 (4.1)	0 (0.0)	
7	4 (8.5)	4 (8.3)	4 (8.2)	2 (2.5)	
8	1 (2.1)	6 (12.5)	5 (10.2)	3 (3.8)	
9	5 (10.6)	13 (27.1)	8 (16.3)	10 (12.7)	
10	35 (74.5)	22 (45.8)	27 (55.1)	64 (81.0)	
Antibiotic treatment duration in days, median (IQR)	2.87 (2.38, 5.28)	4.50 (2.55, 6.57)	2.71 (2.29, 6.53)	NA	0.42
Antibiotics 1-4 days (vs. > 4 days) (%)	27 (57.4)	21 (43.8)	27 (55.1)	NA	0.36
Hospital stay duration in days, median (IQR)	5.00 (3.00, 7.00)	5.00 (3.00, 7.00)	5.00 (3.00, 7.00)	1.00 (1.00, 2.00)	*<0.001*
Number of siblings, median (IQR)	0.00 (0.00, 1.00)	0.00 (0.00, 1.00)	0.00 (0.00, 1.00)	1.00 (0.00, 1.00)	*0.001*
Presence of siblings <5 years of age (%)	11 (23.9)	11 (22.9)	12 (24.5)	43 (53.8)	*<0.001*
Pets (%)					0.35

Continue

Continued

	Amoxicillin + cefotaxime	Co-amoxiclav + gentamicin	Penicillin + gentamicin	Controls	*p*
None	24 (53.3)	24 (53.3)	34 (70.8)	44 (55.0)	
Cat(s)	14 (31.1)	14 (31.1)	10 (20.8)	17 (21.2)	
Dog(s)	4 (8.9)	1 (2.2)	2 (4.2)	7 (8.8)	
Cat(s) + dog(s)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.5)	
Other	3 (6.7)	6 (13.3)	2 (4.2)	10 (12.5)	
Inhouse smoking (%)	0 (0.0)	1 (2.1)	0 (0.0)	1 (1.2)	0.62
Parents finished higher education (%)	32 (71.1)	35 (74.5)	36 (73.5)	65 (81.2)	0.57
Breastfeeding in days, median (IQR)	37.50 (4.00, 153.00)	72.00 (11.75, 239.25)	94.00 (22.50, 350.50)	126.00 (4.50, 277.25)	0.08
Breastfeeding at 1 month (%)	25 (54.3)	32 (68.1)	33 (70.2)	55 (68.8)	0.32
Breastfeeding at 4 months (%)	13 (28.3)	18 (39.1)	23 (48.9)	41 (51.2)	0.07
Breastfeeding at 12 months (%)	2 (4.3)	9 (19.6)	12 (25.5)	12 (15.0)	0.05
Exclusive formula feeding (%)	11 (23.9)	8 (17.0)	7 (14.9)	12 (15.0)	0.59
Age start solid food in days, median (IQR)	142.50 (121.25, 159.50)	131.50 (120.25, 154.00)	135.00 (122.00, 159.00)	130.50 (119.00, 163.50)	0.82
Solid food at 4 months (%)	16 (34.8)	19 (41.3)	15 (31.9)	26 (32.5)	0.75
Pacifier use at 1 month of age (%)	37 (80.4)	33 (70.2)	35 (74.5)	57 (71.2)	0.65
Daycare since (%)					0.83
1 month	1 (2.2)	0 (0.0)	1 (2.2)	0 (0.0)	
4 months	20 (43.5)	20 (44.4)	22 (47.8)	35 (43.8)	
6 months	8 (17.4)	6 (13.3)	4 (8.7)	10 (12.5)	
10 months	6 (13.0)	1 (2.2)	2 (4.3)	10 (12.5)	
12 months	1 (2.2)	4 (8.9)	1 (2.2)	1 (1.2)	
>12 months	10 (21.7)	14 (31.1)	16 (34.8)	24 (30.0)	
Antibiotic use during follow-up (%)	10 (21.7)	7 (15.6)	5 (11.4)	21 (26.2)	0.20
Age at first antibiotic course during follow-up in days, mean (SD)	204.5 (82.3)	245.3 (78.3)	226.4 (61.7)	236.9 (89.0)	0.73
Antibiotics prior to 1 month time point (%)	1 (2.2)	0 (0.0)	0 (0.0)	1 (1.2)	0.64
Antibiotics prior to 4 months time point (%)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.3)	0.62
Antibiotics prior to 12 months time point (%)	9 (19.6)	7 (15.2)	5 (11.1)	20 (25.0)	0.27

Continue

Continued

< Baseline characteristics stratified by antibiotic regimen. Categorical variables are shown in absolute numbers with percentages (%); continuous, normally distributed variables as means with standard deviations (SD); continuous, non-normally distributed variables as medians with interquartile ranges (IQR). One-way analysis of variance tests were used to compare means of normally distributed continuous variables; Kruskal-Wallis tests were applied to compare medians of non-normally distributed continuous variables; significant differences between categorical variables were tested with chi-square tests. The p-values of variables that differed significantly between the two groups are italicized for clarity. NA value means variable is not applicable to group (mothers of control infants did not receive antepartum antibiotics and control infants did not receive antibiotics in their first week of life, so this treatment duration is NA).

treatment after the first week of life was uncommon (43 out of all 225 participants), administered relatively late in the first year (average age 8 months), and did not differ between sEONS infants and controls at baseline. As expected, baseline characteristics did not differ between the infants randomized to each of the three antibiotic regimens, except for antepartum maternal antibiotics exposure. In total, there were only two sEONS infants with culture-proven sepsis.

Effects of early-life broad-spectrum antibiotics on infant gut microbial composition development

Antibiotic treatment led to a decrease in α-diversity (median Shannon index 0.77 before, versus 0.59 after antibiotic treatment in the sEONS infants, Wilcoxon test, p=0.01), which was most outspoken immediately after antibiotic treatment (median Shannon index 0.59 in treated infants, versus 1.21 in controls at timepoint 2, Wilcoxon test, p<0.001; eFigure 5). Following, the α-diversity gradually recovered, though remained significantly lower throughout the first year of life (linear mixed model, p=0.01). When stratified, we observed that the co-amoxiclav + gentamicin group had the lowest α-diversity immediately after treatment (median Shannon index 0.43; Figure 1) which persisted throughout the first year of life when compared with controls (linear mixed model, p=0.002). In inter-regimen comparisons, no significant difference in α-diversity was observed.

Before start of antibiotics, the overall microbial community composition did not significantly differ between neonates with sEONS and controls (PERMANOVA-test, R^2=1.2%, p.adj=0.14; Figure 2a), however, a large shift in composition was observed directly following treatment (R^2=9.5%, adjusted p-value [p.adj]=0.001; Figure 2b). Despite gradual recovery over time, a small but significant difference in composition remained at 12 months of age (R^2=1.1%, p.adj=0.03, Figure 2c). The largest effect was found for the amoxicillin + cefotaxime group immediately after treatment (R^2=14.7%, p.adj=0.003; Figure 2e), which was also significantly different from

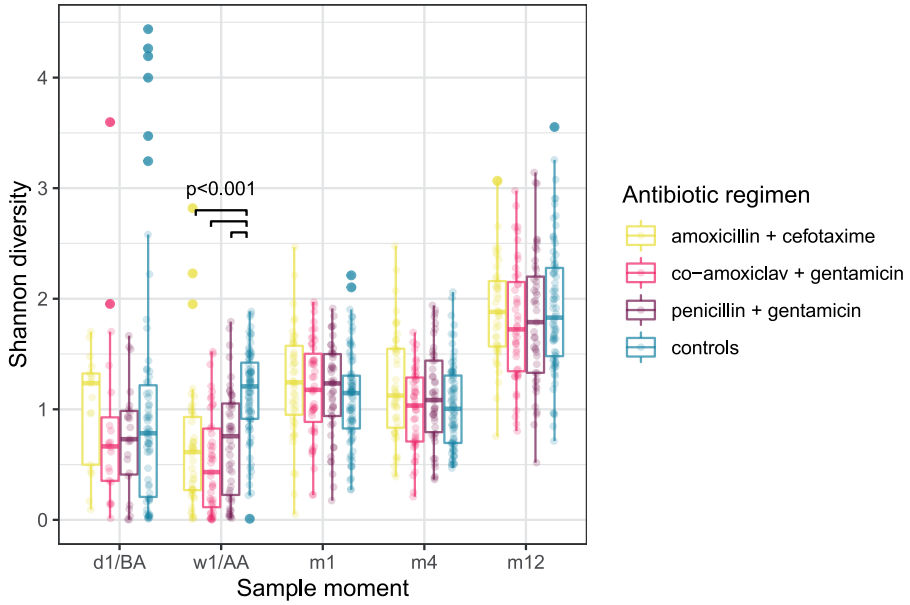
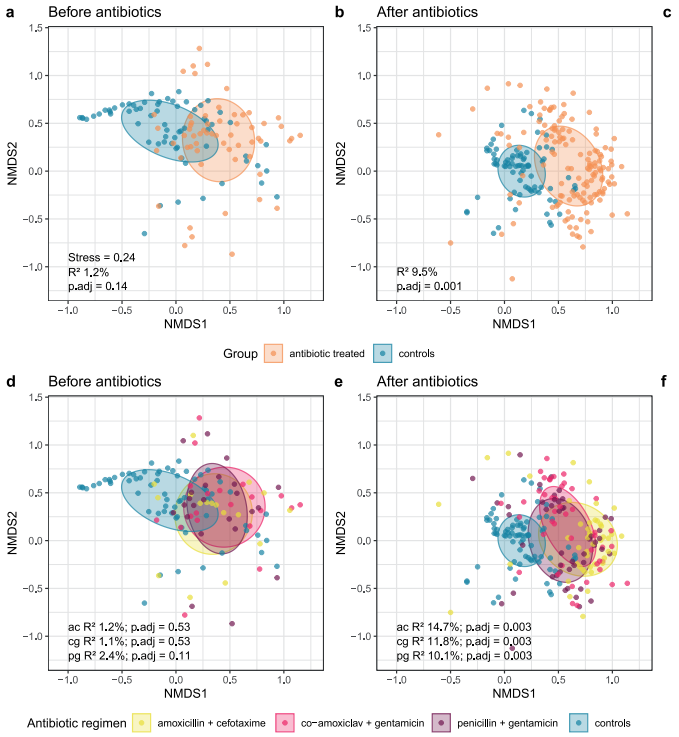


Figure 1. α-diversity. Differences in Shannon diversity of the faecal microbiota between infants treated with the three studied antibiotic regimens and controls plotted per timepoint. Group differences were calculated using Wilcoxon tests. Boxplots with medians are shown; the lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles); the upper and lower whiskers extend from the hinge to the largest and smallest value no further than 1.5*IQR from the hinge; outliers are plotted individually by opaque circles; translucent circles visualize all data points. D = day, w = week, BA = before antibiotics, AA = after antibiotics, m = month.

the other two regimens (versus co-amoxiclav + gentamicin: R^2=7.6%, p.adj=0.003; versus penicillin + gentamicin: R^2=4.9%, p.adj=0.004). Also, at age 1 month, the difference in composition for the amoxicillin + cefotaxime group versus controls was most outspoken, though subsequently, similar recovery patterns were observed as for the other two regimens (Figure 2f). To exclude the potential background noise of antibiotic treatment during follow-up (after the first week of life), we performed a similar post-hoc analysis on a subset of sEONS and control infants not receiving antibiotics after the first week. The effect size of antibiotics administered in the first week of life in this otherwise antibiotic-naïve subset of infants was, although no longer significant at the last two time points, larger at all time points than in the analysis performed on the overall groups (eTable 2).

The stability of microbial community development within individuals over time was lower in antibiotic treated infants compared to controls during the first 4 months of life (eFigure 6). Again, this was most pronounced for infants in the amoxicillin +



Sample moment	R² (%)	p.adj
month 1	3.0	0.001
month 4	1.4	0.02
month 12	1.1	0.03

Sample moment	R² (%)	p.adj
month 1 ac	6.1	0.003
month 4 ac	2.7	0.005
month 12 ac	1.5	0.14
month 1 cg	2.0	0.09
month 4 cg	2.3	0.06
month 12 cg	1.9	0.06
month 1 pg	3.5	0.004
month 4 pg	1.4	0.19
month 12 pg	1.0	0.40

Figure 2. NMDS plots of overall gut microbiota composition stratified per regimen and timepoint. Non-metric multidimensional scaling (nMDS), based on Bray-Curtis (BC) dissimilarity between samples, visualising the overall gut microbial community composition stratified for antibiotic treated infants and controls (**a**, **b**) and per regimen compared with controls (**d**, **e**), for the time points before and immediately after cessation of antibiotic treatment. Each data point represents the microbial community composition of one sample. The ellipses represent the standard deviation of data points belonging to each group, with the centre points of the ellipses calculated using the mean of the coordinates per group. The stress of the ordination, effect sizes (R^2) calculated by permutational multivariate analysis of variance (PERMANOVA)-tests and corresponding adjusted p-values (p.adj, calculated using the Benjamini-Hochberg method) are shown in the plots. In panels **c** and **f** the results of the PERMANOVA-tests for the later time points are summarized. Ac = amoxicillin + cefotaxime, cg = co-amoxiclav + gentamicin, pg = penicillin + gentamicin.

cefotaxime group, with a significant difference compared to both the control group and the penicillin + gentamicin group (after antibiotics vs. month 1 comparison, Wilcoxon test, p<0.001 and p=0.001, respectively).

Clinical covariates associated with infant gut microbial composition development

Besides antibiotic treatment, covariates significantly associated with overall gut microbial community composition during follow-up were: age (R^2 4.1%), day care attendance (R^2 1.0%), breastfeeding at time of sampling (R^2 0.7%) and mode of delivery (R^2 0.1%, all p.adj=0.001). Covariates that differed at baseline between the antibiotic treated infants and controls, but that did not have a significant association with composition were: gravidity of mothers, gestational age, duration of ruptured membranes, hospital stay duration after birth, presence of siblings <5 years of age and inhouse smoking. As mode of delivery was the only variable that differed between groups at baseline and was associated with microbial community composition over the first year of life, this variable was corrected for in all downstream microbiota analyses.

No significant association was found between overall microbial community composition and antibiotic treatment duration in the first week of life nor with maternal antepartum antibiotics when analysed per timepoint (cross-sectionally). In a temporal, multivariable analysis (sEONS group only), we found that both variables were, although significantly, only modestly associated with composition (antibiotic treatment duration R^2 0.3%; maternal antepartum antibiotics R^2 0.1%, both p.adj<0.001), when compared with antibiotic regimen (R^2 0.9%), age (R^2 5.2%), day care attendance (R^2 1.0%), and breastfeeding at time of sampling (R^2 0.8%, all p.adj<0.001). There were only two infants with culture-proven sEONS, so as an approximation of this clinical variable, we categorized the duration of antibiotic treatment into short (1-4 days) and long (>4 days) and studied its effect on overall microbial community composition in a cross-sectional manner (eTable 3). No significant difference could be found at any time point between the infants receiving short or long antibiotic treatment in their first week of life. As maternal antepartum antibiotics was the only variable that differed between the sEONS groups at baseline and was associated with microbial community composition over the first year of life, this variable was corrected for in all downstream inter-regimen microbiota analyses.

Differences in microbial colonization patterns between antibiotic treated infants and controls

The colonization patterns of bacterial taxa in the control group were similar to those described in previous healthy infant studies (Figure 3).[7,50] Facultative anaerobic genera such as *Escherichia*, and *Staphylococcus* were highly abundant in the first sample collected soon after birth, rapidly followed by the predominance of the

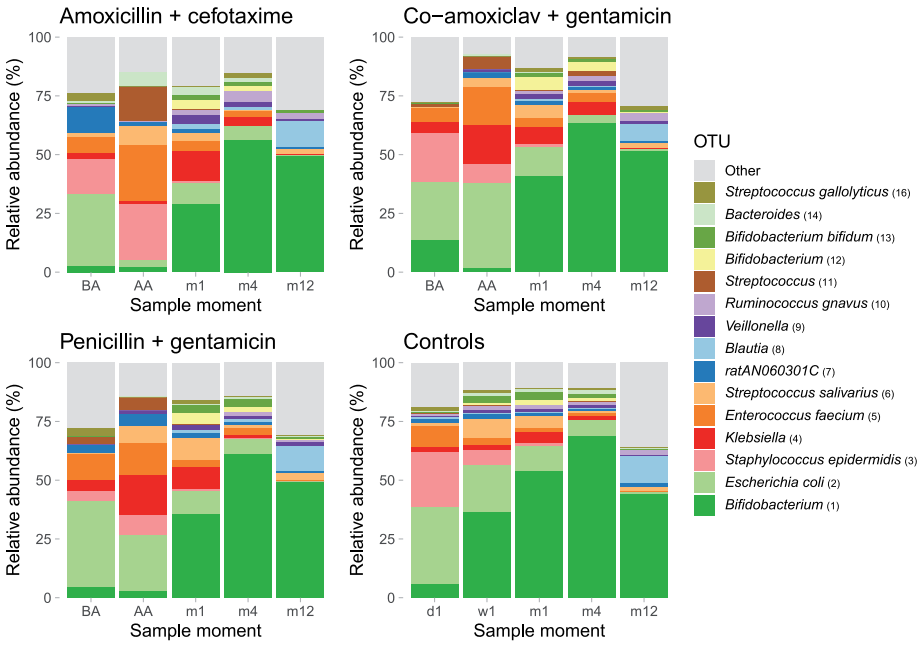


Figure 3. Mean relative abundance of most abundant OTUs. Mean relative abundances of the 15 most abundant OTUs are depicted for all samples per timepoint, stratified by antibiotic regimen. BA = before antibiotics, AA = after antibiotics, d = day, w = week, m = month. In some cases, multiple OTUs of individual bacterial species were identified, so OTUs are referred to by their taxonomical annotations and a rank number (shown in parentheses), which is based on the abundance of each given OTU in the overall dataset.

genus *Bifidobacterium*. Temporal analysis showed a total of 251 differentially abundant OTUs between controls and antibiotic treated infants. Among others, *Bifidobacterium* spp. were more abundant in the controls than in antibiotic treated infants from day 1 until day 36, also after adjusting for mode of delivery (p.adj=0.005; eTable 4). Likewise, *Escherichia* and *Staphylococcus* spp. were more abundant in controls from 1-181 and 1-229 days, respectively. Also, 28 taxa belonging to the genus *Bacteroides* were more abundant in controls compared with the antibiotic treated infants. In contrast, antibiotic treated infants showed, among others, higher abundances of *Klebsiella* (day 1-122) and *Enterococcus* spp. (day 1-121, p.adj=0.02 and 0.005, respectively).

MGS of a subset of samples (post antibiotic time point samples from sEONS infants n=20, week 1 time point samples from controls n=12) confirmed species annotation of *B. adolescentis*, *B. breve and B. longum* (*Bifidobacterium* OTU), *E.*

coli (*Escherichia coli* OTU), *S. epidermidis* (*Staphylococcus epidermidis* OTU), *K. pneumoniae* and *K. oxytoca* (*Klebsiella* OTU), and *E. faecium* (*Enterococcus faecium* OTU) (eFigure 7 and eTable 5). MGS also confirmed the effects of antibiotics on the overall microbial community composition, with a similar effect size (PERMANOVA test, R^2 8.0%, p=0.01) as for the 16S-based sequencing data (R^2 9.5%).

Stratified analyses for the three antibiotic regimens showed that *Bifidobacterium* was decreased in all three treatment groups for similar periods of time (eTable 4). Furthermore, *Enterococcus* spp. were increased in abundance for similar periods of time in all three treatment groups. However, other taxa were more variably affected in abundance by the three different regimens. For example, *E. coli* abundance was reduced in both the amoxicillin + cefotaxime and penicillin + gentamicin regimens (day 1-142 and day 1-177, respectively), but not in the co-amoxiclav + gentamicin regimen. *Akkermansia* spp. were generally only reduced in the amoxicillin + cefotaxime group. In contrast, *Klebsiella* abundance was higher in both the co-amoxiclav + gentamicin and penicillin + gentamicin groups, relative to the amoxicillin + cefotaxime group in early life (day 3-17 and 2-16 respectively), whereas this reversed at a later age (eTables 6-8). Moreover, *Acinetobacter* spp. abundance was increased in the amoxicillin + cefotaxime and penicillin + gentamicin groups, but not the co-amoxiclav + gentamicin group. Comparisons between the three different regimens, adjusted for prepartum maternal antibiotics, showed, among others, that the abundance of *Bifidobacterium* genera was more reduced in the amoxicillin + cefotaxime group compared to the penicillin + gentamicin group (with the exception of *Bifidobacterium animalis*, eTables 6-8).

Effects of early-life broad-spectrum antibiotics on infant gut antimicrobial resistance gene profiles

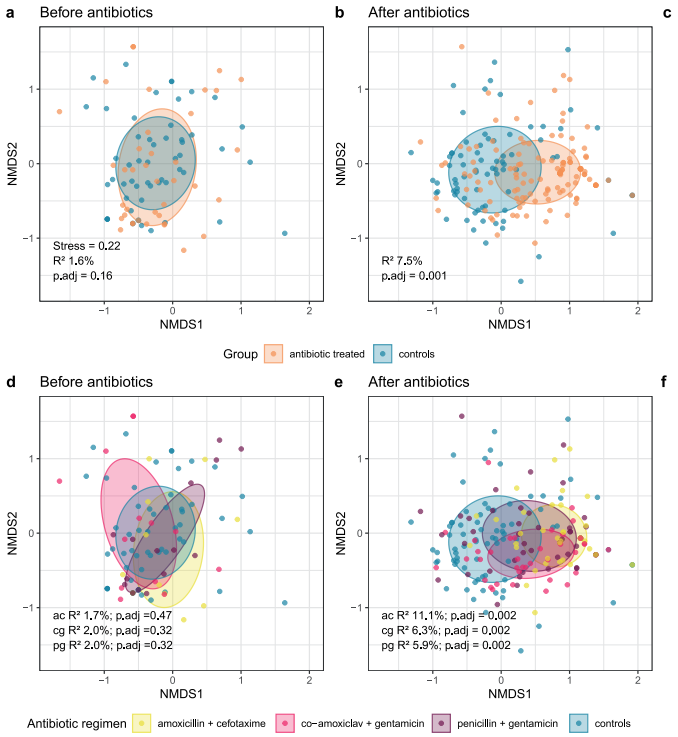
All 31 AMR genes included in the Fluidigm assay were detected in this study. The AMR genes least often detected were the aminoglycoside resistance gene *aph(2")-I(de)* and the vancomycin resistance gene *vanA*, which were both present in only 6/939 samples, while the gene that was most commonly present was the multidrug efflux pump unit *acrA*, found in Gram-negative bacteria (708/939 samples).[51] Other commonly observed genes, present in >80% of all infants, were the aminoglycoside resistance genes *aac(3')-Ii(acde)*, *aac(6')-Ii*, *aph(3')-III*, and the *aadE-like* gene, the beta-lactamase encoding genes bla_{ampC} and $bla_{CTX\text{-}M}$, the macrolide resistance gene *ermB* and the tetracycline resistance gene *tetQ*.

The AMR gene diversity, i.e. number of observed AMR genes detected within a sample, only differed significantly between the antibiotic treated infants and controls at 1 month of age (Wilcoxon test, median observed number of genes in antibiotic treated infants 9, versus 7.5 in controls, p=0.02). Temporal analysis showed no significant difference in the diversity of AMR genes between the antibiotic treated group and controls over the first year of life. Also, no significant inter-regimen differences were found in cross-sectional nor temporal analyses.

On the level of AMR gene composition, or AMR gene profile, significant differences were found at multiple time points between antibiotic treated infants and controls and also between the separate regimens compared with controls. Importantly, before start of antibiotic treatment in the infants with sEONS, the gene profiles largely overlapped between the to-be treated neonates and controls (PERMANOVA, R^2 1.6%, p.adj=0.16; Figure 4a). After antibiotic treatment, however, a major shift was observed in AMR gene profile between the treated infants and controls (R^2 7.5%, p.adj=0.001, Figure 4b). The AMR gene profile slowly normalized over time (1 and 4 months; R^2 5.9% and 2.4%, respectively, both p.adj=0.001), and did not show a significant difference at 12 months of age (R^2 0.6%, p.adj=0.23, Figure 4c). The regimen amoxicillin + cefotaxime showed the biggest effect on AMR gene profile, with an R^2 of 11.1% (p.adj=0.002, Figure 4e) immediately after antibiotic treatment versus R^2 of 6.3 and 5.9 for the co-amoxiclav + gentamicin and penicillin + gentamicin, respectively. The co-amoxiclav + gentamicin regimen had most persistent effects (at 4 months R^2 3.7%, p.adj=0.002, versus R^2 2.5% for penicillin + gentamicin and 2.4% for amoxicillin + cefotaxime, Figure 4f). Again, to exclude the potential background noise caused by antibiotic treatment during follow-up, we performed a post-hoc analysis on a subset of sEONS and control infants not receiving antibiotics after the first week. Similarly to the post-hoc analysis performed on the overall microbial community composition, the effect size of antibiotics administered in the first week of life on AMR gene composition was in this subset comparable or larger at all time points, and differed significantly at the same time points as in the overall analysis (eTable 9).

Clinical covariates associated with antimicrobial resistance gene profiles

Covariates that were significantly associated with AMR gene profile during follow-up are summarized in eTable 10. Gravidity and duration of ruptured membranes before birth were both associated with mode of delivery (Kruskal-Wallis test, p=0.02 and Wilcoxon test p<0.001, respectively). Consequently, as mode of delivery and the presence of siblings <5 years of age were the only variables that



Sample moment	R² (%)	p.adj
month 1	5.9	0.001
month 4	2.4	0.001
month 12	0.6	0.23

Sample moment	R² (%)	p.adj
month 1 ac	6.5	0.002
month 4 ac	2.4	0.002
month 12 ac	1.2	0.32
month 1 cg	5.8	0.002
month 4 cg	3.7	0.002
month 12 cg	1.1	0.37
month 1 pg	6.7	0.002
month 4 pg	2.5	0.005
month 12 pg	1.2	0.29

Figure 4. NMDS plots of AMR gene profile stratified per regimen and timepoint. Non-metric multidimensional scaling (nMDS) plots, based on Jaccard index of binary (presence/absence) data, visualising the antimicrobial resistance (AMR) gene profiles stratified for antibiotic treated infants and controls (**a**, **b**) and per regimen compared with controls (**d**, **e**), for the time points before and immediately after cessation of antibiotic treatment. Each data point represents the composition of AMR genes of one sample. The ellipses represent the standard deviation of data points belonging to each group, with the centre points of the ellipses calculated using the mean of the coordinates per group. The stress of the ordination, effect sizes (R^2) calculated by permutational multivariate analysis of variance (PERMANOVA)-tests and corresponding adjusted p-values (p.adj, calculated using the Benjamini-Hochberg method) are shown in the plots. In panels **c** and **f** the results of the PERMANOVA-tests for the later time points are summarized. Ac = amoxicillin + cefotaxime, cg = co-amoxiclav + gentamicin, pg = penicillin + gentamicin.

both differed between the groups at baseline and had a significant effect on AMR gene composition, these variables were corrected for in all downstream resistome analyses when comparing antibiotic treated infants and controls. In the inter-regimen comparisons, we adjusted for prepartum maternal antibiotics. When studying the effect of antibiotic treatment duration categorized into a short (1-4 days) or long (>4 days) course, we found no significant difference in AMR gene

composition at any time point between the infants receiving a short or long course in their first week of life (eTable 11).

Differences in antimicrobial resistance gene abundances between antibiotic treated infants and controls

In a temporal analysis, 16 AMR genes were found to be significantly differentially abundant between the antibiotic treated infants and controls (eTable 12). Ten of these were enriched in the antibiotic treated group, and consisted of the aminoglycoside resistance genes *aac(6')-aph(2")*, *aac(6')-Ib*, *aac(6')-Ii* and *aph(3')-III*, the beta-lactamases bla_{CMY-2}, bla_{TEM}, the penicillin binding protein 2A gene *mecA*, macrolide resistance genes *ermB* and *ermC* and the colistin resistance gene *mcr-1*. The time intervals in which the differences existed varied from 68 days (*aac(6')-Ib*) to 347 days (371-25*,* bla_{CMY-2}). Six AMR genes, such as the commonly present *tetQ* gene, were more abundant in controls.[52]

When comparing each separate antibiotic regimen with the control group, we found that the penicillin + gentamicin group had, out of the tested panel, the lowest amount of enriched AMR genes (5), versus 10 in both the amoxicillin + cefotaxime and co-amoxiclav + gentamicin groups (eTables 13-15). The genes *aac(6')-aph(2)*, bla_{CMY-2}, *ermB*, *ermC* and *mecA* were consistently more abundant in each of the three regimens compared to the controls. In the inter-regimen comparisons, again the penicillin + gentamicin regimen resulted in the enrichment of the lowest number of AMR genes compared to the other two regimens (2 AMR genes, versus 9 for the amoxicillin + cefotaxime and 5 for the co-amoxiclav + gentamicin group, eTables 16-18). Unexpectedly, the amoxicillin + cefotaxime group showed an enrichment of the aminoglycoside resistance genes *aac(6')-Ii* and *aph(3)-Ia, -Ic* compared to the penicillin + gentamicin group and of the *aadE* gene compared to the co-amoxiclav + gentamicin group.

When studying the correlation between the relative abundance of OTUs and AMR genes (eTable 19), we found that the abovementioned gene *aac(6')-Ii* had the strongest positive correlation with *E. faecium* (Pearson's r -0.50, p.adj<0.001), which was highly abundant in the infants treated with amoxicillin + cefotaxime (negative coefficients here indicate a positive correlation between OTU and gene abundance, as a low Ct value indicates high abundance of an AMR gene). In this way, we could also positively correlate the *aac(6')-aph(2)* gene, which was consistently more abundant in all three regimens compared to controls, to *E. faecium* (Pearson's r -0.42, p.adj<0.001). As expected, *mecA* was positively correlated with *S. epidermidis* (Pearson's r -0.58, p.adj<0.001). *E. coli* was most

strongly positively correlated with *bla*$_{ampC}$ and *acrA* (Pearson's r -0.69 and -0.65, both p.adj<0.001), all of which were consistently more abundant in the controls.

MGS resulted in 1504 AMR genes in this study. Using a cross-sectional differential abundance analysis, 262 genes were found to differ significantly in abundance between the antibiotic treated and control infants, which were all more abundant in the treated infants (eTable 20). MGS confirmed the Fluidigm results of, among others, *aac(6')-aph(2)*, *ermB*, *ermC* and *mecA* genes being enriched in the antibiotic treated infants (zero-inflated Gaussian mixture model, log2 fold change (log2FC): 8.56, p.adj<0.001; 6.10, p.adj=0.001; 4.95, p.adj=0.001; 6.49, p.adj<0.001). Additionally, MGS identified statistically significant enrichment of AMR genes in antibiotic treated infants not included in our Fluidigm panel, including a broad set of genes coding for resistance against rifampin, fluoroquinolones, aminocoumarins, daptomycin, elfamycins, trimethoprim, sulfonamides, fusidic acid and bacitracin (eTable 20).

DISCUSSION

Suspicion of EONS is a highly common indication for broad-spectrum antibiotic treatment in neonates in the first days of life.[12,13] However, in this early period of life, the neonate is also becoming seeded with microbes, assembling its very own unique microbial ecosystem. Therefore, broad-spectrum antibiotics in this life phase might have a significant impact on the development and ultimate composition of the infant microbiome, with potential short and long-term consequences. We studied the ecological effects of antibiotic treatment in early life, and also attempted to identify the regimen which causes least ecological harm. To this purpose we randomized sEONS infants over three different broad-spectrum antibiotic regimens.

We show major effects of broad-spectrum antibiotics on microbial diversity, community composition and AMR gene selection. Although in most neonates antibiotics were already aborted after 48 hours, the impact was still measurable at 12 months of life. We also observed marked differences between the three different antibiotic regimens, suggesting that the combination of penicillin + gentamicin causes least ecological damage.

We observe much more outspoken and prolonged effects of antibiotics than anticipated based on previous studies.[53–55] However, the current body of evidence mostly originates from older children and adults, in whom a more established, and

thus stable and resilient microbiome is to be expected. Therefore, we hypothesize that through the early interruption of microbiome development, our treated infant population did not have a "normal" state to return to.[54] Following, these young infants may have difficulty in regaining typical commensal microbiota from their environment in order to restore a natural developmental trajectory. Relevant infant studies are scarce, and usually performed in preterm infants, with short follow-up.[56,57]

Given the extensive and prolonged effects of the studied early-life antibiotic regimens upon the infants' microbiome, the consequences for the natural process of immune priming and maturation might also be considerable. In the last decade, microbiome-based studies have shown a clear relationship between microbiome perturbation, especially in early-life, and inflammation-driven health problems.[58] For example, a reduced α-diversity of faecal microbiota has been associated with the development of allergic disease and diabetes later in life.[59–62] Given the extent of microbiota perturbations observed in our study, this warrants further research.

On the level of individual taxa, we found especially various beneficial *Bifidobacterium* spp. to be heavily affected by antibiotic treatment. Bifidobacteria are known to promote gut health and provide defence against pathogens.[63] These bacteria are also essential for the digestion of (human) milk oligosaccharides,[64] which in the first 4-6 months of life are the sole food source for infants. Hence, potential effects on infant growth and development are imaginable when the abundance of these bacteria is decreased. Also, the extensive outgrowth of potential pathogenic bacteria such as *Klebsiella* and *Enterococcus* spp. warrants further studies into the susceptibility to infections following early-life antibiotic treatment. In a previous study, antibiotics in early-life have been associated with increased rates of diarrhoea in early childhood, which we now hypothesize might be a consequence of persistent microbial dysbiosis.[65]

Importantly, when comparing the antibiotic regimens, we found that all three showed significant ecological side-effects. However, the amoxicillin + cefotaxime regimen showed most outspoken ecological effects, especially on overall microbial community composition, stability of microbiota development, and AMR gene profile shift directly following treatment. The co-amoxiclav + gentamicin regimen showed the most obvious reduction in microbial diversity and most persistent effects on AMR gene profile. Importantly, the penicillin + gentamicin regimen showed the least detrimental effects on all these parameters. In part, this can be explained by the minimal penetration into the gut lumen of aminoglycosides given intravenously.[66]

While this regimen might clinically be the least popular due to the frequency of penicillin administration and the need for gentamicin serum level monitoring, we think this regimen deserves reconsideration for the treatment of sEONS on neonatal wards given the uniformly lower ecological side-effects observed, and considering the three regimens compared are equally adequate treatments for this indication.[14] Interestingly, some AMR genes were more present in controls. While this may be counterintuitive, it is known that many commensals carry AMR genes, which we confirmed by correlating the AMR genes enriched in controls with the presence and abundance of commensal bacteria, such as the bla_{AMPC} gene in *E. coli*.[52,67]

There is a rapid increase in awareness regarding side-effects of antibiotics on selection and development of AMR in pathogenic bacteria on population level. Most hospitals currently enforce antimicrobial stewardship programmes that ensure appropriate antibiotic therapy. However, these programmes do not consider ecological side-effects yet, largely because information is lacking. Presently, the main focus is to shorten the duration of broad-spectrum antibiotic treatment as much as possible. However, in our study, although a modest effect was observed for antibiotic treatment duration in association with microbiota composition, the effect was negligible compared to the initiation of antibiotics in the first place. Currently, antibiotics are prescribed in 4-10% of all neonates, whereas only an estimated 1 in 1000 will develop a proven infection, likely resulting in unnecessary treatment of >90% of all treated children.[12,13,68] For this reason, more emphasis on improving the diagnostic accuracy of EONS is crucial, as the principle of "when in doubt, there's no harm in treating" appears far from true. Reducing the number of neonates in whom we initiate broad-spectrum antibiotic treatment is feasible, given the fact that prescription rates vary widely even between European countries with similar infection statistics.[69,70] This strongly suggests that guidelines underpinning these national differences should be compared to identify factors that could lead to more reserved prescription practices. Moreover, ongoing efforts to improve the prediction of EONS, for instance through the application of the EONS calculator, could further help to reduce unnecessary antibiotic treatment, even in countries with already low prescription rates.[71] Altogether, this would ideally result in – first and foremost – safe treatment guidelines, whilst simultaneously preserving early-life microbiome development at this critical stage of life.

Strengths of this study include analysing samples before the start of antibiotic treatment to eliminate baseline differences as potential confounders. Importantly, albeit modest differences in overall microbiota composition and AMR gene profiles existed in our study between neonates with sEONS and controls at enrolment

(before antibiotics), these were both non-significant and of limited relevance compared to the observed antibiotic effects. Furthermore, we used a randomized study design to compare the impact of three commonly used broad-spectrum regimens. Additionally, we validated our results through an independent MGS technique. The results of this study are highly applicable to other hospitals and countries, as generally common antibiotic regimens were compared.[14]

A potential weakness of this study is its molecular or epidemiological nature, and as such, causality between antibiotic treatment and observed ecological effects is not proven. However, the likelihood that our observations are a direct biological consequence of antibiotics is high. This is supported by the fact that our findings are in line with known microbial susceptibility patterns of observed bacterial species. Furthermore, because of ethical reasons, we had to include a separate control group, instead of integrating controls in our randomized design. Also, due to clinical, practical and safety reasons we could not perform blinding. Potential selection bias may have occurred due to health-minded parents that have finished higher education being more motivated to participate in our study. Possibly this has resulted in the high percentage of participating parents that have finished higher education and a low percentage of inhouse smoking, and may limit the generalizability of our results.

In conclusion, we found significant long-term effects of broad-spectrum antibiotic treatment for sEONS. We believe our data suggest that more emphasis should be put on reducing the number of neonates that receive broad-spectrum antibiotics for sEONS, and if needed, to preferably prescribe penicillin + gentamicin, as this regimen causes least ecological side-effects.

ACKNOWLEDGEMENTS

The authors are grateful for the participation of all the children and their families. We wish to acknowledge all the members of the research team of the Spaarne Gasthuis Academy and the Neonatology Departments of the Spaarne Gasthuis Hoofddorp and Haarlem, Diakonessenhuis Utrecht and Tergooiziekenhuis Blaricum for the help in participant recruitment, and the laboratory staff, with special thanks to Raiza Hasrat of the University Medical Center Utrecht. We would also like to acknowledge Edinburgh Genomics for executing the MGS sequencing and Katherine Emelianova for the bioinformatic processing of the MGS sequences. This research was funded by the Netherlands Organisation for Health Research and Development

(ZonMw Priority Medicines Antimicrobiële Resistentie project number 205300001), Chief Scientist Office (SCAF/16/03) and the Spaarne Gasthuis. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

REFERENCES

1. Oosterloo, B. C. *et al.* Wheezing and infantile colic are associated with neonatal antibiotic treatment. *Pediatr. Allergy Immunol.* **29**, 151–158 (2018).
2. Alm, B. *et al.* Neonatal antibiotic treatment is a risk factor for early wheezing. *Pediatrics* **121**, 697–702 (2008).
3. Low, J. S. Y. *et al.* Ratio of Klebsiella/Bifidobacterium in early life correlates with later development of paediatric allergy. *Benef. Microbes* **8**, 681–695 (2017).
4. Salvatore, S. *et al.* Neonatal Antibiotics and Prematurity Are Associated with an Increased Risk of Functional Gastrointestinal Disorders in the First Year of Life. *J. Pediatr.* **212**, 44–51 (2019).
5. Korpela, K. *et al.* Childhood BMI in relation to microbiota in infancy and lifetime antibiotic use. *Microbiome* **5**, 26 (2017).
6. Oosterloo, B. C. *et al.* Neonatal Antibiotic Treatment Is Associated With an Altered Circulating Immune Marker Profile at 1 Year of Age. *Front. Immunol.* **10**, (2020).
7. Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **8**, 343ra82-343ra82 (2016).
8. Reyman, M. *et al.* Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. *Nat. Commun.* **10**, 4997 (2019).
9. Gasparrini, A. J. *et al.* Persistent metagenomic signatures of early-life hospitalization and antibiotic treatment in the infant gut microbiota and resistome. *Nat. Microbiol.* **4**, 2285–2297 (2019).
10. Buelow, E. *et al.* Comparative gut microbiota and resistome profiling of intensive care patients receiving selective digestive tract decontamination and healthy subjects. *Microbiome* **5**, 88 (2017).
11. Fjalstad, J. W., Esaiassen, E., Juvet, L. K., van den Anker, J. N. & Klingenberg, C. Antibiotic therapy in neonates and impact on gut microbiota and antibiotic resistance development: a systematic review. *J. Antimicrob. Chemother.* **73**, 569–580 (2018).
12. Mukhopadhyay, S., Eichenwald, E. C. & Puopolo, K. M. Neonatal early-onset sepsis evaluations among well-appearing infants: projected impact of changes in CDC GBS guidelines. *J. Perinatol.* **33**, 198–205 (2013).
13. Escobar, G. J. *et al.* Stratification of Risk of Early-Onset Sepsis in Newborns ≥ 34 Weeks' Gestation. *Pediatrics* **133**, 30–36 (2014).
14. Muller-Pebody, B. *et al.* Empirical treatment of neonatal sepsis: Are the current guidelines adequate? *Arch. Dis. Child. Fetal Neonatal Ed.* **96**, (2011).
15. De Man, P., Verhoeven, B. A. N., Verbrugh, H. A., Vos, M. C. & Van Den Anker, J. N. An antibiotic policy to prevent emergence of resistant bacilli. *Lancet* **355**, 973–978 (2000).
16. Nederlandse Vereniging voor Kindergeneeskunde. Richtlijn Koorts in de tweede lijn bij kinderen van 0-16 jaar. (2013). Available at: https://www.nvk.nl/Portals/0/richtlijnen/koorts/koortsrichtlijn.pdf.
17. Bosch, A. A. T. M. *et al.* Development of Upper Respiratory Tract Microbiota in Infancy is Affected by Mode of Delivery. *EBioMedicine* **9**, 336–345 (2016).
18. Bogaert, D. *et al.* Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. *PLoS One* **6**, e17035 (2011).

19. Mattiello, F. *et al.* A web application for sample size and power calculation in case-control microbiome studies. *Bioinformatics* **32**, 2038–2040 (2016).
20. Reyman, M., van Houten, M. A., Arp, K., Sanders, E. A. M. & Bogaert, D. Rectal swabs are a reliable proxy for faecal samples in infant gut microbiota research based on 16S-rRNA sequencing. *Sci. Rep.* **9**, 16072 (2019).
21. Zaura, E., Keijser, B. J. F., Huse, S. M. & Crielaard, W. Defining the healthy core microbiome of oral microbial communities. *BMC Microbiol.* **9**, 259 (2009).
22. Biesbroek, G. *et al.* Deep sequencing analyses of low density microbial communities: working at the boundary of accurate microbiota detection. *PLoS One* **7**, e32942 (2012).
23. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–20 (2013).
24. 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.* 4516–22 (2011). doi:10.1073/pnas.1000080107
25. Bosch, A. A. T. M. *et al.* Maturation of the Infant Respiratory Microbiota, Environmental Drivers, and Health Consequences. A Prospective Cohort Study. *Am. J. Respir. Crit. Care Med.* **196**, 1582–1590 (2017).
26. Joshi, N. A. & Fass, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). (2011). Available at: https://github.com/najoshi/sickle.
27. Nikolenko, S. I., Korobeynikov, A. I. & Alekseyev, M. A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* **14**, S7 (2013).
28. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* **13**, 31 (2012).
29. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–6 (2010).
30. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
31. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
32. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
33. Subramanian, S. *et al.* Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**, 417–21 (2014).
34. Davis, N. M., Proctor, D., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *bioRxiv* 221499 (2017). doi:10.1101/221499
35. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).

36. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
37. Segata, N. *et al.* Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* **9**, 811–814 (2012).
38. Lakin, S. M. *et al.* MEGARes: An antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res.* **45**, D574–D580 (2017).
39. Broad Institute. Picard Tools. *Broad Institute, GitHub Repos.* https://broadinstitute.github.io/picard/ (2018).
40. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
41. Buelow, E. *et al.* Limited influence of hospital wastewater on the microbiome and resistome of wastewater in a community sewerage system. *FEMS Microbiol. Ecol.* **94**, (2018).
42. R Core Team. R: A language and environment for statistical computing. https://www.r-project.org/. (2017).
43. RStudio Team. RStudio: Integrated Development for R. http://www.rstudio.com/. (2016).
44. Wickham, B. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2009).
45. Kassambara, A. ggpubr: 'ggplot2' Based Publication Ready Plots. https://CRAN.R-project.org/package=ggpubr. (2018).
46. Benjamini, Y. & Hochberg, Y. Controlling The False Discovery Rate–A Practical And Powerful Approach To Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
47. Oksanen, J. *et al.* Community Ecology Package. https://CRAN.R-project.org/package=vegan. (2017).
48. Paulson, J. N., Stine, O. C., Bravo, H. C. & Pop, M. metagenomeSeq: Statistical analysis for sparse high-throughput sequencing. http://cbcb.umd.edu/software/metagenomeSeq. (2016).
49. Paulson, J. N., Stine, O. C., Bravo, H. C. & Pop, M. Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* **10**, 1200–2 (2013).
50. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–7 (2012).
51. Blair, J. M. & Piddock, L. J. Structure, function and inhibition of RND efflux pumps in Gram-negative bacteria: an update. *Current Opinion in Microbiology* **12**, 512–519 (2009).
52. van Schaik, W. The human gut resistome. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370**, 20140087 (2015).
53. 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.* **8**, 343ra81 (2016).
54. Dethlefsen, L. & Relman, D. A. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4554–4561 (2011).

55. Rashid, M.-U. *et al.* Determining the Long-term Effect of Antibiotic Administration on the Human Normal Intestinal Microbiota Using Culture and Pyrosequencing Methods. *Clin. Infect. Dis.* **60**, S77–S84 (2015).
56. Greenwood, C. *et al.* Early Empiric Antibiotic Use in Preterm Infants Is Associated with Lower Bacterial Diversity and Higher Relative Abundance of Enterobacter. *J. Pediatr.* **165**, 23–29 (2014).
57. Gibson, M. K. *et al.* Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat. Microbiol.* **1**, 16024 (2016).
58. Zeissig, S. & Blumberg, R. S. Life at the beginning: perturbation of the microbiota by antibiotics in early life and its role in health and disease. *Nat. Immunol.* **15**, 307–10 (2014).
59. 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.* **128**, 646-652.e5 (2011).
60. Kostic, A. D. *et al.* The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. *Cell Host Microbe* **17**, 260–73 (2015).
61. Vatanen, T. *et al.* The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. *Nature* **562**, 589–594 (2018).
62. 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* **158**,1584-1596 (2020).
63. Tamburini, S., Shen, N., Wu, H. C. & Clemente, J. C. The microbiome in early life: implications for health outcomes. *Nat. Med.* **22**, 713–722 (2016).
64. Marcobal, A. *et al.* Consumption of human milk oligosaccharides by gut-related microbes. *J. Agric. Food Chem.* **58**, 5334–5340 (2010).
65. Rogawski, E. T. *et al.* The Effect of Early Life Antibiotic Exposures on Diarrheal Rates among Young Children in Vellore, India. *Pediatr. Infect. Dis. J.* **34**, 583–588 (2015).
66. Pitt, H. A., Roberts, R. B. & Johnson, W. D. Gentamicin Levels in the Human Biliary Tract. *J. Infect. Dis.* **127**, 299–302 (1973).
67. Flórez, A. B., Ammor, M. S., Álvarez-Martín, P., Margolles, A. & Mayo, B. Molecular analysis of tet(W) gene-mediated tetracycline resistance in dominant intestinal Bifidobacterium species from healthy humans. *Appl. Environ. Microbiol.* **72**, 7377–7379 (2006).
68. Weston, E. J. *et al.* The burden of invasive early-onset neonatal sepsis in the united states, 2005-2008. *Pediatr. Infect. Dis. J.* **30**, 937–941 (2011).
69. Thaulow, C. M. *et al.* Can We Optimize Antibiotic Use in Norwegian Neonates? A Prospective Comparison Between a University Hospital and a District Hospital. *Front. Pediatr.* **7**, (2019).
70. Kerste, M. *et al.* Application of sepsis calculator in newborns with suspected infection. *J. Matern. Fetal. Neonatal Med.* **29**, 3860–5 (2016).
71. Achten, N. B. *et al.* Association of Use of the Neonatal Early-Onset Sepsis Calculator with Reduction in Antibiotic Therapy and Safety: A Systematic Review and Meta-analysis. *JAMA Pediatr.* **173**, 1032–1040 (2019).

SUPPLEMENTARY INFORMATION

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	4-5
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	6
	2b	Specific objectives or hypotheses	6
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	NA
Participants	4a	Eligibility criteria for participants	7
	4b	Settings and locations where the data were collected	7
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	7-8
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	7
	6b	Any changes to trial outcomes after the trial commenced, with reasons	NA
Sample size	7a	How sample size was determined	Supp. Data page 2
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	Supp. Data page 2
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Supp. Data page 2

Continue

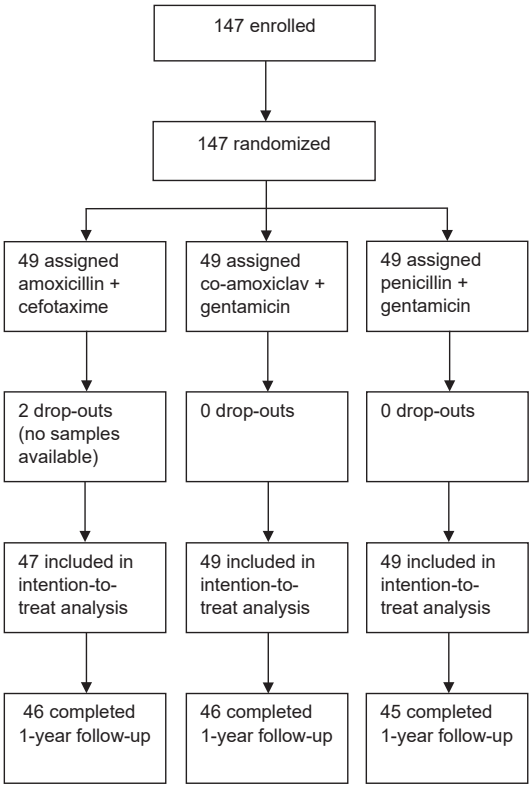
Continued

Section/Topic	Item No	Checklist item	Reported on page No
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Supp. Data page 2
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Supp. Data page 2
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	NA
	11b	If relevant, description of the similarity of interventions	NA
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	8-9 and Supp. Data page 4-5
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	NA
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	10
	13b	For each group, losses and exclusions after randomisation, together with reasons	10
Recruitment	14a	Dates defining the periods of recruitment and follow-up	10
	14b	Why the trial ended or was stopped	NA
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	24-25
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	10
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	10-17
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	10-17
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	NA
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	NA

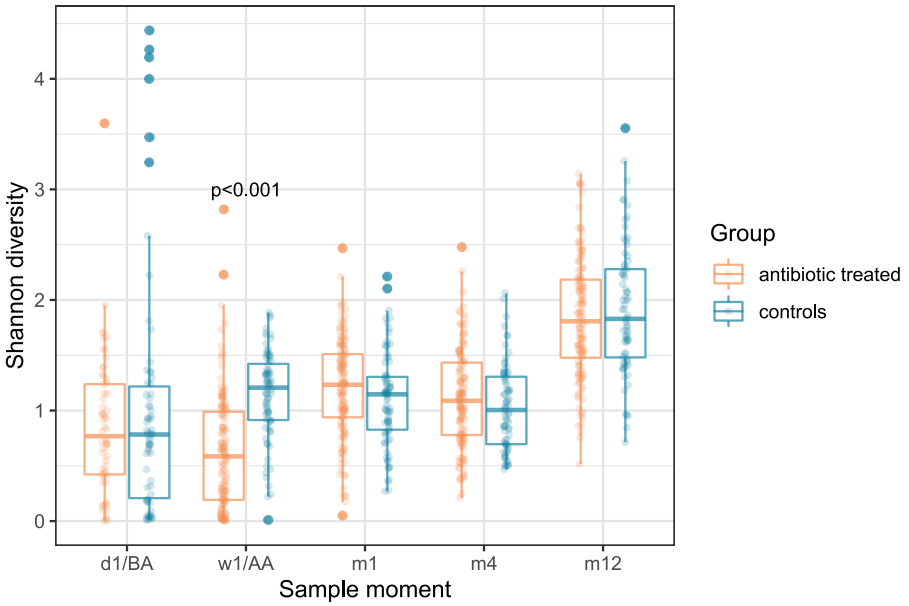
Continue

Continued

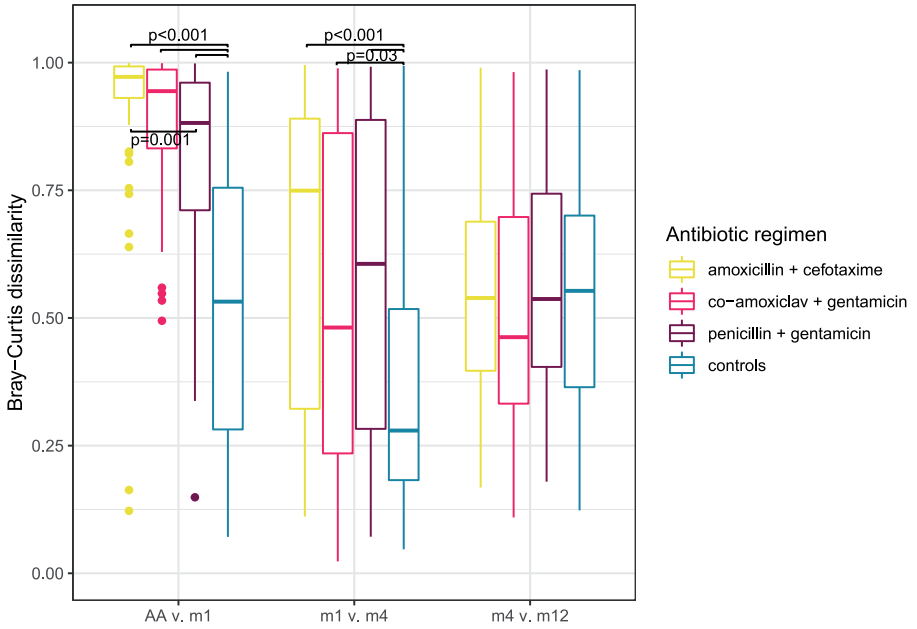
Section/Topic	Item No	Checklist item	Reported on page No
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	21
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	21
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	18-21
Other information			
Registration	23	Registration number and name of trial registry	5
Protocol	24	Where the full trial protocol can be accessed, if available	NA
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	22



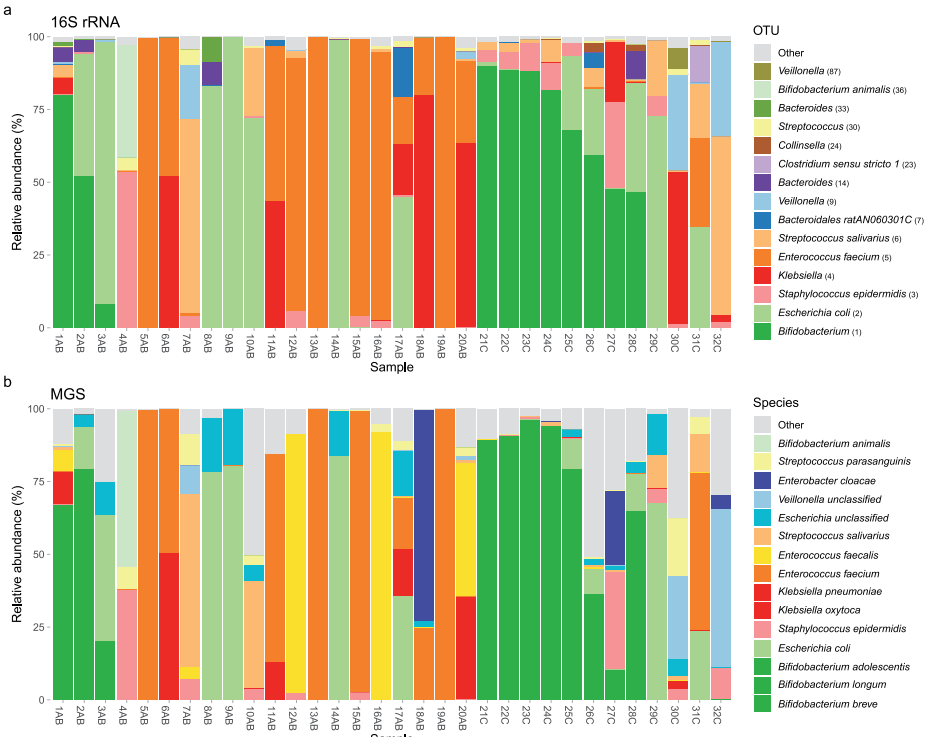
eFigure 1. Recruitment flowchart. Flowchart showing the number of children enrolled and finally included in the intention-to-treat analysis.



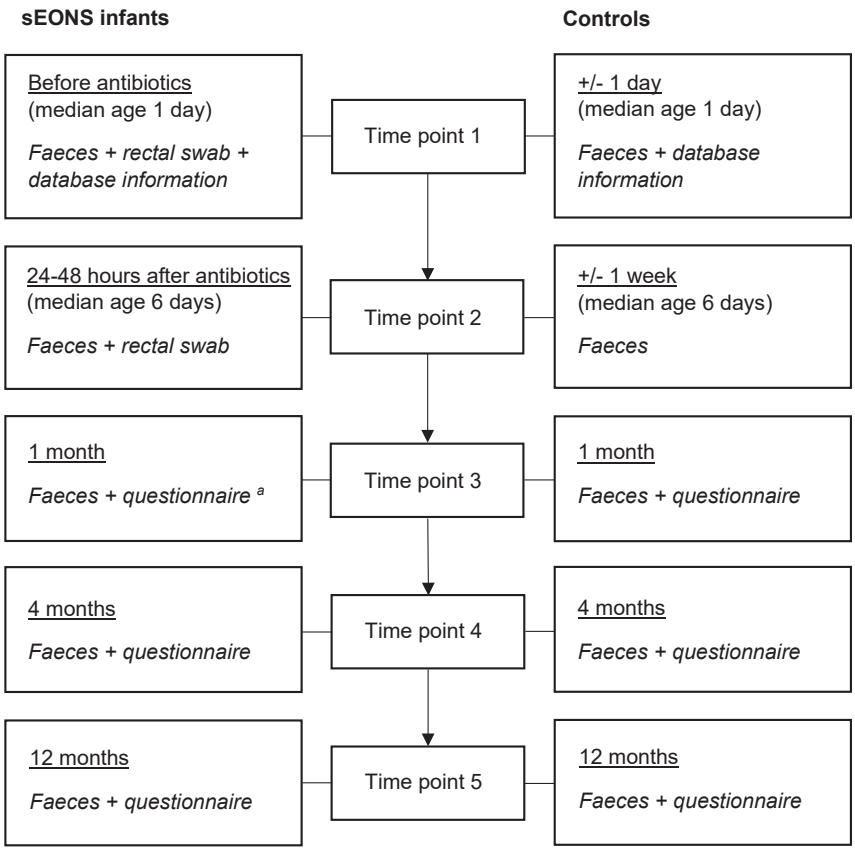
eFigure 2. α-diversity in antibiotic treated infants and controls. Differences in Shannon diversity of the faecal microbiota between antibiotic treated children and controls plotted per timepoint. Group differences were calculated using Wilcoxon tests. Boxplots with medians are shown; the lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles); the upper and lower whiskers extend from the hinge to the largest and smallest value no further than 1.5*IQR from the hinge; outliers are plotted individually by opaque circles; translucent circles visualize all data points. D = day, w = week, BA = before antibiotics, AA = after antibiotics, m = month.



eFigure 3. Microbiota stability over time. As a measure of microbiota stability, we calculated the Bray-Curtis distance between consecutive sample pairs belonging to each individual per time interval, i.e. between stop of antibiotics and month 1, between month 1 and month 4, and between month 4 and month 12. Boxplots with medians are shown; the lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles); the upper and lower whiskers extend from the hinge to the largest and smallest value no further than 1.5*IQR from the hinge; outliers are plotted individually. AA = after antibiotics, m = month.

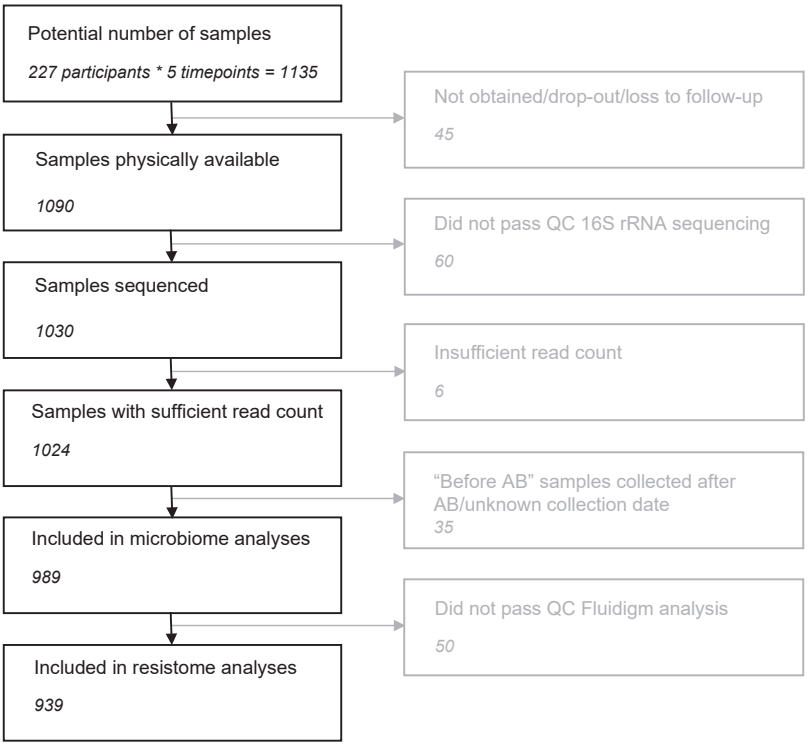


eFigure 4. Comparison of 16S rRNA OTUs with species found by MGS. A random subset of 32 samples was analysed by metagenomic shotgun sequencing (MGS). The mean relative abundances of the 15 most abundant 16S rRNA OTUs were compared to the species found by MGS sequencing. Taxa approaching a similar annotation in both methods are coloured similarly to allow for easier visual comparison. AB = antibiotic treated, C = control.

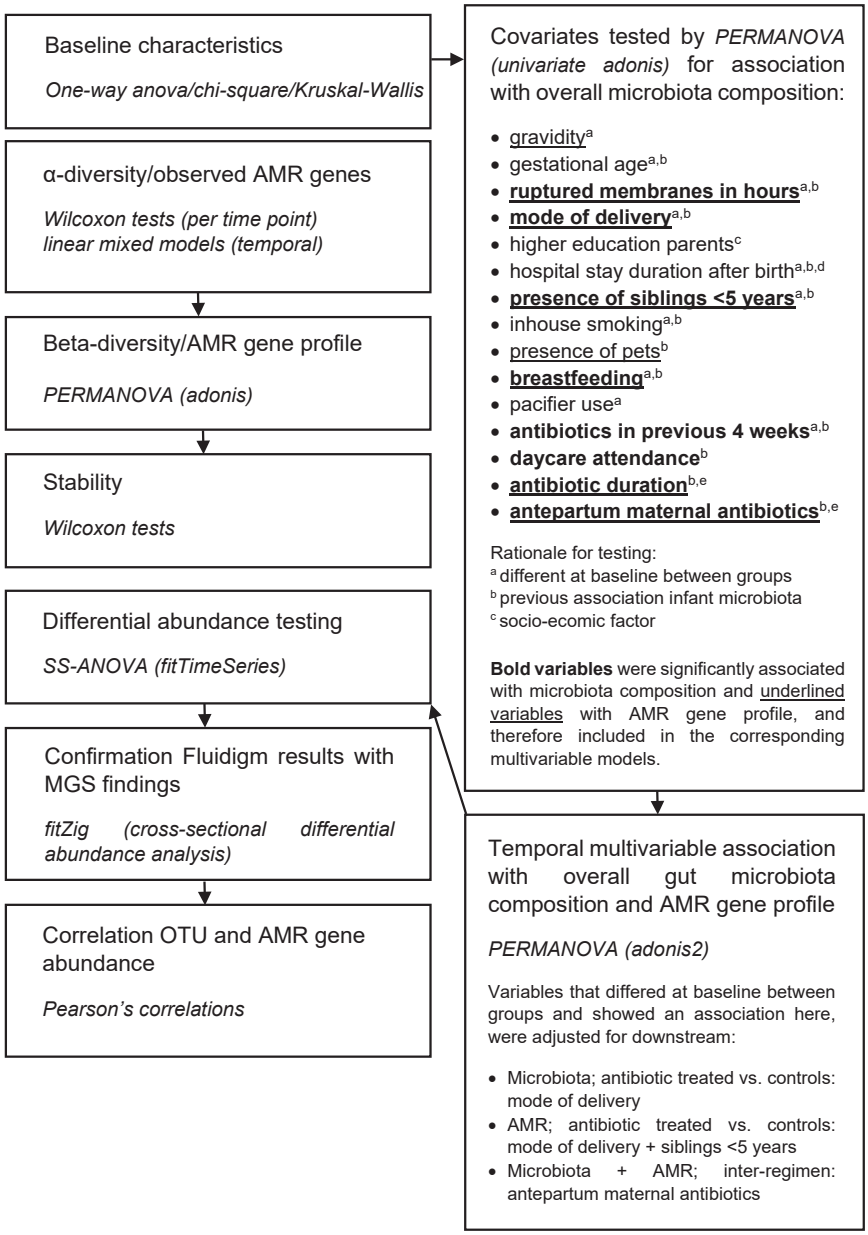


eFigure 5. Sample moments

[a]For the sEONS infants additional questionnaire information was collected at 6, 8 and 10 months of age; for the control infants this corresponded with questionnaires collected at 6 and 9 months of age.



eFigure 6. Sample flowchart. Flowchart showing number of samples aimed for, obtained, and of sufficient quality for analyses. QC = Quality Control, AB = antibiotics.



eFigure 7. Statistical analysis scheme

[d]Tested in stratified manner for the antibiotic treated infants and controls, because of collinearity with antibiotic treatment in the first week of life. [e]Variables only present in antibiotic treated infants. AMR = antimicrobial resistance, anova = analysis of variance, PERMANOVA = permutational multivariate ANOVA, SS-ANOVA = smoothing spline ANOVA, MGS = metagenomic shotgun sequencing, OTU = Operational Taxonomical Unit.

eTable 1. Effect of antibiotic treatment in the first week of life on overall gut microbiota composition in infants not receiving antibiotics later in life

Time point	R^2 (%)	Adjusted p-value
Before antibiotics	1.9	0.09
After antibiotics	11.2	0.001
Month 1	4.8	0.001
Month 4	1.5	0.05
Month 12	1.4	0.06

Effect sizes (R^2) and adjusted p-values are shown for the effect of antibiotics in the first week of life on overall gut microbiota composition of the subset of participants not receiving antibiotics later in life.

eTable 2. Effect of short (1-4 days) versus long (>4 days) antibiotic treatment in the first week of life on overall microbiota composition

Time point	R^2 (%)	Adjusted p-value
Before antibiotics	1.3	0.63
After antibiotics	1.3	0.17
Month 1	1.3	0.17
Month 4	0.7	0.54
Month 12	1.3	0.17

Effect sizes (R^2) and adjusted p-values are shown for the effect of short versus long antibiotic treatment in the first week of life on overall microbiota composition.

eTable 3. FitTimeSeries results of all significantly differentially abundant taxa between antibiotic treated infants and controls

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Bifidobacterium_1	1	1	36	-285	0.0054	1-43	1-31	1-38
Bifidobacterium_1	2	119	123	-5	0.0054	NA	115-119	NA
Escherichia_coli_2	1	1	181	-236	0.0117	1-142	NA	1-177
Enterococcus_faecium_5	1	1	121	274	0.0054	1-117	1-106	1-119
Staphylococcus_epidermidis_3	1	1	229	-225	0.0206	12-175	1-29	1-211
Klebsiella_4	1	1	122	306	0.0161	16-127	1-256	1-106
Streptococcus_salivarius_subsp_thermophilus_6	1	1	30	-106	0.0054	1-27	1-30	1-27
Blautia_8	1	359	375	-22	0.0377	NA	*54-95	NA
Blautia_8	2	NA	NA	NA	NA	NA	neg 126-162	NA
Blautia_8	3	NA	NA	NA	NA	NA	neg 358-411	NA
ratAN060301C_7	1	1	388	-712	0.008	1-343	1-310	NA
Ruminococcus_gnavus_CC55_001C_10	1	1	195	-309	0.0117	1-64	1-151	1-238
Streptococcus_11	1	1	411	768	0.0195	1-384	1-411	1-403
Bifidobacterium_12	1	1	19	-51	0.0054	1-18	1-17	1-19
Bifidobacterium_12	2	35	88	133	0.0311	NA	NA	NA
Bifidobacterium_bifidum_NCIMB_41171_13	1	1	39	-149	0.0054	1-213	1-34	1-43
Bifidobacterium_bifidum_NCIMB_41171_13	2	97	128	-94	0.0054	NA	99-121	NA
Veillonella_9	1	2	15	-62	0.0054	4-17	5-18	1-21
Veillonella_9	2	117	128	25	0.0054	119-127	118-127	118-123
Bacteroides_14	1	1	255	-425	0.0117	NA	1-244	1-169
Clostridium_sensu_stricto_1_15	1	1	10	-20	0.008	1-10	1-10	1-2
Clostridium_sensu_stricto_1_15	2	21	136	458	0.0054	23-127	25-130	23-142
Lactobacillus_paracasei_18	1	1	166	-201	0.0285	NA	NA	NA

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Fusicatenibacter_saccharivorans_20	1	NA	NA	NA	NA	pos 78-384	NA	pos 364-403
Blautia_25	1	33	411	-433	0.0317	NA	34-411	NA
Akkermansia_28	1	71	411	-338	0.0377	NA	NA	NA
Collinsella_24	1	1	356	-1172	0.013	1-334	1-248	1-354
Clostridium_sensu_stricto_1_23	1	1	12	-20	0.0054	1-7	NA	NA
Clostridium_sensu_stricto_1_23	2	25	92	172	0.0054	26-123	NA	NA
Clostridium_sensu_stricto_1_23	3	NA	NA	NA	NA	pos 364-369	NA	NA
Veillonella_29	1	97	197	-296	0.0054	26-289	104-180	52-208
Streptococcus_pyogenes_17	1	1	43	-17	0.0394	1-168	NA	NA
Clostridium_butyricum_26	1	1	385	690	0.013	5-241	1-411	1-363
Streptococcus_30	1	1	11	-16	0.0285	NA	NA	NA
Streptococcus_30	2	30	96	142	0.0098	NA	NA	NA
Peptostreptococcaceae_32	1	NA	NA	NA	NA	pos 51-135	NA	NA
Bacteroides_31	1	1	411	-606	0.008	1-384	1-220	27-242
Parabacteroides_37	1	NA	NA	NA	NA	NA	neg 83-411	NA
Bacteroides_33	1	1	411	-974	0.0054	1-384	1-411	1-403
Bifidobacterium_animalis_36	1	1	23	-51	0.0054	1-24	1-20	1-24
Bifidobacterium_animalis_36	2	47	97	89	0.0161	NA	46-117	NA
Clostridium_butyricum_35	1	1	14	-15	0.0098	1-5	NA	1-14
Clostridium_butyricum_35	2	40	140	135	0.0452	37-146	NA	87-205
Clostridium_paraputrificum_41	1	56	153	148	0.0117	51-369	75-113	NA
Clostridium_paraputrificum_41	2	353	376	31	0.0054	NA	363-374	NA
Lachnospiraceae_39	1	38	411	-411	0.0222	NA	56-411	174-403
Haemophilus_38	1	1	8	-10	0.0145	1-13	NA	NA
Haemophilus_38	2	44	86	70	0.0054	43-75	NA	NA
Lactobacillus_51	1	1	290	-126	0.013	26-289	1-262	1-182
Rothia_43	1	1	6	-6	0.0417	NA	1-17	*24-125
Rothia_43	2	27	125	189	0.0054	NA	59-102	NA

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Lactobacillus_fermentum_49	1	1	120	-58	0.0285	NA	1-231	NA
Streptococcus_anginosus_subsp_whileyi_55	1	1	88	-98	0.0054	1-79	1-65	1-105
Akkermansia_64	1	1	306	-124	0.0145	34-384	NA	31-148
Bacteroides_58	1	8	411	-335	0.0054	81-384	8-340	110-403
Bacillus_cereus_68	1	1	141	21	0.0206	NA	NA	NA
Moryella_60	1	NA	NA	NA	NA	NA	neg 142-378	NA
Corynebacterium_47	1	1	144	57	0.0377	1-170	NA	NA
Bifidobacterium_63	1	1	307	-411	0.0098	1-361	1-260	1-162
Gardnerella_vaginalis_0288E_71	1	1	177	-40	0.0098	NA	1-150	NA
Roseburia_72	1	NA	NA	NA	NA	NA	neg 356-369	neg 243-364
Roseburia_72	2	371	411	284	0.0394	NA	376-411	371-403
Bacteroides_70	1	1	371	-464	0.0098	1-335	1-260	1-301
Lachnospiraceae_86	1	NA	NA	NA	NA	pos 203-384	NA	NA
Megamonas_46	1	NA	NA	NA	NA	NA	neg 164-411	pos 1-257
Lactococcus_lactis_78	1	NA	NA	NA	NA	pos 1-167	pos 74-310	NA
Peptostreptococcaceae_75	1	1	138	-71	0.0161	NA	1-96	52-155
Erysipelotrichaceae_74	1	2	411	-255	0.0098	80-384	78-411	67-293
Actinomyces_73	1	17	170	-73	0.046	43-197	NA	NA
Ruminococcaceae_90	1	NA	NA	NA	NA	NA	NA	neg 319-355
Ruminococcaceae_90	2	NA	NA	NA	NA	NA	NA	pos 383-403
Lachnospiraceae_103	1	NA	NA	NA	NA	pos 120-384	NA	NA
Veillonella_76	1	1	25	-34	0.008	1-51	*81-139	1-42
Tepidimonas_95	1	1	27	24	0.0054	1-15	NA	NA
Peptostreptococcus_98	1	1	183	-154	0.0098	1-166	1-155	1-140

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Peptostreptococcus_98	2	364	411	36	0.0311	NA	396-411	NA
Clostridium_nexile_89	1	127	220	-610	0.008	NA	NA	NA
Clostridium_nexile_89	2	257	352	-749	0.0145	NA	NA	NA
Clostridium_nexile_89	3	366	411	266	0.0054	NA	NA	NA
Eubacterium_sp_CS1_Van_96	1	NA	NA	NA	NA	NA	NA	pos 1-279
Corynebacterium_striatum_77	1	105	411	135	0.0054	61-384	NA	NA
Stenotrophomonas_maltophilia_91	1	1	139	28	0.0195	1-185	NA	NA
Blautia_112	1	1	11	-3	0.0098	NA	NA	*223-403
Blautia_112	2	176	411	124	0.0275	NA	NA	NA
Enhydrobacter_119	1	1	42	20	0.0054	1-33	NA	NA
Clostridium_sensu_stricto_1_124	1	14	120	39	0.0357	1-171	NA	NA
Veillonella_87	1	1	37	-29	0.0054	1-35	1-5	1-121
Coprococcus_99	1	110	411	-237	0.023	NA	NA	193-403
Atopobium_138	1	1	411	-88	0.0054	97-206	95-208	95-209
Bacteroides_48	1	1	228	-98	0.008	46-315	1-95	1-189
Lachnospira_108	1	297	364	711	0.0054	NA	318-363	NA
Lachnospira_108	2	370	388	-28	0.0325	NA	367-374	NA
Parasutterella_88	1	151	411	-163	0.043	NA	NA	NA
Comamonadaceae_130	1	1	26	15	0.0054	1-27	NA	1-27
Nitriliruptor_126	1	1	26	14	0.0054	NA	NA	NA
Varibaculum_117	1	NA	NA	NA	NA	neg 64-214	NA	NA
Anaerococcus_105	1	NA	NA	NA	NA	NA	neg 17-51	pos 127-168
Bacteroides_94	1	NA	NA	NA	NA	NA	pos 1-127	NA
Lachnospiraceae_145	1	NA	NA	NA	NA	NA	neg 146-411	NA
Bilophila_wadsworthia_3_1_6_123	1	1	411	-489	0.0054	1-384	NA	1-403

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Bifidobacteriaceae_122	1	1	134	-92	0.0336	NA	NA	1-143
Veillonella_115	1	1	13	-11	0.0098	1-13	*89-134	NA
Veillonella_115	2	95	136	40	0.0222	NA	NA	NA
Acinetobacter_143	1	1	28	19	0.0054	1-27	NA	1-34
Erysipelotrichaceae_148	1	1	7	-1	0.018	NA	NA	NA
Ruminococcus_sp_14531_180	1	NA	NA	NA	NA	pos 1-255	NA	NA
Subdoligranulum_136	1	NA	NA	NA	NA	NA	neg 40-150	NA
Acinetobacter_157	1	1	151	61	0.0098	1-178	NA	1-82
Aquabacterium_152	1	1	25	16	0.0054	NA	NA	NA
Bifidobacterium_146	1	1	193	-173	0.0054	1-148	1-172	1-173
Dorea_170	1	242	365	-581	0.0145	268-363	NA	NA
Dorea_170	2	373	411	135	0.0377	369-384	NA	NA
Coriobacteriaceae_bacterium_WAL_18889_151	1	1	373	-248	0.023	NA	1-411	NA
Phascolarctobacterium_165	1	1	153	-71	0.0054	NA	1-206	1-89
Phascolarctobacterium_165	2	NA	NA	NA	NA	NA	NA	neg 298-403
Sutterella_134	1	1	349	-197	0.023	1-381	NA	NA
Akkermansia_153	1	15	127	-35	0.0161	59-308	NA	NA
Burkholderia_156	1	1	27	9	0.0394	1-27	NA	NA
Escherichia_Shigella_154	1	1	25	-50	0.0054	1-29	1-16	1-39
Pseudomonas_fluorescens_144	1	1	28	10	0.0098	NA	NA	NA
Coprococcus_sp_DJF_B005_176	1	128	411	-102	0.0472	NA	NA	NA
Bifidobacterium_161	1	NA	NA	NA	NA	NA	NA	neg 88-234
Lachnospiraceae_188	1	82	411	-201	0.0262	NA	82-411	NA
Bifidobacteriaceae_193	1	NA	NA	NA	NA	NA	pos 30-74	NA
Eggerthella_178	1	1	190	-232	0.0098	1-160	1-150	4-189
Eggerthella_178	2	326	411	156	0.008	320-384	320-411	345-403
Bacteroides_194	1	99	411	-153	0.046	139-384	NA	NA
Gemella_120	1	1	38	-83	0.0054	1-32	1-40	1-32

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Gemella_120	2	98	129	31	0.013	113-124	103-132	NA
Bifidobacterium_dentium_Bd1_179	1	1	8	-5	0.008	NA	*30-135	NA
Bifidobacterium_dentium_Bd1_179	2	39	131	86	0.0195	NA	NA	NA
Bacteroides_214	1	49	133	-39	0.0325	NA	13-122	NA
Bifidobacterium_233	1	1	206	-109	0.0117	1-189	NA	1-128
Streptococcus_192	1	1	17	-27	0.0054	1-18	1-18	1-9
Streptococcus_192	2	32	110	135	0.0098	49-62	105-121	34-103
Alistipes_215	1	7	358	-189	0.008	31-351	NA	NA
Lactobacillus_delbrueckii_subsp_bulgaricus_190	1	NA	NA	NA	NA	neg 155-384	NA	NA
Bifidobacterium_dentium_Bd1_198	1	NA	NA	NA	NA	NA	pos 54-134	NA
Bifidobacterium_animalis_202	1	1	25	-21	0.0054	1-51	NA	1-137
Turicibacter_220	1	307	368	-101	0.0417	NA	NA	NA
Turicibacter_220	2	397	411	30	0.0442	NA	NA	NA
Bacteroides_222	1	1	177	-91	0.023	NA	1-175	1-97
Akkermansia_196	1	16	126	-29	0.0161	71-350	NA	NA
Enterococcus_186	1	12	74	65	0.008	1-70	NA	1-123
Collinsella_tanakaei_163	1	1	184	-56	0.0206	32-132	NA	37-117
Actinobaculum_schaalii_FB123_CNA_2_239	1	22	156	-83	0.008	23-143	NA	21-137
Klebsiella_200	1	13	129	233	0.0054	15-132	16-130	12-113
Bifidobacterium_228	1	1	211	-145	0.0195	1-245	1-140	1-130
Bifidobacterium_223	1	1	146	-90	0.0394	NA	NA	1-109
Bacteroides_249	1	NA	NA	NA	NA	neg 277-384	NA	NA
Bifidobacterium_251	1	1	347	-288	0.023	1-239	1-270	1-336
Streptococcus_241	1	1	45	-17	0.013	NA	NA	NA
Ruminococcaceae_243	1	1	141	-75	0.013	NA	1-199	31-89
Bacteroides_210	1	1	411	-309	0.008	1-384	1-200	NA
Epulopiscium_208	1	NA	NA	NA	NA	NA	pos 83-411	NA
Paracoccus_266	1	1	25	7	0.0275	NA	NA	NA

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Anaerostipes_269	1	NA	NA	NA	NA	NA	NA	neg 269-355
Anaerostipes_269	2	NA	NA	NA	NA	NA	NA	pos 366-403
Bacteroides_236	1	11	317	-129	0.0161	108-384	1-136	NA
Parabacteroides_distasonis_207	1	NA	NA	NA	NA	neg 1-292	NA	NA
Lactococcus_189	1	1	93	42	0.008	1-25	NA	1-130
Actinomyces_292	1	NA	NA	NA	NA	pos 149-271	NA	NA
Ruminococcaceae_287	1	209	411	-103	0.0347	187-384	NA	289-403
Collinsella_227	1	1	203	-68	0.013	NA	NA	1-229
Bacteroides_247	1	1	361	-220	0.013	1-312	1-216	NA
Alistipes_255	1	124	411	-133	0.0317	142-384	NA	NA
Collinsella_237	1	1	342	-793	0.0054	1-320	1-262	1-357
Barnesiella_276	1	NA	NA	NA	NA	neg 97-384	NA	NA
Veillonella_sp_DNF00869_201	1	1	21	-17	0.0054	1-22	1-19	1-15
Veillonella_sp_DNF00869_201	2	51	112	44	0.0054	106-120	NA	NA
Streptococcus_338	1	4	8	1	0.046	1-25	362-411	NA
Ruminococcaceae_306	1	NA	NA	NA	NA	NA	NA	neg 277-403
Roseburia_304	1	262	369	-352	0.0262	NA	NA	269-363
Roseburia_304	2	NA	NA	NA	NA	NA	NA	pos 373-403
Dermabacter_335	1	NA	NA	NA	NA	neg 1-163	NA	neg 31-143
ratAN060301C_268	1	1	296	-178	0.0117	1-364	1-197	NA
Streptococcus_303	1	1	162	77	0.0161	NA	1-144	1-109
Bacteroides_fragilis_CL03T00C08_291	1	1	329	-289	0.0098	13-384	6-267	24-183
Pasteurella_pneumotropica_301	1	NA	NA	NA	NA	neg 1-14	NA	NA
Ruminococcaceae_277	1	230	411	-132	0.0054	NA	215-411	222-403
Finegoldia_164	1	1	199	109	0.0347	1-229	NA	1-178

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Acinetobacter_ursingii_ANC_3649_388	1	35	204	52	0.008	NA	NA	NA
Bifidobacteriaceae_290	1	1	28	-14	0.013	NA	NA	1-41
Lachnospiraceae_294	1	NA	NA	NA	NA	NA	neg 109-411	NA
Bacteroides_248	1	NA	NA	NA	NA	NA	neg 1-171	NA
Bacteroides_324	1	NA	NA	NA	NA	NA	neg 1-110	NA
Prevotella_361	1	95	411	-88	0.0145	NA	NA	NA
Oscillospiraceae_bacterium_VE202_24_316	1	NA	NA	NA	NA	NA	NA	neg 93-403
Dialister_365	1	1	397	-98	0.0336	77-384	75-411	NA
Gardnerella_332	1	1	305	-191	0.013	1-384	NA	1-189
Bifidobacteriaceae_320	1	98	219	-237	0.0098	94-175	104-198	51-227
Bifidobacterium_333	1	1	23	-41	0.0054	1-26	1-28	1-26
Streptococcus_gallolyticus_subsp_macedonicus_350	1	1	20	-6	0.0452	NA	NA	NA
Corynebacterium_328	1	1	7	4	0.0394	NA	NA	NA
Lachnospiraceae_326	1	NA	NA	NA	NA	NA	neg 106-411	NA
Escherichia_Shigella_331	1	1	18	-13	0.0054	1-9	NA	NA
Streptococcus_gallolyticus_subsp_macedonicus_386	1	1	184	-46	0.0145	1-173	NA	1-139
Ruminococcus_gnavus_CC55_001C_364	1	1	49	-22	0.0117	1-4	NA	1-130
Blautia_hydrogenotrophica_262	1	95	411	-108	0.0311	NA	NA	NA
Blautia_378	1	NA	NA	NA	NA	NA	neg 118-411	NA
Lactobacillus_371	1	1	106	-35	0.0054	1-140	1-95	1-90
Roseburia_383	1	246	370	-474	0.0206	NA	355-371	250-367
Roseburia_383	2	NA	NA	NA	NA	NA	pos 376-393	pos 376-403

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Bifidobacterium_dentium_Bd1_395	1	22	73	31	0.0317	NA	NA	NA
Bifidobacteriaceae_344	1	1	15	-13	0.0054	1-10	1-6	1-9
Bifidobacteriaceae_344	2	34	71	36	0.0247	*120-134	NA	NA
Bifidobacteriaceae_344	3	124	143	-17	0.0206	NA	NA	NA
Bifidobacterium_368	1	1	225	-145	0.0054	2-211	1-143	1-216
Ruminococcus_gnavus_CC55_001C_366	1	1	220	-107	0.0275	NA	1-130	1-194
Veillonella_300	1	89	132	29	0.013	96-129	99-116	NA
Dialister_302	1	13	349	-79	0.0145	NA	NA	NA
Clostridium_sensu_stricto_1_341	1	1	3	-1	0.0054	NA	NA	*21-124
Clostridium_sensu_stricto_1_341	2	18	84	67	0.0054	NA	NA	NA
Lachnospiraceae_362	1	149	411	-97	0.0317	NA	NA	202-403
Fusobacterium_369	1	1	32	-12	0.0498	NA	NA	NA
Bifidobacterium_336	1	NA	NA	NA	NA	neg 45-216	NA	NA
Clostridium_sensu_stricto_1_355	1	1	7	-3	0.0098	*28-129	NA	NA
Clostridium_sensu_stricto_1_355	2	28	85	35	0.0098	NA	NA	NA
Leuconostoc_275	1	1	20	7	0.008	1-26	NA	*271-372
Leuconostoc_275	2	364	366	-1	0.0145	NA	NA	NA
Bifidobacterium_347	1	25	411	-173	0.0222	40-384	NA	NA
Coprobacter_443	1	1	252	-53	0.0371	49-210	NA	NA
Lachnospiraceae_419	1	NA	NA	NA	NA	pos 204-384	neg 312-366	neg 299-360
Lachnospiraceae_419	2	NA	NA	NA	NA	NA	NA	pos 366-400
Bacteroides_400	1	1	411	-205	0.0054	96-384	1-262	1-294
Streptococcus_gallolyticus_subsp_macedonicus_464	1	1	20	-12	0.0054	1-20	1-16	1-17
Leuconostoc_mesenteroides_282	1	1	22	12	0.0054	1-26	NA	1-204

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Corynebacterium_propinquum_408	1	1	56	-27	0.0285	1-73	NA	NA
Lactobacillus_390	1	64	213	-85	0.0347	NA	NA	NA
Actinomyces_sp_oral_clone_DR002_406	1	NA	NA	NA	NA	pos 18-117	NA	NA
Subdoligranulum_479	1	NA	NA	NA	NA	NA	NA	neg 211-403
Plesiomonas_485	1	58	171	35	0.0247	39-200	NA	NA
Corynebacterium_freneyi_429	1	NA	NA	NA	NA	pos 1-143	NA	NA
Staphylococcus_427	1	1	46	-27	0.008	1-70	1-24	1-72
Clostridium_difficile_630_345	1	NA	NA	NA	NA	neg 1-11	NA	NA
Clostridium_difficile_630_359	1	NA	NA	NA	NA	NA	neg 25-120	NA
Clostridiales_bacterium_20_2a_461	1	27	395	-121	0.0117	NA	84-345	NA
Pseudobutyrivibrio_497	1	294	367	-178	0.0325	NA	NA	249-367
Bifidobacterium_breve_473	1	1	82	-40	0.0054	1-65	NA	1-166
Staphylococcaceae_452	1	1	42	15	0.0317	1-64	NA	NA
Bifidobacterium_466	1	NA	NA	NA	NA	NA	NA	pos 300-348
Streptococcus_521	1	1	175	-43	0.0301	1-178	1-125	NA
Escherichia_Shigella_469	1	19	226	-57	0.0417	1-168	NA	NA
Bacteroides_445	1	1	386	-190	0.008	1-384	NA	1-266
Clostridium_sensu_stricto_1_458	1	NA	NA	NA	NA	NA	NA	pos 117-161
Propionibacterium_434	1	3	67	-32	0.0357	NA	NA	8-91
Prevotella_buccae_ATCC_33574_586	2	366	374	4	0.0325	NA	NA	NA
Streptococcus_447	1	1	180	-61	0.023	NA	NA	NA
Bacteroides_493	1	1	338	-138	0.0117	38-247	6-185	1-267
Veillonella_502	1	NA	NA	NA	NA	neg 1-6	pos 1-228	NA
Barnesiella_516	1	70	411	-65	0.023	NA	NA	NA
Veillonella_454	1	NA	NA	NA	NA	NA	pos 74-344	NA
Blautia_484	1	90	411	-106	0.0336	NA	NA	93-403

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Staphylococcus_507	1	1	18	-11	0.023	1-19	1-18	NA
Ruminococcaceae_532	1	NA	NA	NA	NA	NA	NA	neg 180-403
Bacteroides_482	1	41	411	-111	0.0206	37-111	NA	NA
Bacteroides_482	2	NA	NA	NA	NA	neg 165-384	NA	NA
Pseudobutyrivibrio_570	1	296	365	-261	0.008	NA	NA	246-367
Pseudobutyrivibrio_570	2	372	397	28	0.0311	NA	NA	376-403
Streptococcus_gallolyticus_subsp_macedonicus_554	1	19	106	-15	0.0325	NA	NA	3-201
Staphylococcus_489	1	1	23	-17	0.0054	NA	1-20	1-21
Actinomyces_europaeus_ACS_120_V_Col10b_541	1	36	212	-53	0.0317	NA	NA	NA
Bacilli_471	1	1	36	-33	0.0054	1-69	1-31	1-32
Campylobacter_646	1	19	113	40	0.0054	1-131	NA	NA
Bifidobacterium_animalis_581	1	NA	NA	NA	NA	NA	NA	neg 1-70
Bacteroides_603	1	47	282	-54	0.0377	125-384	NA	NA
Bacteroides_645	1	NA	NA	NA	NA	pos 1-118	NA	NA
Veillonella_555	1	34	63	12	0.039	NA	NA	NA
Anaerococcus_418	1	1	120	15	0.0357	NA	NA	NA
Bacillales_615	1	NA	NA	NA	NA	NA	neg 1-21	NA
Clostridium_sensu_stricto_1_110	1	81	271	60	0.0285	118-136	67-352	NA
Prevotella_melaninogenica_639	1	1	53	-15	0.0317	NA	NA	NA
Bacteroides_535	1	72	411	-100	0.013	89-384	NA	143-403
Peptococcus_664	1	NA	NA	NA	NA	NA	neg 52-135	NA
Ruminococcus_sp_CE2_596	1	102	411	-124	0.0206	107-384	NA	115-403
Corynebacteriaceae_504	1	38	131	-45	0.008	*1-13	NA	NA
Corynebacteriaceae_504	2	NA	NA	NA	NA	neg 39-122	NA	NA
Staphylococcus_565	1	1	21	-18	0.0054	1-23	1-18	1-22
Staphylococcus_565	2	NA	NA	NA	NA	NA	pos 50-65	NA

Continue

Continued

OTU	Int. no	Interval start	Interval end	Area	p.adj	ac	Ag	pg
Methylobacterium_radiotolerans_437	1	1	273	-79	0.013	1-217	1-222	56-196
Lachnospiraceae_633	1	67	411	-77	0.0117	NA	NA	NA
Bifidobacterium_animalis_647	1	NA	NA	NA	NA	pos 9-264	NA	NA
Clostridium_sensu_stricto_1_648	1	42	154	37	0.0054	NA	45-144	NA
Bacteroides_658	1	70	366	-77	0.0206	64-384	NA	NA
Blautia_679	1	NA	NA	NA	NA	NA	neg 110-411	NA
Dialister_sp_oral_taxon_502_636	1	150	411	42	0.0117	NA	NA	NA
Bacteroides_635	1	28	299	-41	0.0275	NA	1-238	NA
Bacteroides_663	1	121	411	-52	0.018	NA	NA	227-403
Lachnospiraceae_652	1	124	137	-5	0.018	NA	NA	115-403
Lachnospiraceae_652	2	194	365	-467	0.0098	NA	NA	NA
Lachnospiraceae_652	3	373	411	77	0.0145	NA	NA	NA
Bifidobacterium_breve_634	1	NA	NA	NA	NA	NA	NA	pos 122-403
Akkermansia_671	1	49	215	-20	0.0394	NA	NA	NA
Lachnospiraceae_660	1	NA	NA	NA	NA	NA	neg 122-411	NA
Bacteroides_657	1	1	411	-120	0.008	1-314	1-310	1-274
Enterobacteriaceae_672	1	43	108	18	0.013	NA	50-175	NA
Clostridium_sensu_stricto_1_673	1	77	239	36	0.0195	42-242	NA	NA

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in OTUs existed between the groups, adjusted for mode of delivery. The OTUs are arranged in descending order based on their relative abundance in this dataset. A positive Area value indicates that the abundance of a specific OTU is higher in the antibiotic treated group, while a negative area value indicates that the abundance of that OTU is higher in the controls. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown. The last three columns show the results of the comparisons between each regimen (ac = amoxicillin + cefotaxime, cg = co-amoxiclav + gentamicin, pg = penicillin + gentamicin) and controls, with the values representing interval in days where a statistically significant difference in abundance exists. NA indicates that for this comparison a specific OTU was not differentially abundant. An asterisk stands for an opposite direction of the effect compared to the value shown under Area. If the value for the antibiotic treated versus controls

comparison is NA, but a statistically significant differential interval is found for an antibiotic regimen versus controls comparison, pos indicates that the abundance of a specific OTU is higher in the antibiotic regimen group, while neg indicates the abundance of that OTU is higher in the controls. Int.no = interval number. In some cases, multiple OTUs of individual bacterial species were identified, so OTUs are referred to by their taxonomical annotations and a rank number (shown in parentheses), which is based on the abundance of each given OTU in the overall dataset.

eTable 4. Correlation between top five most abundant 16S rRNA OTUs and MGS species

16S rRNA OTUs	MGS species	Pearson's r	p.adj
Bifidobacterium (1)	*Bifidobacterium adolescentis* *Bifidobacterium breve* *Bifidobacterium longum*	0.95	<0.001
Escherichia coli (2)	*Escherichia coli*	0.89	<0.001
Staphylococcus epidermidis (3)	*Staphylococcus epidermidis*	0.92	<0.001
Klebsiella (4)	*Klebsiella oxytoca* *Klebsiella pneumoniae*	0.60	<0.001
Enterococcus faecium (5)	*Enterococcus faecium*	0.80	<0.001

The top five most abundant OTUs of the 16S rRNA dataset were correlated with the corresponding metagenomic shotgun sequencing (MGS) sequencing species using Pearson correlations. An adjusted p-value (p.adj, calculated using the Benjamini-Hochberg method) of <0.05 stands for a significant correlation (not a significant difference).

eTable 5. FitTimeSeries results of the significantly differentially abundant taxa between the amoxicillin + cefotaxime and co-amoxiclav + gentamicin regimens

OTU	Int.no	Interval start	Interval end	Area	p.adj
Escherichia_coli_2	1	3	54	175	0.0228
Klebsiella_4	1	3	17	33	0.0426
Peptostreptococcaceae_32	1	52	151	-198	0.0466
Lactobacillus_40	1	3	19	-11	0.0228
Corynebacterium_47	1	3	20	-20	0.0466
Tepidimonas_95	1	3	26	-32	0.0426
Clostridium_sensu_stricto_1_124	1	3	111	-85	0.0426
Comamonadaceae_130	1	3	25	-20	0.0492
Acinetobacter_143	1	3	25	-18	0.0466
Ruminococcus_sp_14531_180	1	19	138	-54	0.0466
Lachnospiraceae_175	1	254	368	-944	0.0466
Lachnospiraceae_175	2	375	411	235	0.0228
Bifidobacteriaceae_193	1	23	69	62	0.0426
Bifidobacteriaceae_193	2	374	410	105	0.0492
Bacteroides_236	1	252	411	125	0.0466
Parabacteroides_distasonis_207	1	3	178	118	0.0228
Finegoldia_164	1	3	29	-27	0.0466
Anaerococcus_tetradius_329	1	3	47	-23	0.0466
Bacteroides_248	1	3	127	-48	0.0228
Leuconostoc_275	1	3	30	-21	0.032
Bacteroides_380	1	3	66	-19	0.0466
Actinomyces_sp_oral_clone_DR002_406	1	15	70	-59	0.0228
Lachnospiraceae_415	1	376	403	-127	0.0426
Dorea_531	1	164	411	-149	0.032
Veillonella_454	1	309	411	89	0.0228
Bacteroides_482	1	80	411	130	0.032

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in OTUs existed between the groups, adjusted for prepartum maternal antibiotics. The OTUs are arranged in descending order based on their relative abundance in this dataset. A positive Area value indicates that the abundance of a specific OTU is higher in the co-amoxiclav + gentamicin regimen, while a negative area value indicates that the abundance of that OTU is higher in the amoxicillin + cefotaxime regimen. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown. Int.no = interval number. In some cases, multiple OTUs of individual bacterial species were identified, so OTUs are referred to by their taxonomical annotations and a rank number (shown in parentheses), which is based on the abundance of each given OTU in the overall dataset.

eTable 6. FitTimeSeries results of the significantly differentially abundant taxa between the amoxicillin + cefotaxime and penicillin + gentamicin regimens

OTU	Int.no	Interval start	Interval end	Area	p.adj
Klebsiella_4	1	2	16	35	0.0457
Klebsiella_4	2	68	121	-126	0.0426
Ruminococcus_gnavus_ CC55_001C_10	1	39	273	-495	0.047
Prevotella_27	1	2	185	147	0.016
Peptostreptococcaceae_32	2	90	195	-243	0.0457
Subdoligranulum_45	1	302	403	-185	0.0457
Lactobacillus_42	1	89	160	87	0.0457
Peptostreptococcaceae_50	1	121	403	-357	0.0457
Bifidobacterium_63	1	14	320	283	0.0457
Blautia_79	1	280	367	-549	0.0426
Ruminococcaceae_90	1	269	366	-714	0.0457
Ruminococcaceae_90	2	376	403	136	0.047
Eubacterium_sp_CS1_Van_96	1	2	360	231	0.016
Bilophila_wadsworthia_3_1_6_123	1	2	166	69	0.0492
Bifidobacteriaceae_122	1	22	43	-21	0.0457
Lachnospiraceae_162	1	106	403	-179	0.0457
Eubacterium_desmolans_125	1	253	368	-938	0.032
Sutterella_134	1	2	16	9	0.0457
Roseburia_185	1	372	403	88	0.0457
Escherichia_Shigella_154	1	2	17	19	0.0457
Gemella_120	1	201	403	-165	0.047
Lactobacillales_142	1	2	52	-28	0.0457
Lachnospiraceae_253	1	281	361	-286	0.032
Lachnospiraceae_253	2	370	403	100	0.0457
Bacteroides_210	1	124	403	206	0.016
Sutterella_197	1	71	210	-112	0.0457
Anaerostipes_269	1	273	368	-734	0.0284
Anaerostipes_269	2	376	403	98	0.016
Parabacteroides_distasonis_207	1	2	166	87	0.0457
Actinomyces_292	1	13	189	-81	0.047
Ruminococcaceae_306	1	134	403	-221	0.0457
ratAN060301C_268	1	44	231	84	0.0476
Bifidobacterium_animalis_293	1	52	403	-214	0.0457
Gardnerella_332	1	326	403	66	0.047

Continue

Continued

OTU	Int.no	Interval start	Interval end	Area	p.adj
Bifidobacterium_336	1	43	289	107	0.047
Leuconostoc_275	1	2	25	-15	0.0393
Bacteroides_400	1	2	21	-6	0.0457
Subdoligranulum_479	1	147	403	-129	0.0457
Clostridium_difficile_630_345	1	2	86	65	0.016
Clostridium_difficile_630_359	1	2	39	23	0.016
Pseudobutyrivibrio_497	1	259	370	-456	0.047
Streptococcus_521	1	36	111	513	0.016
Streptococcus_521	2	123	124	-1	0.016
Veillonella_465	1	125	403	130	0.0476
Lachnospiraceae_415	2	366	403	-149	0.0457
Peptostreptococcaceae_396	1	2	205	93	0.0457
Veillonella_454	1	107	129	-12	0.0457
Veillonella_454	2	345	403	53	0.032
Lachnospiraceae_512	1	106	403	-200	0.0457
Fusicatenibacter_sacchari-vorans_606	1	290	370	-181	0.0457
Fusicatenibacter_sacchari-vorans_606	2	394	403	16	0.0457
Bifidobacterium_animalis_591	1	87	303	-60	0.0457

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in OTUs existed between the groups, adjusted for prepartum maternal antibiotics. The OTUs are arranged in descending order based on their relative abundance in this dataset. A positive Area value indicates that the abundance of a specific OTU is higher in the penicillin + gentamicin regimen, while a negative area value indicates that the abundance of that OTU is higher in the amoxicillin + cefotaxime regimen. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown. Int.no = interval number. In some cases, multiple OTUs of individual bacterial species were identified, so OTUs are referred to by their taxonomical annotations and a rank number (shown in parentheses), which is based on the abundance of each given OTU in the overall dataset.

eTable 7. FitTimeSeries results of the significantly differentially abundant taxa between the co-amoxiclav + gentamicin and penicillin + gentamicin regimens

OTU	Int.no	Interval start	Interval end	Area	p.adj
Escherichia_coli_2	1	2	43	-74	0.0389
Megamonas_46	1	26	411	291	0.0214
Clostridium_sensu_stricto_1_65	1	31	282	316	0.0267
Lachnospiraceae_175	1	344	357	-66	0.0267
Lachnospiraceae_175	3	373	411	-241	0.0214
Lactobacillus_delbrueckii_subsp_bulgaricus_190	1	255	411	155	0.0321
Oscillospiraceae_bacterium_VE202_24_316	1	81	411	-318	0.0214
Streptococcus_gallolyticus_subsp_macedonicus_350	1	102	177	-53	0.0493
Lachnospiraceae_362	1	124	411	-167	0.0499
Ruminococcus_sp_CE2_417	1	109	411	299	0.0321
Subdoligranulum_479	1	149	411	-193	0.0267
Pseudobutyrivibrio_497	1	344	358	-37	0.0214
Pseudobutyrivibrio_497	2	366	374	10	0.0214
Bifidobacterium_breve_473	1	2	189	-82	0.0445
Clostridium_sensu_stricto_1_458	1	102	185	31	0.0499

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in OTUs existed between the groups, adjusted for prepartum maternal antibiotics. The OTUs are arranged in descending order based on their relative abundance in this dataset. A positive Area value indicates that the abundance of a specific OTU is higher in the penicillin + gentamicin regimen, while a negative area value indicates that the abundance of that OTU is higher in the co-amoxiclav + gentamicin regimen. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown. Int.no = interval number. In some cases, multiple OTUs of individual bacterial species were identified, so OTUs are referred to by their taxonomical annotations and a rank number (shown in parentheses), which is based on the abundance of each given OTU in the overall dataset.

eTable 8. Effect of antibiotic treatment in the first week of life on AMR gene composition in infants not receiving antibiotics later in life

Time point	R^2 (%)	Adjusted p-value
Before antibiotics	1.8	0.23
After antibiotics	7.0	0.001
Month 1	5.9	0.001
Month 4	2.7	0.001
Month 12	1.0	0.15

Effect sizes (R^2) and adjusted p-values are shown for the effect of antibiotics in the first week of life on AMR gene composition of the subset of participants not receiving antibiotics later in life.

eTable 9. Covariates significantly associated with AMR gene profile during follow-up

Covariate	R^2 (%)	p.adj
Age	4.9	0.001
Antibiotic treatment in 1st week of life (main study outcome)	2.1	0.001
Pets in the household	1.0	0.001
Mode of delivery	0.5	0.001
Breastfeeding at time of sampling	0.5	0.001
Gravidity	0.2	0.001
Siblings <5 years of age	0.2	0.001
Duration of ruptured membranes before birth	0.2	0.01

A temporal, multivariable permutational multivariate analysis of variance (PERMANOVA) test was performed with all covariates that were univariately associated with antimicrobial resistance (AMR) gene profile in cross-sectional analyses included in the model. The covariates that ensued as being associated with AMR gene profile during follow-up are shown here with their effect sizes (R^2) and adjusted p-values (p.adj, using Benjamini-Hochberg method to correct for multiple testing).

eTable 10. Effect of short (1-4 days) versus long (>4 days) antibiotic treatment in the first week of life on AMR gene composition

Time point	R^2 (%)	Adjusted p-value
Before antibiotics	1.4	0.92
After antibiotics	1.0	0.42
Month 1	0.8	0.52
Month 4	1.0	0.42
Month 12	1.2	0.42

Effect sizes (R^2) and adjusted p-values are shown for the effect of short versus long antibiotic treatment in the first week of life on AMR gene composition.

eTable 11. FitTimeSeries results of differentially abundant AMR genes between antibiotic treated infants and controls

Gene	Interval start	Interval end	Area	p.adj
aac6_aph2	1	140	-807	0.002
aac6_Ib	1	68	-35	0.04
aac6_Ii	1	185	-355	0.02
acrA	1	86	288	0.002
aph3_III	1	116	-205	0.03
blaAMPC	1	53	212	0.002
blaCMY_2	25	371	-423	0.02
blaCTX_M	1	196	252	0.03
blaNDM	1	180	216	0.02
blaTEM	1	112	-177	0.04
cblA	1	411	1085	0.004
ermB	1	213	-617	0.002
ermC	1	83	-263	0.002
MCR1	19	149	-73	0.002
mecA	1	147	-715	0.002
tetQ	1	306	1009	0.002

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in antimicrobial resistance (AMR) genes existed between the antibiotic treated infants and controls, adjusted for mode of delivery and siblings <5 years of age. Here, a negative Area value indicates that the abundance of a specific gene is higher in the antibiotic treated group, while a positive Area value indicates that the abundance of gene is higher in the controls. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown.

eTable 12. List of significantly differentially abundant AMR genes between amoxicillin + cefotaxime treated children and controls as studied by FitTimeSeries

Gene	Int.no	Interval start	Interval end	Area	p.adj
aac6_aph2	1	1	124	-815	0.0026
aac6_Ii	1	1	200	-536	0.0072
acrA	1	1	39	258	0.0026
aph3_Ia_Ic	1	32	53	-47	0.028
aph3_III	1	361	384	68	0.013
aph3_III	2	1	99	-229	0.029
blaAMPC	1	1	40	270	0.0026
blaCMY_2	1	25	139	-209	0.015
cblA	1	1	384	1283	0.0040
ermB	1	1	123	-494	0.0040
ermC	1	1	44	-225	0.0026
MCR1	1	95	153	-43	0.019
mecA	1	1	132	-822	0.0026
tetQ	1	1	316	1154	0.0026
vanA	1	1	223	-57	0.0026

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in AMR genes existed between children treated with amoxicillin + cefotaxime and controls. Here, a negative Area value indicates that the abundance of a specific gene is higher in the amoxicillin + cefotaxime group, while a positive Area value indicates that the abundance of gene is higher in the controls. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown. Int.no = interval number.

eTable 13. List of significantly differentially abundant AMR genes between co-amoxiclav + gentamicin treated children and controls as studied by FitTimeSeries

Gene	Interval start	Interval end	Area	p.adj
aac6_aph2	1	312	-1615	0.0047
aac6_Ii	1	178	-442	0.012
aadE	199	411	373	0.031
aadE_like_gene	1	254	477	0.031
aph3_Ia_Ic	1	218	-396	0.027
aph3_III	1	130	-273	0.033
blaCMY_2	33	347	-470	0.027
blaCTX_M	1	166	249	0.033
blaNDM	2	191	241	0.027
blaTEM	1	259	-820	0.012
cblA	1	248	652	0.015
ermB	1	193	-528	0.011
ermC	1	160	-326	0.0047
MCR1	15	180	-151	0.012
mecA	1	162	-682	0.0047
tetQ	1	299	1187	0.0047

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in antimicrobial resistance (AMR) genes existed between children treated with co-amoxiclav + gentamicin and controls. Here, a negative Area value indicates that the abundance of a specific gene is higher in the co-amoxiclav + gentamicin, while a positive Area value indicates that the abundance of gene is higher in the controls. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown.

eTable 14. List of significantly differentially abundant AMR genes between penicillin + gentamicin treated children and controls as studied by FitTimeSeries

Gene	Interval start	Interval end	Area	p.adj
aac6_aph2	1	141	-789	0.0026
acrA	1	73	234	0.0087
aph3_Ia_Ic	33	159	240	0.023
blaAMPC	1	72	247	0.0026
blaCMY_2	46	403	-394	0.028
blaCTX_M	1	179	317	0.0087
blaNDM	4	137	445	0.0026
cblA	31	223	338	0.028
ermB	1	182	-503	0.0087
ermC	1	88	-270	0.0026
mecA	1	169	-753	0.0026
tetQ	1	194	520	0.0087

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in antimicrobial resistance (AMR) genes existed between children treated with penicillin + gentamicin and controls. Here, a negative Area value indicates that the abundance of a specific gene is higher in the penicillin + gentamicin group, while a positive Area value indicates that the abundance of gene is higher in the controls. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown.

eTable 15. List of differentially abundant AMR genes between amoxicillin + cefotaxime and co-amoxiclav + gentamicin treated children as studied by FitTimeSeries

Gene	Int.no	Interval start	Interval end	Area	p.adj
aadE	1	340	411	167	0.035
acrA	1	3	85	-342	0.021
aph2_Ide	3	280	364	5-12	0.035
aph2_Ide	4	28	36	-8E-14	0.032
aph2_Ide	5	129	253	-9E-12	0.035
blaAMPC	1	3	81	-343	0.021
blaKPC	1	345	411	93	0.036
blaTEM	1	3	355	-1292	0.021
ermB	1	3	18	57	0.032
ermC	1	3	23	60	0.032
mecA	1	3	62	153	0.035

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in antimicrobial resistance (AMR) genes existed between children treated with amoxicillin + cefotaxime and co-amoxiclav + gentamicin. Here, a negative Area value indicates that the abundance of a specific gene is higher in the co-amoxiclav + gentamicin group, while a positive Area value indicates that the abundance of gene is higher in the amoxicillin + cefotaxime group. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown. Int.no = interval number.

eTable 16. List of differentially abundant AMR genes between amoxicillin + cefotaxime and penicillin + gentamicin treated children as observed by FitTimeSeries

Gene	Interval start	Interval end	Area	p.adj
aac6_Ii	2	157	384	0.042
acrA	2	24	-86	0.026
aph3_Ia_Ic	18	136	460	0.0070
blaAMPC	2	25	-84	0.026
blaNDM	2	17	21	0.042

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in antimicrobial resistance (AMR) genes existed between children treated with amoxicillin + cefotaxime and penicillin + gentamicin. Here, a negative Area value indicates that the abundance of a specific gene is higher in the penicillin + gentamicin group, while a positive Area value indicates that the abundance of gene is higher in the amoxicillin + cefotaxime group. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown.

eTable 17. FitTimeSeries results of differentially abundant AMR genes between co-amoxiclav + gentamicin and penicillin + gentamicin treated children

Gene	Interval start	Interval end	Area	p.adj
aac3_Iiacde	2	191	577	0.030

Differential abundance testing by smoothing spline analysis of variance (SS-ANOVA) was executed to test in which specific intervals significant differences in antimicrobial resistance (AMR) genes existed between children treated with co-amoxiclav + gentamicin and penicillin + gentamicin. Here, a positive Area value indicates that the abundance of gene is higher in the co-amoxiclav + gentamicin group. To correct for multiple testing, the Benjamini-Hochberg method was applied, and the adjusted p-values (p.adj) are shown.

eTable 18. Significant correlations observed between OTUs and AMR genes

Gene	OTU	Pearson's r	p.adj
aac6_aph2	Bifidobacterium_1	0.42	0
aac6_IIa	Bifidobacterium_1	0.56	0.00012
ermC	Bifidobacterium_1	0.42	0
mecA	Bifidobacterium_1	0.55	0
acrA	Escherichia_coli_2	-0.65	0
aph2_Ib	Escherichia_coli_2	0.41	0.018
blaAMPC	Escherichia_coli_2	-0.69	0
blaTEM	Escherichia_coli_2	-0.51	0
vanB	Escherichia_coli_2	-0.42	0
aac6_aph2	Enterococcus_faecium_5	-0.42	0
aac6_Ii	Enterococcus_faecium_5	-0.50	0
aac6_IIa	Staphylococcus_epidermidis_3	-0.49	0.0026
ermC	Staphylococcus_epidermidis_3	-0.55	0
mecA	Staphylococcus_epidermidis_3	-0.58	0
blaCMY_2	Klebsiella_4	-0.40	0.0026
aac6_Ib	Clostridium_butyricum_35	-0.47	0.032
aac6_IIa	Lactobacillus_40	-0.47	0.0043
aac6_IIa	Corynebacterium_47	-0.45	0.0088
aac6_IIa	Erysipelotrichaceae_74	-0.46	0.0064
aac6_IIa	Tepidimonas_95	-0.50	0.0016
aac6_IIa	Comamonadaceae_130	-0.46	0.0056
aac6_IIa	Faecalibacterium_132	-0.46	0.0061
aac6_IIa	Blautia_141	-0.46	0.0064
aph2_Ide	Veillonella_115	-0.98	0.035
aac6_IIa	Ruminococcaceae_158	-0.46	0.0064
aac6_IIa	Lachnospiraceae_187	-0.46	0.0064

Continue

Gene	OTU	Pearson's r	p.adj
aac6_Ila	Bacteroides_314	-0.44	0.011
aac6_Ila	Corynebacterium_328	-0.46	0.0064
aac6_Ila	Lachnospiraceae_342	-0.46	0.0064
aac6_Ila	Coriobacteriaceae_377	-0.46	0.0064
aac6_Ila	Ruminococcaceae_387	-0.46	0.0064
aac6_Ila	Phascolarctobacterium_317	-0.46	0.0064
aac6_Ila	Bacteroides_coprophilus_402	-0.46	0.0064
aac6_Ila	Leuconostoc_mesenteroides_282	-0.43	0.013
aac6_Ila	Bacteroides_242	-0.46	0.0064
aac6_Ila	Oscillibacter_444	-0.46	0.0064
aac6_Ila	Ruminococcaceae_424	-0.46	0.0064
aac6_Ila	Prevotellaceae_449	-0.46	0.0064
aac6_Ila	Prevotella_509	-0.46	0.0064
aac6_Ila	Christensenellaceae_517	-0.46	0.0064
aac6_Ila	Anaerococcus_418	-0.46	0.0064
aac6_Ila	Faecalibacterium_477	-0.46	0.0064
aac6_Ila	Odoribacter_589	-0.46	0.0064
aac6_Ila	Staphylococcus_588	-0.41	0.025
aac6_Ila	Prevotella_558	-0.46	0.0064
aac6_Ib	Clostridium_sensu_stricto_1_621	-0.51	0.016
aac6_Ib	Clostridium_butyricum_685	-0.48	0.026

We correlated all 16S OTUs with all AMR genes found with the Fluidigm platform using the Pearson's correlation coefficient. Depicted are all correlations with an adjusted p-value <0.05 and correlation coefficient of ≤-0.40 or ≥0.40. The OTUs are arranged in descending order based on their relative abundance in the overall dataset. A positive coefficient stands for a negative correlation between OTU and gene abundance, as a high Ct value indicates low abundance of an AMR gene. We applied the Benjamini-Hochberg method to correct for multiple testing and adjusted p-values (p.adj) are shown.

eTable 19. FitZig results of significantly differentially abundant AMR genes as found by MGS between antibiotic treated children and controls

Gene	log2FC	p.adj
Rif|CP003583.1|gene2745|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	10.04607	1.71E-06
Flq|CP003583.1|gene1174|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	8.921015	1.21E-05
ACou|CP003583.1|gene1173|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	8.653329	1.48E-05
AGly|Aac6-Aph2|M13771|304-1743|1440|Aminoglycosides|Aminoglycoside_N-acetyltransferases|AAC6-PRIME	8.564101	3.08E-06
CARD|phgb|CP003583|2403484-2405254|ARO:3000776|adeC|Multi--drug_resistance|Multi-drug_efflux_pumps|ADEC	8.518329	2.18E-05
736|JX560992.1|JX560992|Aminoglycosides|Aminoglycoside_N-acetyltransferases|AAC6-PRIME	8.442804	5.2E-06
AGly|NC_005024.2598370|Aminoglycosides|Aminoglycoside_N--acetyltransferases|AAC6-PRIME	8.397573	5.2E-06
CARD|pvgb|CP003583|1050805-1052287|ARO:3003092|Enterococcus|Lipopeptides|Daptomycin-resistant_cls|CLS|RequiresSNPConfirmation	8.274826	2.65E-05
MLS|MsrC|AY004350|496-1974|1479|MLS|Macrolide_resistance_efflux_pumps|MSRC	8.017561	2.45E-05
Rif|AE016830.1|gene3155|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	7.921589	1.11E-05
CARD|pvgb|CP003583|913745-914813|ARO:3003079|Enterococcus|Lipopeptides|Daptomycin-resistant_liaFSR|LIAFSR|RequiresSNPConfirmation	7.66265	5.43E-05
Flq|AE016830.1|gene6|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	7.519387	1.45E-05
CARD|pvgb|HE999704|2567249-2568437|ARO:3001312|elfamycin|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	7.403174	1.43E-05
CARD|pvgb|NC_007622|559336-560521|ARO:3001312|elfamycin|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	7.368328	9.39E-08
940|AF028812.1|AF028812|Trimethoprim|Dihydrofolate_reductase|DFRF	7.329183	4.98E-05
Elf|NC_007622.3795131|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	7.206572	1.31E-07
Tet|tetM_13_AM990992|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	7.163641	2.45E-05
Rif|NC_003098.1.933647|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	7.145893	1.74E-06
Elf|HE999704.1|gene2984|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	7.124943	1.48E-05

Continue

Continued

Gene	log2FC	p.adj
Flq|AE015929.1|gene5|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	7.073008	2.65E-05
Elf|NC_002745.1123311|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	7.02243	1.31E-07
CARD|pvgb|NC_002952|594518-598142|ARO:3003291|Staphylococcus|Lipopeptides|Daptomycin-resistant_beta-subunit_of_RNA_polymerase_RpoC|RPOC|RequiresSNPConfirmation	6.978523	1.21E-05
Tet|tetM_8_X04388|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	6.974866	3.37E-05
Elf|NC_002951.3236433|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	6.872335	5.77E-07
CARD|pvgb|AP009048|3760295-3762710|ARO:3003303|Escherichia|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRB|RequiresSNPConfirmation	6.807278	5.61E-05
Tet|tetM_6_M21136|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	6.777552	4.4E-05
Mdr|AB566411.1|gene1|Multi-drug_resistance|Multi-drug_efflux_pumps|QACAB	6.749747	9.71E-05
Rif|NC_012469.1.7686402|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	6.738756	3.4E-06
CARD|pvgb|NC_003197|4038854-4041269|ARO:3003307|Salmonella|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRB|RequiresSNPConfirmation	6.698354	6.98E-05
Rif|CP000647.1|gene4402|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	6.691371	5.61E-05
Elf|NC_002952.2860408|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	6.568222	3.92E-06
Bla|mecA_15_AB505628|betalactams|Penicillin_binding_protein|MECA	6.486463	0.0002
Elf|CP000647.1|gene4394|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	6.463796	5.61E-05
Tet|tetM_1_X92947|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	6.358008	9.56E-05
Rif|FN543093.2|gene314|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	6.320669	5.43E-05
Rif|NC_009487.5169226|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	6.312594	7.68E-05
605|V01547.1|V01547|Aminoglycosides|Aminoglycoside_O-phosphotransferases|APH3-PRIME	6.309507	0.000322
CARD|pvgb|AE014075|3901532-3902762|ARO:3003438|Escherichia|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	6.279452	0.000207
CARD|phgb|NC_012469|1800927-1802391|ARO:3000616|mel|MLS|Macrolide_resistance_efflux_pumps|MEL	6.272894	2.45E-05

Continue

Continued

Gene	log2FC	p.adj
CARD|phgb|AB037671|24420-26427|ARO:3001209|mecC|betalactams|Penicillin_binding_protein|MECC	6.269044	0.000247
Flq|CP000647.1|gene3437|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	6.263472	0.000969
CARD|pvgb|CP003583|914819-915452|ARO:3003078|Enterococcus|Lipopeptides|Daptomycin-resistant_liaFSR|LIAFSR|RequiresSNPConfirmation	6.217033	0.000793
Flq|AE016830.1|gene1598|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	6.192162	0.000217
Bla|KC243783.1|gene1|betalactams|Penicillin_binding_protein|MECA	6.163537	0.000249
Tet|tetM_7_FN433596|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	6.143436	0.000142
1463|M81802.1|STAMSRB|MLS|Macrolide_resistance_efflux_pumps|MSRA	6.141506	0.000306
Flq|OqxBgb|EU370913|47851-51003|3153|Multi-drug_resistance|Multi-drug_efflux_pumps|OQXB	6.123797	0.000945
CARD|phgb|AB091338|173-1355|ARO:3003551|emeA|Multi-drug_resistance|Multi-drug_efflux_pumps|EMEA	6.117836	0.000243
MLS|ermB_6_AF242872|MLS|23S_rRNA_methyltransferases|ERMB	6.103928	0.000534
Flq|CP001918.1|gene3562|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	6.072613	0.001581
Bla|AF336096.1|gene1|betalactams|Mutant_porin_proteins|OMP36|RequiresSNPConfirmation	6.061827	1.5E-05
Bla|NC_002745.1122813|betalactams|Penicillin_binding_protein|MECA	6.048143	0.000249
CARD|phgb|L12710|0-549|ARO:3002556|AAC6-PRIME-Ii|Aminoglycosides|Aminoglycoside_N-acetyltransferases|AAC6-PRIME	6.048094	0.000916
Flq|CP001138.1|gene2385|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	5.951057	0.000348
Bla|NC_009487.5167689|betalactams|Penicillin_binding_protein|MECA	5.949511	0.000422
Mdr|oqxB_1_EU370913|Multi-drug_resistance|Multi-drug_efflux_pumps|OQXB	5.947138	0.001235
Mdr|AY661734.1|gene2|Multi-drug_resistance|Multi-drug_efflux_pumps|MEPA	5.937923	0.011589
Elf|CP000647.1|gene3761|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	5.934923	0.000138
11|DQ679966.1|DQ679966|Multi-drug_resistance|Multi-drug_efflux_pumps|ACRB	5.933073	0.000348
MLS|ermB_15U48430|MLS|23S_rRNA_methyltransferases|ERMB	5.915585	0.000533
Bla|AF336097.1|gene1|betalactams|Mutant_porin_proteins|OMP36|RequiresSNPConfirmation	5.89168	0.000207

Continue

Continued

Gene	log2FC	p.adj
Flq|CP001138.1|gene3329|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	5.881479	0.001782
MLS|MsrD|AF274302|2462-3925|1464|MLS|Macrolide_resistance_efflux_pumps|MSRD	5.88022	3.52E-05
Tet|tetM_12_FR671418|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	5.87791	0.000249
Tet|tetM_5_U58985|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	5.859626	0.000312
Flq|FN543093.2|gene2948|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	5.857616	0.001913
ACou|CP000647.1|gene3444|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	5.816394	0.001497
ACou|AE016830.1|gene1599|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	5.816186	0.000457
CARD|phgb|U00096|2155262-2158385|ARO:3000793|mdtB|Multi-drug_resistance|Multi-drug_efflux_pumps|MDTB	5.768583	0.000312
Sul|AE016830.1|gene3181|Sulfonamides|Sulfonamide-resistant_dihydropteroate_synthases|FOLP|RequiresSNPConfirmation	5.760223	0.000797
Rif|CP000034.1|gene3741|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	5.688878	0.000312
MLS|LsaA|AY225127|41-1537|1497|MLS|Streptogramin_resistance_ATP--binding_cassette_ABC_efflux_pumps|LSA	5.672168	0.000536
MLS|mefA_10_AF376746|MLS|Macrolide_resistance_efflux_pumps|MEFA	5.66807	8.54E-05
Tet|tetM_2_X90939|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	5.653387	0.000249
Tet|tetM_11_JN846696|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	5.638708	0.000109
CARD|phgb|U00096|2812615-2814154|ARO:3000074|emrB|Multi-drug_resistance|Multi-drug_efflux_pumps|EMRB	5.638329	0.000275
Rif|CP001918.1|gene250|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	5.634411	0.000945
CARD|phgb|NC_007779|2586250-2589364|ARO:3000491|acrD|Aminoglycosides|Aminoglycoside_efflux_pumps|ACRD	5.598443	0.000457
1021|AJ243209.1|AJ243209|MLS|Macrolide_resistance_efflux_pumps|MSRC	5.597685	0.000793
23|U58210.1|STU58210|betalactams|Penicillin_binding_protein|PBP2B	5.544856	4.4E-05
Elf|CP001918.1|gene242|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	5.520286	0.000945
CARD|pvgb|AE016830|2790820-2791552|ARO:3003077|Enterococcus|Lipopeptides|Daptomycin-resistant_liaFSR|LIAFSR|RequiresSNPConfirmation	5.490529	0.000679

Continue

Continued

Gene	log2FC	p.adj
Rif|NC_007622.3792928|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	5.487949	0.000519
Rif|AP009048.1|gene3341|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	5.470726	0.000681
MLS|lsaA_3_AY737526|MLS|Streptogramin_resistance_ATP-binding_cassette_ABC_efflux_pumps|LSA	5.419232	0.00054
Flq|FN543093.2|gene3588|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	5.382396	0.003009
Bla|mecA_7_BA000018|betalactams|Penicillin_binding_protein|MECA	5.377954	0.000533
Bla|blaZ_34_AP003139|betalactams|Class_A_betalactamases|BLAZ	5.342823	0.000237
Bla|AF336095.1|gene1|betalactams|Mutant_porin_proteins|OMP36|RequiresSNPConfirmation	5.339979	0.000312
CARD|pvgb|CP000647|4764664-4765849|ARO:3001312|elfamycin|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	5.337423	0.000249
Tet|TetM|U08812|1981-3900|1920|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	5.29587	0.000793
Flq|CP000647.1|gene2640|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	5.293169	0.001025
Bla|MECA|AB221124|91-2097|2007|betalactams|Penicillin_binding_protein|MECA	5.290229	0.000818
1454|J01764.1|PT1CG|Tetracyclines|Tetracycline_resistance_major_facilitator_superfamily_MFS_efflux_pumps|TETK	5.288726	0.002392
Tet|tetM_4_X75073|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	5.248432	0.000716
Rif|CP001138.1|gene4362|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	5.221324	0.00094
Tet|tetM_9_X56353|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	5.212076	0.000614
AGly|AY712687.1|gene1|Aminoglycosides|Aminoglycoside_O-nucleotidyltransferases|ANT6	5.207654	0.001759
CARD|pvgb|U00096|3163714-3165973|ARO:3003308|Escherichia|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	5.190758	0.001405
179|M16217.1|PNS1CG|Tetracyclines|Tetracycline_resistance_major_facilitator_superfamily_MFS_efflux_pumps|TETK	5.178832	0.004669
CARD|phgb|AM180355|2319373-2320111|ARO:3000375|ErmB|MLS|23S_rRNA_methyltransferases|ERM	5.170926	0.000716
CARD|pvgb|CP001918|235165-236350|ARO:3001312|elfamycin|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	5.160506	0.001913
Bla|mecA_10_AB512767|betalactams|Penicillin_binding_protein|MECA	5.159533	0.000298

Continue

Continued

Gene	log2FC	p.adj
AGly|Ant6-Ia|AF330699|22-930|909|Aminoglycosides|Aminoglycoside_O-nucleotidyltransferases|ANT6	5.142961	0.001268
Bla|AmpH|CP003785|4208384-4209544|1161|betalactams|Penicillin_binding_protein|AMPH	5.125804	0.003508
CARD|pvgb|CP001918|4816613-4817537|ARO:3001312|elfamycin|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	5.111733	0.000768
Flq|NC_003197.1.1253794|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	5.107397	0.001759
Bla|NC_010066.5774791|betalactams|Class_A_betalactamases|BLAZ	5.088592	0.000237
Rif|NC_007793.3914836|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	5.079887	0.000207
Rif|NC_002745.1123305|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	5.074613	0.000519
MLS|ermB_11_M19270|MLS|23S_rRNA_methyltransferases|ERMB	5.033314	0.000922
Elf|CP001918.1|gene4764|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	5.022946	0.001619
Flq|NC_003197.1.1254697|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	5.020705	0.00428
Tet|tetM_10_EU182585|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETM	4.96337	0.000573
MLS|ermC_13_M13761|MLS|23S_rRNA_methyltransferases|ERMC	4.948993	0.000945
738|JX560992.1|JX560992|Aminoglycosides|Aminoglycoside_O-nucleotidyltransferases|ANT6	4.931574	0.00224
Bla|mecA_4_AB033763|betalactams|Penicillin_binding_protein|MECA	4.927409	0.000916
Bla|OXY6-2|AJ871875|1-873|873|betalactams|Class_A_betalactamases|OXY	4.927085	0.00128
MLS|mphC_1_AB013298|MLS|Macrolide_phosphotransferases|MPHC	4.862321	0.001963
MLS|ermB_9_AF299292|MLS|23S_rRNA_methyltransferases|ERMB	4.839503	0.002544
Flq|M58408|gene|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	4.83048	0.001782
Bla|OXY6-4|AJ871877|1-873|873|betalactams|Class_A_betalactamases|OXY	4.82262	0.003385
MLS|lsaA_2_AY58982|MLS|Streptogramin_resistance_ATP-binding_cassette_ABC_efflux_pumps|LSA	4.776619	0.002408
AGly|Sat4A|X92945|38870-39412|543|Aminoglycosides|Aminoglycoside_N-acetyltransferases|SAT	4.750136	0.003009
CARD|phgb|NC_010066|9682-10528|ARO:3000621|PC1|betalactams|-Class_A_betalactamases|BLAZ	4.744185	0.000861
CARD|pvgb|NC_002952|590830-594382|ARO:3003287|Staphylococcus|Lipopeptides|Daptomycin-resistant_beta-subunit_of_RNA_polymerase_RpoB|rpoB|RequiresSNPConfirmation	4.732193	0.000772

Continue

Continued

Gene	log2FC	p.adj
gi|149167|gb|M88143.1|KPNBETALAC|betalactams|Class_A_betalactamases|TEM	4.730381	0.001782
CARD|phgb|NC_023287|60996-61860|ARO:3002626|ANT6-Ia|Aminoglycosides|Aminoglycoside_O-nucleotidyltransferases|ANT6	4.726548	0.003919
CARD|phgb|AF028811|0-462|ARO:3002875|dfrE|Trimethoprim|Dihydrofolate_reductase|DFRE	4.719341	0.001778
199|AY566250.1|AY566250|Fluoroquinolones|Quinolone_active_efflux|NORA	4.715596	0.003745
CARD|phgb|NC_002695|1737553-1737967|ARO:3000676|H-NS|Multi-drug_resistance|MDR_regulator|HNS	4.697152	0.000818
CARD|phgb|U00096|3415032-3418137|ARO:3000502|acrF|Aminoglycosides|Aminoglycoside_efflux_pumps|ACRF	4.679166	0.003379
Fos|fosA_8_ACHE01000077|Fosfomycin|Fosfomycin_thiol_transferases|FOSA	4.657483	0.001497
AGly|APH-Stph|HE579073|1778413-1779213|801|Aminoglycosides|Aminoglycoside_O-phosphotransferases|APH3-PRIME	4.657181	0.001633
CARD|phgb|U00096|2158385-2161463|ARO:3000794|mdtC|Multi-drug_resistance|Multi-drug_efflux_pumps|MDTC	4.65583	0.001997
CARD|pvgb|NC_002695|4990140-4994169|ARO:3003288|Escherichia|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	4.645964	0.001268
246|DQ679966.1|DQ679966|Multi-drug_resistance|Multi-drug_efflux_pumps|MEXE	4.626894	0.002908
1066|unknown_id|unknown_name|MLS|23S_rRNA_methyltransferases|ERMB	4.608049	0.003764
MLS|ermB_12_U18931|MLS|23S_rRNA_methyltransferases|ERMB	4.603798	0.004169
CARD|phgb|AP009048|4153663-4154296|ARO:3000518|CRP|Multi-drug_resistance|MDR_regulator|CRP	4.601957	0.001127
Flq|NC_012469.1.7686721|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	4.589629	3.37E-05
Bla|X04121.1|gene1|betalactams|Class_A_betalactamases|BLAZ	4.558406	0.001759
Rif|NC_002951.3236234|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	4.550078	0.00071
CARD|pvgb|NC_007779|3172159-3174052|ARO:3003316|Escherichia|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|parE|RequiresSNPConfirmation	4.538684	0.001997
554|FN806789.1|FN806789|MLS|23S_rRNA_methyltransferases|ERMB	4.527378	0.003248
Rif|NC_003197.1.1255679|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	4.505091	0.001633
MLS|msrA_2_AB013298|MLS|Macrolide_resistance_efflux_pumps|MSRA	4.479036	0.009354
CARD|phgb|U00096|3660413-3663527|ARO:3000796|mdtF|Multi-drug_resistance|Multi-drug_efflux_pumps|MDTC	4.465299	0.005238

Continue

Continued

Gene	log2FC	p.adj
Rif|NC_002695.1.914942|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	4.445286	0.002337
Tet|tetW_1_DQ060146|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETW	4.436131	0.022933
AGly|aph3-prime-Ia_1_V00359|Aminoglycosides|Aminoglycoside_O--phosphotransferases|APH3-PRIME	4.427153	0.003692
Bla|mecA_14_AB505630|betalactams|Penicillin_binding_protein|MECA	4.427001	0.00269
Mdr|NC_002695.1.910541|Multi-drug_resistance|MDR_regulator|CPXAR	4.423297	0.004103
Bla|blaZ_40_DQ269019|betalactams|Class_A_betalactamases|BLAZ	4.42069	0.000629
Rif|NC_013450.8613267|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	4.418607	0.001099
CARD|pvgb|FQ312006|705198-706383|ARO:3001312|elfamycin|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	4.417459	0.004487
MLS|ermB_17_X64695|MLS|23S_rRNA_methyltransferases|ERMB	4.401039	0.004573
Tet|tetW_6_FN396364|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETW	4.392852	0.004103
CARD|pvgb|NC_002952|1419759-1422162|ARO:3003312|Staphylococcus|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	4.348429	0.004077
13|L11616.1|PSEENVCD|Multi-drug_resistance|Multi-drug_efflux_pumps|MEXB	4.324791	0.001266
Bla|PBP|Ecoli|CP002291|664439-666340|1902|betalactams|Penicillin_binding_protein|PBP2	4.316706	0.00493
MLS|ermB_1_JN899585|MLS|23S_rRNA_methyltransferases|ERMB	4.31464	0.004104
Flq|OqxA|EU370913|46652-47827|1176|Multi-drug_resistance|Multi--drug_efflux_pumps|OQXA	4.30986	0.017512
AGly|U01945.1|gene1|Aminoglycosides|Aminoglycoside_N-acetyltransferases|SAT	4.286232	0.006611
CARD|phgb|U00096|2484373-2487967|ARO:3000833|evgS|Multi-drug_resistance|MDR_regulator|EVGS	4.257267	0.006704
MLS|MphC|AF167161|5665-6564|900|MLS|Macrolide_phosphotransferases|MPHC	4.240403	0.003919
CARD|phgb|AE015929|1128933-1129419|ARO:3002865|dfrC|Trimethoprim|Dihydrofolate_reductase|DFRC|RequiresSNPConfirmation	4.221985	0.003666
Bla|BlaZ|AB245469|2235-3080|861|betalactams|Class_A_betalactamases|BLAZ	4.212092	0.005429
CARD|phgb|CP000034|2662758-2663289|ARO:3000516|emrR|Multi--drug_resistance|MDR_regulator|EMRR	4.207096	0.006385
Mdr|AF535087.1|gene1|Multi-drug_resistance|Multi-drug_efflux_pumps|-QACAB	4.206097	0.008229

Continue

Continued

Gene	log2FC	p.adj
Mdr|AP009048.1|gene2135|Multi-drug_resistance|Multi-drug_efflux_pumps|MDTD	4.197468	0.004173
CARD|pvgb|NC_002952|590830-594382|ARO:3003285|Staphylococcus|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	4.143949	0.001782
Bla|Z_35_AJ302698|betalactams|Class_A_betalactamases|BLAZ	4.12885	0.005563
MLS|ermB_2_K00551|MLS|23S_rRNA_methyltransferases|ERMB	4.127583	0.005999
CARD|pvgb|AL450380|2273268-2276805|ARO:3003284|Mycobacterium|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	4.125903	0.017476
CARD|phgb|AY047358|1710-2352|ARO:3003552|fusB|Fusidic_acid|Ribosomal_zinc-binding_protein|FUSB	4.07809	0.00478
Rif|NC_002758.1120515|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	4.077947	0.001997
672|AF181950.1|AF181950|betalactams|Penicillin_binding_protein|MECA	4.070819	0.001468
MLS|ermB_18_X66468|MLS|23S_rRNA_methyltransferases|ERMB	4.064816	0.005741
Bla|AF336098.1|gene1|betalactams|Mutant_porin_proteins|OMP36|RequiresSNPConfirmation	4.057089	0.003132
ACou|NC_003197.1.1254704|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	4.04722	0.048887
ACou|CP000034.1|gene3210|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	4.006142	0.002408
Flq|NC_003098.1.933198|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	3.990149	0.000247
CARD|pvgb|CP000647|489264-489915|ARO:3003373|Klebsiella|Multi-drug_resistance|MDR_regulator|ACRR|RequiresSNPConfirmation	3.949866	0.007895
Tet|tetO_3_Y07780|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETO	3.932456	0.00259
Bla|AY077483.1|gene1|betalactams|Class_A_betalactamases|OXY	3.926918	0.016
Bla|Z_32_AP004832|betalactams|Class_A_betalactamases|BLAZ	3.922653	0.002746
Bla|NC_010063.5774822|betalactams|Class_A_betalactamases|BLAZ	3.921497	0.000888
Flq|NC_003098.1.934295|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	3.896854	0.000237
252|FJ744595.1|FJ744595|Multi-drug_resistance|Multi-drug_efflux_pumps|EMRD	3.885592	0.018497
Sul|CP000647.1|gene3624|Sulfonamides|Sulfonamide-resistant_dihydropteroate_synthases|FOLP|RequiresSNPConfirmation	3.866427	0.014305
CARD|phgb|AP009048|2810082-2811255|ARO:3000027|emrA|Multi-drug_resistance|Multi-drug_efflux_pumps|EMRA	3.866067	0.008436
ACou|FN543093.2|gene3601|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	3.866029	0.016082

Continue

Continued

Gene	log2FC	p.adj
MLS|ermB_7_AF368302|MLS|23S_rRNA_methyltransferases|ERMB	3.865604	0.008653
1268|AB976602.1|AB976602|betalactams|Class_A_betalactamases|CTX	3.822986	0.0145
Flq|NC_002952.2859941|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	3.808601	0.004573
MLS|msrA_1_X52085|MLS|Macrolide_resistance_efflux_pumps|MSRA	3.778458	0.00578
CARD|phgb|AP009048|4304505-4306557|ARO:3003549|mdtO|Multi--drug_resistance|Multi-drug_efflux_pumps|MDTO	3.773864	0.016082
Mdr|AP009048.1|gene3415|Multi-drug_resistance|MDR_regulator|CPXAR	3.766044	0.013282
MLS|msrC_2_AF313494|MLS|Macrolide_resistance_efflux_pumps|MSRC	3.757821	0.008803
CARD|pvgb|AF335467|36-1164|ARO:3003385|Enterobacter|Multi-drug_resistance|MDR_mutant_porin_proteins|omp36|RequiresSNPConfirmation	3.751259	0.010853
Sul|NC_012469.1.7686560|Sulfonamides|Sulfonamide-resistant_dihydropteroate_synthases|FOLP|RequiresSNPConfirmation	3.744433	0.000749
1737|AY033516.1|AY033516|betalactams|Class_A_betalactamases|CTX	3.740823	0.009108
CARD|phgb|NC_003098|1886035-1888501|ARO:3003044|PBP1b|betalactams|Penicillin_binding_protein|PBP1B	3.702681	0.000888
CARD|phgb|AP009048|4338624-4340268|ARO:3003576|PmrC|Cationic_antimicrobial_peptides|Lipid_A_modification|PMRC	3.691995	0.017673
CARD|pvgb|U00096|2336792-2339420|ARO:3003294|Escherichia|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	3.675845	0.011589
Flq|NC_012469.1.7685406|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	3.663717	0.00016
CARD|phgb|U00096|3203309-3204131|ARO:3002986|bacA|Bacitracin|Undecaprenyl_pyrophosphate_phosphatase|BACA	3.662531	0.009354
Elf|FQ312006.1|gene760|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	3.651592	0.013375
ACou|NC_003098.1.934298|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	3.644373	0.000243
Tmt|DfrC|Z48233|337-822|486|Trimethoprim|Dihydrofolate_reductase|DFRC|RequiresSNPConfirmation	3.642526	0.006927
CARD|phgb|U00096|2368038-2370021|ARO:3002985|arnA|Cationic_antimicrobial_peptides|Lipid_A_modification|ARNA	3.62985	0.017612
CARD|pvgb|AE005672|800699-803171|ARO:3003311|Streptococcus|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	3.617412	0.000242
MLS|MsrA|AY591760|274-1740|1467|MLS|Macrolide_resistance_efflux_pumps|MSRA	3.615451	0.016082
Flq|NC_002952.2859949|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	3.609137	0.033068

Continue

Continued

Gene	log2FC	p.adj
Bla|AmpC1_Ecoli|FN649414|2765051-2766355|1302|betalactams|Penicillin_binding_protein|PBP4B	3.602782	0.026057
Flq|CP000034.1|gene2423|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	3.591779	0.016082
Mdr|CP000034.1|gene2198|Multi-drug_resistance|MDR_regulator|ASMA	3.555453	0.020306
Rif|NC_002953.2862188|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	3.554997	0.004103
CARD|phgb|NC_007779|1745170-1746544|ARO:3001327|mdtK|Multi--drug_resistance|Multi-drug_efflux_pumps|MDTK	3.540711	0.014911
MLS|ermB_16_X82819|MLS|23S_rRNA_methyltransferases|ERMB	3.522125	0.0109
Tet|tetW_5_AJ427421|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETW	3.511888	0.036779
CARD|phgb|AP009048|4303042-4304509|ARO:3003550|mdtP|Multi--drug_resistance|Multi-drug_efflux_pumps|MDTP	3.480891	0.026559
Mdr|AP009048.1|gene3616|Multi-drug_resistance|Multi-drug_efflux_pumps|MDTL	3.478173	0.028324
Bac|AP009048.1|gene3103|Bacitracin|Undecaprenyl_pyrophosphate_phosphatase|BACA	3.445745	0.020148
ACou|NC_012469.1.7686068|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	3.434298	0.000249
CARD|phgb|AP009048|3980025-3981183|ARO:3000795|mdtE|Multi--drug_resistance|Multi-drug_efflux_pumps|MDTE	3.400992	0.029937
Tet|EU434751.1|gene2|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETW	3.396945	0.025711
CARD|pvgb|NC_002952|5036-6968|ARO:3003301|Staphylococcus|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|gyrB|RequiresSNPConfirmation	3.392899	0.015558
222|JQ394987.1|JQ394987|Multi-drug_resistance|Multi-drug_efflux_pumps|MDFA	3.356031	0.025796
CARD|phgb|U00096|4567286-4568519|ARO:3001214|mdtM|Multi-drug_resistance|Multi-drug_efflux_pumps|MDTC	3.319964	0.032239
CARD|phgb|FJ768952|0-1488|ARO:3000237|tolC|Multi-drug_resistance|Multi-drug_efflux_pumps|TOLC	3.31457	0.029937
Flq|CP003275.1|gene5239|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	3.281213	0.006611
Flq|CP000034.1|gene3218|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	3.245899	0.009998
1090|DQ303459.3|DQ303459|betalactams|Class_A_betalactamases|CTX	3.238795	0.023877
ACou|NC_007622.3794232|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	3.23569	0.008921
MLS|msrD_3_AF227520|MLS|Macrolide_resistance_efflux_pumps|MSRD	3.165335	0.007355

Continue

Continued

Gene	log2FC	p.adj
CARD|phgb|U00096|1124117-1125326|ARO:3001216|mdtH|Multi-drug_resistance|Multi-drug_efflux_pumps|MDTH	3.157944	0.034433
CARD|phgb|D78168|1591-3130|ARO:3000254|emrY|Multi-drug_resistance|Multi-drug_efflux_pumps|EMRY	3.137179	0.040869
Bla|AMPH_Ecoli|AP012030|395554-396711|1158|betalactams|Penicillin_binding_protein|AMPH	3.127537	0.042188
25|JN645706.1|JN645706|betalactams|Penicillin_binding_protein|PBP2X	3.127129	0.003825
CARD|phgb|AP009048|4306556-4307588|ARO:3003548|mdtN|Multi--drug_resistance|Multi-drug_efflux_pumps|MDTN	3.113903	0.043479
CARD|pvgb|NC_002952|1417762-1419760|ARO:3003315|Staphylococcus|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|parE|RequiresSNPConfirmation	3.113369	0.005429
Tet|TetW|AJ222769|3687-5606|1920|Tetracyclines|Tetracycline_resistance_ribosomal_protection_proteins|TETW	3.10601	0.022918
CARD|phgb|U00096|1188315-1189776|ARO:3000835|phoQ|Multi-drug_resistance|MDR_regulator|PHOQ	3.069706	0.043953
CARD|phgb|D78168|536-1592|ARO:3000206|emrK|Multi-drug_resistance|Multi-drug_efflux_pumps|EMRK	3.050985	0.040245
Rif|CP003248.2|gene696|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	3.013207	0.014656
Fos|fosA_2_AGDM01000012|Fosfomycin|Fosfomycin_thiol_transferases|-FOSA	3.001212	0.026297
CARD|pvgb|NC_007779|986315-987404|ARO:3003390|Escherichia|Multi-drug_resistance|MDR_mutant_porin_proteins|OMPF|RequiresSNPConfirmation	2.941379	0.040096
Mdr|AP009048.1|gene3414|Multi-drug_resistance|MDR_regulator|CPXAR	2.936078	0.033647
Tmt|DfrA5|X12868|1306-1779|474|Trimethoprim|Dihydrofolate_reductase|DHFR|RequiresSNPConfirmation	2.872087	0.016082
Flq|NC_002695.1.916822|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	2.855839	0.037008
Mdr|CP000034.1|gene3834|Multi-drug_resistance|MDR_regulator|CPXAR	2.847657	0.043953
ACou|NC_002952.2859942|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	2.747109	0.043934
CARD|phgb|NC_007779|1621287-1621671|ARO:3000263|marA|Multi--drug_resistance|MDR_regulator|MARA	2.666231	0.048077
Mdr|NC_002695.1.914983|Multi-drug_resistance|MDR_regulator|CPXAR	2.665751	0.043109
Mdr|CP000647.1|gene2517|Multi-drug_resistance|MDR_regulator|ASMA	2.662098	0.047145
Flq|NC_002745.1124026|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	2.606064	0.043953
gi|41816|emb|X57972.1|betalactams|Class_A_betalactamases|TEM	2.598093	0.036797

Continue

Continued

Gene	log2FC	p.adj
CARD|pvgb|AL123456|759806-763325|ARO:3003283|Mycobacterium|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	2.515975	0.016082
Rif|AE000516.2|gene708|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	2.36511	0.02685
Phe|CatB4|EU935739|59054-59602|549|Phenicol|Chloramphenicol_acetyltransferases|CATB|RequiresSNPConfirmation	2.285494	0.038788

Differentially abundant genes identified using metagenomic shotgun sequencing (MGS) sequencing between antibiotic treated children and controls. To validate our Fluidigm results we performed a fitZig (zero-inflated Gaussian mixture model) analysis on the 32 samples obtained at week 1 that were analysed using shotgun metagenomics sequencing. Positive log2 fold change values indicate a gene was more abundant in the antibiotic treated children. The Benjamini-Hochberg method was applied to correct for multiple testing and adjusted p-values (p.adj) are shown.

eTable 20. List of Fluidigm primers used

AMR gene	Accession number	Forward primer	Reverse primer	Resistance to	Organism
aac(3')-Ii(acde)	HQ246166.1	TGACGTATGAGATGCCGATG	GAGAATGCCGTTTGAATCGT	aminoglycosides	*S. flexneri*
aac(6')-aph(2")	ABY79711.1	TCCAAGAGCAATAAGGGCATA	TGCCCTCGTGTAATTCATGT	aminoglycosides	***Synthetic M. genitalium JCVI-1.0*** *Enterococcus Macrococcus Staphylococcus Streptococcus*
aac(6')-Ib	KM387722.1	TTGCAATGCTGAATGGAGAG	TGGTCTATTCCGCGTACTCC	aminoglycosides	*S. senftenberg*
aac(6')-Ii	WP_002293989.1	AGACAGCTCGGCAGAAGAAG	ACCGTATTGAGGGATTGCAC	aminoglycosides	*E. faecium*
aac(6')-IIa	ACR24243.1	GGAACACTACCTGCCCAGAG	GCGACGTACGACTGAGCATA	aminoglycosides	***P. aeruginosa***
aadA	ADW23165.1	CAGCGGAGGAATTCTTTGAC	GCTGCGAGTTCCATAGCTTC	aminoglycosides	***P. aeruginosa*** *E. coli*
aadE	CAZ55809.1	TGTGCCGCAAAGAGATACTG	AACCTTCCACGACATCATCC	aminoglycosides	***S. suis*** *S. epidermidis*
aadE-like gene	AAW34138.1	GCATGATTTCCTGGCTGATT	CCACAATTCCTCTGGGACAT	aminoglycosides	***C. jejuni*** *Enterococcus*
acrA	ACI36997.1	GAAGGTAGCGACATCGAAGC	CTTTCGCCAGATCACCTTTC	aminoglycosides and beta-lactams	***E. coli O157:H7*** *Citrobacter Enterobacter Escherichia Klebsiella Salmonella Shigella*
aph(2")-Ib	AF207840.1	ATCAAATCCCTGCGGTAGTG	CAAGGGCATCCTTTTCCTTT	aminoglycosides	***E. faecium*** *E. coli*
aph(2")-Ii(de)	AAC14693.1	CGGAGGTGGTTTTTACAGGA	TTGCTTCGGCAGATTATTGA	aminoglycosides	***E. casseliflavus***

Continue

Continued

AMR gene	Accession number	Forward primer	Reverse primer	Resistance to	Organism
aph(3')-Ia, -Ic	CAQ58482.1	ATTCTCACCGGATTCAGTCG	GATTCCGACTCGTCCAACAT	aminoglycosides	*E. coli*
aph(3')-III	ACB90577.1	CCGGTATAAAGGGACCACCT	CTTTGGAACAGGCAGCTTTC	aminoglycosides	***S. pneumoniae*** *Bacillus* *Bacteroides* *Campylobacter* *Enterococcus* *Escherichia* *Roseburia* *Staphylococcus* *Streptococcus*
bla_{AMPC}	ABF06289.1	ACCGCTAAACAGTGGAATGG	GCAAGTCGCTTGAGGATTTC	beta-lactams	***S. flexneri***
bla_{CMY-2}	AAZ99133.1	CGATCCGGTCACGAAATACT	CCTGCCGTATAGGTGGCTAA	beta-lactams	***E. coli***
bla_{CTX-M}	ABG46354.1	ACTATGGCACCACCAACGAT	GGTTGAGGCTGGGTGAAGTA	beta-lactams	***E. coli***
bla_{KPC}	AEL12451.1	TGGCTAAAGGGAAACACGAC	TAGTCATTTGCCGTGCCATA	beta-lactams	***P. aeruginosa*** *Acinetobacter spp*
bla_{NDM}	CAZ39946.1	TGGATCAAGCAGGAGATCAA	ATTGGCATAAGTCGCAATCC	beta-lactams	***K. pneumonia*** *E. cloacae* *E. coli* *Proteus spp* *C. freundii* *M. morganii* *Providencia spp* *K. oxytoca*
bla_{OXA}	AAP70012.1	GTGGCATCGATTATCGGAAT	AGAGCACAACTACGCCCTGT	beta-lactams	***K. pneumoniae***
bla_{TEM}	NP_775035.1	AAGCCATACCAAACGACGAG	TTGCCGGGAAGCTAGAGTAA	beta-lactams	***C. freundii***
cblA	AAA66962.1	TGCCTGCGACATCTTGATAG	CCGTCTTCTGTTTCCGAGAG	beta-lactams	***B. uniformis***
cfxA	AY769933.1	GCGCAAATCCTCCTTTAACA	ACAATAACCGCCACACCAAT	beta-lactams	***B. fragilis***

Continue

Continued

AMR gene	Accession number	Forward primer	Reverse primer	Resistance to	Organism
ermB	BAH18720.1	GGTTGCTCTTGCACACTCAA	CTGTGGTATGGCGGGTAAGT	macrolides	***M. caseolyticus***
ermC	BAE05991.1	TGAAATCGGCTCAGGAAAAG	GGTCTATTTCAATGGCAGT-TACG	macrolides	***S. haemolyticus***
mcr-1	Liu 2016 Lancet Infect Dis	5′-TCGGACTCAAAAGGCGT GAT-3′	5′-GACATCGCGGCATTCGT-TAT-3′	colistin	***E.coli*** *K. pneumoniae* *P.aeruginosa*
mecA	YP_184944.1	TCCAGGAATGCAGAAAGACC	GGCCAATTCCACATTGTTTC	beta-lactams/ marker for MRSA	***S. aureus***
spc	AAL05549.1	TGACGAACGCAATGTGATTT	TCAGCTGCCAGATCTTTTGA	aminoglycosides	***E. faecalis***
strB	CAJ77026.1	GGCGATTATAGCCGATCAAA	CGCGACTGGAGAACATGATA	aminoglycosides	***A. baumannii***
tetQ	Y08615.1	GCAAAGGAAGGCATACAAGC	AAACGCTCCAAATTCACACC	tetracyclines	***B. fragilis***
vanA	ACP19236.1	GTGCGGTATTGGGAAACAGT	TGCGTTTTCAGAGCCTTTTT	vancomycin	***E. faecium***
vanB	WP_032489746.1	CCTGCCTGGTTTTACATCGT	GCTGTCAATCAGTGCAGGAA	vancomycin	***E. faecalis***
16S rRNA	Gloor 2010 PloS ONE	CAACGCGARGAACCTTACC	ACAACACGAGCTGACGAC		

Organisms in bold are listed under the specified accession number in the Antibiotic Resistance Database.

CHAPTER 5

Microbial community networks across body sites are associated with susceptibility to respiratory infections in infants

Marta Reyman*, Melanie Clerc*, Marlies A. van Houten, Kayleigh Arp, Mei Ling J.N. Chu, Raiza Hasrat, Elisabeth A.M. Sanders, Debby Bogaert

* Both authors contributed equally to this manuscript

Submitted

ABSTRACT

Respiratory tract infections (RTIs) are a major cause of morbidity and mortality worldwide in young children. Concepts such as the gut-lung axis have highlighted the impact of microbial communities at distal sites in mediating disease locally. However, little is known about the extent to which microbial communities from multiple body sites are linked, and how this relates to disease susceptibility. Here, we combined 16S-based rRNA sequencing data from 112 healthy, term born infants, spanning three body sites (oral cavity, nasopharynx, gut) and the first six month of life. Using a cross-niche microbial network approach, we show that, already from the first week of life on, there is a strong association between both network structure and species essential to these structures (keystone species), and consecutive susceptibility to RTIs in this cohort. Our findings underline the crucial role of cross-niche microbial connections in respiratory health.

Glossary

Term	Explanation
Niche unique species	A species/Operational Taxonomical Unit (OTU) that is unique to a certain site, i.e. does not occur at any other site.
Niche indicator species	A species/OTU that is significantly more abundant in one site or group of sites following the method of De Cáceres *et al.* (Oikos 2010).
Keystone species	A species/OTU that plays a significant role in determining the structure of a community network as defined by Banerjee *et al.* (Nat. Rev. Microbiol. 2018) and Berry and Widder (Front. Microbiol. 2014).
Network degree	The number of connections each OTU has within a network.
Network betweenness centrality	The number of shortest paths that pass through a specific OTU within a network.
Network closeness centrality	The reciprocal of the sum of the length of the shortest paths between an OTU and all other OTUs in a network.

INTRODUCTION

The human microbiome is widely recognised as an important mediator of health and disease, making it the subject of extensive study. The microbiome is highly variable between individuals, but also within one individual when studied over time or across body sites.[4] Exposures to a wide range of environmental factors, such as delivery mode, antibiotics and diet have been shown to contribute to this variation.[5] In general, microbiome research is focused on the relation between the microbial community composition of a single anatomical niche and health or disease parameters.

However, it is becoming increasingly clear that direct or indirect effects of microbial communities at distal body sites can also have major implications for both development as well as severity of a number of diseases.[6] One well-established example of cross-niche microbial interaction is the gut-lung axis. It has been shown that microbial components and metabolites in both the gut and lung are capable of modulating immunity not only locally, but also systemically.[7] Additionally, specific taxa in both the gastrointestinal and respiratory tract are associated with lung diseases such as asthma, chronic obstructive pulmonary disease and respiratory tract infections (RTIs).[7]

Next to variation in species presence/absence and abundance within niches, other features such as the microbial network structure may also contribute to resilience or susceptibility to RTIs. Commonly, studies of microbiome networks have been restricted to microbial networks within a single niche.[8,9] For instance, a study performed in patients suffering from inflammatory bowel disease (IBD) showed that the community structure of the microbial network within the gut was distorted in IBD patients.[10] However, it has been theorized that a microbe in a specific niche might not only affect the network structure within that niche, but also influence the wider human microbiome network.[11] Hence, exploring the relationships between bacteria and the microbial networks they form, both locally and across body niches, may help to gain insight into their health-promoting or pathological properties.

We therefore investigate the structure of microbial networks across different body sites, in order to identify signatures in the overall microbial community network in infancy that can be associated with susceptibility to RTIs in the first year of life. To do so, we used 16S rRNA sequencing data from samples of 112 healthy, term born infants collected longitudinally over the first six months of life and spanning three essential body sites (oral cavity, nasopharynx and gastrointestinal tract). We

aimed to 1) compare the development of the microbial communities per niche over time, 2) build cross-niche microbial networks, 3) identify keystone species within the networks, and 4) study network structure and keystone species in relation to RTI susceptibility during the first year of life.

METHODS

Data Collection

A detailed description of the study design and inclusion criteria of the birth cohort can be found elsewhere.[13] In short, we used microbiota data from samples obtained from 112 healthy infants at four different time points (week 1, month 2, month 4 and month 6) and from three different niches, namely the gastrointestinal tract (faecal samples), the upper respiratory tract (nasopharyngeal swabs) and the oral cavity (saliva samples). Metadata was available for the first year of life, but matched samples collected in parallel were only available until month 6. Written informed consent was obtained from parents of all children and the study was approved by METC Noord-Holland (M012-015, NTR3986).

16S rRNA gene sequencing

DNA extraction and library preparation of the bacterial 16S rRNA gene V4 region was performed as described previously.[14] Samples were spread over 27 individual library pools, each containing DNA extraction and qPCR blanks as negative controls, and a mock community (positive control). All pools were sequenced on the Illumina MiSeq platform (Illumina Inc., San Diego, CA, USA). The raw sequencing reads were trimmed using Sickle v1.33,[15] error corrected using BayesHammer (SPAdes v3.8.1)[16] and assembled using PANDAseq v2.10.[17] Following removal of chimeric sequences, reads were grouped into Operational Taxonomic Units (OTUs) using VSEARCH v2.0.3 with a 97% similarity threshold.[18] Taxonomic annotation was performed using QIIME v1.9.1 based on the SILVA database v119.[19,20] We removed spurious OTUs by filtering at a minimum relative abundance of 0.1% and presence in at least two samples.[21] We combined each OTU name with a number representing its rank in the OTU table, based on the relative abundance in the overarching dataset to discriminate OTUs with the same taxonomic annotation.

Statistical analysis

Analyses regarding the microbiota composition have been performed in R version 3.4.3, while network analyses have been performed in R versions 3.5.0 and 3.6.2.[22]

Niche comparison

Results of analyses of observed species richness are based on raw read counts. Similar results were namely obtained when rarefying read counts to a sequencing depth of 3000 reads (data not shown). We used the *lmer* function (lme4 package)[23] to study the temporal changes of species richness within and across niches, with participant added as random effect to take repeated measures into account. To compare compositions of microbial communities across niches, we used nonmetric multidimensional scaling (nMDS) plots based on ordinations that used the Bray-Curtis (BC) dissimilarity matrix of relative abundance data as input (function *ordinate* with parameter trymax 10,000; vegan package).[24] Stability of the microbiota composition in the first 6 months of life within each niche was visualised by measuring the BC dissimilarities between consecutive samples of each participant, and we tested differences across niches using Wilcoxon's signed rank test.

We used permutational multivariate analysis of variance (PERMANOVA) with 1999 permutations to investigate the associations between microbiota composition (response variable) and environmental covariates using the function *adonis2* (vegan package).[24] The covariates tested were mode of delivery (vaginal vs. caesarean section), season of birth (Summer, Autumn, Winter, Spring), hospital stay duration after birth in day parts, presence of siblings <5 years of age (yes vs. no), presence of pets in the household (none, cat(s), dog(s), cat(s) and dog(s), other), breastfeeding at sampling moment (yes vs. no), attendance of day care at sampling moment (yes vs. no), use of pacifier at sampling moment (yes vs. no) and use of antibiotics in the month prior to sampling moment (yes vs. no). Covariates that were significantly associated in univariate models (per time point per niche) with microbiota community composition at a minimum of 1 time point were included in a multivariate model that included niche as well as age, and subject to control for repeated measures. P-values were adjusted for multiple testing using the Benjamini-Hochberg method.[25]

Network construction and cluster definition

Infants were divided into three groups based on the distribution of the number of RTIs they experienced over the first year of life (0-2 RTIs, 3-4 RTIs, 5-7 RTIs), as reported by the parents.[14] Prior to network construction, we split the data by RTI group and filtered the OTU tables to only include the 100 most abundant OTUs within each niche within each RTI group in order to avoid overestimation of the impact of very rare taxa on the overall network structure. Because of overlapping OTUs between niches, this resulted in a final dataset that included 220 OTUs for

the 0-2 RTI group, 214 OTUs for the 3-4 RTI group and 228 OTUs for the 5-7 RTI group. 149 OTUs appeared in all three groups (Supplementary Figure S1).

We defined niche indicator OTUs as OTUs that were characteristic for a particular niche. To identify such OTUs, we ran a niche indicator species analysis using the function *multipatt* (indicspec package).[1] This function calculates an indicator value for each OTU at each niche, taking its total abundance per niche into account. An OTU was deemed a niche indicator when it had an association value of >0.5 with a p-value <0.05 for one specific niche.

For network construction, we used the SpiecEasi pipeline with Meinshausen-Buhlmann estimation.[26] This method is based on the concept of conditional independence rather than correlation, making it less likely to detect spurious connections between taxa that are indirectly connected, but not directly connected. Further, it uses centred log-ratio transformation of the data to overcome the compositionality of microbiome data. In total, we constructed 12 networks (3 RTI groups*4 time points) using the methods described above. We used random walks (number of steps = 5) to group OTUs into clusters that were closely associated using the function *walktrap.community* (igraph package).[27] We considered clusters with 10 or more OTUs as "main clusters", as smaller clusters usually contained only 1 or 2 OTUs (Supplementary Figure S2). The OTUs within a cluster could either be niche indicator OTUs (saliva, faeces or nasopharynx) or non-niche indicator OTUs. Therefore, if the composition of a cluster was random, we would expect an equal proportion of either of these four OTU classes within a cluster, i.e. each OTU class would have a prevalence of 25%. We considered a cluster to be niche-specific if one of the OTU classes exceeded that threshold value of 25%. In addition, niche indicator OTUs only made up for between 5.7% and 11.5% of OTUs within an RTI group, making this threshold even more conservative. If there were two classes of niche indicator OTUs that met this threshold within a single cluster, then the largest set determined the niche annotation. If there were not enough niche indicator OTUs that met this threshold, the cluster was defined as a "mixed (niche) cluster".

Cluster stability analysis

We compared the composition of all clusters from each time-specific network per RTI group using an n*n matrix, calculating the difference in cluster composition from a scale from 0 to 1, where 1 indicates complete similarity (i.e. two clusters share 100% of their OTUs). We used ANOVA to assess whether cluster similarity was different across niches.

Keystone species analysis

To identify keystone species (i.e. OTUs playing a significant role in determining the structure of a community network), we calculated three important network metrics for each OTU within a network at each time point: 1) degree (the number of connections each OTU has within a network), 2) betweenness centrality (number of shortest paths that pass through a specific OTU within a network) and 3) closeness centrality (the reciprocal of the sum of the length of the shortest paths between an OTU and all other OTUs in a network). We followed the definitions from Banerjee *et al*. and Berry and Widder by defining keystone species as those OTUs with a high degree, high closeness centrality and low betweenness centrality.[2,3] We defined high degree or closeness centrality as the top 10% of the distribution of each of those metrics, and low betweenness centrality as the bottom 10% of the distribution for betweenness. By doing so, we ensured a selection of OTUs that were at the extremes of the distributions for all three values (see Supplementary Figure S3). We selected OTUs within RTI groups that met at least one of those conditions as keystone species.

RESULTS

Niche comparison

Of the 1250 available samples obtained from the three niches at 1 week, 2, 4 and 6 months of life of 112 healthy infants, 1248 samples fulfilled our quality threshold (433 faecal samples, 430 nasopharyngeal samples and 385 saliva samples). Sequencing of these samples resulted in a total of 58,608,834 high quality reads with a minimum Good's coverage of 99.47% (median 99.97%). The overall OTU-table, including the samples of the three niches collected at the four time points, contained 1148 bacterial OTUs distributed over 18 bacterial phyla, with *Firmicutes* being the most abundant phylum and *Streptococcus* the most abundant genus.

In the oral cavity, the most abundant genus over the first 6 months of life was *Streptococcus,* while this was *Moraxella* for the nasopharynx and *Bifidobacterium* for the gut. Observed species richness was highest in the nasopharyngeal samples with 895 OTUs identified, compared to 746 OTUs in the saliva and 595 OTUs in the faecal samples. In addition, the nasopharynx also contained the highest number of unique OTUs (defined as observed in a single niche only) namely 232, followed by 121 in the gut and only 38 in the oral cavity. In other words, the oral niche contained a relatively high number of observed species, but few unique OTUs, and so it seems to be a reservoir sharing many OTUs with the nasopharynx and gut. Although

highly variable in abundance, a total of 331 overlapping OTUs were observed in all three niches, including the top 10 most abundant OTUs in the overall dataset of the combined niches: *Streptococcus* (1), *Bifidobacterium* (2), *Moraxella* (3), *Staphylococcus* (4), *Corynebacterium propinquum* (5), *Streptococcus salivarius* (6), *Dolosigranulum* (7), *Escherichia coli* (9), *Veillonella* (10) and *Haemophilus* (8).

Supplementary Figure S4 shows the succession patterns of the 15 most abundant OTUs in each niche, highlighting the gradual increase of *Bifidobacterium* (2) over time in the gut (Supplementary Figure S4A). In the nasopharynx, the initial high abundance of *Staphylococcus* (4) at 1 week of life was gradually replaced by an increasing abundance of *C. propinquum* (5), *Dolosigranulum* (7), *Moraxella* (3) and *Haemophilus* (8) (Supplementary Figure S4B). Lastly, the figure highlights an overall dominance of *Streptococcus* (1) in the oral cavity (Supplementary Figure S4C). Observed species richness increased over time in both faecal and saliva samples, though this effect was most pronounced in saliva (linear mixed effect model including age and diversity per niche: $p<0.0001$ for both niches; Supplementary Figure S5).

Microbial community development over time (shown in Figure 1A) was most stable for the oral cavity in early-life, although this was overtaken by a higher stability in the gut at the later timepoints (median BC dissimilarity between month 4 and 6 for faecal samples 0.18 versus 0.29 for saliva, Wilcoxon test, $p<0.0001$; Figure 1B). The microbial community composition in the nasopharynx showed the lowest temporal stability compared to the other two niches over the first 6 months of life. As expected, the community composition differed significantly between the three niches at all time points, although niche composition was most similar at the earliest time point, in line with the mutual origin of initial microbial seeding at birth. This was followed by a gradual deviation into niche-specific communities over time (calculating the association of niche with composition using PERMANOVA, week 1: R^2 31.8%, month 2: R^2 47.8%, month 4; R^2 57.5%, month 6: R^2 56.3%, all p-values <0.0001).

Characteristics explaining the overall community composition across niches and time points were in order of importance, niche itself (R^2 42.4%, adjusted p-value 0.00079), followed by age with a much smaller effect size (R^2 1.8%, adjusted p-value 0.00079), and presence of siblings <5 years, pets, breastfeeding at time of sampling, season of birth, mode of delivery, exposure to antibiotics in the month prior to sampling, day care attendance, pacifier use, and duration of hospital stay after birth (Supplementary Table S1). When studying the niches separately and

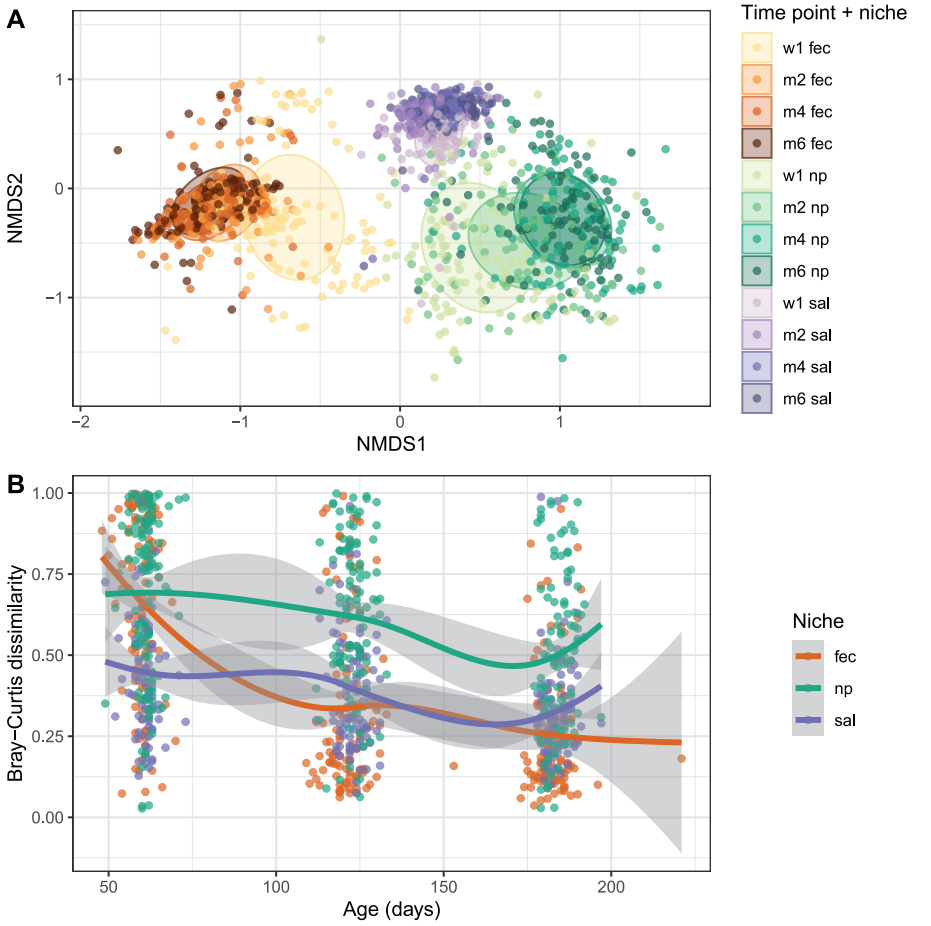


Figure 1. Microbiota composition development per niche over time. Non-metric multidimensional scaling (nMDS) plot (**A**), based on Bray-Curtis (BC) dissimilarity between samples, visualising the overall gut microbial community composition stratified for each niche per time point. Each data point represents the microbial community composition of one sample. The ellipses represent the standard deviation of data points belonging to each group, with the centre points of the ellipses calculated using the mean of the coordinates per group. The stress of the ordination was 0.18. In panel **B** the temporal community stability for each niche is shown. Data points represent the BC dissimilarity between consecutive sample pairs belonging to each individual per time interval, i.e. between week 1 and month 2, between month 2 and month 4, and between month 4 and month 6. Loess lines were fitted over the data points per niche, and the gray areas represent the 0.95 confidence intervals. Temporal stability was highest for the saliva samples in early-life, which was later overtaken by a higher stability in consecutive faecal samples (median BC dissimilarity between month 4 and 6 for faecal samples 0.18 versus 0.29 for saliva, Wilcoxon test, $p<0.0001$).

cross-sectionally per time point, we observed different effect sizes of associations between environmental variables and microbial community composition (Supplementary Table S2). In total, we identified eight covariates to be associated with the nasopharyngeal microbiome (among others presence of siblings <5 years, breastfeeding at sampling and season of birth), seven covariates for the faecal microbiome (among others mode of delivery, breastfeeding at sampling and day care attendance), and only three covariates for the salivary microbiome composition (breastfeeding at sampling, antibiotic use 1 month prior to sampling and season of birth).

Network analysis

We identified a total of 107 niche-indicator OTUs (defined as characteristic, but not exclusive, for a particular niche), with 18 niche indicator OTUs specific for the nasopharynx, 37 for the gut and 52 for the oral cavity. Participants were divided into three groups based on the distribution of the number of RTIs they experienced over the first year of life (0-2 RTIs, 3-4 RTIs, 5-7 RTIs). After selecting the 100 most abundant OTUs per niche and per RTI group, we constructed for each of the three RTI groups a cross-niche microbial network per time point using the SpiecEasi pipeline.[26] We studied the formation of bacterial clusters within these cross-niche networks, and whether clusters were indicative of a specific niche or not, based on the proportion of niche-indicator OTUs the clusters contained. In the results below, we will further refer to the cross-niche networks as networks.

We found that at 1 week of age, the network of the least susceptible (0-2 RTI) group was structured into 6 clusters (1 indicative of nasopharynx, 1 saliva, 2 faeces and 2 mixed), while the network of the average susceptible (3-4 RTI) group contained 7 clusters (1 nasopharynx, 1 saliva, 2 faeces and 3 mixed) and the network of the most susceptible (5-7 RTI group) contained 9 clusters (2 nasopharynx, 2 saliva, 3 faeces and 2 mixed; Supplementary Figure S6). The number of clusters per network generally decreased over time with cluster sizes increasing (Figure 2A and B). However, though the networks of the lowest and average susceptible groups were ultimately defined by 4 clusters each, the network from the most susceptible group still contained 6 clusters at the age of 6 months, suggesting more fragmented networks for the latter group. The two extra clusters identified in the most susceptible group (1 extra saliva and 1 extra faecal cluster) thereby split the oral and gut communities into two communities each (Supplementary Figure S6).

We then studied the similarity of clusters over time, as a measure of temporal network stability (Figure 3). The mean similarity over time differed between

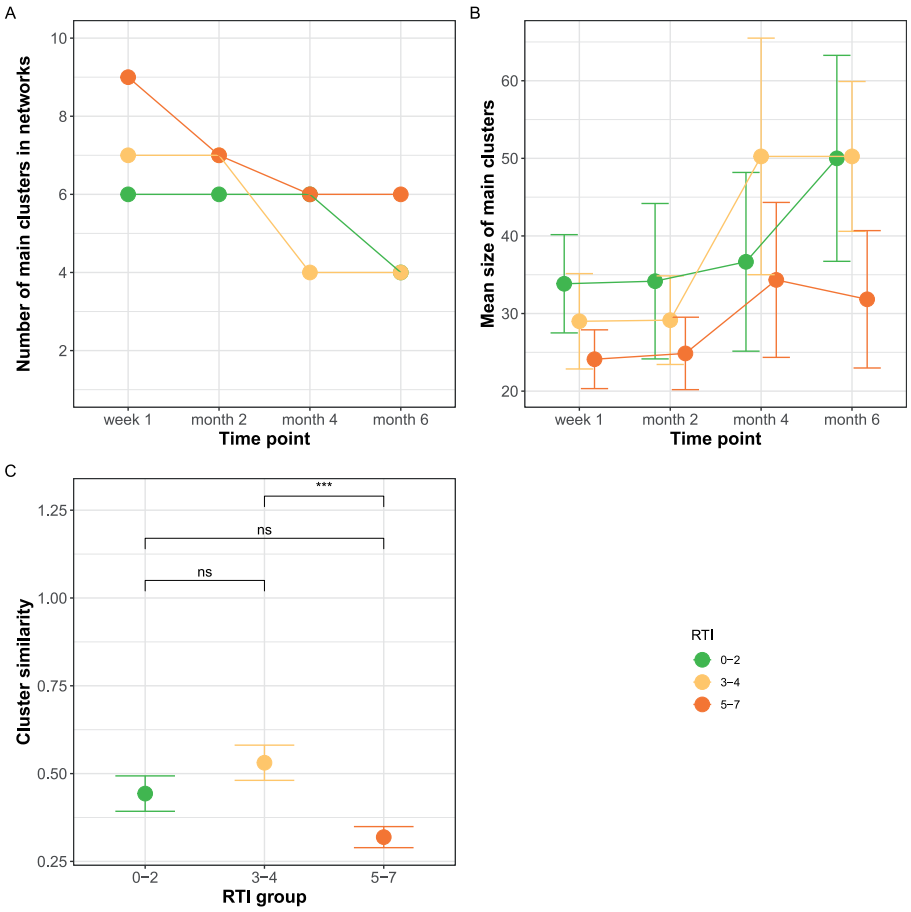


Figure 2. Network characteristics of the cross-niche microbial networks built per RTI group. Number of main clusters (**A**) and mean size of main clusters (**B**) over time. (**C**) Mean cluster similarity for all clusters per RTI group across time. Whiskers represent minimum and maximum values.

RTI groups, with 44.3% ± 5.0 SE similarity over time for the least susceptible group, 53.1% ± 5.0 SE for the average susceptible group and 31.9% ± 3.0 SE for the most susceptible group (Kruskal-Wallis test: X2=11.7, p=0.003, Figure 2C). Specifically, the similarity of clusters in the networks of the most susceptible group was lower compared to the other two RTI groups (Dunn post-hoc test, 0-2 vs 3-4 RTIs: p=0.173, 0-2 vs 5-7 RTIs: p=0.097, 3-4 vs 5-7 RTIs: p=0.002; Figure 2C). In stratified analyses per niche, we found significant differences in cluster similarity between RTI groups for saliva clusters (similarity over time 44.0% ± 9.9 SE in the 0-2 RTI group, 78.9% ±6.7 SE in the 3-4 RTI group, and 26.6% ± 6.3 SE in the 5-7 RTI group, Kruskal-Wallis test X2=14.0, p=0.0009, Supplementary Figure

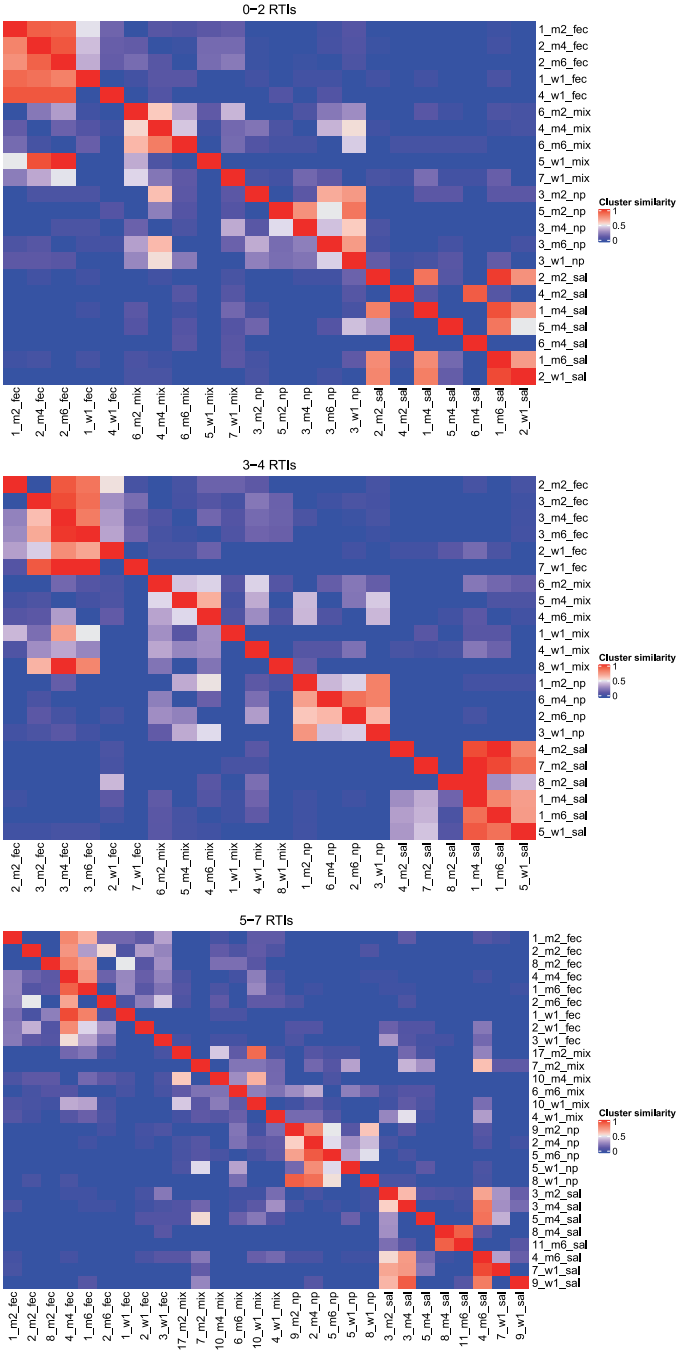


Figure 3. Cluster stability. Heatmaps displaying the similarity between the cluster compositions of all clusters at all time points for networks for the 0-2 RTI group, 3-4 RTI group and 5-7 RTI group, respectively.

S7). In addition, we found non-significant trends for the association between RTI susceptibility and loss in cluster similarity for faecal clusters (similarity over time 49.9% ± 11.1 SE in the 0-2 RTI group, 49.4% ± 9.3 SE in the 3-4 RTI group, and 34.0% ± 4.5 SE in the 5-7 RTI group, Kruskal-Wallis test X^2=3.27, p=0.195, Supplementary Figure S7) and nasopharyngeal clusters (similarity over time 58.9% ± 7.4 SE in the 0-2 RTI group, 65.6% ±7.9 SE in the 3-4 RTI group, and 44.2% ± 6.4 SE in the 5-7 RTI group, Kruskal-Wallis test X2=4.34, p=0.114). Overall, this suggests reduced cluster stability over time in networks from children from the 5-7 RTI group.

Keystone species

To identify OTUs with an important role for the network structure, we calculated the degree, betweenness centrality and closeness centrality for each of the OTUs in each network. Table 1 lists the 134 keystone species found, also reporting whether they were niche indicator species or not, and for which RTI network (least, average, or most susceptible) they were identified as being important. We identified only 5 OTUs (*Akkermansia* (280), *Moraxella* (190), *Actinomyces* (58), *Veillonella* (10) and *Prevotella melaninogenica* (22)) as keystone species in all three RTI groups, indicating that different OTUs are central to the network structure of the different RTI groups.

To assess whether the identified keystone species were high or low abundant species, we calculated summary statistics for all keystone species split by the RTI group(s) for which they were identified as keystone species (Supplementary Table S3). The mean abundance of keystone species in the 0-2 RTI group networks was 0.0150% ± 0.54 SD (range 0 – 55.6%), 0.34% ± 3.8 SD for the 3-4 RTI group (range 0–99.5%) and 0.12% ± 1.8 SD (range 0 – 96.5%) for the 5-7 RTI group, underlining keystone OTUs are often low abundant species. Furthermore, most keystone species were not identified as niche-indicator OTUs. This suggests that niche-indicator OTUs are less relevant for the microbial network structure, and vice versa, that keystone species are less niche-specific.

The keystone species unique to the 0-2 RTI networks included previously reported respiratory health-associated OTUs such as *Neisseria lactamica* (47), a low abundant lactic acid producing *Dolosigranulum pigrum* (147), and *Corynebacterium* (162 and 111) OTUs. Also, OTUs associated with the production of butyrate, such as *Ruminococcus bromii* (112), *Megasphaera* (133 and 383) and *Anaerostipes* (142) were found to be keystone species in the 0-2 RTI networks. Alternatively, keystone species in the 5-7 RTI networks included OTUs previously associated

Table 1. Keystone species for all RTI cohorts

OTU	Niche Indicator	Keystone ID
Akkermansia_280	none	all
Moraxella_190	none	all
Actinomyces_58	saliva	all
Prevotella_melaninogenica_22	saliva	all
Veillonella_10	saliva	all
7B_8_687	none	0-2
Acinetobacter_calcoaceticus_146	none	0-2
Aggregatibacter_354	none	0-2
Alloscardovia_316	none	0-2
Altererythrobacter_748	none	0-2
Anaerostipes_142	none	0-2
Blastococcus_281	none	0-2
Blastococcus_707	none	0-2
Blautia_212	none	0-2
Chroococcidiopsis_338	none	0-2
Chroococcidiopsis_593	none	0-2
Chryseobacterium_299	none	0-2
Corynebacterium_111	none	0-2
Corynebacterium_162	none	0-2
Craurococcus_sp_HM28_1_726	none	0-2
Cupriavidus_31	none	0-2
Curvibacter_286	none	0-2
Dorea_86	none	0-2
Erysipelotrichaceae_149	none	0-2
Lachnospiraceae_176	none	0-2
Limnobacter_245	none	0-2
Megasphaera_133	none	0-2
Megasphaera_sp_TrE9262_383	none	0-2
Moraxella_131	none	0-2
Neisseria_lactamica_47	none	0-2
Pseudobutyrivibrio_195	none	0-2
Rubellimicrobium_596	none	0-2
Ruminococcus_bromii_L2_63_112	none	0-2
Sphingomonas_573	none	0-2
Subdoligranulum_71	none	0-2
Varibaculum_199	none	0-2
Escherichia_Shigella_267	faeces	0-2

Continue

Continued

OTU	Niche Indicator	Keystone ID
Comamonadaceae_118	np	0-2
Dolosigranulum_pigrum_ATCC_51524_147	np	0-2
Enhydrobacter_138	np	0-2
Pseudomonas_158	np	0-2
Haemophilus_211	saliva	0-2
Helcococcus_179	none	0-2 and 3-4
Scardovia_wiggsiae_F0424_127	none	0-2 and 3-4
Bifidobacterium_2	faeces	0-2 and 3-4
Rothia_259	saliva	0-2 and 3-4
Streptococcus_1	saliva	0-2 and 3-4
Actinomyces_100	none	0-2 and 5-7
Blautia_308	none	0-2 and 5-7
Moryella_119	none	0-2 and 5-7
Gallibacterium_salpingitidis_141	saliva	0-2 and 5-7
Akkermansia_210	none	3-4
Akkermansia_40	none	3-4
boneC3G7_70	none	3-4
Clostridium_colinum_487	none	3-4
Clostridium_nexile_166	none	3-4
Corynebacterium_500	none	3-4
Corynebacterium_aurimucosum_ATCC_700975_260	none	3-4
Haemophilus_226	none	3-4
Lachnoanaerobaculum_115	none	3-4
Lactobacillus_plantarum_337	none	3-4
Modestobacter_323	none	3-4
Moraxella_204	none	3-4
Morganella_294	none	3-4
Neisseria_345	none	3-4
Peptococcus_like_sp_oral_clone_I070_371	none	3-4
Propionibacterium_312	none	3-4
Sphingomonas_228	none	3-4
Turicella_264	none	3-4
Zymomonas_335	none	3-4
Bifidobacteriaceae_213	faeces	3-4
Clostridium_sensu_stricto_1_33	faeces	3-4
Collinsella_220	faeces	3-4
Eggerthella_189	faeces	3-4
Klebsiella_13	faeces	3-4

Continue

Continued

OTU	Niche Indicator	Keystone ID
Corynebacterium_350	np	3-4
Moraxella_3	np	3-4
Alloprevotella_23	saliva	3-4
Bergeyella_108	saliva	3-4
Lachnoanaerobaculum_128	saliva	3-4
Lactobacillales_63	saliva	3-4
Rothia_126	saliva	3-4
Rothia_16	saliva	3-4
Rothia_192	saliva	3-4
Rothia_287	saliva	3-4
Solobacterium_moorei_232	saliva	3-4
Streptococcus_20	saliva	3-4
Veillonella_66	saliva	3-4
Collinsella_110	none	3-4 and 5-7
Lachnospiraceae_98	none	3-4 and 5-7
Campylobacter_123	saliva	3-4 and 5-7
Veillonella_sp_DNF00869_15	saliva	3-4 and 5-7
Actinobaculum_schaalii_FB123_CNA_2_408	none	5-7
Bacteroides_84	none	5-7
Capnocytophaga_196	none	5-7
Erysipelotrichaceae_102	none	5-7
Erysipelotrichaceae_231	none	5-7
Flavonifractor_200	none	5-7
Fusicatenibacter_saccharivorans_44	none	5-7
Haemophilus_haemolyticus_663	none	5-7
Lachnospiraceae_144	none	5-7
Lachnospiraceae_230	none	5-7
Lachnospiraceae_302	none	5-7
Lactobacillus_fermentum_91	none	5-7
Moraxella_153	none	5-7
Moraxella_412	none	5-7
Moraxella_494	none	5-7
Neisseria_1000	none	5-7
Neisseria_469	none	5-7
Neisseria_527	none	5-7
Neisseria_949	none	5-7
Neisseria_meningitidis_830	none	5-7
Parascardovia_denticolens_F0305_486	none	5-7

Continue

Continued

OTU	Niche Indicator	Keystone ID
Peptoniphilus_sp_S7MS8_148	none	5-7
Peptostreptococcaceae_117	none	5-7
Peptostreptococcaceae_75	none	5-7
Prevotella_sp_oral_clone_ID019_94	none	5-7
Ruminococcaceae_160	none	5-7
Ruminococcaceae_358	none	5-7
Streptococcus_salivarius_subsp_thermophilus_364	none	5-7
Subdoligranulum_249	none	5-7
Bacteroides_74	faeces	5-7
Bifidobacterium_dentium_Bd1_28	faeces	5-7
Blautia_19	faeces	5-7
Collinsella_26	faeces	5-7
Erysipelotrichaceae_78	faeces	5-7
Parabacteroides_distasonis_82	faeces	5-7
ratAN060301C_27	faeces	5-7
Ruminococcus_gnavus_CC55_001C_24	faeces	5-7
Haemophilus_8	np	5-7
Actinomyces_sp_oral_clone_DR002_81	saliva	5-7
Haemophilus_215	saliva	5-7
Haemophilus_413	saliva	5-7
Veillonella_156	saliva	5-7

The column Niche Indicator shows whether an OTU is also a niche indicator OTU. Keystone ID indicates for which RTI group (or combination of RTI groups) an OTU is a keystone species. Rows are ordered firstly by Keystone ID, then Niche Indicator (none, faeces, np or saliva), then on OTU in alphabetical order. Np = nasopharynx.

with poorer respiratory health, such as the proteobacterial OTUs *Haemophilus* (8, 215, 413), *Haemophilus haemolyticus* (663), and *Neisseria* (469, 527, 949, 1000), and the anaerobic *Lachnospiraceae* (144, 230, 302). Additionally, *Actinobaculum schaalii* (408) often associated with (invasive) infections, the proinflammatory *Ruminococcus gnavus* (24) and the faecal pH increasing *Peptostreptococcaceae* (75 and 117) were keystone species for the networks of infants who experienced 5-7 RTIs. Also, cariogenic species such as *Bifidobacterium dentium* (28) and *Parascardovia denticolens* (486) were identified as keystone species in the 5-7 RTI networks.

DISCUSSION

Previous studies from this healthy birth cohort showed already that the nasopharyngeal, oral and gut microbiota individually develop along specific trajectories, and that various environmental factors are associated with the microbial community composition and development of these three niches in the first months of life. Also, associations between the microbiome development of each of these separate niches and respiratory health have been observed.[14,28,29] Although this suggests an interplay between the microbial communities within these niches, this phenomenon has remained unstudied. Here, we studied cross-niche infant microbiota composition and development, and its relationship with susceptibility to RTIs. To our knowledge, this is the first study using a cross-niche microbial network strategy to study the composition of the wider human microbiome in relation to (respiratory) health.

In the rapidly developing early-life microbiome, we observed that the anatomical niche was by itself the most important explanatory variable for the microbial community composition, suggesting that niche environment is the main driver of the human microbiome. We furthermore confirmed additional environmental drivers of microbiota composition and development, among others mode of delivery, feeding type and the presence of siblings <5 years in the household.[13,28,29]

Building cross-niche microbial networks allowed us to additionally investigate the connections between local and distal microbial communities across the body, and how this related to respiratory health. In doing so, we were not only able to identify clusters of OTUs in the cross-niche networks that were niche-specific, but also clusters of bacteria that were niche-independent. The network clusters found to be dominated by niche-specific bacteria are in line with previous results from human microbiome network studies.[11]

When analysing the cross-niche network structure in relation to RTI susceptibility, we observed that the number of clusters within a network was highest for the most susceptible group, while the temporal stability of clusters was lowest. This demonstrates more fragmented and less stable networks in the first 6 months of life of infants with higher susceptibility to RTIs. This finding of fragmentation is in line with a study focusing on the structure of (niche-specific) gut microbial co-occurrence networks in IBD patients, which was more distorted when compared to networks in healthy individuals.[10]

When studying keystone OTUs central to the respective cross-niche networks,[2,3] we observed little overlap between RTI groups. We found OTUs previously associated with low susceptibility to, and low severity of, respiratory infections, such as *Corynebacterium, D. pigrum* and *N. lactamica*, as keystone species in our least susceptible group. Both *Corynebacterium* and *D. pigrum* have been consistently associated with an adequate microbiota maturation and a decreased risk of developing RTIs in later life.[28,30,31] Furthermore, *N. lactamica* has been shown in human challenge studies to inhibit colonization by the pathogenic *Neisseria meningitidis,* which supports that *N. lactamica* plays a role in a resilient microbial community network.[32] Inversely, we observed keystone species that were previously associated with recurrent respiratory infections, such as *Haemophilus* and *Lachnospiraceae,* in the networks of the group with most RTIs, but generally not in the group with few RTI episodes.[28,33–35] Therefore, our data suggests that the presence of these bacteria is not incidental, but potentially central to a less beneficial bacterial community structure associated with more RTI episodes.

Not only did we identify keystone species that were previously associated with respiratory health, but we also identified OTUs that were previously associated with gut health, oral health, immunological diseases and infections as keystone species in the networks of infants experiencing either 0-2 or 5-7 RTIs. With respect to gut health, OTUs known to produce or enhance the production of butyrate, such as *R. bromii* (112), *Megasphaera* (133 and 383) and *Anaerostipes* (142) were found to be keystone species in the 0-2 RTI networks.[36–38] The short chain fatty acid butyrate is a microbial end-product of the human gut fermentation process and an essential metabolite in the gut environment, being the preferred energy source for colon epithelial cells. It has anti-inflammatory properties and lowers the pH of the gut, in this way inhibiting growth of pathogens.[36] Conversely, *Peptostreptococcaceae* spp. (75 and 117), previously associated with increased faecal pH, were identified as keystone species for the 5-7 RTI networks.[39]

Additionally, in the 5-7 RTI networks, the mucin degrader *R. gnavus* (24) was identified as a keystone species. This OTU has previously been associated with a broad scale of immunological disorders, such as paediatric allergy, IBD and psoriatic arthritis, and also with failure of faecal microbiota transplantation.[40–45] Also in the most susceptible RTI networks, we observed *A. schaalii* (408), involved in urinary tract infections, as keystone species, as well as the cariogenic *B. dentium* (28) and *P. denticolens* (486).[46–48] Altogether, this suggests that cross-niche bacterial networks may stand at the basis of overall systemic susceptibility to inflammation-driven health problems. It further implies that the overall construction of microbial

networks, rather than merely the presence/absence or abundance of specific commensals or pathogens, is the driving force behind inflammation-mediated disease.

Of note, in our study, although the more predominant OTUs were generally niche-specific, most keystone network OTUs were actually lower abundant taxa, and not associated with a specific niche. This supports the "rare taxa" concept, which postulates that the abundance of a species is not necessarily the best determinant for its importance within the microbial community structure.[49] Studying the human microbiome in a more generic context therefore might provide insight into the importance and roles of lesser known microbes.

One key strength of our study is the integrated use of microbiota data across different body sites to build cross-niche microbial networks in a longitudinal fashion, using samples collected from a prospectively followed birth cohort. By using a cross-niche network approach with an algorithm based on conditional independence rather than correlation, we showed that besides the known associations between niche-specific microbiota and respiratory health, there is likely also an association between the overall human microbiome network features and respiratory health. Furthermore, we were able to identify keystone species that were important for the network structure, which were often niche-independent, suggesting these microbes might be missed or their value might be underestimated, when the microbiota is studied within a single niche. The most important limitation of our study is inherent to its observational design, namely the associative rather than causative nature of our findings. Furthermore, we are potentially lacking statistical power when comparing network cluster similarities across RTI groups within niches, as there were only between one and four niche-specific clusters in any one network.

In conclusion, we observed cohesive and stable microbial networks across body sites already from early-life on, which in turn were associated with a lower susceptibility for RTIs in the first year of life. In contrast, we found more fragmented and unstable cross-niche networks over the first 6 months of life in infants with higher susceptibility to RTIs. Finally, we identified bacteria that were central to each of the cross-niche microbial network structures, though these differed between RTI groups. These bacteria were often not indicative for a specific niche within the body, and represented low abundant species, underlining the potential importance of low abundant bacteria for microbiome function.

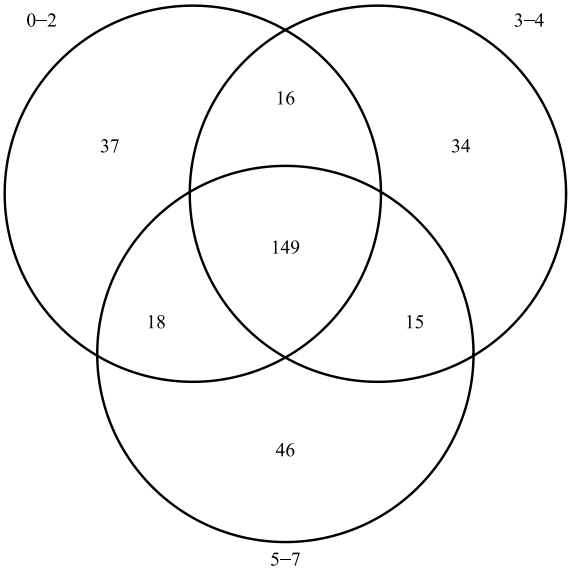
REFERENCES

1. De Cáceres, M., Legendre, P. & Moretti, M. Improving indicator species analysis by combining groups of sites. *Oikos* **119**, 1674–1684 (2010).
2. Banerjee, S., Schlaeppi, K. & van der Heijden, M. G. A. Keystone taxa as drivers of microbiome structure and functioning. *Nature Reviews Microbiology* **16**, 567–576 (2018).
3. Berry, D. & Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front. Microbiol.* **5**, 219 (2014).
4. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220–230 (2012).
5. Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **8**, 343ra82-343ra82 (2016).
6. Zhang, D. *et al.* The Cross-Talk Between Gut Microbiota and Lungs in Common Lung Diseases. *Frontiers in Microbiology* **11**, (2020).
7. Budden, K. F. *et al.* Emerging pathogenic links between microbiota and the gut–lung axis. *Nat. Rev. Microbiol.* **15**, 55–63 (2016).
8. Hall, C. V. *et al.* Co-existence of Network Architectures Supporting the Human Gut Microbiome. *iScience* **22**, 380–391 (2019).
9. Soret, P. *et al.* Respiratory mycobiome and suggestion of inter-kingdom network during acute pulmonary exacerbation in cystic fibrosis. *Sci. Rep.* **10**, 3589 (2020).
10. Baldassano, S. N. & Bassett, D. S. Topological distortion and reorganized modular structure of gut microbial co-occurrence networks in inflammatory bowel disease. *Sci. Rep.* **6**, 1–14 (2016).
11. Faust, K. *et al.* Microbial co-occurrence relationships in the Human Microbiome. *PLoS Comput. Biol.* **8**, (2012).
12. Hoen, A. G. *et al.* Associations between gut microbial colonization in early life and respiratory outcomes in cystic fibrosis. in *Journal of Pediatrics* **167**, 138-147.e3 (Mosby Inc., 2015).
13. Bosch, A. A. T. M. *et al.* Development of Upper Respiratory Tract Microbiota in Infancy is Affected by Mode of Delivery. *EBioMedicine* **9**, 336–345 (2016).
14. Bosch, A. A. T. M. *et al.* Maturation of the Infant Respiratory Microbiota, Environmental Drivers, and Health Consequences. A Prospective Cohort Study. *Am. J. Respir. Crit. Care Med.* **196**, 1582–1590 (2017).
15. Joshi, N. A. & Fass, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). (2011). Available at: https://github.com/najoshi/sickle.
16. Nikolenko, S. I., Korobeynikov, A. I. & Alekseyev, M. A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* **14**, S7 (2013).
17. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* **13**, 31 (2012).
18. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).

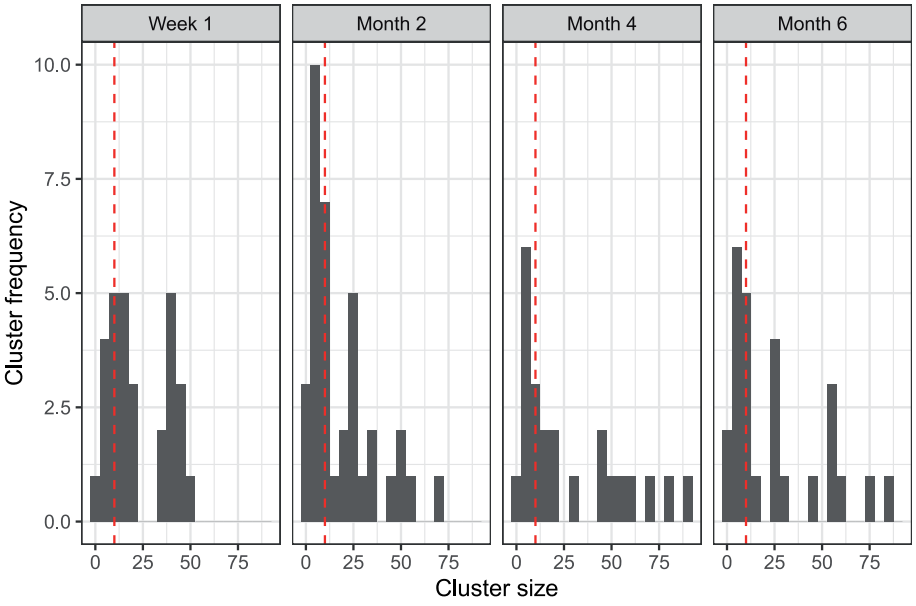
19. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–6 (2010).
20. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
21. Subramanian, S. *et al.* Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**, 417–21 (2014).
22. R Core Team. R: A language and environment for statistical computing. https://www.r-project.org/. (2017).
23. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
24. Oksanen, J. *et al.* Community Ecology Package. https://CRAN.R-project.org/package=vegan. (2017).
25. Benjamini, Y. & Hochberg, Y. Controlling The False Discovery Rate–A Practical And Powerful Approach To Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
26. Kurtz, Z. D. *et al.* Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLoS Comput. Biol.* **11**, (2015).
27. Csárdi G & Nepusz T. The igraph software package for complex network research. *InterJournal, Complex Syst.* **1695**, 1–9 (2006).
28. Man, W. H. *et al.* Loss of microbial topography between oral and nasopharyngeal microbiota and development of respiratory infections early in life. *Am. J. Respir. Crit. Care Med.* **200**, 760–770 (2019).
29. Reyman, M. *et al.* Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. *Nat. Commun.* **10**, 4997 (2019).
30. Biesbroek, G. *et al.* Early respiratory microbiota composition determines bacterial succession patterns and respiratory health in children. *Am. J. Respir. Crit. Care Med.* **190**, 1283–1292 (2014).
31. Teo, S. M. *et al.* The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe* **17**, 704–715 (2015).
32. Deasy, A. M. *et al.* Nasal inoculation of the commensal neisseria lactamica inhibits carriage of neisseria meningitidis by young adults: A controlled human infection study. *Clin. Infect. Dis.* **60**, 1512–1520 (2015).
33. Man, W. H. *et al.* Bacterial and viral respiratory tract microbiota and host characteristics in children with lower respiratory tract infections: a matched case-control study. *Lancet Respir. Med.* **7**, 417–426 (2019).
34. Dickson, R. P. *et al.* Lung microbiota predict clinical outcomes in critically ill patients. *Am. J. Respir. Crit. Care Med.* **201**, 555–563 (2020).
35. Langevin, S. *et al.* Early nasopharyngeal microbial signature associated with severe influenza in children: a retrospective pilot study. *J. Gen. Virol.* **98**, 2425–2437 (2017).
36. Baxter, N. T. *et al.* Dynamics of human gut microbiota and short-chain fatty acids in response to dietary interventions with three fermentable fibers. *MBio* **10**, (2019).
37. Yuille, S., Reichardt, N., Panda, S., Dunbar, H. & Mulder, I. E. Human gut bacteria as potent class I histone deacetylase inhibitors in vitro through production of butyric acid and valeric acid. *PLoS One* **13**, e0201073 (2018).

38. Rivière, A., Selak, M., Lantin, D., Leroy, F. & De Vuyst, L. Bifidobacteria and butyrate-producing colon bacteria: Importance and strategies for their stimulation in the human gut. *Frontiers in Microbiology* **7**, 979 (2016).
39. Henrick, B. M. *et al.* Elevated Fecal pH Indicates a Profound Change in the Breastfed Infant Gut Microbiome Due to Reduction of Bifidobacterium over the Past Century . *mSphere* **3**, (2018).
40. Chua, H. H. *et al.* Intestinal Dysbiosis Featuring Abundance of Ruminococcus gnavus Associates With Allergic Diseases in Infants. *Gastroenterology* **154**, 154–167 (2018).
41. Png, C. W. *et al.* Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *Am. J. Gastroenterol.* **105**, 2420–2428 (2010).
42. Tailford, L. E. *et al.* Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. *Nat. Commun.* **6**, 1–12 (2015).
43. Henke, M. T. *et al.* Ruminococcus gnavus, a member of the human gut microbiome associated with Crohn's disease, produces an inflammatory polysaccharide. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 12672–12677 (2019).
44. Scher, J. U. *et al.* Decreased bacterial diversity characterizes the altered gut microbiota in patients with psoriatic arthritis, resembling dysbiosis in inflammatory bowel disease. *Arthritis Rheumatol.* **67**, 128–139 (2015).
45. Fuentes, S. *et al.* Microbial shifts and signatures of long-term remission in ulcerative colitis after faecal microbiota transplantation. *ISME J.* **11**, 1877–1889 (2017).
46. Lotte, R., Lotte, L. & Ruimy, R. Actinotignum schaalii (formerly Actinobaculum schaalii): A newly recognized pathogen-review of the literature. *Clinical Microbiology and Infection* **22**, 28–36 (2016).
47. Ventura, M. *et al.* The Bifidobacterium dentium Bd1 Genome Sequence Reflects Its Genetic Adaptation to the Human Oral Cavity. *PLoS Genet.* **5**, e1000785 (2009).
48. Gueimonde, M. *et al.* Genome sequence of Parascardovia denticolens IPLA 20019, isolated from human breast milk. *Journal of Bacteriology* **194**, 4776–4777 (2012).
49. Lynch, M. D. J. & Neufeld, J. D. Ecology and exploration of the rare biosphere. *Nature Reviews Microbiology* **13**, 217–229 (2015).

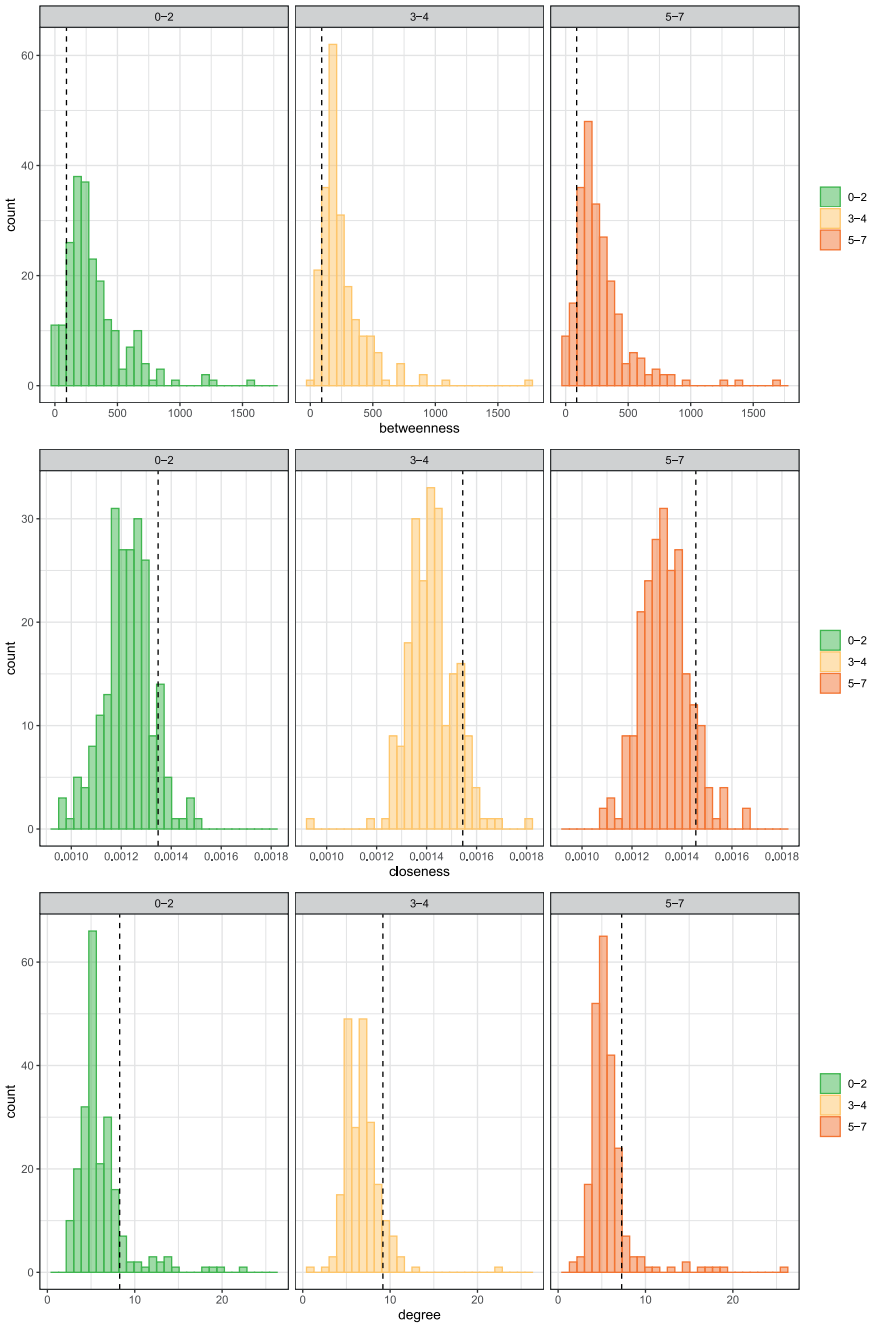
SUPPLEMENTARY INFORMATION



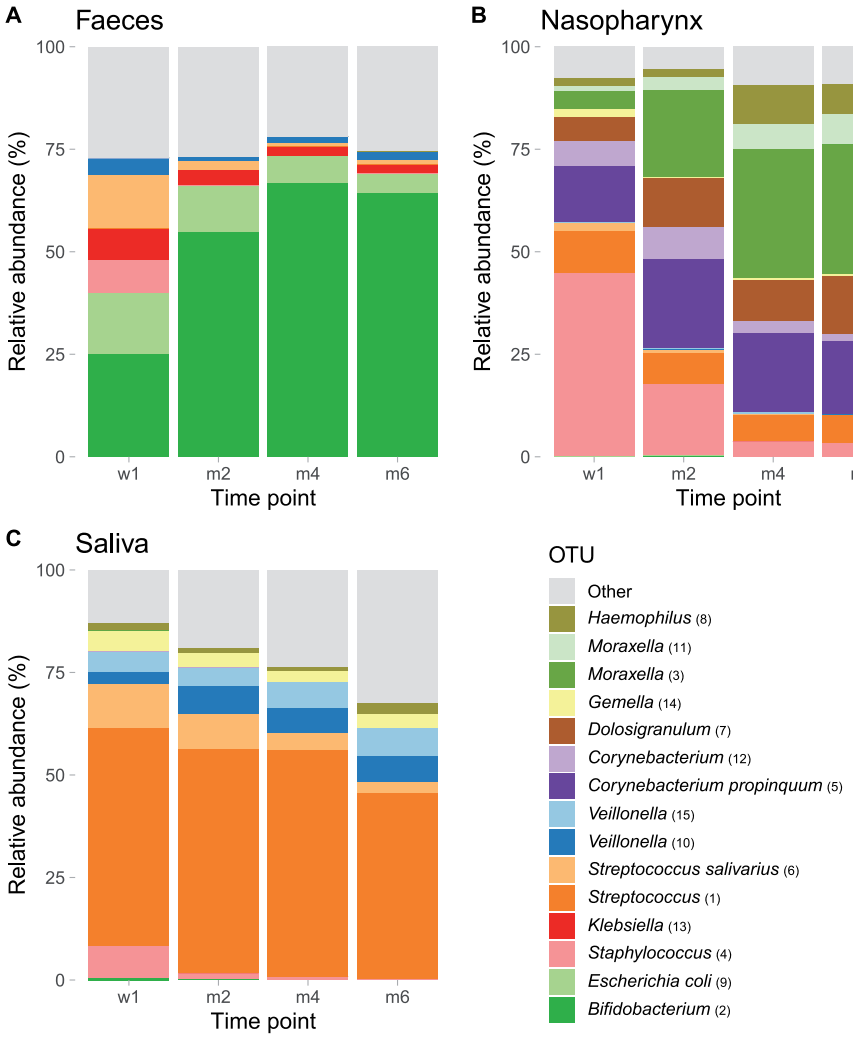
Supplementary Figure S1. Overlap of most abundant OTUs between RTI groups



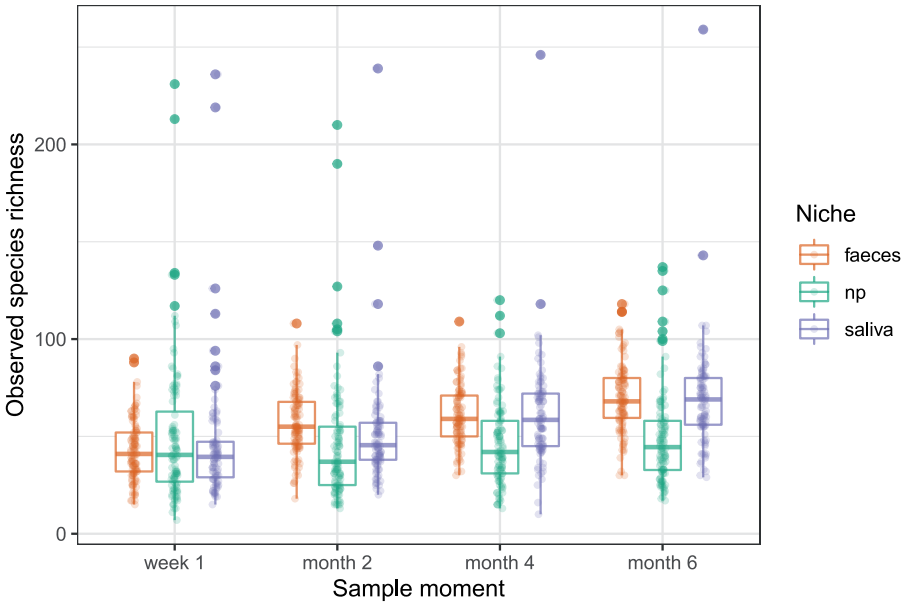
Supplementary Figure S2. Histogram of network cluster sizes. Walktrap.community() was used with step length = 5.



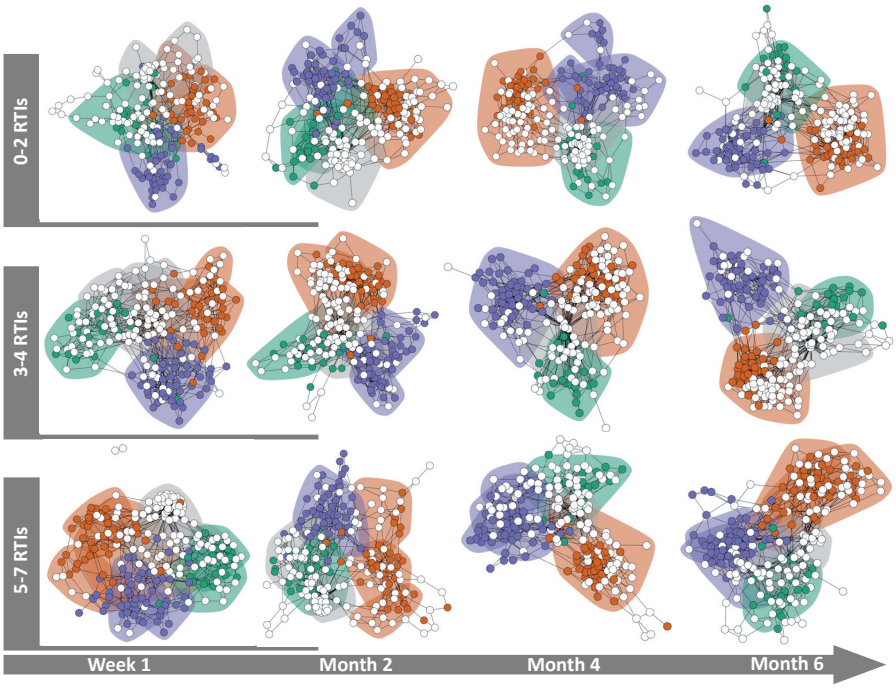
Supplementary Figure S3. Distribution of keystone network metrics. Distribution of betweenness centrality values (top row), closeness centrality values (middle row) and degree centrality (bottom row) split by RTI group.



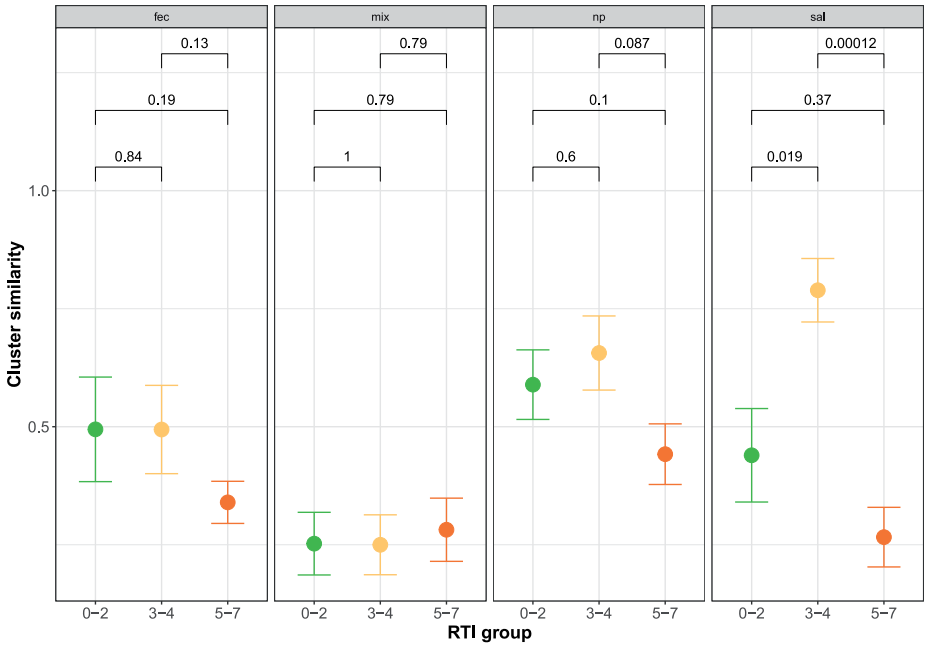
Supplementary Figure S4. Relative abundance of the most abundant OTUs. Stacked bar plots depicting mean relative abundance of the top 15 OTUs at each of 4 time points per niche. The number in parentheses refers to the rank of an OTU in the OTU table and is based on its overall relative abundance in the dataset.



Supplementary Figure S5. Observed species richness. Observed species richness (number of species per sample) is shown per time point and niche. Boxplots with medians are shown; the lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles); the upper and lower whiskers extend from the hinge to the largest and smallest value no further than 1.5*IQR from the hinge; outliers are plotted individually by opaque circles; translucent circles visualise all data points.



Supplementary Figure S6. Cross-niche networks for each RTI group per time point. Nodes represent individual OTUs and were coloured depending on their indicator species identity (orange = faeces-specific, green = nasopharynx-specific, blue = saliva specific, white = not niche-specific). Shaded areas around groups of nodes represent clusters defined by walktrap community analysis and are shaded depending on their niche-identity (orange = faeces-specific, green = nasopharynx-specific, blue = saliva specific, gray = mixed cluster/not niche-specific).



Supplementary Figure S7. Mean cluster similarity. The mean cluster similarity is shown for all clusters per RTI group across time split by cluster specificity. Whiskers represent standard errors.

Supplementary Table S1. Results of multivariate, longitudinal PERMANOVA analysis

Covariate	R^2 (%)	adjusted *p*
Niche	42.4	0.00079
Age	1.8	0.00079
Siblings <5 years present	0.3	0.00079
Pets	0.3	0.00079
Breastfeeding at sampling	0.3	0.00079
Season of birth	0.2	0.00079
Mode of delivery	0.2	0.00079
Antibiotics 1 month prior to sampling	0.1	0.041
Day care attendance	0.1	0.041
Pacifier use	0.1	0.041
Hospital duration postpartum	0.1	0.019

We performed permutational multivariate analysis of variance (PERMANOVA) with 1999 permutations to investigate the associations between microbiota composition (response variable) and environmental covariates. Covariates that were significantly associated in univariate models (per time point per niche) with microbiota community composition at a minimum of 1 time point were included in a multivariate model that included niche as well as age, and subject to control for repeated measures. The effect sizes (R^2) and adjusted p-values (using Benjamini-Hochberg method to correct for multiple testing) are shown.

Supplementary Table S2. Results of univariate, cross-sectional PERMANOVA analyses performed per niche

Niche	Time point	Covariate	R^2 (%)	*p*	adjusted *p*
Faeces	Week 1	Mode of delivery	14.4	0.0005	0.0020
		Hospital duration pp	11.9	0.0005	0.0020
		Breastfeeding at sampling	2.7	0.0095	0.019
		Antibiotics 1 month prior	2.6	0.0055	0.015
	Month 2	Breastfeeding at sampling	5.0	0.001	0.009
		Siblings <5 years present	4.1	0.002	0.009
		Mode of delivery	2.0	0.038	0.11
	Month 4	Day care attendance	3.4	0.007	0.036
		Breastfeeding at sampling	2.9	0.008	0.036
		Siblings <5 years present	2.1	0.042	0.13
	Month 6	Siblings <5 years present	3.0	0.012	0.10
		Day care attendance	2.5	0.023	0.10
Nasopharynx	Week 1	Pets	7.5	0.0085	0.023
		Mode of delivery	3.6	0.0080	0.023
		Hospital duration pp	3.4	0.0050	0.023
		Breastfeeding at sampling	2.9	0.015	0.029
		Siblings <5 years present	2.8	0.018	0.029
	Month 2	Siblings <5 years present	4.4	0.0010	0.0090
		Breastfeeding at sampling	2.5	0.024	0.071
		Antibiotics 1 month prior	1.8	0.013	0.056
	Month 4	Season of birth	4.7	0.050	0.150
		Siblings <5 years present	3.3	0.0065	0.059
		Pacifier use	2.5	0.024	0.11
	Month 6	Season of birth	5.2	0.022	0.066
		Siblings <5 years present	3.6	0.0015	0.014
		Day care attendance	2.9	0.0065	0.029
Saliva	Week 1	Antibiotics 1 month prior	3.9	0.0085	0.04
		Breastfeeding at sampling	3.1	0.010	0.04
	Month 2	Breastfeeding at sampling	5.9	0.0005	0.0045
	Month 4	Breastfeeding at sampling	7.3	0.0005	0.0045
		Season of birth	6.2	0.0075	0.034
	Month 6	Season of birth	6.0	0.011	0.038
		Antibiotics 1 month prior	3.8	0.013	0.038
		Breastfeeding at sampling	3.7	0.0045	0.038

We performed permutational multivariate analysis of variance (PERMANOVA) with 1999 permutations to investigate the associations between microbiota composition (response variable) and environmental covariates. The effect sizes (R^2), p-values and adjusted p-values (using Benjamini-Hochberg method to correct for multiple testing) are shown. Pp = postpartum, antibiotics 1 month prior = antibiotics administered in the month prior to sampling (for the 1 week time point this refers to antibiotics administered in the first week of life).

Supplementary Table S3. Summary statistics of relative abundance of keystone species split by RTI group specificity

RTI group specificity	mean relative abundance	SD relative abundance	median relative abundance	minimum relative abundance	maximum relative abundance
0-2	0.00015	0.00535	0	0	0.556
3-4	0.00342	0.0376	0	0	0.995
5-7	0.00124	0.0187	0	0	0.965
0-2 and 3-4	0.0743	0.204	0	0	0.995
0-2 and 5-7	0.000209	0.00459	0	0	0.296
3-4 and 5-7	0.00477	0.0245	0	0	0.398
all groups	0.00687	0.0282	0	0	0.508

Shown are mean relative abundance, standard deviation of relative abundance, median relative abundance, minimum relative abundance and maximum relative abundance.

CHAPTER 6

Summarizing discussion and future perspectives

In this thesis, we have described the development of the infant gut microbiome and resistome, identified important environmental and nutritional drivers of this development, and studied the gut microbiome and cross-niche microbial network development in relation to health and disease. We have shown that mode of delivery impacts the gut microbiome composition, with vaginally delivered (VD) infants having a more stable microbiome development and a higher abundance of beneficial bacteria in their gut than infants born with cesarean section (CS). Contrarily, we found that CS born infants carry a higher abundance of potential pathogens such as *Klebsiella* and *Enterococcus* species (spp.) in their gut. Interestingly, we could associate an increased abundance of these gut bacteria at one week of life with a higher susceptibility to respiratory tract infections (RTIs) in the first year of life. Next, we established that even a short course of early-life broad-spectrum antibiotics has long-term effects on the developing infant gut microbiome, selecting for potential pathogenic bacteria and leading to a higher abundance of antimicrobial resistance (AMR) genes. In case of early-onset neonatal sepsis (EONS), we strongly advise treatment with the antibiotic combination of penicillin + gentamicin, as this affects the microbial and AMR gene composition of the infant gut the least, when compared to co-amoxiclav + gentamicin or amoxicillin + cefotaxime. While the penicillin + gentamicin regimen might clinically be the least popular for this indication due to the frequency of penicillin administration and the need for gentamicin serum level monitoring, we think it deserves reconsideration on the neonatal ward, given the uniformly lower ecological side-effects observed, and considering the three regimens compared are equally adequate treatments for EONS. Finally, letting go of the single-pathogen principle in the pathogenesis of disease, and studying the human microbiome in a more holistic fashion, we built cross-niche microbial networks and associated these with susceptibility to RTIs. We found that children suffering from more frequent RTIs in their first year of life had more fragmented cross-niche microbial networks than their less susceptible counterparts.

Delivery mode affects infant gut microbiome development, independent of maternal antibiotics

After an infant is born, it is flooded with microbes from its environment, which nestle themselves on all its body surfaces, initiating the development of its unique microbiome.[1] The developing microbiome is sensitive to external influences such as antibiotics, diet and exposure to siblings.[2–4] Before an infant is exposed to such factors, however, its microbiome has already been influenced through the mode of delivery into this microbial world, be it by the vaginal route or by CS. It is standard

procedure for mothers delivering by CS to receive intravenous antibiotics to prevent infection following the procedure.[5]

A large body of evidence supports the findings that CS birth leads to an aberrant microbial development, a decreased abundance of health-associated bifidobacteria and an increased abundance of potential pathogens in the gut.[6–11] However, in the field of microbiome research, a longstanding discussion persists whether changes to an infant's microbiome can be attributed to mode of delivery itself, or whether these changes can be mostly explained by the antibiotics administered ante-partum.[12] Recently, one large study sparked controversy by suggesting microbial changes after CS birth were principally driven by intrapartum antibiotic prophylaxis (IAP), rather than mode of delivery.[12] In this study, however, IAP was colinear with CS birth, so these factors could not be effectively studied independently. In **chapter 2** we described the results of our healthy birth cohort study (Microbiome Utrecht Infant Study, abbreviated to MUIS), including 74 VD and 46 CS born infants. This study was specifically set-up so that mothers for whom antibiotic treatment was indicated, received this only after clamping of the umbilical cord. This ensured that no transfer of antibiotics took place between mother and child, enabling us to study the effect of delivery mode independently of IAP.

We compared VD and CS born infants based on the overall microbial community composition of their gut samples (a measure which takes the relative abundances of the different microbial species which are present in a sample into account) and found this differed significantly between the two groups. This difference was most prominent one week after birth, but was still significant at two months of life. These findings were confirmed through a sub-analysis performed on exclusively formula fed infants (11 VD and 11 CS born), eliminating the potential effects of maternal antibiotic exposure through breastmilk. The results of this sub-analysis additionally debunk the hypothesis that differences in gut microbiome composition between VD and CS born infants can be mainly ascribed to suboptimal breastfeeding practices in CS infants.[13,14] Next to less successful breastfeeding, it has been called into question whether other environmental factors inherent to CS birth, such as the absence of ruptured membranes and lower gestational age, are not actually the drivers of the microbial phenotypes that have been associated with delivery mode.[13] Importantly, in our healthy birth cohort, we could not detect an association between the duration of ruptured membranes or gestational age and the gut microbial community composition. Furthermore, we found that the community composition of infants born by emergency CS (n=10), who in general have undergone physiological labor

akin to vaginal delivery, was more similar to that of planned CS infants than VD infants.

In **chapter 2,** we described that the microbial perturbations in infants born by CS are typified by a delay in the acquisition of health-associated *Bifidobacterium*. Bifidobacteria are microbes well-known for their use as probiotics.[15] They are considered to promote gut health and provide defense against pathogens.[16] In VD infants, the abundance of *Bifidobacterium* was higher than in CS born infants over the entire first month of life. The CS born infants, on the other hand, on acquiring the potential pathogenic and pro-inflammatory *Enterococcus* after birth,[17] carried a higher abundance of this bacterium in this same period of time. Possibly, this could be explained by a weaker microbial resilience through a lack of protective *Bifidobacterium* in this timeframe in CS born infants. Despite the convergence of overall gut microbial community composition between VD and CS born infants after two months, differences in abundance of specific bacteria were seen for much longer periods of time. For instance, the abundance of *Klebsiella* spp., responsible for nosocomial infections such as pneumonia,[17] was increased for over four months after birth in CS born infants. Interestingly, both *Klebsiella* and *Enterococcus* abundance were associated with respiratory health in our cohort, discussed in more detail in the section below. In previous, relatively small cohorts, the colonization of *Klebsiella* and *Enterococcus* spp. in CS born infants seems to have been under-reported.[9,12] However, Shao *et al.* published a major paper (314 VD and 282 CS born infants) on the effects of delivery mode shortly after our manuscript was published,[18] and their findings parallel our own, showing a high-level of colonization by *Klebsiella* and *Enterococcus* spp. in CS born infants.

The primary microbial seeding of CS born infants may be attributed to operating room microbes, as suggested by the results of a study in which multiple sites in operating rooms were sampled.[36] The bacterial content of operation room dust corresponded to human skin bacteria, and was dominated by especially *Staphylococcus*. This skin bacterium genus was also found to be more abundant in the first days of life of CS infants in our study. Possibly, the highly abundant *Klebsiella* in our studied CS infants originates from the operating room as well. In the past, an outbreak of infection with a multiresistant *Klebsiella pneumoniae* strain in a Dutch regional teaching hospital was associated with contaminated material in the operating room. In the future, it would therefore be valuable to assess firstly whether it is acquisition from the hospital environment rather than horizontal transfer from the mother of *Klebsiella* that is causing this microbial pattern in CS

born infants, and secondly, if it is the acquisition or the outgrowth through a lack of commensal competition that plays a more important role.

Early-life gut microbiome composition is associated with respiratory health

Although, as mentioned above, the overall differences found between VD and CS born infants in gut microbial community composition converged over the first year of life, the observed perturbances in early-life may have long-lasting consequences for health. The crucial development of the human microbiome is believed to occur in the narrow timeframe of the first 1000 days of life, starting from conception.[19] Early-life microbial colonization coincides with an important timeframe in life, during which the immune system is receptive to microbial training. In the course of this so-called window of opportunity, the interaction between the microbiome and host is suggested to play a pivotal role in the priming of the host's immunity.[20] Early-life gut microbiota interact, primarily via the production and release of microbial metabolites and degradation products, with the developing immune system and contribute to immune programming and sensitization.[21–23] In this way, a perturbed microbial development could form the basis of susceptibility to disease later in life. This is supported by a vast range of epidemiological studies linking CS birth to an increased risk of childhood asthma, atopic disease, allergies, obesity, type 1 diabetes, and inflammatory bowel disease.[13]

Additionally, ample epidemiological data links delivery mode to respiratory illness,[24,25] but so far few studies have assessed the role of delivery mode-induced gut microbial changes in the development of respiratory disease. Remarkably, we found in **chapter 2** that the gut microbiome composition at a very early age (one week of life) was associated with the number of RTIs experienced in the first year of life. The MUIS cohort could, based on the normal distribution of RTI incidence, be divided in infants experiencing few (0-2), average (3-4) and a high number of (5-7) RTIs.[27] The taxa strongly associated with higher susceptibility to RTIs were *Klebsiella* and *Enterococcus* spp. As mentioned above, these species were more abundant in CS children, providing a possible link between mode of delivery and susceptibility to respiratory infections, and perhaps also other infectious problems.

Evidence for fecal seeding from mother to infant

The question remains what aspect of delivery mode drives the differences in microbial development trajectories. A possible explanation for the dissimilarity in gut microbiome composition between VD and CS born infants is the difference in birthing route: while a VD infant passes through its mother's birthing canal,

an infant born by CS delivery enters the world through its mother's stomach. Consequently, the infants are exposed to different body surfaces and so, different microbes are physically transferred from mother to child. Dominguez-Bello *et al.* reported previously that the early-life microbial communities across multiple body sites of VD infants were more similar to their mother's vaginal microbiome, while CS infants harbored microbiomes most similar in composition to their mother's skin.[7] In an intervention study by the same author, CS born infants were seeded with their mother's vaginal flora, and subsequently their anal, oral and skin samples were compared to those of VD and non-seeded CS infants.[32] The samples originating from vaginally seeded CS infants were more often classified as samples collected from VD infants than from non-seeded CS infants. However, this was mainly the case for the skin and oral samples, and less so for the anal samples. We therefore hypothesized that in order to reach a stronger resemblance in microbial communities between VD and CS infants, it is not only necessary for CS infants to be exposed to their mothers' vaginal flora, but also their intestinal flora.

In our study, we found evidence for fecal seeding from mother to infant in VD, but not in CS born infants. When studying the concordance of fecal samples from mother-infant pairs, we found that an infant's fecal microbiome composition was significantly more similar to that of its own mother than to that of other mothers, but only in VD infants. This substantiates our hypothesis that perhaps not merely vaginal microbiota seeding, but also fecal microbiota seeding during vaginal delivery, is instrumental in shaping the newborn's gut microbial environment. The delayed acquisition of *Bifidobacterium* in CS infants described in **chapter 2,** suggested that maternal transmission during vaginal delivery is essential in acquiring these bacterial species in early life. Evidence for vertical transmission of gut bacteria from mother to child has been previously established using single-nucleotide variation to identify transmission of specific strains from mother to child. Vertical transmission has been demonstrated for a number of species, including *Bifidobacterium longum* and *Bifidobacterium adolescentis*.[33,34] These two particular species especially made up the most abundant *Bifidobacterium* taxon in our study, as confirmed by metagenomic shotgun sequencing. A recent proof-of-concept study further corroborates the importance of fecal seeding: in CS born infants who were administered a diluted fecal sample from their own mothers, the development of the gut microbiota composition showed significant similarity to that of VD infants.[35] The health benefits of this approach still need to be studied.

Breastfeeding may not correct for the existing lack of *Bifidobacterium* in CS born infants

Similarly to mode of delivery, feeding type is a frequently studied variable in relation to gut microbiome development. The World Health Organization recommends exclusive breastfeeding for the first 6 months of life.[37] In CS infants, however, breastfeeding is often delayed and suboptimal.[14] Therefore, guidance in breastfeeding practices in mothers that have delivered by CS seems crucial. Even more so, because IAP-induced gut microbial changes in infants born by emergency CS have been observed to be modified by breastfeeding.[38] Bacteroidaceae were significantly less, and Clostridiales more abundant in infants exposed to IAP during emergency CS than in vaginally born infants with no IAP exposure, but only in the absence of exclusive breastfeeding at 3 months. In other words, the IAP-induced changes were not detected in breastfed infants. In **chapter 2** we studied whether breastfeeding modulated the delivery-mode induced microbial changes we found. However, we found only a modest difference in mode of delivery effect on overall gut microbial community composition when comparing the results of analyses performed on all infants versus those receiving exclusive formula feeding. Moreover, though breastfeeding could be associated with increased *Bifidobacterium* abundance, CS infants receiving breastfeeding still had a lower *Bifidobacterium* abundance in their fecal samples than VD infants receiving formula feeding. This suggests that only after abovementioned proper initial (vaginal-)fecal seeding takes place, the growth of beneficial groups such as *Bifidobacterium* can be promoted through the prebiotic oligosaccharides present in breast milk.[39] This leads us to believe that the stimulation of breastfeeding in women that have delivered by CS is not enough to ameliorate the reduced *Bifidobacterium* abundance in their children. Perhaps in future a combination of vaginal and fecal seeding practices could prove useful, or colonization through probiotic administration together with breastfeeding.

Even a short course of antibiotics in early-life majorly impacts the gut microbiome and resistome

As discussed in **chapter 1**, antibiotics are well-known drivers of microbial change and AMR development in the human population and in the environment. Broad-spectrum antibiotics are nonetheless potentially life-saving in cases of EONS. The antibiotic prescription rate in neonates is high, with 4-10% of all newborns receiving a course of antibiotics in early-life.[40,41] In our ZEBRA trial presented in **chapter 4**, which stands for "Zuigelingen En Bacteriële Resistentie na Antibiotica", we included 147 infants with (suspected) EONS for whom broad-spectrum antibiotic treatment was indicated in their first week of life. We randomized them over three regimens representing the most commonly prescribed intravenous antibiotic combinations

for this indication in the Netherlands, namely penicillin + gentamicin, co-amoxiclav + gentamicin and amoxicillin + cefotaxime.[42] As controls, we included a subset of 80 age-matched infants from the healthy MUIS cohort described in **chapter 2,** who did not receive antibiotics in their first week of life.

The common causative pathogens of EONS are group B streptococcus (GBS), *Listeria monocytogenes* and *Escherichia coli*. These organisms are susceptible to antibiotic regimens consisting of a combination of penicillin with gentamicin or amoxicillin with cefotaxime, making these treatments equally effective.[43] However, already in the year 2000, de Man *et al.* highlighted the issue of amoxicillin driving the overgrowth of β-lactamase producing bacteria, such as *Klebsiella* spp. Another concern raised was that third-generation cephalosporins, such as cefotaxime, select for resistant *Enterobacter* spp. strains.[44] *Klebsiella* and *Enterobacter* spp., along with *Enterococcus faecium*, *Staphylococcus aureus*, *Acinetobacter baumanii* and *Pseudomonas aeruginosa*, belong to the ESKAPE pathogen group, known for causing hospital and community acquired infections and their efficacy at "escaping" effects of antibiotics.[17] Thus, it is not only important to choose an effective antibiotic combination for EONS that keeps the neonatal commensal gut microbes intact, but which also selects the least AMR.

Antibiotics and AMR genes conferring resistance to antibiotics namely go hand in hand. As discussed in **chapter 1,** antibiotics have existed in nature for billions of years as microbial products, providing competitive advantage to bacteria by killing or inhibiting the growth of others competing for scarce resources in the same environment.[45] Under antibiotic pressure, a previously susceptible bacterial population may become resistant, thereby increasing its chance of survival. Resistance is acquired through mutations occurring in the bacterial genome or by horizontal transfer of AMR genes on small, extrachromosomal DNA molecules, called plasmids.[46] Ultimately, functional resistance ensues through various mechanisms, such as the modification of the antibiotic target, the enzymatic inactivation of the antibiotic or the reduction of the lethal intracellular concentration of the antibiotic by efflux pumps.[47] In some cases, genes capable of conferring antibiotic resistance actually have a different function in the original host. For instance, the Gammaproteobacterium *Providencia stuartii* contains a chromosomal gene encoding for the enzyme 2'-N-acetyltransferase, which plays a role in the bacterium's autolytic system by modifying peptidoglycan, a cell wall component.[48] As aminoglycoside antibiotics are structurally similar to peptidoglycan, this enzyme can inactivate aminoglycosides, making *P. stuartii* intrinsically resistant to this class of antibiotics.

In our ZEBRA study, even a short course (on average 48 hours) of early-life broad-spectrum antibiotic treatment led to an altered gut microbial community composition that could still be measured at the first birthday of the participants in our study. Also, twelve months after antibiotic administration, specific bacterial species, for example species belonging to the generally anti-inflammatory genus *Bacteroides*,[49] still had a lower abundance in antibiotic treated children compared to controls. Antibiotic treatment was associated with a higher abundance of the same potential pathogens found to be more abundant in CS children: *Enterococcus* and *Klebsiella*. The biggest impact of antibiotic treatment could be measured directly following treatment, with major shifts in both gut microbiome composition as well as resistome profile at that same moment. We defined the resistome profile as the composition of 31 tested, clinically relevant,[50,51] bacterial AMR genes that were either present or absent in a fecal sample as identified by quantitative polymerase chain reaction (qPCR). After antibiotic administration, the resistome profile of the suspected EONS (sEONS) infants showed a large and significant shift (R^2 [effect size] 7.5%) from the healthy situation.

Importantly, the resistome profiles did not differ between the sEONS infants and controls before start of antibiotic treatment. Both groups of infants already carried AMR genes at the first tested time point, which was on average the first day of life. This is in line with the intrinsic resistant potential of commensals and with previous studies, which have reported the presence of a diverse set of AMR genes in healthy infants in early-life.[52–54] We hereby re-established that AMR gene carriage in healthy infants is normal, but additionally found in **chapter 4** that the fecal samples of sEONS infants were more enriched with AMR genes post antibiotic treatment. Of great importance, the AMR genes found to be more enriched in the antibiotic treated children conferred resistance to clinically essential antibiotics, such as aminoglycosides (of which gentamicin is an example), beta-lactams (to which penicillin, amoxicillin and cefotaxime belong) and colistin (a last-resort antibiotic prescribed to treat multidrug-resistant Gram negative infections beyond infancy).[55]

Given that our qPCR approach selectively targeted a pre-specified panel of AMR genes, we also performed metagenomic shotgun sequencing (MGS) to evaluate the full spectrum of AMR genes present in our participants. MGS was executed on a subset of samples collected directly following antibiotic treatment, which corresponded with the 1-week time point of the controls. In this way, we identified a total of 1504 AMR genes in the tested samples. Of these, 262 AMR genes were found to be differentially abundant between the antibiotic treated infants and controls, all of which were more abundant in the treated infants. Besides confirming the qPCR

results, metagenomic data also showed enrichment of genes coding for resistance against, among others, trimethoprim, sulfonamides, bacitracin and fusidic acid, though none of these antibiotics had been administered to the infants yet at that time. This finding is in accordance with previous research, where broad-spectrum antibiotic treatment also showed enrichment of AMR genes without known activity against the specific antibiotic driver.[56] Likely, this collateral enrichment can be explained by the selection of multi-drug resistant organisms, which carry genetic material encoding for multiple AMR genes. Clinically, acquired resistance against trimethoprim and sulfonamides has potential implications in pediatric medicine, as this combination is commonly used as prophylactic treatment in children with vesicoureteral reflux after urinary tract infection to prevent recurrence[57], whereas the antimicrobial agents bacitracin and fusidic acid are frequently used for the topical treatment of skin infections.[58]

So far, the focus of antibiotic stewardship programs lies especially on shortening the antibiotic course duration.[59,60] We described in **chapter 4**, however, that the association between the duration of antibiotic treatment and the overall gut microbial composition was relatively small. The calculated effect size, or R^2, of antibiotic treatment duration was only 0.3%. When compared to other covariates associated with gut microbiome composition, such as antibiotic regimen (R^2 0.9%), age (R^2 5.2%), day care attendance (R^2 1.0%) and breastfeeding at time of sampling (R^2 0.8%), the effect size of antibiotic treatment duration in our study was minor.

Of course, it still remains to be seen whether the microbial dysbiosis or enrichment of the resistance genes found in our study leads to a higher incidence of bacterial infections (by resistant organisms) later in life. In any case, in animal studies it has already been found that antibiotic-induced alterations of the healthy gut microbiome made mice more prone to bacterial infection later in life.[61,62] Furthermore, in a pediatric study,[63] antibiotics in early-life have been associated with increased rates of diarrhea in early childhood, which we now hypothesize might be a consequence of persistent microbial dysbiosis. The participants of the MUIS and ZEBRA cohorts are currently being followed until the age of 5, so these issues would be valuable to address.

Penicillin combined with gentamicin is the preferred combination for treatment of EONS

Our primary aim in **chapter 4** was to compare the effects on gut microbiome and resistome composition of three intravenous antibiotic regimens most commonly prescribed for EONS in the Netherlands, namely penicillin + gentamicin, co-

amoxiclav + gentamicin and amoxicillin + cefotaxime. As mentioned previously, these three regimens are equally effective in combatting the causative pathogens responsible for EONS.[43] In practice, the choice of regimen prescribed on the neonatal ward is also based on ease of administration, possible side effects and costs. The frequency of administration for both the penicillin + gentamicin and co-amoxiclav + gentamicin regimens is the same (penicillin and co-amoxiclav are administered three times daily and gentamicin once daily), while amoxicillin is administered three times daily combined with a twice daily dose of cefotaxime. When administering a regimen including gentamicin, an extra blood test is necessary to monitor the gentamicin serum level, because levels that are too high are associated with nephrotoxicity and ototoxicity.[64]

Of the three antibiotic regimens that we compared, we found that the combination of penicillin + gentamicin had the smallest impact on all study outcomes compared to the co-amoxiclav + gentamicin and amoxicillin + cefotaxime regimens. First of all, the penicillin + gentamicin regimen resulted in the smallest decrease in microbial species diversity directly following treatment (median Shannon alpha diversity 0.76), when compared to healthy controls (median alpha diversity 1.21). In comparison, the co-amoxiclav + gentamicin led to the largest dip in median alpha diversity of 0.43 at this time point. The relevance of a reduced alpha diversity of the early-life gut microbiome becomes clear in light of its association with the development of allergic disease and diabetes later in life.[65–68] Secondly, the penicillin + gentamicin regimen affected the overall gut microbial community composition the least, with an effect size, or R^2, of 10.1%, while amoxicillin + cefotaxime showed the biggest effect (R^2 14.7%). In microbiome research this magnitude of effect is major, and interestingly, the latter effect size was very similar to that of mode of delivery on overall gut microbial community composition at 1 week of life in the healthy MUIS cohort (R^2 14.2%). By calculating the differences in overall gut microbiome composition between consecutive samples within a participant we could compare the stability of gut microbiome development following each antibiotic regimen. A stable state of the gut microbial community is thought to promote resilience of the microbiome by resisting colonization by microbes adapted to a disturbed equilibrium.[69] In the nasopharyngeal niche, a stable microbiome profile has been associated with a reduced number of consecutive RTIs.[70] In our study, the gut microbiome showed the most stable development in the infants treated with penicillin + gentamicin. Regarding specific bacterial taxa, we observed that especially *Bifidobacterium* genera were more abundant in the penicillin + gentamicin treated infants (with the exception of *Bifidobacterium animalis*), compared with amoxicillin + cefotaxime treated infants. In part, we hypothesize

these overall lower ecological effects of the penicillin + gentamicin regime might be explained by the minimal penetration into the gut lumen of aminoglycosides given intravenously, combined with the narrow spectrum of penicillin.[71]

Also, the gut AMR gene composition and enrichment based on our prespecified panel was least affected by the penicillin + gentamicin regimen. After treatment, an R^2 of 5.9% was found for penicillin + gentamicin, followed by 6.3% for co-amoxiclav + gentamicin and 11.1% for amoxicillin + cefotaxime treatment. Additionally, when comparing each separate antibiotic regimen with the control group, we found that the penicillin + gentamicin group had the lowest number of enriched AMR genes (5 out of 31 in total) versus 10 out of 31 for both the co-amoxiclav + gentamicin and amoxicillin + cefotaxime groups. Finally, in the inter-regimen comparisons, again the penicillin + gentamicin regimen resulted in the enrichment of the lowest number of AMR genes compared to the other two regimens (2 AMR genes, versus 9 for the amoxicillin + cefotaxime and 5 for the co-amoxiclav + gentamicin group). In conclusion, to help reduce the problem of AMR, penicillin + gentamicin is also the preferred choice for antibiotic treatment for EONS in our youngest patients.

Building cross-niche microbial networks and studying their relation to health

In EONS, as is currently often the case in many diseases, a single microbe (i.e. GBS, *L. monocytogenes* or *E. coli*) is identified as the causative pathogen. However, as discussed previously, overgrowth by such a pathogen is likely precluded by a perturbation of the microbial community's stable state, disabling the community's resistance to colonization. Through insights like these, the focus of microbiome research is shifting from studying individual pathogens in isolation, towards a more holistic approach, and interactions are being studied between microbes within a community and their host.[72] In a microbial community, ecological principles apply: competition for limited resources result in competitive exclusion, and conversely, cooperation exists between microbes when the product of one microbe becomes the substrate for another.[73] In this way, a microbe may influence niche community structure and function, but it has even been theorized that a single microbe might also influence the wider human microbiome network.[74] Hence, exploring the relationships between bacteria and the microbial networks they form, both locally and across the human host, may help to gain insight into their health-promoting and pathological properties.

In earlier studies based on the MUIS cohort, our research group already showed that the development of the upper respiratory tract microbiome (nasopharynx and

oropharynx) was associated with respiratory health. Bosch *et al.* provided evidence that an accelerated maturation of the nasopharyngeal microbiome is associated with the number of RTIs experienced in the first year of life.[27] Man *et al.* illustrated how a loss of topography of the upper respiratory microbiome appears to precede RTI episodes.[28] This loss of topography was characterized by an influx of oral taxa in the nasopharynx.

Anatomically speaking, it is not surprising that the microbial community of the upper respiratory tract directly affects respiratory health. However, the notion that the gut microbiome might be associated with respiratory health is also gaining traction, leading to the concept of the so-called gut-lung axis. Microbial components and metabolites in both the gut and lung have been shown to modulate immunity locally and systemically, whereas specific taxa at both sites are associated with lung diseases such as asthma, chronic obstructive pulmonary disease and RTIs.[29] In addition, translocation of microbes between the gut and lung has been observed: the lung microbiome was found to be enriched with gut-associated bacteria in acute respiratory distress syndrome and sepsis.[30] Moreover, gut-associated bacteria in the lungs were predictive of clinical outcomes in critically ill patients, suggesting that translocation of gut bacteria to the lungs contributes to the pathogenesis of lung injury.[31]

Typically, microbial network studies are still restricted to microbial networks within one niche.[75,76] In the MUIS cohort we collected samples from three niches, namely the nasopharynx, oropharynx and the gut, giving us the unique position to study the cross-niche infant microbiome in relation to respiratory health. By combining the microbial data of three niches, we could build cross-niche microbial networks to investigate the interaction between local microbial communities across the body, and relate the network structure to respiratory health. We based network construction on conditional independence, rather than correlation, to decrease the chance of spurious (false) connections between taxa.

Strikingly, we found that infants with a high susceptibility to RTIs (5-7 RTIs) had more fragmented and less stable networks throughout the first 6 months of life. Fragmented networks in relation to disease have been described before, as in inflammatory bowel disease patients, where patients had a distorted gut microbial co-occurrence network compared to healthy individuals.[77] Following our finding in **chapter 2** that an increased abundance of the potential pathogens *Klebsiella* and *Enterococcus* spp. in early-life can be associated with impaired respiratory health,

we next tried to identify if specific taxa important for the network structure could be associated to RTIs in a similar manner.

Keystone species in cross-niche microbial networks can be associated with RTIs

Species that play a central role in an overall network structure can be identified as keystone species. In the definition by Banerjee *et al.* and Berry and Widder, a keystone species is an OTU having a high degree, high closeness centrality and low betweenness centrality in a network.[78,79] Network degree means the number of connections each OTU has within the network, betweenness centrality means the number of shortest paths that pass through an OTU in the network, and closeness centrality stands for the reciprocal sum of the length of the shortest paths between an OTU and all other OTUs in the network.

The keystone species for the 0-2 RTI networks included well-known health-associated OTUs, not only with respect to respiratory health, but also to gut health. In the microbial networks of this least susceptible group of infants we found *Corynebacterium*, *Dolosigranulum pigrum* and *Neisseria lactamica* to be keystone species. Our research group has previously established, using nasopharyngeal and saliva sample data, that both *Corynebacterium* and *D. pigrum* are associated with a healthy microbiota maturation process and with a decreased risk of developing RTIs in the first year of life.[28,70] Our microbial network study, using nasopharyngeal, saliva and fecal samples combined, thereby suggests that *Corynebacterium* and *D. pigrum* are not only important players in the nasopharynx and oral cavity, but also in the overall infant microbial network. *N. lactamica* is thought to make the microbial community network more resilient, as it was shown to inhibit the colonization of the pathogenic *Neisseria meningitidis* in human challenge studies.[80] This OTU being a keystone species in our 0-2 RTI networks corroborates this theory.

We also found *Ruminococcus bromii*, *Megasphaera* and *Anaerostipes* to be keystone species in the 0-2 RTI networks. These OTUs are associated with (the enhanced) production of butyrate, which is a short chain fatty acid and microbial end-product of the human gut fermentation process.[81–83] Butyrate is an essential metabolite in the gut environment, being the preferred energy source for the epithelial cells of the colon. It has anti-inflammatory properties and lowers the pH of the gut, inhibiting the growth of pathogens.[81]

While keystone species identified in the 0-2 RTI networks have been previously associated with health, the opposite was also true, with keystone species found in the

5-7 RTI networks being associated in literature with respiratory and inflammatory diseases, or with unfavorable properties such as cariogenicity. For instance, the 5-7 RTI keystone species *Haemophilus* and *Lachnospiraceae* have repeatedly been linked with recurrent respiratory infections.[28,31,84,85] Also a keystone species in the 5-7 RTI networks, the pro-inflammatory *Ruminococcus gnavus* has been reported to play a role in a broad scale of immunological disorders, namely pediatric allergy, Crohn's disease and psoriatic arthritis.[86–88] Furthermore, the cariogenic *Bifidobacterium dentium* and *Parascardovia denticolens* were keystone species in the most susceptible RTI networks.[89,90] Contrarily to gut pH-lowering OTUs being keystone species in the 0-2 RTI networks, two species belonging to the genus *Peptostreptococcaceae*, which is associated with increased fecal pH, were identified as keystone species for the 5-7 RTI networks.[91] Altogether, this suggests that, next to commensal and pathogenic bacteria influencing their local niche environments, cross-niche bacterial networks may stand at the basis of overall systemic susceptibility.

Future perspectives

The role of the human microbiome in health and disease is becoming increasingly clear. Early-life disturbances of the microbial development could form the basis of susceptibility to inflammatory and infectious disease later in life. As discussed in this thesis, there are multiple environmental factors that have the potential of disrupting the infant microbiome in its developmental phase, an important one being mode of delivery. It is the role of (medical) scientists and health care workers to inform expecting parents of the effects of CS delivery on the microbiome and, following, the health of their future child. Obstetrician-gynecologists recognize the increased demand for CS delivery on maternal request, and while they believe that the risks of this procedure outweigh the benefits, evidence regarding risks is crucial to guide policy making regarding delivery practices.[92] Evidence that CS delivery negatively affects infant outcomes with respect to microbiome development, immunological priming and health problems later in life, could help in the coaching of parents towards vaginal delivery when safe for mother and child. When there is no choice in the matter, however, for instance in case of maternal or infant complications necessitating CS delivery, it might be desirable to replace the beneficial microbes the infant might lack. Although efforts are being made in this direction through studying the effects of vaginal and fecal seeding, more research is needed to assess the effectiveness of probiotics and the advantageous combination of vaginal-fecal seeding along with an adequate administration route to achieve long-term colonization.

Next, with the aim to preserve the healthy infant microbiome after birth, and to reduce the selection of potential pathogens and AMR genes, action must be taken to prevent unnecessary antibiotic prescription. As discussed in detail, we have shown that even a short course (48 hours) of broad-spectrum antibiotics can have long-lasting and detrimental effects on the infant gut microbiome and AMR gene pool. It is undoubtedly alarming that antibiotics are prescribed in 4-10% of all neonates in Western countries, while only an estimated 1 in 1000 neonates develops a proven infection, meaning that >90% of treatments are likely unnecessary.[40,41,93] We strongly advocate initiatives improving the identification of neonates at risk for developing EONS, such as the use of the sepsis calculator, which estimates the EONS risk based on five objective maternal risk factors and four clinical neonatal risk factors.[94] In a meta-analysis, the use of this tool was associated with a substantial reduction in the use of empirical antibiotics for suspected EONS.[95] Recently, the monitoring of serum procalcitonin levels has been shown to safely guide decision-making in reducing the duration of antibiotic treatment.[96] Ideally, in future, microbial biomarkers, or patterns in overall (gut) microbiome composition, will be identified to help decide whether antibiotic treatment should be initiated in the first place. To that end, real-time bacterial DNA sequencing could potentially test for such favorable or unfavorable patterns and help guide clinical decision making.[97]

Notwithstanding, antibiotic treatment will remain absolutely necessary for proven EONS. In that case, we fully promote the prescription of the combination of penicillin and gentamicin over co-amoxiclav with gentamicin or amoxicillin with cefotaxime. Of the three antibiotic regimens we compared, this is the most narrow-spectrum antibiotic combination. An ideal therapeutic candidate for the perinatal prevention and treatment of EONS would be one that is highly specific for its target, or "super" narrow-spectrum, sparing the surrounding microbiota. Up and coming agents that fulfill this criterion are bacteriophages, viruses that kill specific bacteria.[98]

To replace antibiotic-induced depletion of beneficial bacteria, there might be a role for probiotics. Both public interest and interest in this field of research is rising. Probiotics can be found containing various bacterial strains and are sold as powders or capsules in health shops, or even yoghurts and drinks in supermarkets. Probably, the most well-known is the Yakult probiotic fermented milk beverage, containing *Lactobacillus casei*.[99] Many health-beneficial aspects have been ascribed to probiotics and an increasing number of studies is currently being performed to study whether probiotics can be administered for the prevention of specific disease.[100] It would be interesting to test whether the administration of probiotics,

especially of health-associated bacteria that are affected by antibiotics, might be helpful in reversing overall gut microbial perturbation. The main problem in altering infant intestinal dysbiosis via the administration of probiotics is the transient character of colonization of the administered bacteria.[101] However, in a study performed in breastfed infants supplemented with *Bifidobacterium infantis,* it was shown that they had higher population levels of fecal *Bifidobacteriaceae*, in particular *B. infantis*, than the control group for more than 30 days after cessation of supplementation.[102] In a double blind, placebo-controlled, randomized trial studying the effects of multispecies probiotic supplementation to mothers and infants on gut dysbiosis caused by CS birth and antibiotic use, it was found that probiotics had a strong impact on microbiota composition, depending on the infants' diet.[103] Only breastfed infants showed an increase in bifidobacteria. In the probiotic group, the effects of antibiotics and CS birth on the gut microbiota were reduced or even completely eliminated.

Thus, multispecies probiotic supplementation seems promising. However, the question remains how to select for the ideal "microbial cocktail" in order to achieve a balanced gut microbiome. It is not necessarily the case that the most abundant bacterial species perform the most important biological functions. We learned that the keystone species in the microbial networks we built, and which we could associate with respiratory health, were generally low abundant taxa. This matches the "rare taxa" concept, which postulates that the abundance of a species is not the best determinant for its importance within the microbial community structure.[104] Studying the human microbiome in a more generic context therefore might also provide insights into the importance and roles of the lesser-known microbes and their potential to promote health.

REFERENCES

1. Koenig, J. E. *et al.* Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci.* **108**, 4578–4585 (2011).
2. Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **8**, 343ra82-343ra82 (2016).
3. Martin, R. *et al.* Early-Life Events, Including Mode of Delivery and Type of Feeding, Siblings and Gender, Shape the Developing Gut Microbiota. *PLoS One* **11**, e0158498 (2016).
4. Yasmin, F. *et al.* Cesarean Section, Formula Feeding, and Infant Antibiotic Exposure: Separate and Combined Impacts on Gut Microbial Changes in Later Infancy. *Front. Pediatr.* **5**, 200 (2017).
5. NICE. *Caesarean section clinical guideline [CG132].* (2011).
6. Penders, J. *et al.* Factors Influencing the Composition of the Intestinal Microbiota in Early Infancy. *Pediatrics* **118**, 511–521 (2006).
7. 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.* **107**, 11971–5 (2010).
8. Azad, M. B. *et al.* Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. *CMAJ* **185**, 385–94 (2013).
9. Bäckhed, F. *et al.* Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**, 690–703 (2015).
10. Stokholm, J. *et al.* Cesarean section changes neonatal gut colonization. *J. Allergy Clin. Immunol.* **138**, 881-889.e2 (2016).
11. Hill, C. J. *et al.* Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort. *Microbiome* **5**, 4 (2017).
12. 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**: 314-26.
13. Stinson, L. F., Payne, M. S. & Keelan, J. A. A Critical Review of the Bacterial Baptism Hypothesis and the Impact of Cesarean Delivery on the Infant Microbiome. *Front. Med.* **5**, 135 (2018).
14. Dewey, K. G., Nommsen-Rivers, L. A., Heinig, M. J. & Cohen, R. J. Risk factors for suboptimal infant breastfeeding behavior, delayed onset of lactation, and excess neonatal weight loss. *Pediatrics* **112**, 607–19 (2003).
15. Allen, S. J., Martinez, E. G., Gregorio, G. V & Dans, L. F. Probiotics for treating acute infectious diarrhoea. *Cochrane database Syst. Rev.* CD003048 (2010). doi:10.1002/14651858.CD003048.pub3
16. Tamburini, S., Shen, N., Wu, H. C. & Clemente, J. C. The microbiome in early life: implications for health outcomes. *Nat. Med.* **22**, 713–722 (2016).
17. Boucher, H. W. *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **48**, 1–12 (2009).
18. Shao, Y. *et al.* Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* **574**, 117–121 (2019).

19. Robertson, R. C., Manges, A. R., Finlay, B. B. & Prendergast, A. J. The Human Microbiome and Child Growth – First 1000 Days and Beyond. *Trends in Microbiology* **27**, 131–147 (2019).
20. Gensollen, T., Iyer, S. S., Kasper, D. L. & Blumberg, R. S. How colonization by microbiota in early life shapes the immune system. *Science*. **352**, 539–544 (2016).
21. Rautava, S. & Walker, W. A. Commensal bacteria and epithelial cross talk in the developing intestine. *Current Gastroenterology Reports* **9**, 385–392 (2007).
22. Smith, P. M. *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic T reg cell homeostasis. *Science*. **341**, 569–573 (2013).
23. Renz, H. *et al.* An exposome perspective: Early-life events and immune development in a changing world. *J. Allergy Clin. Immunol.* **140**, 24–40 (2017).
24. Kristensen, K. *et al.* Caesarean section and hospitalization for respiratory syncytial virus infection: a population-based study. *Pediatr. Infect. Dis. J.* **34**, 145–8 (2015).
25. Baumfeld, Y. *et al.* Elective cesarean delivery at term and the long-term risk for respiratory morbidity of the offspring. *Eur. J. Pediatr.* **177**, 1653–1659 (2018).
26. Patrick, D. M. *et al.* Decreasing antibiotic use, the gut microbiota, and asthma incidence in children: evidence from population-based and prospective cohort studies. *Lancet Respir Med* 2020; **11**: 1094-1105.
27. Bosch, A. A. T. M. *et al.* Maturation of the Infant Respiratory Microbiota, Environmental Drivers, and Health Consequences. A Prospective Cohort Study. *Am. J. Respir. Crit. Care Med.* **196**, 1582–1590 (2017).
28. Man, W. H. *et al.* Loss of Microbial Topography between Oral and Nasopharyngeal Microbiota and Development of Respiratory Infections Early in Life. *Am J Respir Crit Care Med* 2019; **200**: 760–70.
29. Budden, K. F. *et al.* Emerging pathogenic links between microbiota and the gut–lung axis. *Nat. Rev. Microbiol.* **15**, 55–63 (2016).
30. Dickson, R. P. *et al.* Enrichment of the lung microbiome with gut bacteria in sepsis and the acute respiratory distress syndrome. *Nat. Microbiol.* **1**, 16113 (2016).
31. Dickson, R. P. *et al.* Lung microbiota predict clinical outcomes in critically ill patients. *Am. J. Respir. Crit. Care Med.* **201**, 555–563 (2020).
32. Dominguez-Bello, M. G. *et al.* Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. *Nat. Med.* **22**, 250–3 (2016).
33. Yassour, M. *et al.* Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. *Cell Host Microbe* **24**, 146-154.e4 (2018).
34. Makino, H. *et al.* Mother-to-Infant Transmission of Intestinal Bifidobacterial Strains Has an Impact on the Early Development of Vaginally Delivered Infant's Microbiota. *PLoS One* **8**, e78331 (2013).
35. Korpela, K. *et al.* Maternal Fecal Microbiota Transplantation in Cesarean-Born Infants Rapidly Restores Normal Gut Microbial Development: A Proof-of-Concept Study. *Cell* **183**, 324-334.e5 (2020).
36. Shin, H. *et al.* The first microbial environment of infants born by C-section: the operating room microbes. *Microbiome* **3**, 59 (2015).
37. World Health Organization. *Protecting, promoting and supporting breastfeeding in facilities providing maternity and newborn services. World Health Organisation* (2017).

38. Azad, M. B. *et al.* Impact of maternal intrapartum antibiotics, method of birth and breastfeeding on gut microbiota during the first year of life: a prospective cohort study. *BJOG* **123**, 983–93 (2016).
39. Zivkovic, A. M., German, J. B., Lebrilla, C. B. & Mills, D. A. Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl 1**, 4653–8 (2011).
40. Mukhopadhyay, S., Eichenwald, E. C. & Puopolo, K. M. Neonatal early-onset sepsis evaluations among well-appearing infants: projected impact of changes in CDC GBS guidelines. *J. Perinatol.* **33**, 198–205 (2013).
41. Escobar, G. J. *et al.* Stratification of Risk of Early-Onset Sepsis in Newborns ≥ 34 Weeks' Gestation. *Pediatrics* **133**, 30–36 (2014).
42. Nederlandse Vereniging voor Kindergeneeskunde. Richtlijn Koorts in de tweede lijn bij kinderen van 0-16 jaar. (2013). Available at: https://www.nvk.nl/Portals/0/richtlijnen/koorts/koortsrichtlijn.pdf.
43. Muller-Pebody, B. *et al.* Empirical treatment of neonatal sepsis: Are the current guidelines adequate? *Arch. Dis. Child. Fetal Neonatal Ed.* **96**, (2011).
44. De Man, P., Verhoeven, B. A. N., Verbrugh, H. A., Vos, M. C. & Van Den Anker, J. N. An antibiotic policy to prevent emergence of resistant bacilli. *Lancet* **355**, 973–978 (2000).
45. Davies, J. & Davies, D. Davies J, Davies D. Origins and Evolution of Antibiotic Resistance. *Microbiol. Mol. Biol. Rev.* **74**, 417–433 (2010).
46. van Schaik, W. The human gut resistome. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370**, 20140087 (2015).
47. Martinez, J. L. General principles of antibiotic resistance in bacteria. *Drug Discovery Today: Technologies* **11**, 33–39 (2014).
48. Macinga, D. R. & Rather, P. N. The chromosomal 2'-N-acetyltransferase of Providencia stuartii: physiological functions and genetic regulation. *Frontiers in bioscience : a journal and virtual library* **4**, (1999).
49. Macfarlane, S. & Macfarlane, G. T. Regulation of short-chain fatty acid production. *Proc. Nutr. Soc.* **62**, 67–72 (2003).
50. Buelow, E. *et al.* Comparative gut microbiota and resistome profiling of intensive care patients receiving selective digestive tract decontamination and healthy subjects. *Microbiome* **5**, 88 (2017).
51. Buelow, E. *et al.* Limited influence of hospital wastewater on the microbiome and resistome of wastewater in a community sewerage system. *FEMS Microbiol. Ecol.* **94**, (2018).
52. Moore, A. M. *et al.* Pediatric fecal microbiota harbor diverse and novel antibiotic resistance genes. *PLoS One* **8**, e78822 (2013).
53. Gosalbes, M. J. *et al.* High frequencies of antibiotic resistance genes in infants' meconium and early fecal samples. *J. Dev. Orig. Health Dis.* **7**, 35–44 (2016).
54. von Wintersdorff, C. J. H. *et al.* The gut resistome is highly dynamic during the first months of life. *Future Microbiol.* **11**, 501–10 (2016).
55. Falagas, M. E. & Michalopoulos, A. Polymyxins: Old antibiotics are back. *Lancet* **367**, 633–634 (2006).
56. Gibson, M. K. *et al.* Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat. Microbiol.* **1**, 16024 (2016).

57. Hoberman, A. *et al.* Antimicrobial prophylaxis for children with vesicoureteral reflux. *N. Engl. J. Med.* **370**, 2367–2376 (2014).
58. Williamson, D. A., Carter, G. P. & Howden, B. P. Current and emerging topical antibacterials and antiseptics: Agents, action, and resistance patterns. *Clinical Microbiology Reviews* **30**, 827–860 (2017).
59. Hyun, D. Y. *et al.* Antimicrobial stewardship in pediatrics: how every pediatrician can be a steward. *JAMA Pediatr.* **167**, 859–66 (2013).
60. Rooney, A. M. *et al.* Each Additional Day of Antibiotics is Associated with Lower Gut Anaerobes in Neonatal Intensive Care Unit Patients. *Clin. Infect. Dis.* (2019). doi:10.1093/cid/ciz698
61. Roubaud-Baudron, C. *et al.* Long-term effects of early-life antibiotic exposure on resistance to subsequent bacterial infection. *MBio* **10**, (2019).
62. Singer, J. R. *et al.* Preventing dysbiosis of the neonatal mouse intestinal microbiome protects against late-onset sepsis. *Nat. Med.* **25**, 1772–1782 (2019).
63. Rogawski, E. T. *et al.* The Effect of Early Life Antibiotic Exposures on Diarrheal Rates among Young Children in Vellore, India. *Pediatr. Infect. Dis. J.* **34**, 583–588 (2015).
64. Perletti, G. *et al.* Prevention and modulation of aminoglycoside ototoxicity (Review). *Mol. Med. Rep.* **1**, 3–13
65. 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.* **128**, 646-652.e5 (2011).
66. Kostic, A. D. *et al.* The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. *Cell Host Microbe* **17**, 260–73 (2015).
67. Vatanen, T. *et al.* The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. *Nature* **562**, 589–594 (2018).
68. 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* **158**,1584-1596 (2020).
69. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220–230 (2012).
70. Biesbroek, G. *et al.* Early respiratory microbiota composition determines bacterial succession patterns and respiratory health in children. *Am. J. Respir. Crit. Care Med.* **190**, 1283–1292 (2014).
71. Pitt, H. A., Roberts, R. B. & Johnson, W. D. Gentamicin Levels in the Human Biliary Tract. *J. Infect. Dis.* **127**, 299–302 (1973).
72. Layeghifard, M., Hwang, D. M. & Guttman, D. S. Disentangling Interactions in the Microbiome: A Network Perspective. *Trends in Microbiology* **25**, 217–228 (2017).
73. Ley, R. E., Peterson, D. A. & Gordon, J. I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**, 837–848 (2006).
74. Faust, K. *et al.* Microbial co-occurrence relationships in the Human Microbiome. *PLoS Comput. Biol.* **8**, (2012).
75. Hall, C. V. *et al.* Co-existence of Network Architectures Supporting the Human Gut Microbiome. *iScience* **22**, 380–391 (2019).

76. Soret, P. *et al.* Respiratory mycobiome and suggestion of inter-kingdom network during acute pulmonary exacerbation in cystic fibrosis. *Sci. Rep.* **10**, 3589 (2020).
77. Baldassano, S. N. & Bassett, D. S. Topological distortion and reorganized modular structure of gut microbial co-occurrence networks in inflammatory bowel disease. *Sci. Rep.* **6**, 1–14 (2016).
78. Banerjee, S., Schlaeppi, K. & van der Heijden, M. G. A. Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* **16**, 567–576 (2018).
79. Berry, D. & Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front. Microbiol.* **5**, 219 (2014).
80. Deasy, A. M. *et al.* Nasal inoculation of the commensal neisseria lactamica inhibits carriage of neisseria meningitidis by young adults: A controlled human infection study. *Clin. Infect. Dis.* **60**, 1512–1520 (2015).
81. Baxter, N. T. *et al.* Dynamics of human gut microbiota and short-chain fatty acids in response to dietary interventions with three fermentable fibers. *MBio* **10**, (2019).
82. Yuille, S., Reichardt, N., Panda, S., Dunbar, H. & Mulder, I. E. Human gut bacteria as potent class I histone deacetylase inhibitors in vitro through production of butyric acid and valeric acid. *PLoS One* **13**, e0201073 (2018).
83. Rivière, A., Selak, M., Lantin, D., Leroy, F. & De Vuyst, L. Bifidobacteria and butyrate-producing colon bacteria: Importance and strategies for their stimulation in the human gut. *Frontiers in Microbiology* **7**, 979 (2016).
84. Man, W. H. *et al.* Bacterial and viral respiratory tract microbiota and host characteristics in children with lower respiratory tract infections: a matched case-control study. *Lancet Respir. Med.* **7**, 417–426 (2019).
85. Langevin, S. *et al.* Early nasopharyngeal microbial signature associated with severe influenza in children: a retrospective pilot study. *J. Gen. Virol.* **98**, 2425–2437 (2017).
86. Chua, H. H. *et al.* Intestinal Dysbiosis Featuring Abundance of Ruminococcus gnavus Associates With Allergic Diseases in Infants. *Gastroenterology* **154**, 154–167 (2018).
87. Henke, M. T. *et al.* Ruminococcus gnavus, a member of the human gut microbiome associated with Crohn's disease, produces an inflammatory polysaccharide. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 12672–12677 (2019).
88. Scher, J. U. *et al.* Decreased bacterial diversity characterizes the altered gut microbiota in patients with psoriatic arthritis, resembling dysbiosis in inflammatory bowel disease. *Arthritis Rheumatol.* **67**, 128–139 (2015).
89. Ventura, M. *et al.* The Bifidobacterium dentium Bd1 Genome Sequence Reflects Its Genetic Adaptation to the Human Oral Cavity. *PLoS Genet.* **5**, e1000785 (2009).
90. Gueimonde, M. *et al.* Genome sequence of Parascardovia denticolens IPLA 20019, isolated from human breast milk. *Journal of Bacteriology* **194**, 4776–4777 (2012).
91. Henrick, B. M. *et al.* Elevated Fecal pH Indicates a Profound Change in the Breastfed Infant Gut Microbiome Due to Reduction of Bifidobacterium over the Past Century . *mSphere* **3**, (2018).
92. Bettes, B. A. *et al.* Cesarean delivery on maternal request: Obstetrician-gynecologists' knowledge, perception, and practice patterns. *Obstet. Gynecol.* **109**, 57–66 (2007).

93. Weston, E. J. *et al.* The burden of invasive early-onset neonatal sepsis in the united states, 2005-2008. *Pediatr. Infect. Dis. J.* **30**, 937–941 (2011).
94. Kuzniewicz, M. W. *et al.* A quantitative, risk-based approach to the management of neonatal early-onset sepsis. *JAMA Pediatr.* **171**, 365–371 (2017).
95. Achten, N. B. *et al.* Association of Use of the Neonatal Early-Onset Sepsis Calculator with Reduction in Antibiotic Therapy and Safety: A Systematic Review and Meta-analysis. *JAMA Pediatr.* **173**, 1032–1040 (2019).
96. Stocker, M. *et al.* Procalcitonin-guided decision making for duration of antibiotic therapy in neonates with suspected early-onset sepsis: a multicentre, randomised controlled trial (NeoPIns). *Lancet* **390**, 871–881 (2017).
97. Leggett, R. M. *et al.* Rapid MinION profiling of preterm microbiota and antimicrobial-resistant pathogens. *Nat. Microbiol.* **5**, 430–442 (2020).
98. Furfaro, L. L., Chang, B. J. & Payne, M. S. Applications for bacteriophage therapy during pregnancy and the perinatal period. *Frontiers in Microbiology* **8**, (2018).
99. Yakult. (2018). Available at: https://yakult.nl/.
100. Guo, Q., Goldenberg, J. Z., Humphrey, C., El Dib, R. & Johnston, B. C. Probiotics for the prevention of pediatric antibiotic-associated diarrhea. *Cochrane Database Syst. Rev.* **4**, CD004827 (2019).
101. Petschow, B. W. *et al.* Effects of feeding an infant formula containing Lactobacillus GG on the colonization of the intestine: A dose-response study in healthy infants. *J. Clin. Gastroenterol.* **39**, 786–790 (2005).
102. Frese, S. A. *et al.* Persistence of Supplemented Bifidobacterium longum subsp. infantis EVC001 in Breastfed Infants. *mSphere* **2**, e00501-17 (2017).
103. Korpela, K. *et al.* Probiotic supplementation restores normal microbiota composition and function in antibiotic-treated and in caesarean-born infants. *Microbiome* **6**, 182 (2018).
104. Lynch, M. D. J. & Neufeld, J. D. Ecology and exploration of the rare biosphere. *Nature Reviews Microbiology* **13**, 217–229 (2015).

APPENDICES

Review committee

Prof. dr. M.J.N.L. Benders
Prof. dr. L.J. Bont
Prof. dr. E.J. Kuijper
Prof. dr. J.H.H.M. van de Wijgert
Prof. dr. R.J.L. Willems

Nederlandse samenvatting

Nadat een baby is geboren, wordt het overspoeld door micro-organismen uit de omgeving die zich vervolgens nestelen op alle lichaamsoppervlakten. Zo wordt de ontwikkeling van een unieke bacteriegemeenschap, het microbioom, ingeluid. Het wordt verondersteld dat de cruciale ontwikkeling van het microbioom zich afspeelt in de eerste 1000 dagen vanaf de bevruchting. Deze vroege microbiële kolonisatie valt samen met een belangrijke periode in het leven waarin het immuunsysteem ontvankelijk is voor microbiële training. Gedurende deze "window of opportunity" speelt de interactie tussen het microbioom en de gastheer een fundamentele rol in het trainen van diens immuunsysteem. Bacteriën kunnen door middel van het produceren van microbiële metabolieten en afbraakproducten communiceren met het ontwikkelende immuunsysteem en helpen dit zo te programmeren. Zodoende kan een verstoorde microbiële ontwikkeling de basis vormen van vatbaarheid voor ziekten later in het leven. Het ontwikkelende microbioom is gevoelig voor invloeden van buitenaf, zoals antibiotica, het dieet en blootstelling aan broers en zussen. Nog voordat een kind aan dit soort factoren wordt blootgesteld, is het microbioom echter al beïnvloed door de wijze van geboorte. Dit kan een vaginale geboorte zijn, maar steeds vaker ook geboorte middels keizersnede. Daarbij is het in het ziekenhuis vaak standaardprocedure om moeders die met een keizersnede bevallen intraveneuze (via het infuus) antibiotica te geven om een postoperatieve infectie te voorkomen. Deze antibiotica bereiken dan vaak ook het kind en hebben zo invloed op het microbioom.

De geboortewijze beïnvloed de ontwikkeling van het darmmicrobioom in zuigelingen, onafhankelijk van maternale antibiotica

Er bestaat omvangrijk wetenschappelijk bewijs dat geboorte middels keizersnede leidt tot 1) een afwijkende microbiële ontwikkeling, 2) een afname van de met gezondheid geassocieerde bifidobacteriën en 3) een toename van potentieel ziekmakende bacteriën in de darmen. In de onderzoekswereld speelt al geruime tijd een discussie of de door keizersnedegeboorte veroorzaakte veranderingen in het microbioom van de zuigeling verklaard kunnen worden door de geboortewijze zelf, of door de antibiotica die vooraf aan de moeder zijn gegeven. Recent veroorzaakte een grote studie ophef door te suggereren dat de gevonden microbiële veranderingen na keizersnede hoofdzakelijk gedreven worden door de antibioticaprofylaxe (IAP) die moeders en zuigelingen krijgen, en niet de wijze van geboorte. In de betreffende studie kregen alle moeders die met een keizersnede bevielen IAP, dus konden deze

twee factoren niet onafhankelijk van elkaar worden bestudeerd. In **hoofdstuk 2** beschreven we de resultaten van ons gezonde geboortecohort (Microbiome Utrecht Infant Study, afgekort: MUIS), waaraan 74 vaginaal geboren en 46 middels keizersnede geboren kinderen deelnamen. Deze studie was specifiek opgezet zodat moeders die antibioticabehandeling nodig hadden dit pas kregen nadat de navelstreng was doorgeknipt. Dit verzekerde ons ervan dat er geen antibiotica werd overgedragen van moeder naar kind, zodat we de effecten van geboortewijze onafhankelijk van IAP konden testen.

We verzamelden ontlastingmonsters van de vaginale en middels keizersnede geboren kinderen en onderzochten de algehele darmmicrobioom samenstelling hiervan. Dit is een methode die rekening houdt met de relatieve aanwezigheid van de verschillende bacteriesoorten die zich in een monster kunnen bevinden. We vonden dat de algehele darmmicrobioom samenstelling verschilde tussen de twee groepen kinderen. Dit verschil was een week na de geboorte het meest opvallend, maar na twee maanden was dit nog steeds significant. Onze bevindingen werden onderbouwd door de resultaten van een subanalyse die we verrichtten onder kinderen die exclusieve flesvoeding kregen (11 vaginaal en 11 middels keizersnede geboren kinderen), waarmee we de potentiële effecten van antibioticaoverdracht via moedermelk konden elimineren. De resultaten van deze subanalyse hielpen ook de hypothese te ontkrachten dat verschillen in darmmicrobioom samenstelling tussen vaginaal en middels keizersnede geboren kinderen voornamelijk veroorzaakt worden door de veelal suboptimale borstvoeding na een keizersnede. Naast minder succesvolle borstvoeding zijn ook andere aan keizersnedegeboorte gerelateerde factoren, zoals de afwezigheid van gebroken vliezen en kortere zwangerschapsduur, aangewezen als oorzakelijke factoren van een afwijkende samenstelling van het microbioom. In ons gezonde geboortecohort konden wij echter geen associatie vinden tussen de duur van de gebroken vliezen of zwangerschapsduur en het darmmicrobioom. Bovendien zagen wij dat het darmmicrobioom van zuigelingen die waren geboren met een spoedkeizersnede, die een vergelijkbaar fysiek geboorteproces doormaakten als kinderen bij een vaginale bevalling, toch meer leek op dat van kinderen die met een geplande keizersnede waren geboren, dan dat van vaginaal geboren kinderen.

In **hoofdstuk 2** beschrijven we dat middels keizersnede geboren zuigelingen een vertraagde verwerving hebben van *Bifidobacterium*. Bifidobacteriën staan bekend om hun gebruik als probiotica. Ze bevorderen de gezondheid van de darm en bieden bescherming tegen ziekteverwekkers. Gedurende de eerste levensmaand was de aanwezigheid van *Bifidobacterium* hoger in de darmen van vaginaal

geboren kinderen dan in de middels keizersnede geboren kinderen. In dezelfde periode was de potentieel ziekmakende *Enterococcus* hoger aanwezig in de darmen van met keizersnede geboren kinderen. Mogelijk kan dit verklaard worden door een zwakkere microbiële weerstand, vanwege een gebrek aan de beschermende *Bifidobacterium* in de middels keizersnede geboren kinderen. Ondanks het feit dat de algehele samenstelling van het darmmicrobioom van de twee groepen kinderen na twee maanden vergelijkbaar was, bleven verschillen in specifieke bacteriën veel langer bestaan. Bijvoorbeeld de aanwezigheid van *Klebsiella*, die infecties kan veroorzaken zoals een longontsteking, was vier maanden na de geboorte nog steeds verhoogd aanwezig in de keizersnedegroep. Zowel de aanwezigheid van *Klebsiella* als *Enterococcus* kon in ons cohort geassocieerd worden met longgezondheid (zie het volgende hoofdstuk).

De primaire microbiële kolonisatie van kinderen geboren middels keizersnede wordt waarschijnlijk deels gestuurd door microben afkomstig uit de operatiekamer. Uit een studie waarin verschillende oppervlakten van operatiekamers werden bemonsterd, bleek dat de inhoud van operatiekamerstof overeenkwam met menselijke huidbacteriën, en voornamelijk bestond uit *Staphylococcus*. Deze huidbacterie was in onze studie hoger aanwezig in de eerste levensdagen van kinderen geboren middels keizersnede. Mogelijk dat de *Klebsiella* die ook meer aanwezig was in deze groep kinderen eveneens afkomstig is uit de operatiekamer. In het verleden is een uitbraak met een multiresistente *Klebsiella pneumoniae* stam in een Nederlands perifeer ziekenhuis geassocieerd met besmet operatiekamermateriaal. Het zou waardevol zijn om in de toekomst in ziekenhuizen te beoordelen of operatiekameroppervlakten in het algemeen meer besmet zijn met *Klebsiella* dan de verloskamers, om meer inzicht te krijgen in de wijze van verwerving van deze bacterie direct na keizersnedegeboorte.

De samenstelling van het darmmicrobioom vroeg in het leven is geassocieerd met longgezondheid

Zoals eerder beschreven, namen de verschillen in algehele darmmicrobioom samenstelling tussen de vaginale en middels keizersnede geboren kinderen af in de loop van hun eerste levensjaar, maar de vastgestelde verstoringen in het ontwikkelende darmmicrobioom kunnen langdurige gevolgen hebben voor de gezondheid. Deze hypothese wordt ondersteund door meerdere epidemiologische studies die keizersnedegeboorte verbinden met een verhoogde kans op astma, allergieën, overgewicht, suikerziekte en inflammatoire darmziekten zoals de ziekte van Crohn en colitis ulcerosa.

Daarbij bestaat er epidemiologische data dat geboortewijze met longaandoeningen verbindt, maar slechts weinig studies hebben de rol bestudeerd van veranderingen in het darmmicrobioom, die door geboortewijze gedreven zijn, en longziekten. In **hoofdstuk 2** vonden wij echter dat de samenstelling van het darmmicrobioom op heel jonge leeftijd (een week) geassocieerd was met het aantal doorgemaakte luchtweginfecties (LWI's) in het eerste levensjaar. Het MUIS-cohort kon verdeeld worden op basis van de hoeveelheid doorgemaakte LWI's, namelijk weinig (0-2), gemiddeld (3-4) en veel (5-7). De bacteriën die een sterke associatie lieten zien met een hoge vatbaarheid voor LWI's waren *Klebsiella* en *Enterococcus*. Zoals hierboven vermeld, waren deze soorten meer aanwezig in kinderen geboren met keizersnede, en vormen zo een mogelijke link tussen geboortewijze en vatbaarheid voor infecties.

Er vindt bacteriële overdracht van moeder naar kind plaats, via de ontlasting

Het blijft de vraag welke aspecten van geboortewijze de verschillen in de ontwikkeling van het microbioom verklaren. Een mogelijke verklaring voor het verschil in darmmicrobioom samenstelling tussen vaginaal en middels keizersnede geboren kinderen is het verschil in geboorteroute. Een vaginaal geboren kind wordt door het baringskanaal van de moeder geperst, terwijl een kind die middels keizersnede geboren wordt uit de buik van de moeder wordt geplukt door een operateur. Het gevolg hiervan is dat de zuigelingen aan verschillende lichaamsoppervlakten worden blootgesteld en op die manier worden verschillende microben fysiek overgedragen van moeder naar kind. In een kleine interventiestudie werden met keizersnede geboren kinderen ingesmeerd met de vaginale flora van hun moeder. Vervolgens werden hun anale, orale en huidmonsters vergeleken met die van vaginaal geboren kinderen, alsook met die van middels keizersnede geboren kinderen die niet waren ingesmeerd. De monsters van de ingesmeerde keizersnede kinderen werden vaker geclassificeerd als zijnde monsters verzameld van vaginaal geboren kinderen, dan als zijnde monsters van niet-ingesmeerde keizersnede kinderen. Dit was hoofdzakelijk het geval voor de orale en huidmonsters, minder vaak voor de anale monsters. Wij veronderstellen dat het voor een gezonde microbiële ontwikkeling noodzakelijk is dat een kind zowel aan de vaginale als de intestinale flora van moeder wordt blootgesteld. In onze studie vonden we daadwerkelijk bewijs voor microbiële overdracht van moeder naar kind via ontlasting in de vaginaal geboren, maar niet in de met keizersnede geboren kinderen.

De vertraagde verwerving van *Bifidobacterium* bij de middels keizersnede geboren kinderen die we in **hoofdstuk 2** beschreven, impliceert dat de overdracht van deze bacterie van moeder naar kind tijdens een vaginale geboorte essentieel is voor

de oorspronkelijke kolonisatie. Verticale transmissie (van moeder naar kind) van meerdere *Bifidobacterium* soorten is in eerdere studies reeds gedemonstreerd, zoals van *Bifidobacterium longum* en *Bifidobacterium adolescentis*. Deze twee specifieke soorten vormden overigens samen het grootste aandeel van alle bifidobacteriën in onze studie. Een recente "proof-of-concept" studie bevestigt nogmaals het belang van blootstelling van kinderen aan de ontlasting van hun moeder: kinderen geboren middels keizersnede kregen een verdunde ontlastingmonster van hun moeder toegediend, waarna hun darmmicrobioom een vergelijkbare ontwikkeling doormaakte als dat van vaginaal geboren kinderen.

Borstvoeding herstelt onvolledig het gebrek aan *Bifidobacterium* in kinderen na keizersnede

Net als geboortewijze wordt het dieet veel bestudeerd in relatie tot de ontwikkeling van het darmmicrobioom. De Wereldgezondheidsorganisatie adviseert exclusieve borstvoeding in de eerste 6 maanden van het leven. Na een keizersnede is de (productie) van borstvoeding vaak vertraagd en suboptimaal. Daarom lijkt een goede begeleiding van borstvoeding in moeders die met keizersnede zijn bevallen cruciaal te zijn. Des te meer omdat het geven van borstvoeding de veranderingen die door IAP zijn toegebracht aan het darmmicrobioom lijkt te kunnen herstellen, zoals uit een eerdere studie is gebleken. In **hoofdstuk 2** bestudeerden wij of borstvoeding de veranderingen in het darmmicrobioom na keizersnede zou kunnen moduleren. We vonden slechts een bescheiden verschil in effectgrootte van geboortewijze op darmmicrobioom samenstelling wanneer we de resultaten vergeleken van analyses verricht met de gehele groep kinderen of alleen die exclusieve flesvoeding hadden gekregen (en dus nooit borstvoeding hadden gehad). Daarnaast hadden kinderen geboren met keizersnede die borstvoeding kregen nog steeds een lager gehalte van deze bacterie in hun ontlastingmonsters dan vaginaal geboren kinderen die exclusieve flesvoeding kregen. Onze theorie is dat er eerst een adequate bacteriële overdracht van vaginale en darmbacteriën moet plaatsvinden, voordat de groei van gunstige soorten zoals *Bifidobacterium* kan worden bevorderd door middel van bacteriële voedingsstoffen die in borstvoeding aanwezig zijn. Daarom denken wij dat het stimuleren van borstvoeding bij vrouwen die met een keizersnede zijn bevallen niet voldoende is om het tekort in *Bifidobacterium* bij hun kinderen op te lossen. Mogelijk dat in de toekomst het insmeren van een kind met een combinatie van vaginale en darmbacteriën de oplossing zal zijn, of de toediening van probiotica samen met borstvoeding.

Zelfs een korte antibioticakuur vroeg in het leven beïnvloedt sterk het darmmicrobioom en -resistoom

Antibiotica zijn bekende drijvers van microbiële veranderingen en antimicrobiële resistentie (AMR) in de samenleving. Breedspectrum antibiotica zijn desalniettemin potentieel levensreddend in het geval van early-onset neonatale sepsis (EONS), oftewel een neonatale infectie in de eerste levensdagen. Antibiotica worden veel voorgeschreven aan pasgeborenen: 4-10% van alle pasgeborenen krijgt een antibioticakuur toegediend vroeg in het leven. In onze ZEBRA-studie, dat staat voor Zuigelingen en Bacteriële Resistentie na Antibiotica, includeerden wij 147 pasgeborenen met een (verdenking op een) EONS waarbij breedspectrum antibioticabehandeling geïndiceerd was in hun eerste levensweek. We randomiseerden de deelnemers over drie antibioticaregimes die het meest worden voorgeschreven voor deze indicatie in Nederland, namelijk penicilline + gentamicine, Augmentin + gentamicine en amoxicilline + cefotaxim. Een subset van 80 leeftijd-gematchte controles van het gezonde MUIS-cohort, die geen antibiotica ontvingen in hun eerste levensweek, dienden als controles. De resultaten van de ZEBRA-studie presenteerden we in **hoofdstuk 4** van dit proefschrift.

De meest voorkomende verwekkers van EONS zijn de groep-B-streptokok (GBS), *Listeria monocytogenes* en *Escherichia coli*. Deze organismen zijn gevoelig voor antibioticaregimes die bestaan uit een combinatie van penicilline met gentamicine of amoxicilline met een 3e-generatie cefalosporine (bijvoorbeeld cefotaxim). Daarom zijn deze combinaties even effectief voor de behandeling van EONS. Echter, al in het jaar 2000 werd aan het licht gebracht dat amoxicilline overgroei zou geven van bacteriën die β-lactamase (een enzym dat penicilline inactiveert) produceren, zoals *Klebsiella* soorten. Een ander probleem dat werd aangedragen, was dat 3e-generatie cefalosporines zouden selecteren voor resistente *Enterobacter* soorten. *Klebsiella* en *Enterobacter* soorten behoren samen met *Enterococcus faecium*, *Staphylococcus aureus*, *Acinetobacter baumanii* en *Pseudomonas aeruginosa* tot de ESKAPE-groep. Dit zijn pathogenen die bekend staan om het veroorzaken van infecties die worden opgelopen in het ziekenhuis of daarbuiten en die de effecten van antibiotica effectief kunnen ontsnappen (vandaar “escape”). Daarom is het belangrijk om een werkzame antibioticacombinatie te kiezen voor de behandeling van EONS die de neonatale darmcommensalen intact houdt, en tot het minste AMR leidt.

Antibiotica en bacteriële AMR-genen (die coderen voor de resistentie tegen antibiotica) gaan namelijk hand in hand. Antibiotica bestaan al miljarden jaren in de natuur als microbiële producten. In een omgeving met schaarse middelen geven de antibiotica die bacteriën afscheiden een competitief voordeel. Concurrerende

bacteriën kunnen ermee worden gedood of geremd in hun groei. Onder antibioticadruk kan een voorheen gevoelige bacteriële populatie resistent worden, waardoor de overlevingskansen van die populatie worden vergroot. Resistentie wordt verworven door mutaties die ontstaan in het bacteriële genetische materiaal of door horizontale overdracht van AMR-genen via kleine, extra-chromosomale DNA-moleculen, plasmiden genaamd. Uiteindelijk ontstaat functionele resistentie door verschillende mechanismen, zoals de modificatie van het antibioticum, het enzymatisch inactief maken van het antibioticum of de verlaging van de dodelijke intracellulaire concentratie van het antibioticum door het uit de bacteriecel te pompen. In sommige gevallen heeft een AMR-gen eigenlijk een andere functie in de oorspronkelijke bacterie. Zo beschikt *Providencia stuartii* over een gen dat codeert voor een enzym dat een rol speelt in het autolytisch (zelfvernietiging) systeem. Dit enzym modificeert peptidoglycaan, een bacteriële celwandcomponent. Aminoglycoside antibiotica hebben een vergelijkbare structuur als peptidoglycaan, waardoor *P. stuartii* intrinsiek resistent is voor deze antibioticaklasse.

In onze ZEBRA-studie veroorzaakte zelfs een korte (gemiddeld 48 uur durende), breedspectrum antibioticumkuur vroeg in het leven al een verandering in de samenstelling van het darmmicrobioom dat nog meetbaar was op de eerste verjaardag van de deelnemers. Twaalf maanden na antibioticatoediening waren er nog steeds specifieke bacteriesoorten, zoals die behorend tot de anti-inflammatoire *Bacteroides* genus, verminderd aanwezig in de kinderen die behandeld waren met antibiotica, in vergelijking met de gezonde controles. De toediening van antibiotica was geassocieerd met een hogere aanwezigheid van dezelfde potentieel ziekmakende bacteriën die verhoogd aanwezig waren in keizersnede kinderen: *Enterococcus* en *Klebsiella*. De grootste impact van antibiotica kon direct na de behandeling worden gemeten met grote verschuivingen in zowel darmmicrobioom samenstelling en resistoomprofiel op dat moment. We definieerden het resistoomprofiel als de samenstelling van 31 middels quantitative polymerase chain reaction (qPCR) geteste, klinisch relevante bacteriële AMR-genen die ofwel aan- dan wel afwezig waren in een monster. Na antibioticatoediening was het resistoomprofiel van de kinderen met suspected (verdenking op) EONS significant en in grote mate verschoven ten opzichte van de gezonde situatie.

Van belang is dat de resistoomprofielen niet opvallend verschilden tussen de suspected EONS (sEONS) kinderen en gezonde controles vóór het starten van de antibiotica. Beide groepen droegen al AMR-genen bij zich op het eerste testmoment, wat ongeveer overeenkwam met de eerste levensdag. Dit is in lijn met het feit dat sommige commensalen intrinsiek resistent zijn, en met eerdere studies die

de aanwezigheid van diverse AMR-genen van jonge, gezonde kinderen hebben beschreven. Wij konden dus bevestigen dat de gezonde kinderen in onze studie AMR-genen bij zich droegen, maar zagen tevens dat de ontlastingmonsters van sEONS zuigelingen een hoger aantal AMR-genen bevatten na antibiotica. In het bijzonder waren dit AMR-genen die coderen voor klinisch essentiële antibiotica zoals aminoglycosiden (waar gentamicine een voorbeeld van is), beta-lactamasen (waartoe penicilline, amoxicilline en cefotaxim behoren) en colistine (dat als laatste redmiddel voor Gram-negatieve infecties wordt ingezet na de kinderleeftijd).

Naast gerichte qPCR voerden wij ook metagenomic shotgun sequencing (MGS) uit op een subset van monsters verzameld na antibioticatoediening (overeenkomend met 1 week leeftijd van de controles) om het volledige spectrum van aanwezige AMR-genen te kunnen evalueren. Zo konden wij in totaal 1504 AMR-genen identificeren in de onderzochte monsters. Hiervan verschilden 262 AMR-genen in hoeveelheid tussen de kinderen die met antibiotica waren behandeld en controles; deze waren overigens allemaal meer aanwezig in de behandelde kinderen. De MGS-data bevestigden onze qPCR resultaten, en daarnaast konden we hiermee vaststellen dat de met antibiotica behandelde kinderen meer AMR-genen in hun monsters bevatten die codeerden voor resistentie tegen onder andere trimethoprim, sulfonamiden, bacitracine en fusidinezuur. Geen van deze antibiotica waren op het onderzochte moment nog toegediend geweest. Deze bevinding is in overeenstemming met eerder onderzoek, waarin werd aangetoond dat breedspectrum antibiotica ook een verhoging kan geven van AMR-genen die geen bekende activiteit hebben tegen het toegediende antibioticum. Mogelijk komt deze collaterale verrijking doordat het toegediende antibioticum leidt tot overgroei van multiresistente organismen die multipele AMR-genen bevatten. Resistentie tegen trimethoprim en sulfonamiden heeft klinische implicaties in de kindergeneeskunde, aangezien deze combinatie frequent wordt toegepast als profylactische behandeling voor kinderen met vesico-ureterale reflux (waarbij de urine van de blaas terug naar de nieren kan stromen), om recidiverende urineweginfecties te voorkomen. Bacitracine en fusidinezuur worden vaak gebruikt voor de lokale behandeling van huidinfecties bij zowel kinderen als volwassenen.

Overigens ligt de focus van zorgvuldig antibioticabeleid tot nu toe op het verkorten van de duur van de antibioticakuur. Wij beschreven echter in **hoofdstuk 4** dat de associatie tussen de duur van de kuur en de algehele darmmicrobioom samenstelling relatief klein was. De berekende effectgrootte, of R^2, van antibioticaduur was slechts 0.3%. In vergelijking met andere variabelen die geassocieerd waren met de samenstelling van het microbioom, zoals het antibioticaregime (R^2 0.9%), leeftijd

(R^2 5.2%), kinderdagverblijf bezoek (R^2 1.0%), en borstvoeding (R^2 0.8%), was de effectgrootte van antibioticaduur in onze studie minimaal.

Natuurlijk is het vooralsnog onbekend of de verstoring van het darmmicrobioom of de verrijking van de AMR-genen die wij in onze studie vonden, zal leiden tot een hogere incidentie van (antibioticaresistente) bacteriële infecties later in het leven. Een muisstudie heeft al wel laten zien dat door antibiotica veroorzaakte veranderingen in het darmmicrobioom van muizen deze meer vatbaar maakten voor bacteriële infecties later in het leven. In een studie bij kinderen werd vroege antibioticatoediening geassocieerd met een verhoogde incidentie van diarree op de kinderleeftijd, mogelijk als consequentie van persistente microbiële verstoring, hypothetiseren wij nu. De deelnemers van de MUIS- en ZEBRA-cohorten worden gevolgd tot de leeftijd van vijf jaar, dus deze vraagstellingen zullen tegen die tijd waardevol zijn om aan te kaarten.

De combinatie penicilline met gentamicine is de voorkeursbehandeling voor EONS

Ons hoofddoel in **hoofdstuk 4** was het vergelijken van de effecten op het darmmicrobioom en -resistoom van drie intraveneuze antibioticaregimes die het meest worden voorgeschreven voor EONS in Nederland, namelijk penicilline + gentamicine, Augmentin + gentamicine en amoxicilline + cefotaxim. Zoals eerder beschreven, zijn deze drie regimes even effectief in het bestrijden van de meest voorkomende verwekkers van EONS. In de praktijk wordt de keuze van het regime dat op de neonatologieafdeling wordt voorgeschreven ook gebaseerd op toedieningsgemak, mogelijke bijwerkingen en kosten. De frequentie van toediening van de penicilline + gentamicine en Augmentin + gentamicine regimes is hetzelfde (penicilline en Augmentin worden drie keer per dag toegediend en gentamicine een keer per dag), terwijl amoxicilline drie keer per dag wordt toegediend, gecombineerd met twee keer een dosis cefotaxim. Als gentamicine wordt toegediend, is een extra bloedonderzoek nodig om het serumgehalte hiervan te monitoren, aangezien een te hoog gehalte geassocieerd wordt met nier- en gehoorschade.

Van de drie antibioticaregimes die wij vergeleken, vonden we dat de combinatie penicilline + gentamicine de kleinste impact had op alle bestudeerde eindpunten in vergelijking met de Augmentin + gentamicine en amoxicilline + cefotaxim regimes. Ten eerste, beïnvloedde de combinatie penicilline + gentamicine de microbiële soortrijkheid het minste in vergelijking met gezonde controles. Een verlaagde soortrijkheid van het darmmicrobioom vroeg in het leven is geassocieerd met allergische aandoeningen en suikerziekte later in het leven. Ten tweede, tastte de

combinatie penicilline + gentamicine de algehele darmmicrobioom samenstelling het minste aan, terwijl amoxicilline + cefotaxim de grootste verstoring gaf. Vervolgens vergeleken we de stabiliteit van de ontwikkeling van het darmmicrobioom tussen de regimes. Men neemt aan dat stabiliteit van het darmmicrobioom de veerkracht ervan bevordert door weerstand te bieden tegen kolonisatie door microben. In de nasopharynx (neus-keelholte) is een stabiel microbioom door onze onderzoeksgroep reeds eerder geassocieerd met een verminderd aantal van opeenvolgende LWI's. In onze studie liet het darmmicrobioom de meest stabiele ontwikkeling zien in de kinderen behandeld met penicilline + gentamicine. Verder, op het niveau van specifieke bacteriesoorten, zagen we dat alle bifidobacteriën die significant in aanwezigheid verschilden tussen de penicilline + gentamicine en amoxicilline + cefotaxim regimes meer aanwezig waren in de eerste groep (behalve *Bifidobacterium animalis*). De beperkte ecologische bijwerkingen van de penicilline + gentamicine regime kunnen deels worden verklaard doordat aminoglycosiden slechts in geringe mate in de darm doordringen wanneer ze intraveneus worden gegeven.

Ook het resistoomprofiel werd het minste beïnvloed door penicilline + gentamicine, gevolgd door Augmentin + gentamicine en daarna amoxicilline + cefotaxim. Wanneer we elk regime apart vergeleken met de controlegroep, vonden we dat penicilline + gentamicine de kleinste toename gaf van AMR-genen (5 van de 31 in totaal) versus 10 van de 31 voor zowel de Augmentin + gentamicine als de amoxicilline + cefotaxim groepen. Tenslotte, in de inter-regime vergelijkingen, gaf weer de combinatie penicilline + gentamicine de kleinste toename van AMR-genen in vergelijking met de andere twee regimes (2 AMR-genen, versus 9 voor de amoxicilline + cefotaxim en 5 voor de Augmentin + gentamicine groep). Concluderend, ook om het probleem van AMR te verkleinen, bevelen wij penicilline + gentamicine aan voor de behandeling van EONS.

Niche-overkoepelende microbiële netwerken en hun relatie met gezondheid

Bij EONS wordt vaak, zoals bij veel ziektebeelden, een enkele microbe (bijvoorbeeld GBS, *L. monocytogenes* of *E. coli*) aangewezen als de veroorzakende pathogeen. Zoals eerder besproken, wordt overgroei door een pathogeen vaak voorafgegaan door een verstoring van het stabiele evenwicht van het microbioom, waardoor de afweer tegen kolonisatie wordt uitgeschakeld. Door dit soort inzichten is de focus van microbioomonderzoek aan het verschuiven, van het bestuderen van individuele pathogenen naar een meer holistische aanpak waarin de interacties tussen de microbiële gemeenschap en de gastheer worden onderzocht. In zo'n microbiële samenleving gelden ecologische principes: concurrentie om schaarse middelen leidt

tot competitieve uitsluiting, en omgekeerd ontstaat er een samenwerking tussen microben wanneer het product van één microbe het substraat wordt van de ander. Zo kan een microbe invloed uitoefenen op de structuur en functie van de microbiële gemeenschap binnen een anatomische niche. In theorie zou een enkele microbe de gehele microbiële gemeenschap binnen een mens kunnen beïnvloeden. Inzichten in de relaties tussen bacteriën en de netwerken die ze vormen (zowel lokaal als overkoepelend) kunnen helpen om de gezondheidsbevorderende en pathologische karakteristieken van het microbioom beter te begrijpen.

In eerdere studies gebaseerd op het MUIS-cohort heeft onze onderzoeksgroep reeds laten zien dat de ontwikkeling van het microbioom van de bovenste luchtwegen (nasopharynx en mondholte) geassocieerd is met respiratoire gezondheid. Zo is een versnelde rijping van het nasopharyngeale microbioom geassocieerd met het aantal LWI's dat een kind in het eerste levensjaar doormaakt. Daarnaast worden LWI's voorafgegaan door een verlies van zogeheten microbiële topografie in de bovenste luchtwegen. Dit verlies van microbiële topografie werd gekenmerkt door een toevloed van orale bacteriesoorten vanuit de mondholte naar de nasopharynx.

Anatomisch gezien, is het niet verrassend dat de samenstelling van het microbioom in de bovenste luchtwegen de respiratoire gezondheid beïnvloedt. Het concept dat het darmmicrobioom geassocieerd is met longgezondheid wordt ook steeds populairder, en wordt aangeduid met de term "gut-lung axis", of darm-long-as. Microbiële componenten en metabolieten in zowel de darmen en longen kunnen het immuunsysteem lokaal en systemisch beïnvloeden. Specifieke bacteriesoorten op beide locaties zijn geassocieerd met longziekten zoals astma, chronisch obstructieve longziekte en LWI's. Daarbovenop kan er translocatie optreden van microben tussen de darmen en de longen: het longmicrobioom is in geval van sepsis verrijkt met darmbacteriën. Bovendien kan de aanwezigheid van darmbacteriën in de longen de ernst van het beloop voorspellen in acuut zieke patiënten, wat suggereert dat de translocatie van darmbacteriën naar de longen een aandeel vormt in de pathogenese van longschade.

In het algemeen worden microbiële netwerkstudies beperkt binnen een anatomische niche. In het MUIS-cohort hebben we monsters verzameld van drie niches, namelijk de nasopharynx, de mondholte en de darmen. Zo hadden wij de mogelijkheid om het niche-overkoepelende microbioom in relatie tot longgezondheid van kinderen te bestuderen. Door de microbiële data van de drie niches te combineren, konden wij niche-overkoepelende microbiële netwerken bouwen om zo de interacties te onderzoeken tussen bacteriegemeenschappen verspreid door het lichaam. Verder

konden wij de structuur van de microbiële netwerken relateren aan LWI's. Opvallend was dat de kinderen met een hoge vatbaarheid voor LWI's meer gefragmenteerde, en minder stabiele microbiële netwerken hadden in hun eerste zes levensmaanden. Gefragmenteerde netwerken in relatie tot ziekte zijn al eens eerder beschreven, zoals bij de ziekte van Crohn. Crohn patiënten hebben ook veranderde darmmicrobioom netwerken in vergelijking met gezonde mensen. Vervolgens wilden wij in onze studie identificeren welke bacteriesoorten belangrijk waren voor de netwerkstructuur en of deze geassocieerd konden worden met LWI's.

Kernsoorten in niche-overkoepelende microbiële netwerken zijn geassocieerd met LWI's

Bacteriesoorten die een centrale rol spelen in de algehele netwerkstructuur kunnen kernsoorten worden genoemd. Die centrale rol wordt mede bepaald door de hoeveelheid verbindingen die een bacterie vormt met andere bacteriën binnen het netwerk. De kernsoorten in de netwerken van kinderen die weinig LWI's hadden doorgemaakt in hun eerste levensjaar bestonden onder andere uit de met gezondheidsgeassocieerde bacteriën *Corynebacterium*, *Dolosigranulum pigrum* en *Neisseria lactamica*. Onze onderzoeksgroep heeft al eerder vastgesteld dat zowel *Corynebacterium* en *D. pigrum* zijn geassocieerd met een gezonde rijping van het nasopharyngeale microbioom en een verlaagd risico op het ontwikkelen van LWI's in het eerste levensjaar. Onze microbiële netwerkstudie suggereert niet alleen dat *Corynebacterium* en *D. pigrum* belangrijke spelers zijn in de neus-keel en mondholte, maar ook in het algehele microbiële netwerk. Van *N. lactamica* denkt men dat het een microbieel netwerk meer veerkrachtig maakt, omdat eerdere studies hebben gedemonstreerd dat het de kolonisatie van de pathogene *Neisseria meningitidis* tegengaat.

We vonden ook dat *Ruminococcus bromii*, *Megasphaera* en *Anaerostipes* kernsoorten waren in de netwerken van kinderen die weinig vatbaar waren voor LWI's. Deze bacteriën zijn geassocieerd met de (verhoogde) productie van butyraat. Butyraat is een vetzuur en microbieel eindproduct van het fermentatieproces in de menselijke darm. Het is een essentiële metaboliet in de darm, omdat het de voorkeursenergiebron is van de bekledende cellen van de darm. Het heeft anti-inflammatoire eigenschappen en verlaagt de zuurtegraad van de darm, waarmee de groei van pathogenen wordt geremd.

Terwijl de kernsoorten die wij vonden in de netwerken van weinig vatbare kinderen in voorgaande studies geassocieerd zijn met gezondheid, was het omgekeerde ook waar. De kernsoorten in de netwerken van de meest vatbare kinderen zijn in de

literatuur geassocieerd met longziekten en inflammatoire ziekten, of met ongunstige eigenschappen zoals het veroorzaken van cariës. Zo werden *Haemophilus* en *Lachnospiraceae* geïdentificeerd als kernsoorten van de meest vatbare netwerken, welke herhaaldelijk zijn gelinkt aan recidiverende LWI's. De pro-inflammatoire *Ruminococcus gnavus*, die een rol speelt in vele immunologische problemen, zoals allergieën bij kinderen, de ziekte van Crohn en artritis psoriatica, kon ook worden aangewezen als kernsoort in de netwerken van de meest vatbare kinderen. Bovendien waren de cariogene *Bifidobacterium dentium* en *Parascardovia denticolens* ook kernsoorten in de meest vatbare netwerken. Terwijl zuurtegraad verlagende bacteriën een belangrijke rol speelden in de netwerken van de minst vatbare kinderen, werden er twee zuurtegraad verhogende *Peptostreptococcaceae* soorten geïdentificeerd als kernsoorten in de netwerken van de meest vatbare kinderen. Al met al suggereren onze bevindingen dat naar niche-overkoepelende microbiële netwerken als geheel moet worden gekeken voor verklaringen voor de systemische vatbaarheid voor (infectie)ziekten.

Toekomstperspectieven

De rol van het menselijke microbioom in gezondheid en ziekte wordt steeds duidelijker. Een verstoord microbioom vroeg in het leven staat mogelijk aan de basis van vatbaarheid voor inflammatoire ziekten en infectieziekten later in het leven. Er zijn meerdere omgevingsfactoren die het ontwikkelende microbioom kunnen beïnvloeden, bijvoorbeeld geboortewijze, zoals beschreven in **hoofdstuk 2**. Het is de verantwoordelijkheid van wetenschappers en hulpverleners om ouders die in verwachting zijn te informeren over de rol van keizersnedegeboorte op het microbioom, en daaropvolgend, op de gezondheid van hun toekomstige kind. Gynaecologen erkennen de stijgende vraag naar keizersnedebevalling, en hoewel ze geloven dat de risico's van deze procedure de voordelen overstijgen, is hier wetenschappelijk bewijs voor nodig om het beleid rondom geboortewijze aan te kunnen scherpen.

Hard bewijs dat geboorte middels keizersnede een negatieve invloed heeft op microbioom ontwikkeling, immunologische training en ziekten later in het leven zou kunnen helpen bij het coachen van ouders in de richting van een vaginale bevalling. Wanneer er geen keuze mogelijk is, zoals in het geval van complicaties bij moeder of kind die een keizersnede medisch noodzakelijk maken, is het wellicht wenselijk om de ontbrekende, gunstige microben bij het kind te suppleren. Alhoewel er reeds studies zijn verricht naar de effecten van vaginale overdracht van bacteriën, is er aanvullend onderzoek nodig om de effectiviteit van (de toedieningswijze van) probiotica, of de gecombineerde overdracht van vaginale en intestinale

bacteriën van moeder naar kind, te beoordelen om een lange termijn kolonisatie te bewerkstelligen.

Vervolgens, voor behoud van een gezond microbioom na de geboorte en om de selectie van potentiële pathogenen en AMR-genen te verminderen, moet er actie worden ondernomen om onnodige antibioticatoediening te voorkomen. Zoals in detail beschreven, hebben wij laten zien dat zelfs een kortdurende (48 uur) kuur met breedspectrum antibiotica langdurige en nadelige effecten heeft op het darmmicrobioom en op de samenstelling van AMR genen bij het kind. Het is alarmerend dat antibiotica in 4-10% van alle pasgeborenen worden voorgeschreven, terwijl slechts een geschatte 1 in 1000 pasgeborenen een bewezen infectie ontwikkelt, wat betekent dat >90% van alle behandelingen niet direct noodzakelijk is. Wij pleiten sterk voor initiatieven die een verbetering geven van de identificatie van pasgeborenen die een verhoogd risico lopen op het ontwikkelen van EONS, zoals de sepsis calculator. Deze schat het risico op EONS in op basis van vijf objectieve maternale en vier klinische neonatale risicofactoren. In een meta-analyse was het gebruik van de calculator geassocieerd met een substantiële reductie in het gebruik van empirische antibiotica bij sEONS. Kortgeleden is tevens het monitoren van het serumprocalcitonineniveau veilig gebleken om de besluitvorming te sturen rond het verkorten van de duur van antibioticabehandeling. Idealiter zullen er in de toekomst microbiële biomarkers, of patronen in de algehele (darm)microbioom samenstelling, worden geïdentificeerd die kunnen helpen bij de besluitvorming of antibiotica überhaupt gestart moeten worden. Daarbij zou real-time bacteriële DNA sequencing potentieel kunnen helpen om aan het bed te testen voor gunstige of ongunstige patronen.

Desalniettemin zal antibioticabehandeling absoluut noodzakelijk blijven voor pasgeborenen met bewezen EONS. In die gevallen spreken wij sterk de voorkeur uit voor het voorschrijven van de combinatiebehandeling penicilline met gentamicine, boven Augmentin met gentamicine of amoxicilline met cefotaxim. Van de drie antibioticaregimes die wij in **hoofdstuk 4** hebben vergeleken, is dit de meest smalspectrum antibioticacombinatie. Een ideale antibacteriële behandeling voor EONS zou er een zijn die uiterst specifiek is voor de target, oftewel super smalspectrum is, waarbij het microbioom gespaard blijft. Opkomende kandidaten die dit criterium vervullen zijn bacteriofagen, oftewel virussen die specifieke bacteriën doden.

Om gebreken in gunstige bacteriën die door antibiotica zijn opgelopen te vervangen, is er mogelijk een rol voor probiotica. Publieke interesse en wetenschappelijke

interesse op dit gebied nemen toe. Probiotica kunnen verschillende bacteriesoorten bevatten en worden verkocht als poeders of capsules in gezondheidswinkels, of zelfs als yoghurt of dranken in supermarkten. De meest bekende probioticadrank is waarschijnlijk Yakult, die *Lactobacillus casei* bevat. Er worden veel gezondheidsbevorderende aspecten toegewezen aan probiotica en een toenemend aantal onderzoeken wordt verricht naar de potentie van probiotica in de preventie van specifieke aandoeningen. Het zou interessant zijn om te testen of het toedienen van probiotica, in het bijzonder gunstige bacteriën die worden aangetast door antibiotica, behulpzaam zou zijn in het terugdraaien van verstoringen in het darmmicrobioom. De grootste uitdaging in het oplossen van een microbieel disbalans middels probiotica hangt samen met het kortstondige karakter van de kolonisatie van de toegediende bacteriën. Echter, in een studie verricht in kinderen die borstvoeding kregen en gesuppleerd werden met *Bifidobacterium infantis* bleek dat zij een hogere aanwezigheid hadden van deze bacterie in hun ontlasting dan de controlegroep, voor meer dan 30 dagen na het stoppen van de toediening. Uit een dubbelblinde, placebo-gecontroleerde, gerandomiseerde trial naar de effecten van multi-species probioticatoediening op verstoringen van het darmmicrobioom door keizersnedegeboorte en antibiotica bleek dat probiotica een sterke impact hadden op de samenstelling van het microbioom, afhankelijk van het dieet. Alleen de kinderen die borstvoeding kregen lieten een verhoogde aanwezigheid van bifidobacteriën zien. In de probiotica groep waren de effecten van antibiotica en keizersnedegeboorte op het darmmicrobioom gereduceerd, of zelfs geheel geëlimineerd.

Zodoende lijkt multi-species probioticatoediening veelbelovend. Echter, het is nog onbekend hoe te selecteren voor de ideale “microbiële cocktail” om hiermee een gebalanceerd darmmicrobioom te bereiken. Het is niet per definitie zo dat de meest aanwezige bacteriesoorten ook de meest belangrijke biologische functies uitvoeren. In **hoofdstuk 5** zagen wij dat de kernsoorten in de microbiële netwerken die geassocieerd konden worden met longgezondheid in het algemeen laag aanwezig waren. Dit past bij het “rare taxa” (of zeldzame soorten) principe, dat veronderstelt dat de hoeveelheid van een soort niet de beste maat is voor diens belang in het algehele microbiële netwerk. Het bestuderen van het menselijke microbioom in een meer overkoepelende context kan daarom meer inzicht geven in de rol van minder bekende microben en in hun vermogen de gezondheid te bevorderen.

List of publications

Publications in this thesis

Reyman M, Houten MA van, Baarle D van, Bosch AATM, Man WH, Chu MLJN, Arp K, Watson RL, Sanders EAM, Fuentes S, Bogaert D. Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. *Nature Communications* 2019;10:4997.

Reyman M, Houten MA van, Arp K, Sanders EAM, Bogaert D. Rectal swabs are a reliable proxy for faecal samples in infant gut microbiota research based on 16S-rRNA sequencing. *Scientific Reports* 2019;9:16072.

Other publications

Reyman M, Verrijn Stuart A, Summeren M van, Rakhshandehroo M, Nuboer R, Boer F de, Ham H-J van den, Kalkhoven E, Prakken B, Schipper H. Vitamin D deficiency in childhood obesity is associated with high levels of circulating inflammatory mediators and low insulin sensitivity. *Int J Obes* 2014:38;46-52.

Fieten K, Totté J, Levin E, **Reyman M**, Meijer Y, Knulst A, Schuren F, Pasmans S. Fecal microbiome and food allergy in pediatric atopic dermatitis: a cross-sectional pilot study. *Int Arch Allergy Immunol* 2018;175:77-84.

Dankwoord

Een promotietraject is als een avontuurlijke reis die bestaat uit onvergetelijke ervaringen, soms diepe dalen en duizelingwekkende hoogtepunten. Zo'n reis is onmogelijk te voltooien zonder bepaalde mensen aan je zijde. Anderen maken de reis gewoon net dat ene stukje leuker. Iedereen die mij heeft geholpen, gesteund en gemotiveerd tijdens mijn traject wil ik daarom enorm bedanken.

Prof. dr. Sanders, beste Lieke, jouw enthousiasmerende presentatie over vaccinatieonderzoek welke ik per toeval bijwoonde toen ik nog bij het Onderwijscentrum werkte, overtuigde mij om de stap te zetten de wetenschappelijke wereld in. Deze stap overwoog ik al, maar na jouw inspirerende boodschap wist ik zeker dat ik hiervoor bij jou moest zijn. Bedankt dat je aan mij het ZEBRA project hebt toevertrouwd. Ik blijf mij erover verbazen hoe een ogenschijnlijke toevalligheid zo'n onverwachte wending kan geven aan je leven. Mogelijk is het dan niet zo toevallig dat jij naast de kindergeneeskunde ook altijd al geïnteresseerd was in de dermatologie en mijn keuze hiervoor hartsgrondig hebt kunnen steunen.

Prof. dr. Bogaert, beste Debby, ik heb immens veel respect voor je onuitputtelijke wetenschappelijke ambitie. Jouw visie voor de studies van onze onderzoeksgroep(en), analytisch vermogen en kritische blik zijn onevenaarbaar. Ik wil je bedanken dat je mij deze uitdaging aan hebt laten gaan en voor alles wat ik van jou heb mogen leren.

Dr. van Houten, beste Marlies, het is bewonderenswaardig wat jij allemaal voor elkaar kunt krijgen met de vierentwintig uur die een dag telt. Ik hoop dat jij ooit aan mij je geheim zult verklappen. Je intensieve betrokkenheid bij het opzetten en onder de aandacht houden van het ZEBRA onderzoek hebben het tot een prachtige studie gemaakt. Dankjewel dat ik het stokje van jou heb mogen overnemen als antibioticameisje.

Leden van de leescommissie, prof. dr. Benders, prof. dr. Bont, prof. dr. Kuijper, prof. dr. van de Wijgert en prof. dr. Willems, hartelijk dank voor de bereidheid om mijn manuscript te lezen en te beoordelen.

Er was niks te analyseren, presenteren of beschrijven geweest zonder alle kinderen en hun ouders die hebben deelgenomen aan de MUIS en ZEBRA studies, dus aan hen wil ik heel veel dank uitspreken.

Ook zonder mijn collega's van het Wetenschapsbureau van de Spaarne Gasthuis Academie was er geen onderzoek en dus geen proefschrift geweest. Bedankt voor jullie grenzeloze hulp. Jullie hebben mij laten zien wat het betekent om grootschalige prospectieve studies tot een succes te maken. Jullie steken ziel en zaligheid in alle lopende onderzoeken en dat verdient alleen maar lof. Bovendien kennen jullie alle trucen uit de doos om deelnemers en zorgverleners trouw aan jullie te binden, wat super waardevol is. Lieve Greetje, Jacqueline, José, Marjon, Adinda, Mara, Judith, Eva, Corien, Marjolein, Jeannette, Carlinda, Karlijn, Catharina, Alex, Daphne, Anne, Richard en Peter, bedankt ook voor alle bemoedigende woorden en gezellige activiteiten. Joe joe!

Complimenten aan alle artsen en verpleegkundigen van de neonatologieafdelingen van de deelnemende ziekenhuizen, namelijk het Spaarne Gasthuis (voorheen Spaarneziekenhuis en Kennemer Gasthuis), Diakonessenhuis en Tergooiziekenhuis, voor hun inzet in het includeren van de deelnemers. Ik waardeer de tijd die jullie hiervoor hebben genomen naast jullie drukke, klinische werkzaamheden en dat jullie mijn eindeloze telefoontjes en bezoeken hebben getolereerd. In het bijzonder wil ik Franz Plötz, Wouter de Waal en Irene Schiering bedanken voor hun passie voor het uitvoeren van klinisch onderzoek.

De samenwerking met het Streeklaboratorium en TNO maakte ook een onontbeerlijk onderdeel uit van dit traject, waarvoor ik Jacob Bruin, Sjoerd Euser, Paul Badoux, Bjorn Herpers, Bart Keijser en Nynke van Berkum wil bedanken.

Alle nog niet genoemde mede-auteurs, waaronder Debbie van Baarle, Rob Willems en Willem van Schaik, wil ik bedanken voor de oprechte interesse en opbouwende feedback waarmee de stukken naar een hoger niveau zijn getild. De studenten die ik heb begeleid, Maaike en Sascha, wil ik bedanken voor hun energie en nieuwsgierigheid. Dear Susana, thank you for always being available to discuss research and life. Your empowering words have helped me more than you might think. I respect your work ethic that you combine with your joy for life. I am still looking forward to that celebratory dinner (and maybe some more dancing?). ¡Vamos a pasarlo bien!

Mijn onderzoeksgroep, bestaande uit Astrid, Maartje, Wouter, Mariëtte, Sabine, Saskia, Lilly, Anne, Krzysztof, Mei Ling, Raiza, Kayleigh en Jody, wil ik bij dezen graag uitroepen tot de meest empathische, behulpzame en liefste groep mensen ooit. De hoeveelheid werk die jullie verzetten in combinatie met de hoeveelheid taarten die jullie bakken is waanzinnig. Jullie voelden altijd feilloos aan wanneer ik meer

behoefte had aan extra ondersteuning op het gebied van (laboratorium)analyses of een knuffel. Ik hoop dat jullie altijd zo zorgzaam zullen blijven voor elkaar.

Aangezien ik tijdens mijn onderzoek heel wat kilometers heb gemaakt in het kader van woon-werk verkeer en huisbezoeken was het heerlijk om enkele dagen van de week na een korte fietstocht verwelkomd te worden op mijn onofficiële kamer op het RIVM. Lieve Pauline, Bette en Elsbeth, bedankt dat jullie altijd bereid waren om naar mij te luisteren, waar nodig te adviseren en bovenal lekker samen te eten. Daan, wat ben ik blij met jou als collega en vriendin. Hella, ik hoop dat we binnenkort weer samen een mooie stad kunnen gaan bezoeken (Rotterdam na de lockdown is wat mij betreft een prima begin).

Dear Edinburgh group, namely Becky, Melanie, Carlos, Filipa, Alicia and Chikondi, thank you for your enormous hospitality. Although you are all, without a doubt, amazing researchers, you might also want to consider a career as tour guides. I don't think I have ever seen/participated in/drunk/eaten so many local highlights in such a short time on a trip before.

Lieve, lieve Cavers… Hoe zou mijn traject er uit hebben gezien zonder jullie? Ik denk dat er significant minder successen waren geboekt, beduidend minder zou zijn gelachen en exponentieel meer zou zijn gezwolgen in wanhoop. Nina, Mirjam, Nynke en Oviedo – pareltjes zijn jullie, stuk voor stuk. Erik, je bent hilarisch en specialist in het opvrolijken. Roy, ik reken op nog meer muzikale en culinaire tips in de toekomst. Emma, hadden we met Ierland toch mooi een exotische (want tropische) congresbestemming te pakken hè? Wing Ho, bedankt dat ik je altijd, altijd, altijd, altijd, altijd alles kon vragen. Ik prijs mij gelukkig dat ik jullie heb mogen leren kennen en zo veel diepte- en hoogtepunten in de afgelopen zes (!) jaar met jullie heb mogen delen.

Gelukkig kon ik in die zes jaar naast onderzoek doen ook aan de slag bij de afdeling dermatologie van het Diakonessenhuis. Ik wil al mijn collega's daar bedanken voor de ontzettend leerzame tijd. Jullie hebben mij klaargestoomd voor mijn opleidingsplek in het Erasmus MC. Bedankt aan mijn nieuwe A(N)IOS-collega's voor het warme welkom daar; ik heb mega veel zin in mijn opleidingstijd met jullie.

Lieve Joan, Violet, Jolien en Karlijn, bedankt voor jullie vriendschap. Ik denk met plezier terug aan menig meezingmoment. Joan, mede door jou heb ik de keuze gemaakt om te gaan voor Rotterdam en daar ben ik je dankbaar voor. Als het toch tegen blijkt te vallen, kom ik er nog op terug.

Lief Groepje 4, oftewel Danko, Annemarie, Mascha, Ietje, Sarah, Rian, Rianne en Tom, bedankt voor de veelal hysterische feesten, festivals, weekenden weg en vakanties die ik met plezier heb aangegrepen als zeer geslaagde afleidingsmanoeuvres in tijden van promotiestress. Babette, bedankt voor je motiverende speeches en duwtjes in mijn rug.

Hier een oproep aan de Blarries, dat wil zeggen Margot, Sanne, Lieke, Bastiaan + aanhang, om snel weer een spelletjesweekend te organiseren! Spelletjes met jullie maken het leven een stuk leuker.

Maar het leven wordt pas echt leuk wanneer er wordt gedanst. Gelukkig kan ik voor dansjes altijd terecht bij mijn Gekke Chickies, de UDC en Dansjunkies. Een speciale shout-out voor Dhilani, Nadine, Esther, Malou, rode Miriam en blonde Myriam. Laten we alsjeblieft dansend (en drinkend) door het leven blijven gaan.

Aan mijn middelbare schoolvriendinnen die mij soms beter kennen dan ikzelf, waardoor er minimale context of uitleg nodig is, wil ik laten weten: jullie zijn heel speciaal voor mij. Seline, Rebecca, Sanne, Bernadette, Seline, Susanne en Gülsüm, hopelijk worden we samen heel oud (en gezond).

Lieke en Nina, mijn paranimfen, jullie aanmoedigingen tijdens deze reis waren onmisbaar. Om op toevalligheden terug te komen, ben ik eeuwig dankbaar voor het toeval dat ons bij elkaar heeft gebracht als onervaren ANIOS kindergeneeskunde in het Diakonessenhuis. Ik beschouw jullie als vriendinnen voor het leven.

Arie, wat heb jij standvastig in mij geloofd. Je rustige raadgeving kwam altijd op het juiste moment. Annemiek en Alice, bedankt dat jullie mijn familie in Nederland willen zijn en deze rol zo serieus nemen. Aguniu, dziękuję za Twoją bezwarunkową miłość.

Maria, big sister, thank you for always being there for me. Even though we don't exactly live next door to each other, I never experience you as being far away. Howard, thank you for being awesome. Liefste Marloes, je bent als een zus voor mij. Ik houd van jou.

Mamusiu i Tatusiu, wierzę, że ze wszystkich ludzi na świecie Wy jesteście ze mnie najbardziej dumni. Nie wiem jak Wam mogę podziękować za siłę którą mi to daje. Jednym z moich marzeń było napisać książkę. Oto ona! W podziękowaniu dedykuję ją Wam. Kocham Was.

Niels. Ik zal nooit genoeg kunnen benadrukken dat dit alles zonder jou niet was gelukt. Jouw naam had net zo goed op de kaft kunnen staan. Je bent de beste.

Curriculum vitae

Marta Reyman was born on March 5th, 1988, in Vlaardingen. Being the second daughter of native Polish parents, she was raised bilingually in The Netherlands. From age 7 to 10 she lived in the Czech Republic, where she attended the International School of Prague. She finished her secondary education at the Anna Van Rijn college in Nieuwegein in 2006. That same year, she started her medical training at the University Medical Center Utrecht. After obtaining her medical degree in 2012, she gained general clinical experience as a junior doctor in pediatrics at the Diakonessenhuis in Utrecht. Subsequently, she worked in the educational field as a policy officer at the department of Medicine of the University of Utrecht, where she contributed to the revision of the medical curriculum, supervised medical students and achieved her Basis Kwalificatie Onderwijs (BKO). At the end of 2014, she started the endeavor which ultimately resulted in the publication of this thesis, under the supervision of prof. dr. E.A.M. Sanders, prof. dr. D. Bogaert and dr. M.A. van Houten. In the meantime, she was happy to return to the Diakonessenhuis, now working as a junior doctor in dermatology. She has recently moved to Rotterdam, where she has been accepted as a resident at the dermatology department of the Erasmus Medical Center. She lives with her husband Niels, to whom she is enormously grateful to have had by her side throughout her career.