

Accepted Manuscript

Intestinal Microbiota in Infants at High-risk for Allergy: Effects of Prebiotics and Role in Eczema Development

Harm Wopereis, BSc., Kathleen Sim, MRCPCH, Ph.D, Alexander Shaw, Ph.D, John O. Warner, MD, FMedSci, Jan Knol, Ph.D, J. Simon Kroll, MA, FMedSci



PII: S0091-6749(17)31343-X

DOI: [10.1016/j.jaci.2017.05.054](https://doi.org/10.1016/j.jaci.2017.05.054)

Reference: YMAI 12971

To appear in: *Journal of Allergy and Clinical Immunology*

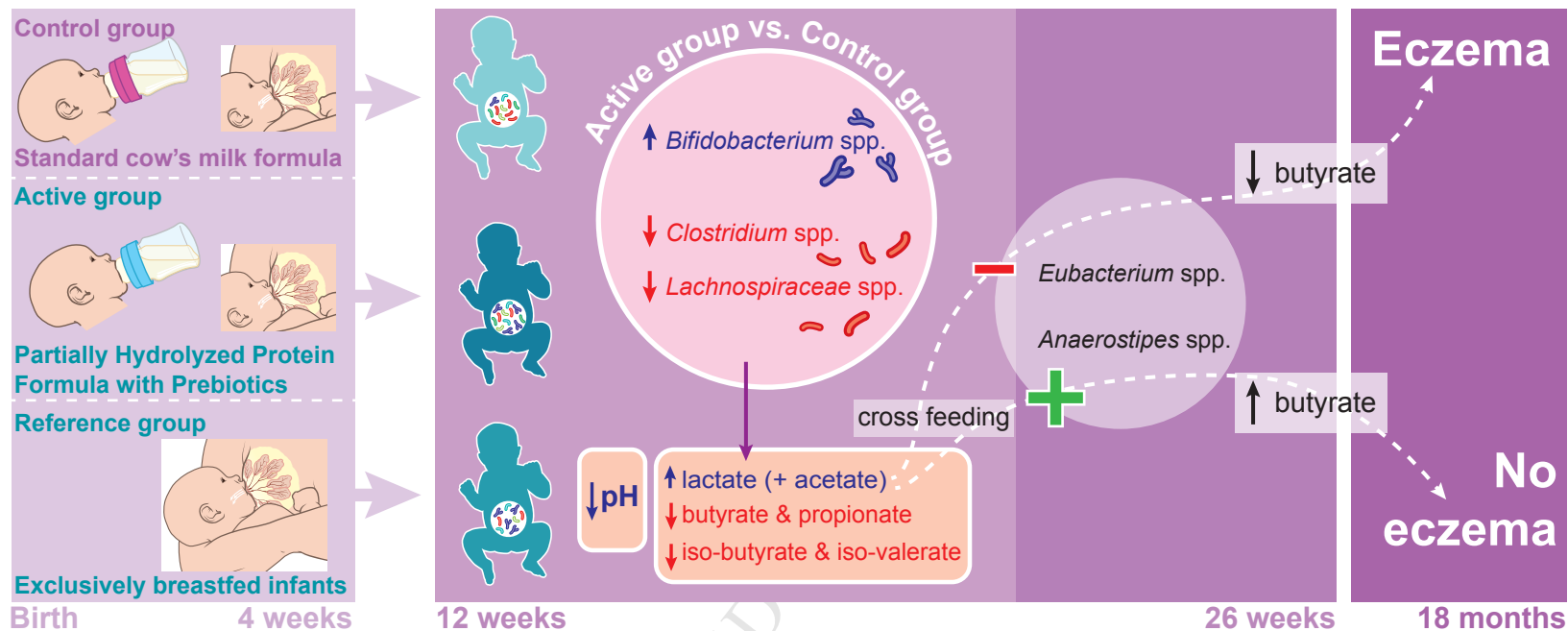
Received Date: 15 December 2016

Revised Date: 17 March 2017

Accepted Date: 4 May 2017

Please cite this article as: Wopereis H, Sim K, Shaw A, Warner JO, Knol J, Kroll JS, Intestinal Microbiota in Infants at High-risk for Allergy: Effects of Prebiotics and Role in Eczema Development, *Journal of Allergy and Clinical Immunology* (2017), doi: 10.1016/j.jaci.2017.05.054.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Intestinal Microbiota in Infants at High-risk for Allergy: Effects of Prebiotics and Role in Eczema Development.

Harm Wopereis^{1,2}, BSc.; Kathleen Sim³, MRCPCH, Ph.D; Alexander Shaw³, Ph.D; John O. Warner³, MD, FMedSci; Jan Knol^{1,2}, Ph.D; J. Simon Kroll³, MA, FMedSci

¹Danone Nutricia Research, Utrecht, the Netherlands

²Laboratory of Microbiology, Wageningen University, Wageningen, the Netherlands

³Department of Medicine, Section of Paediatrics, Imperial College London, London, UK

Corresponding author: Harm Wopereis

Danone Nutricia Research, Uppsalalaan 12, 3584 CT Utrecht, The Netherlands, Phone: +31 (0)30 2095000, e-mail: harm.wopereis@danone.com.

Wageningen University, Laboratory of Microbiology, Stippeneng 4, 6708 WE Wageningen, The Netherlands, e-mail: harm.wopereis@wur.nl

Conflict of interest: This study was funded by Danone Nutricia Research. HW and JK are employees of Danone Nutricia Research. JOW is a member of the global advisory board for Danone Nutricia Research and has received grants and honoraria for presentations at symposia sponsored by Danone Nutricia Research. Work in the Kroll laboratory was supported by funding from the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial Healthcare NHS Trust and Imperial College London.

Abstract

Background: The development of gut microbiota in infancy is important in the maturation of the immune system. Deviations in colonization patterns have been associated with allergic manifestations (e.g. eczema), but exact microbiome dysfunctions underlying allergies remain unclear. We studied the gut microbiota of 138 infants at increased risk of developing allergy, participating in a clinical trial investigating the effectiveness of a partially hydrolyzed protein formula supplemented with non-digestible oligosaccharides (pHF-OS) on the prevention of eczema.

Objective: The effects of the interventions and breastfeeding on fecal microbiota were investigated. Additionally, we aimed to identify microbial patterns associated with the onset of eczema.

Methods: Bacterial taxonomic compositions in the first 26 weeks of life were analyzed using 16S rRNA-gene sequencing. Additionally, fecal pH and microbial metabolites were measured.

Results: Fecal microbial composition, metabolites and pH of infants receiving pHF-OS was closer to breastfed infants than to infants receiving standard cow's milk formula. Infants developing eczema by 18 months showed temporal differences that were marked by decreased relative abundances of *Parabacteroides* and *Enterobacteriaceae* at 4 weeks, and decreased relative abundances of lactate-utilizing bacteria producing butyrate at 26 weeks, namely *Eubacterium* and *Anaerostipes* spp., supported by increased lactate and decreased butyrate levels.

Conclusions: We showed that a pHF with specific prebiotics modulated the gut microbiota closer to that of breastfed infants. Additionally, we identified a potential link between the microbial activity and onset of eczema, which may reflect a suboptimal implementation of gut microbiota at specific developmental stages in infants at high-risk for allergy.

Key messages

- A partially hydrolyzed protein infant formula supplemented with specific prebiotic oligosaccharides modulates the fecal microbiota of high-risk infants for allergy closer to that of breastfed infants, both in taxonomic composition as in metabolite composition.
- Infants developing eczema in the first 18 months showed discordant temporal development of bacterial taxa and metabolites in the first 26 weeks of life. The identified patterns may lead to the development of both preventive and therapeutic targets for eczema and other manifestations of allergy.

Capsule summary

This study investigated the factors influencing the assembly of the intestinal microbiome and reveals links between the functionality of key microbial taxa in the gut and eczema development in infants at high-risk of developing allergy.

Key words

allergy; eczema; gastrointestinal; microbiome; microbiota; short chain fatty acids; butyrate; lactate; prebiotics; infants.

Abbreviations

16S rRNA: 16S ribosomal ribonucleic acid; CCA: Canonical correspondence analysis; LUB: Lactate-utilizing and butyrate producing bacteria; lcFOS: Long chain fructo-oligosaccharides; MCPT: Monte Carlo permutation test; OTU: Operational Taxonomic Unit; pAOS: Pectin-derived oligosaccharides; PCR: Polymerase chain reaction; pHF-OS: Partially hydrolyzed protein infant formula containing oligosaccharides; PRC: Principal Response Curves; SCFA: Short chain fatty acids; scGOS: Short chain galacto-oligosaccharides; Treg: Regulatory T cells

Introduction

Gut microbial development in the first years of life occurs concomitantly to the development of our cognitive, metabolic and immune systems, and form an interactive signaling network (1). The gradual diversification towards a relatively stable adult-like composition is a dynamic process influenced by several environmental factors such as birth mode, gestational age at birth and early life nutrition. Alterations in the early colonization process, e.g. through antibiotic-use or C-section delivery, have been associated with development of allergic manifestations later in life, but the exact microbiome dysfunction underlying this disease remains unclear (2).

Eczema is typically the first allergic manifestation to appear, and its incidence is still increasing in many countries around the world. Considering the concurrent maturation of the immune system, the succession of species into a stable adult-type community, a process naturally guided by human milk, may be critically important. Breastfeeding is thought to protect against the development of allergy, via the presence of allergens and immune mediators in human milk that are absent from artificial milks (3), as well as via the presence of human milk oligosaccharides stimulating a gut microbiota which may favor tolerance induction (4).

Therefore, several studies have been performed with different types of infant formulas (IF) enriched with non-digestible oligosaccharides, known as prebiotics. Prebiotics typically reach the colon intact and selectively stimulate the growth and activity of specific beneficial members of the microbiota (5, 6). Interestingly, significant reductions in eczema risk were observed in intervention trials with IF supplemented with specific prebiotics (7-9).

The present study arises from a parent registered study (PATCH trial) investigating the effects of a partially hydrolyzed formula containing specific oligosaccharides (pHF-OS) on the prevention of eczema in infants at increased risk to develop allergy. The pHF-OS was shown to induce hypo-antigenic and immune-modulatory effects including increased regulatory T cell numbers (Treg), but did not reduce eczema incidence by 12 or 18 months, when compared with infants receiving standard cow's milk formula (10).

In this work, 16S ribosomal RNA (rRNA)-gene sequencing was applied to obtain an in-depth characterization of the microbiota composition of feces collected at 4 and 26 weeks of age in a subset of vaginally born infants, including breastfed infants (n=30) and infants randomized to receive pHF-OS (n=51) or standard cow's milk formula (control, n=57). In addition fecal pH, levels of lactate and short-chain fatty acids (SCFA) were determined in the stool specimens collected at 4, 12 and 26 weeks of age.

There were two primary aims of this study: (i) to investigate whether a pHF-OS could modulate the developing gut microbiota closer to that of breastfed infants, and (ii) to identify patterns in the developing gut microbiota that may be implicated in the onset of eczema. The findings in this study confirm the major influence of early life nutrition on the assembly of the gut microbiota, and provide new insights in how deviations in this assembly are associated with the development of eczema.

Materials and methods

Study design and fecal sample selection

This study arises from a double-blind, randomized, controlled parallel-group nutritional intervention trial in infants with a parental history of allergic disease, conducted in 10 specialist pediatric centers in Australia, Singapore, England and Ireland from April 2006 to March 2011 as described elsewhere (10). In total 1047 infants were recruited, and consisted of three groups of participants. The breastfed reference group consisted of exclusively breastfed infants for the first 18 weeks of life. Participants who chose to stop breastfeeding or to supplement with formula before 18 weeks of age, were randomized to receive either partially hydrolyzed whey protein-dominant infant formula containing short-chain galacto-oligosaccharides and long-chain fructo-oligosaccharides (scGOS/lcFOS) (9:1) (0.68 g/100 ml) and pectin-derived acidic oligosaccharides (pAOS) (0.12 g/100 ml), or standard cow's milk formula (control), up to 26 weeks of age. Information of trial ethics approvals, monitoring and regulatory compliance are summarized in the Online Repository (PATCH trial, study registration ISRCTN65195597 14th February 2006).

Fecal samples for microbial analysis were selected from infants that met the following criteria: infants (I) were randomized to investigational formula before 4 weeks of age, or were part of the breastfed reference group (n=942 infants), (II) vaginally-born (normal or instrumental delivery) (n=673 infants), (III) with stool specimens available at 4 and 26 weeks of age (n=324 infants). A subset of these infants was randomly selected to obtain 60 infants for each formula group and 30 breastfed reference infants (n=150 infants).

Sample collection and preparation

The infants' fecal samples were collected by the parents into 10 ml stool containers (Greiner Bio-One, Kremsmünster, Austria), immediately frozen (-12°C to -20°C) and transported within three months to the hospital. Upon arrival at the hospital and prior to evaluation at the laboratory, samples were kept and transported at -80°C . Frozen stool samples were defrosted on ice and stool pH was measured using a pH meter equipped with a glass-body pH electrode (Mettler-Toledo, Columbus, USA). Fecal samples were 10-fold diluted in PBS buffer (150 mM NaCl, 10 mM Na_2HPO_4 , 20 mM NaH_2PO_4 , pH 7.4), and 5-10 glass beads (3 mm in diameter) were added to homogenize the sample by vortexing for 3 minutes, followed by centrifugation ($300\times g$ for 1 minute). Several 1 ml portions of supernatant were stored at -80°C for downstream processing.

16S rRNA gene sequencing and bioinformatics

Fecal suspensions of 200 μl were mixed with 450 μl DNA extraction buffer (100 mM Tris-HCl, 40 mM EDTA, pH 9.0) and 50 μl of 10% sodium dodecyl sulfate. Phenol-chloroform extractions combined with beat-beating were performed as described by Matsuki, et al. (12) except that extracted DNA was re-suspended in 0.1 ml of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

The V3-V5 regions of the 16S rRNA gene were amplified using forward primer 357F, and a 'bifidobacteria-optimised' reverse primer 926Rb (13). 16S rDNA PCR was carried out in quadruplicate and pyrosequenced in three 454 GS FLX (Roche, Branford, CT, USA) runs as

previously described (13). Raw pyrosequencing data for all samples has been deposited in the Sequence Read Archive under accession number PRJEB19801.

The 'Quantitative Insights Into Microbial Ecology' (QIIME) v1.5.0 package was used to analyse sequence data (14) as previously described (13), except that alignments were carried out using the SILVA rRNA database (SSU_REF108) (15).

Fecal SCFA and lactic acid

Fecal suspensions were thawed on ice and centrifuged for 10 minutes at 14.000xg. Then, 350 µl supernatant was inactivated by heating for 10 minutes at 100°C, followed by centrifugation. A portion of the supernatant was used to quantitatively determine the SCFAs: acetic, propionic, n-butyric, isobutyric and n-valeric acid by gas chromatography, as described previously (16). Another portion of the supernatant was used to enzymatically analyze levels of lactate using a D-/L-lactic acid assay kit (Megazyme, Wicklow, Ireland).

Data handling and statistical analyses

Multivariate statistical analyses of 16S rRNA-gene sequencing data (taxonomic dataset) and levels of SCFAs and lactate (metabolite dataset) were performed by a combination of constrained ordination methods using the Canoco 5 software (17), and differential abundance testing using the R-package MetagenomeSeq (18).

Counts of Operational Taxonomic Units (OTUs) were aggregated at genus level (resulting in 142 genera), normalized by total sum scaling and 2-log-transformed. Genera present in less than 10 samples were discarded to remove sparse taxa, resulting in 58 features that were used as input for statistical analyses.

Monte Carlo Permutation tests (MCPT) were used to evaluate statistical significance ($P \leq 0.05$) of explanatory variables (constraints) in the ordination analyses performed.

Benjamini–Hochberg false-discovery rate was used to account for multiple comparisons (19), with significance for adjusted P-values (P_{adj}) at 0.05, except for taxonomic features associated with eczema development, for which significance was considered at 0.1.

Canonical correspondence analysis (CCA) was applied on the taxonomic dataset, as the unimodal model best fitted the relative abundances of bacterial genera (20). Forward selection was applied to identify sample covariates that best explained the variation in microbial taxonomic composition (21), and these were subsequently used in all multivariate models and comparisons performed, either as explanatory variables or as covariates. Temporal changes of bacterial metabolites were assessed using a linear ordination method, known as principal response curves (PRC) (22), as the linear model best fitted the bacterial metabolite dataset. Univariate data analyses were performed using GraphPad Prism version 6.02 for Windows (GraphPad Software, La Jolla, California, USA), applying Mann-Whitney test for two-group comparisons, and one-way ANOVA with Bonferroni's multiple comparisons test for three-group comparisons, with significance at 0.05.

Results

Effects of PHF-OS on microbial richness and diversity.

In total, 12 of the 150 selected infants did not have stool specimens available at 4 and 26 weeks. The characteristics of the remaining 138 participants are summarized in Table 1. Pyrosequencing was performed on stool specimens collected at 4 and 26 weeks of age, representing the specimens taken after randomization and at the end of the intervention period, respectively. The majority of randomized infants were already receiving formula before 4 weeks of age (Table 1). A total of 8 specimens collected at 4 weeks had insufficient amounts for preparation. The remaining 268 samples were successfully sequenced with a mean sequence depth of 6211 reads and a mean read length of 521 nucleotides (Supplementary Table S1).

Microbial richness and diversity of the different feeding groups was assessed at a sequence depth of 1636 reads. Richness was expressed as the number of unique OTUs (observed species) and diversity was assessed using a measure of phylogenetic diversity (PD whole tree metric). This measure takes both the number of unique OTUs as well as their

phylogenetic distance into account (23). No differences between the randomization groups were observed, but the breastfed reference group showed a significantly lower richness and diversity at 26 weeks when compared to all formula-fed infants combined (1-way ANOVA, $P < 0.01$ and $P < 0.05$ respectively (Figure 1).

pHF-OS modulates the microbial composition closer to breastfed infants.

The majority of OTUs were assigned to 4 phyla, namely Actinobacteria (65.0%), Firmicutes (24.8%), Proteobacteria (8.6%) and Bacteroidetes (1.5%) (Supplementary Table S2). The weighted UniFrac distance metric, which is based on phylogenetic distances and relative abundances of bacterial taxa in a pairwise comparison of samples (24), was used to assess the overall (dis)similarity in bacterial composition comparing pHF-OS with the control group and their respective distances to the breastfed reference group. Gut microbial compositions of infants receiving pHF-OS were found to be significantly more similar to breastfed infants than the control group to breastfed infants, which was most pronounced at the end of the intervention (1-way ANOVA, $P < 0.0001$) (Figure 2a).

In order to identify sample covariates that best explained the taxonomic composition of the fecal samples, canonical correspondence analysis (CCA) was combined with forward selection of variables (21). This analysis identified *time* (age in weeks), *ethnicity* (Asian, Caucasian, other), *feeding group* (control, pH-OS, breastfed) and *having siblings* (yes/no) as the factors significantly explaining the taxonomic variation ($P\text{-adj} < 0.05$, Supplementary Table S3).

MetagenomeSeq (18) was subsequently used to assess which differential abundant bacterial taxa accounted for the differences observed in weighted UniFrac at 26 weeks, while correcting for the covariates identified. Infants receiving pHF-OS were discriminated from control by increased levels of *Bifidobacterium* and decreased levels of *Clostridium* and an unassigned genus of *Lachnospiraceae* (Figure 2b-d).

pHF-OS modulates the microbial activity closer to breastfed infants.

Principal response curves (PRC) were used to assess temporal effects of the feeding groups on fecal compositions of SCFA and lactate in stool specimens collected at 4, 12 and 26 weeks, respectively. The interaction of *feeding group* with *time*, corrected for *ethnicity* and *having siblings*, was significant on the first constrained axis (explained variation = 80.3%, MCPT: $P=0.002$ with 499 permutations). Both the metabolite composition of breastfed infants and infants receiving pHF-OS were characterized by increased proportions of D- and L-lactate at 4, 12 and 26 weeks and decreased proportions of propionate, butyrate and branched-chain SCFA (iso-butyrate and iso-valerate) in contrast to the control group (Figure 3a). Moreover, infants receiving pHF-OS were found to have significantly decreased stool pH at all timepoints compared to infants receiving control formula with ranges similar as observed for breastfed infants (Figure 3b).

Aberrant temporal dynamics in infants developing eczema in the first 18 months of life.

In total, 52 of the 138 infants developed eczema in the first 18 months of life (Table 1). No differences in bacterial richness or diversity were observed at 4 or 26 weeks comparing eczematous to non-eczematous infants (Supplementary Figure S1).

PRC were used to investigate the fecal metabolite composition over *time* for infants developing eczema as compared to infants that did not, while correcting for the covariates identified (*ethnicity*, *feeding group* and *having siblings*). Significant temporal differences on the first constrained axis produced (explained variation = 84.4%, MCPT: $P=0.034$ with 499 permutations) were observed for the interaction (Figure 4a). Differential dynamics were most pronounced from 12 to 26 weeks of age. Infants developing eczema were characterized by decreased levels of both isomers of lactate and increased levels of propionate and butyrate at 12 weeks. This pattern was reversed at 26 weeks of age, with infants developing eczema showing increased levels of lactate and decreased levels of propionate and butyrate. No significant differences were observed in stool pH (Figure 4b).

Decreased microbial conversion of lactic acid into butyrate in infants developing eczema.

MetagenomeSeq was used to assess differential abundances of bacterial taxa over time in infants developing and not developing eczema, and to identify taxa being differential over time comparing the two groups, while correcting for the covariates identified (Figure 5a). Decreases of *Staphylococcus* and *Streptococcus* over time were observed in both healthy infants and infants developing eczema. Two genera of *Lachnospiraceae* (*Blautia* and an unassigned genus designated as "Other") and *Erysipelotrichaceae* were found to increase over time in both groups. The genus *Bifidobacterium* was found to increase over time in healthy infants and to decrease in infants developing eczema, however this pattern was not significantly different comparing the two groups ($P\text{-adj}=0.115$). Differential patterns over time were observed for two genera of *Enterobacteriaceae*, namely *Enterobacter* ($P\text{-adj}<0.001$) and an unassigned genus of *Enterobacteriaceae* ($P\text{-adj}<0.1$). These differences were mostly driven by increased levels of these genera in healthy infants at 4 weeks of age, and a subsequent decrease over time. This pattern was less pronounced in infants developing eczema (Supplementary Figure S2a and S2b). Differential development over time was also observed for *Parabacteroides* ($P\text{-adj}<0.0001$), with a decrease in healthy infants, and an increase in infants developing eczema (Supplementary Figure S2c). Furthermore, the increases over time observed for *Eubacterium* ($P\text{-adj}<0.05$) and *Anaerostipes* spp. ($P\text{-adj}<0.1$), were more pronounced in healthy infants as compared to infants developing eczema (Supplementary Figure S2d and S2e, respectively). Both genera are associated with a specialist group of microbes known to convert lactate together with acetate into mainly butyrate, hence referred to as lactate-utilizing and butyrate producing bacteria (LUB) (25, 26). Moreover at 26 weeks of age, decreased levels of both D- and L-lactate (Figure 5b and 5c) and increased levels of butyrate (Figure 5d) were observed, in healthy infants compared to infants developing eczema.

Discussion

In this study, we found that a partially hydrolyzed protein formula, supplemented with a specific oligosaccharide mixture (pHF-OS), modulates the developing gut microbiota of infants towards a pattern closer to that of breastfed infants, both in bacterial taxonomic composition as in metabolite composition. Infants receiving pHF-OS from 4 to 26 weeks of age, showed increased relative abundances of the genus *Bifidobacterium*, which was contrasted by decreases of *Clostridium* and an unassigned genus of *Lachnospiraceae*, when compared to infants receiving standard cow's milk formula. These modulations were reflected in marked differences in gut physiology, characterized by lower stool pH, increased proportions of lactate and decreased proportions of propionate, butyrate, iso-butyrate and iso-valerate.

Furthermore, we found that infants developing eczema in the first 18 months of life showed aberrant gut microbiota development in the first 26 weeks of life with significant temporal differences of the genus *Parabacteroides* and two genera of *Enterobacteriaceae*. These were found to decrease over time in healthy infants, a pattern which was reversed or less pronounced in infants developing eczema. Additionally, infants developing eczema showed decreased relative abundances over time of lactate-utilizing bacteria known to produce butyrate (LUB), namely *Eubacterium* and *Anaerostipes* spp., supported by significantly increased fecal concentrations of lactate, and decreased concentrations of butyrate at 26 weeks of age, a pattern which was independent of *feeding group*, *ethnicity* or *having siblings*.

In order to assess modulatory effects of nutrition and the role of specific microbes in developing eczema, it is of critical importance to consider the dynamic nature of the infant gut microbiota, and the environmental factors that influence its assembly (27). Hence, we excluded infants born by caesarean section to eliminate its confounding effects (28). In addition, multivariate comparisons were controlled for sample covariates that best explained the fecal compositional variation, which were - next to *time* (age in weeks) and *feeding group* - *ethnicity* and *having siblings*.

Other studies investigating the gut microbiota in relation to developing allergy, typically reported differences at specific points in time, especially in the first weeks of life, with most

commonly reduced diversity and richness associated, but often no specific microbes being identified (2). We did not observe differences in diversity or richness, but rather identified differences in the colonization patterns of infants developing eczema compared to those that did not.

Some of these patterns were driven by early differences at 4 weeks of age, like the increased levels of two genera within the family of *Enterobacteriaceae* in subsequent healthy infants, which was followed by sharp decreases over time. Facultative anaerobic bacteria, like the *Enterobacteriaceae*, are typically high in the early colonization process before replacement with anaerobic bacteria. This pattern being less pronounced in infants developing eczema may indicate a reduced immune-stimulus from species within this family. These include several potential pathogens, which may be needed for adequate development of the adaptive immune system. A recent study by West, et al. (2015) indeed associated low abundance of *Enterobacteriaceae* at 1 month, with an exaggerated immune response at 6 months. This deviating pattern was confirmed in infants subsequently developing allergy (29-31).

It is known, and confirmed in this study, that breastfed infants compared to formula fed infants have a lower stool pH, increased amounts of colonic acetate and lactate, and decreased amounts of propionate and butyrate. This reflects the dominance of *Bifidobacterium* species that produce acetate and L-lactate (11). This specific colonic environment, also observed for infants receiving pHF-OS, is known to form an effective chemical barrier against potential pathogens (32, 33), and is thought to contribute to the protective effects of breastfeeding against infections (34, 35).

We hypothesize that this typical infant-type colonic environment may also be crucial for the establishment of LUB, like *Eubacterium* and *Anaerostipes* spp, which were found to be poorly established at 26 weeks of age in infants developing eczema. Indeed these infants showed first decreased levels of lactate at 12 weeks, followed by increased levels of lactate and decreased levels of butyrate at 26 weeks of age. This may indicate the importance of LUB in guiding the transition around the weaning-period (between 4-6 months of age) from a

lactate and acetate-rich environment, towards a more adult-like butyrogenic milieu (2, 36). Moreover, the consumption of lactate may contribute to the development of a stable and healthy microbial ecosystem (37). Indeed lactate, although commonly detected in infant feces, is undetectable in healthy adults (38), while accumulation of colonic lactate in adults has been associated with ulcerative colitis (39, 40). The weaning-period also exposes the infant to an increasing level of antigens from the diet, which requires the immune system to adequately respond to these harmless substances. A study investigating the correlation of specific microbial signatures with the severity of eczema in 6-month old infants found an inverse correlation with the levels of butyrate-producing bacteria, and suggested a role in the observed alleviation of symptoms (41). The role of SCFA on the host, especially butyrate, has received increased interest due to accumulating evidence from murine studies showing their regulatory effects on host immunity, including anti-inflammatory and anti-allergic effects (42-45). The establishment of butyrate-producing bacteria around 26 weeks of age, possibly specifically those cross-feeding on lactate and acetate, may therefore prove important for establishing and maintaining homeostasis with our immune system during this critical stage of development. In conclusion, this study confirms the impact of early life nutrition on the establishment of the infant gut microbiota. Moreover, it indicates a potential link between the activity of the microbiota and the expression of eczema in early life. It emphasizes the importance of the microbial succession of species and metabolite cross-feeding, to develop a gut physiology that supports gut development, and also supports development of normal immune responses towards environmental triggers. These observations could aid the development of optimal nutritional strategies to support the timely gut colonization of keystone species in the gradually diversifying infant gut.

Acknowledgement

We want to thank all infants and their families who took part in the study and all the PATCH investigators for their collaboration in this study.

References

1. Shanahan F. The gut microbiota - a clinical perspective on lessons learned. *Nat Rev Gastroenterol Hepatol.* 2012;9(10):609-14.
2. Wopereis H, Oozeer R, Knipping K, Belzer C, Knol J. The first thousand days - intestinal microbiology of early life: establishing a symbiosis. *Pediatr Allergy Immunol.* 2014;25(5):428-38.
3. Verhasselt V. Oral tolerance in neonates: from basics to potential prevention of allergic disease. *Mucosal Immunol.* 2010;3(4):326-33.
4. Oozeer R, van Limpt K, Ludwig T, Ben Amor K, Martin R, Wind RD, et al. Intestinal microbiology in early life: specific prebiotics can have similar functionalities as human-milk oligosaccharides. *American journal of Nutrition.* 2013;98(2):561S-71S.
5. Gibson GR, Roberfroid MB. Dietary Modulation of the Human Colonic Microbiota: Introducing the Concept of Prebiotics. *J Nutr.* 1995;125(6):1401-12.
6. Bindels LB, Delzenne NM, Cani PD, Walter J. Towards a more comprehensive concept for prebiotics. *Nat Rev Gastroenterol Hepatol.* 2015;12(5):303-10.
7. Moro G, Arslanoglu S, Stahl B, Jelinek J, Wahn U, Boehm G. A mixture of prebiotic oligosaccharides reduces the incidence of atopic dermatitis during the first six months of age. *Archives of Disease in Childhood.* 2006;91(10):814-9.
8. Arslanoglu S, Moro GE, Boehm G, Wiens F, Stahl B, Bertino E. Early prebiotic oligosaccharide supplementation reduces the incidence of some allergic manifestations in the first 5 years of life. *Journal of Biology Regulation & Homeostatic Agents.* 2012;26(3):49-59.
9. Grüber C, van Stuijvenberg M, Mosca F, Moro G, Chirico G, Braegger CP, et al. Reduced occurrence of early atopic dermatitis because of immunoactive prebiotics among low-atopy-risk infants. *Journal of Allergy and Clinical Immunology.* 2010;126(4):791-7.
10. Boyle RJ, Tang MLK, Chiang WC, Chua MC, Ismail I, Nauta A, et al. Prebiotic-supplemented partially hydrolysed cow's milk formula for the prevention of eczema in high risk infants: a randomised controlled trial. *Allergy.* 2016:in press.

11. Knol J, Scholtens P, Kafka C, Steenbakkers J, Gro S, Helm K, et al. Colon microflora in infants fed formula with galacto- and fructo-oligosaccharides: more like breast-fed infants. *J Pediatr Gastroenterol Nutr.* 2005;40(1):36-42.
12. Matsuki T, Watanabe K, Fujimoto J, Kado Y, Takada T, Matsumoto K, et al. Quantitative PCR with 16S rRNA-gene-targeted species-specific primers for analysis of human intestinal bifidobacteria. *Appl Environ Microbiol.* 2004;70(1):167-73.
13. Sim K, Cox MJ, Wopereis H, Martin R, Knol J, Li M-S, et al. Improved Detection of Bifidobacteria with Optimised 16S rRNA-Gene Based Pyrosequencing. *PLoS One.* 2012;7(3):e32543.
14. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods.* 2010;7(5):335-6.
15. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 2007;35(21):7188-96.
16. Bakker-Zierikzee AM, Alles MS, Knol J, Kok FJ, Tolboom JJ, Bindels JG. Effects of infant formula containing a mixture of galacto- and fructo-oligosaccharides or viable *Bifidobacterium animalis* on the intestinal microflora during the first 4 months of life. *Br J Nutr.* 2005;94(5):783-90.
17. Šmilauer P, Lepš. J. Multivariate Analysis of Ecological Data using CANOCO 5. 2nd ed. ed. Cambridge Books Online: Cambridge University Press; 2014.
18. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Meth.* 2013;10(12):1200-2.
19. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological).* 1995;57(1):289-300.
20. Ter Braak CJ. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology.* 1986;67(5):1167-79.

21. Legendre P, Legendre L. Numerical ecology. Amsterdam: Elsevier; 2012.
22. van den Brink P, den Besten P, bij de Vaate A, ter Braak CF. Principal response curves technique for the analysis of multivariate biomonitoring time series. *Environ Monit Assess.* 2009;152(1-4):271-81.
23. Faith DP. Conservation evaluation and phylogenetic diversity. *Biological Conservation.* 1992;61(1):1-10.
24. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol.* 2005;71(12):8228 - 35.
25. Sato T, Matsumoto K, Okumura T, Yokoi W, Naito E, Yoshida Y, et al. Isolation of lactate-utilizing butyrate-producing bacteria from human feces and in vivo administration of *Anaerostipes caccae* strain L2 and galacto-oligosaccharides in a rat model. *FEMS Microbiology Ecology.* 2008;66(3):528-36.
26. Kanauchi O, Fujiyama Y, Mitsuyama K, Araki Y, Ishii T, Nakamura T, et al. Increased growth of *Bifidobacterium* and *Eubacterium* by germinated barley foodstuff, accompanied by enhanced butyrate production in healthy volunteers. *International journal of molecular medicine.* 1999;3(2):175-9.
27. Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, et al. Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe.* 2015;17(5):690-703.
28. Martin R, Makino H, Cetinyurek Yavuz A, Ben-Amor K, Roelofs M, Ishikawa E, 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(6):e0158498.
29. West CE, Ryden P, Lundin D, Engstrand L, Tulic MK, Prescott SL. Gut microbiome and innate immune response patterns in IgE-associated eczema. *Clin Exp Allergy.* 2015;45(9):1419-29.
30. Prescott SL, Noakes P, Chow BW, Breckler L, Thornton CA, Hollams EM, et al. Presymptomatic differences in Toll-like receptor function in infants who have allergy. *J Allergy Clin Immunol.* 2008;122(2):391-9, 9.e1-5.

- 469 31. Tulic MK, Hodder M, Forsberg A, McCarthy S, Richman T, D'Vaz N, et al. Differences
470 in innate immune function between allergic and nonallergic children: new insights into
471 immune ontogeny. *J Allergy Clin Immunol*. 2011;127(2):470-8.e1.
- 472 32. Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, et al. Bifidobacteria
473 can protect from enteropathogenic infection through production of acetate. *Nature*.
474 2011;469(7331):543-7.
- 475 33. Knol J, Boehm G, Lidestri M, Negretti F, Jelinek J, Agosti M, et al. Increase of faecal
476 bifidobacteria due to dietary oligosaccharides induces a reduction of clinically relevant
477 pathogen germs in the faeces of formula-fed preterm infants. *Acta Paediatrica*.
478 2005;94(s449):31-3.
- 479 34. Li R, Dee D, Li C-M, Hoffman HJ, Grummer-Strawn LM. Breastfeeding and Risk of
480 Infections at 6 Years. *Pediatrics*. 2014;134(Supplement 1):S13-S20.
- 481 35. Ip S, Chung M, Raman G, Chew P, Magula N, DeVine D, et al. Breastfeeding and
482 maternal and infant health outcomes in developed countries. Evidence report/technology
483 assessment. 2007(153):1-186.
- 484 36. Duncan SH, Louis P, Flint HJ. Lactate-Utilizing Bacteria, Isolated from Human Feces,
485 That Produce Butyrate as a Major Fermentation Product. *Appl Environ Microbiol*.
486 2004;70(10):5810-7.
- 487 37. Flint HJ, Duncan SH, Scott KP, Louis P. Links between diet, gut microbiota
488 composition and gut metabolism. *Proceedings of the Nutrition Society*. 2015;74(1):13-22.
- 489 38. Belenguer A, Holtrop G, Duncan SH, Anderson SE, Calder AG, Flint HJ, et al. Rates
490 of production and utilization of lactate by microbial communities from the human colon. 2011
491 2011-07-01 00:00:00. 107-19 p.
- 492 39. Vernia P, Caprilli R, Latella G, Barbetti F, Magliocca FM, Cittadini M. Fecal lactate
493 and ulcerative colitis. *Gastroenterology*. 1988;95(6):1564-8.
- 494 40. Hove H, Nordgaard-Andersen I, Mortensen PB. Faecal DL-lactate concentration in
495 100 gastrointestinal patients. *Scand J Gastroenterol*. 1994;29(3):255-9.

41. Nylund L, Nermes M, Isolauri E, Salminen S, de Vos WM, Satokari R. Severity of atopic disease inversely correlates with intestinal microbiota diversity and butyrate-producing bacteria. *Allergy*. 2015;70(2):241-4.
42. Steinmeyer S, Lee K, Jayaraman A, Alaniz RC. Microbiota Metabolite Regulation of Host Immune Homeostasis: A Mechanistic Missing Link. *Curr Allergy Asthma Rep*. 2015;15(5):1-10.
43. Stefka AT, Feehley T, Tripathi P, Qiu J, McCoy K, Mazmanian SK, et al. Commensal bacteria protect against food allergen sensitization. *Proc Natl Acad Sci U S A*. 2014;111(36):13145-50.
44. Smith PM, Howitt MR, Panikov N, Michaud M, Gallini CA, Bohlooly YM, et al. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science*. 2013;341(6145):569-73.
45. Arpaia N, Campbell C, Fan X, Dikly S, van der Veeken J, deRoos P, et al. Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature*. 2013;504(7480):451-5.

Table 1: Characteristics of the study population. Continuous data are presented as mean (standard deviation). N/A = Not Applicable.

	Control n=57	pHF-OS n=51	Breastfed n=30
Started formula before 4w	54 (95%)	48 (84%)	N/A
Male Sex	23 (40%)	30 (59%)	18 (60%)
Birthweight (g)	3317 (502)	3461 (442)	3606 (393)
Both parents allergic	11 (19%)	6 (12%)	11 (37%)
Only mother allergic	33 (58%)	29 (57%)	11 (37%)
Only father allergic	13 (23%)	16 (31%)	8 (27%)
Pet at home	11 (19%)	22 (43%)	7 (23%)
At least one sibling	28 (49%)	27 (53%)	22 (73%)
Vaginal delivery	46 (81%)	41 (80%)	25 (83%)
Instrumental delivery	11 (19%)	10 (20%)	5 (17%)
Caucasian ethnicity	16 (28%)	28 (55%)	25 (83%)
Asian ethnicity	35 (61%)	21 (41%)	2 (7%)
Other ethnicity	6 (12%)	2 (4%)	3 (10%)
Australia	13 (23%)	14 (27%)	13 (43%)
Ireland	4 (7%)	10 (20%)	5 (17%)
Singapore	35 (61%)	21 (41%)	2 (7%)
UK	5 (9%)	6 (12%)	10 (33%)
Eczema by 12 months	15 (26%)	17 (33%)	12 (40%)
Eczema by 18 months	18 (32%)	20 (39%)	14 (47%)
Antibiotics by 26 weeks	6 (11%)	12 (24%)	3 (10%)

Figure 1: Barplots (mean with 95% confidence interval) summarizing microbial intestinal richness **(a)** and phylogenetic diversity **(b)** at 4 and 26 weeks of age in breastfed infants (B) and infants receiving pHF-OS (A) or control formula (C). Comparisons were done at a

sequencing depth of 1636 reads per sample. Statistics were performed by a one-way ANOVA with Bonferroni's multiple comparison correction (* = $P\text{-adj} < 0.05$, ** = $P\text{-adj} < 0.01$).

Figure 2: (a) Barplots (mean with 95% confidence interval) summarizing weighted UniFrac distances at 4 and 26 weeks in infants receiving pHF-OS or control formula relative to breastfed infants (A-B vs. C-B, respectively). Distances are measured on a scale from 0-1, with 0 meaning 100% identical and 1 meaning 0% identical gut microbial composition. Statistical comparison of weighted UniFrac was performed by a one-way ANOVA with Bonferroni's multiple comparison correction (**** = $P\text{-adj} < 0.0001$). **(b-c)** Boxplots summarizing 2-log-transformed levels of differential abundant genera, present in >15% of the samples and a mean relative abundance >0.1%, when comparing pHF-OS with control formula at 26 weeks, namely: **(b)** *Bifidobacterium*, **(c)** *Clostridium* and **(d)** an unassigned genus of *Lachnospiraceae*. Differential abundances were computed with MetagenomeSeq with correction for effects of *ethnicity* and *having siblings* and adjustment of significance values for multiple comparisons using Benjamini–Hochberg false discovery rate (** = $P\text{-adj} < 0.01$, **** = $P\text{-adj} < 0.0001$).

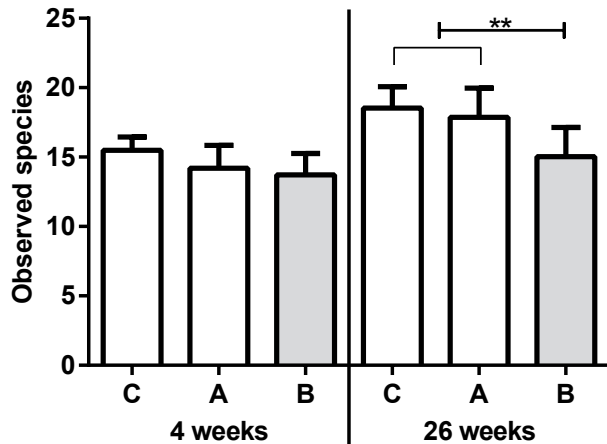
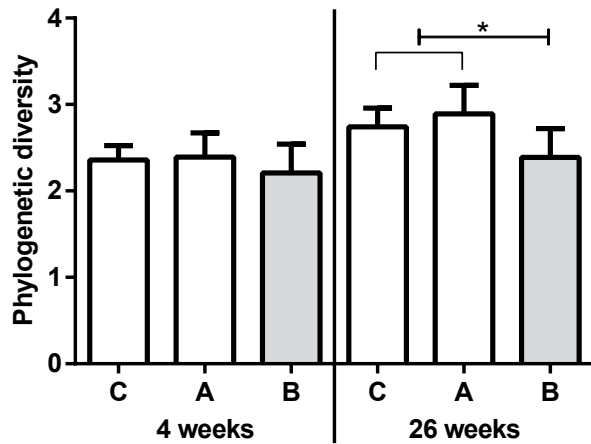
Figure 3: (a) Principal Response Curves showing changes in bacterial metabolites (levels of SCFA and lactic acids) across time and its interaction with the different feeding groups. The horizontal axis represents time and the vertical axis the PRC score values. Fecal metabolite composition of infants receiving control formula (C) was used as reference level and has zero PRC values and so its curve lays over the horizontal axis. Changes for infants being breastfed (B) or receiving pHF-OS (A) are shown as response curves relative to this reference. Metabolite response scores are shown on the separate vertical (one-dimensional) plot. The multiple of the PRC score with the response score provides a quantitative interpretation as well as the direction of the change at the respective timepoints (4, 12 and 26 weeks). Significance of the interaction, corrected for covariates, was tested using MCPT ($P = 0.002$, 499 permutations). **(b)** Boxplots summarizing stool pH at 4, 12 and 26 weeks for

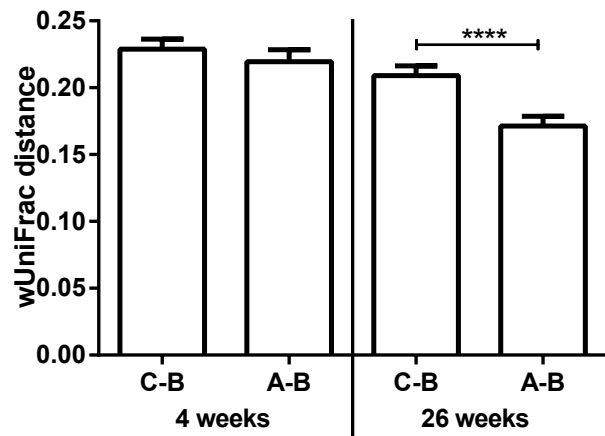
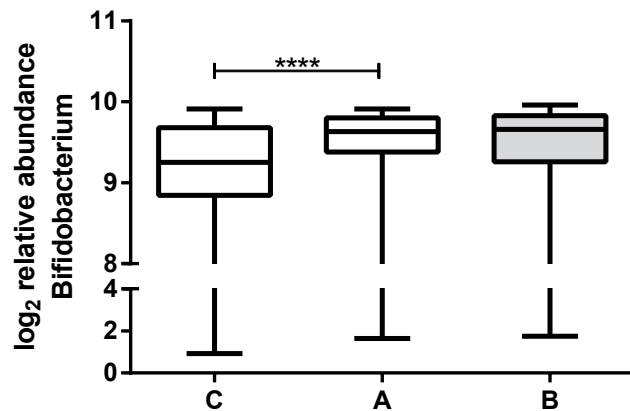
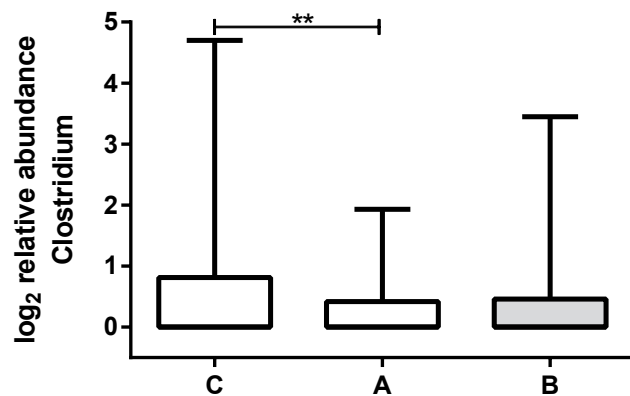
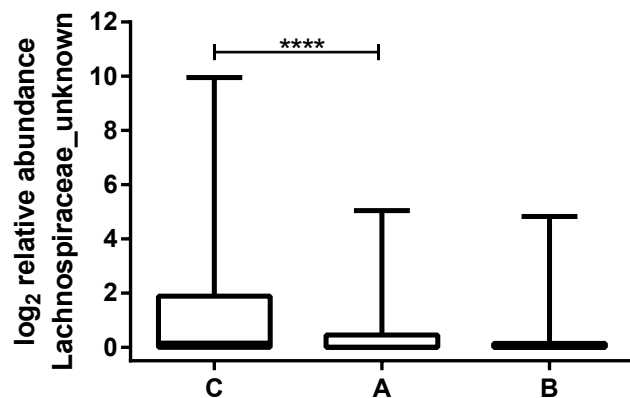
the different feeding groups. Statistical comparison was performed by a one-way ANOVA with Bonferroni's multiple comparison correction comparing pHF-OS to control formula (*= $P\text{-adj} < 0.05$, **= $P\text{-adj} < 0.01$, ***= $P\text{-adj} < 0.001$).

Figure 4: (a) Principal Response Curves showing changes in bacterial metabolites (levels of SCFA and lactic acids) across time and its interaction with developing eczema. The horizontal axis represents time and the vertical axis the PRC score values. The fecal metabolite composition of infants not developing eczema (NOECZ) were used as reference level and has zero PRC values. The change for infants developing eczema in the first 18 months of life (ECZ) is shown as a response curve relative to this reference. Metabolite response scores are shown on the separate vertical (one-dimensional) plot. Significance of the interaction, corrected for the effects of *feeding group*, *ethnicity* and *having siblings*, was tested using MCPT ($P=0.034$, 499 permutations). **(b)** Boxplots summarizing stool pH at 4, 12 and 26 weeks in infants developing eczema vs. healthy infants. Statistical comparison was performed by a one-way ANOVA with Bonferroni's multiple comparison correction (not significantly different).

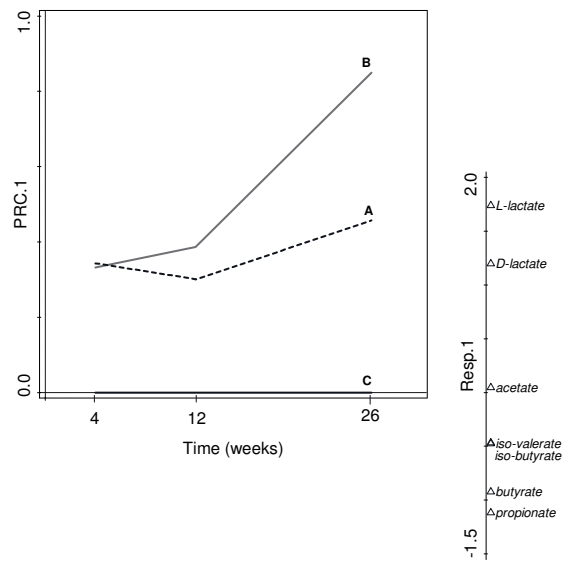
Figure 5: (a) Plot with bacterial taxa that change over time (dt = change from 4 to 26 weeks of age) in healthy infants and in infants developing eczema, as well as taxa that are differential over time in infants developing eczema compared to healthy infants. Data are shown as fold changes with standard deviations of log-transformed data. Taxa are detailed at phylum (B=Bacteroidetes, F=Firmicutes, P=Proteobacteria), family and genus level, respectively. The contrasts were computed by MetagenomeSeq and corrected for the effects of *feeding group*, *ethnicity* and *having siblings*. Adjustment of significance values for multiple comparisons was done using Benjamini–Hochberg false discovery rate. Differential taxa with an adjusted P -value ($P\text{-adj}$) < 0.05 for all contrasts, if present in $>15\%$ of the samples and a mean abundance $>0.1\%$, are shown. Taxa plotted in dark grey background showed no significant (ns) temporal differences when comparing healthy infants with eczematous

574 infants. Taxa plotted in light grey and white background were different with $P\text{-adj}<0.1$ and $P\text{-}$
575 $\text{adj}<0.05$, respectively. **(b)** Fecal concentration of D-lactic acid (D-LA) at 26 weeks in healthy
576 infants (NOECZ) compared to infants developing eczema (ECZ). **(c)** Fecal concentration of
577 L-lactic acid (L-LA) at 26 weeks in healthy versus eczema. **(d)** Fecal concentration of
578 butyrate (BA) at 26 weeks in healthy versus eczema. Significance was tested with a Mann-
579 Whitney two-group comparison (* = $P\text{-value}<0.05$, ** = $P\text{-value}<0.001$).

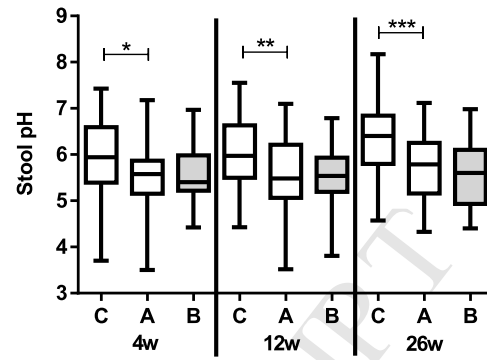
a**b**

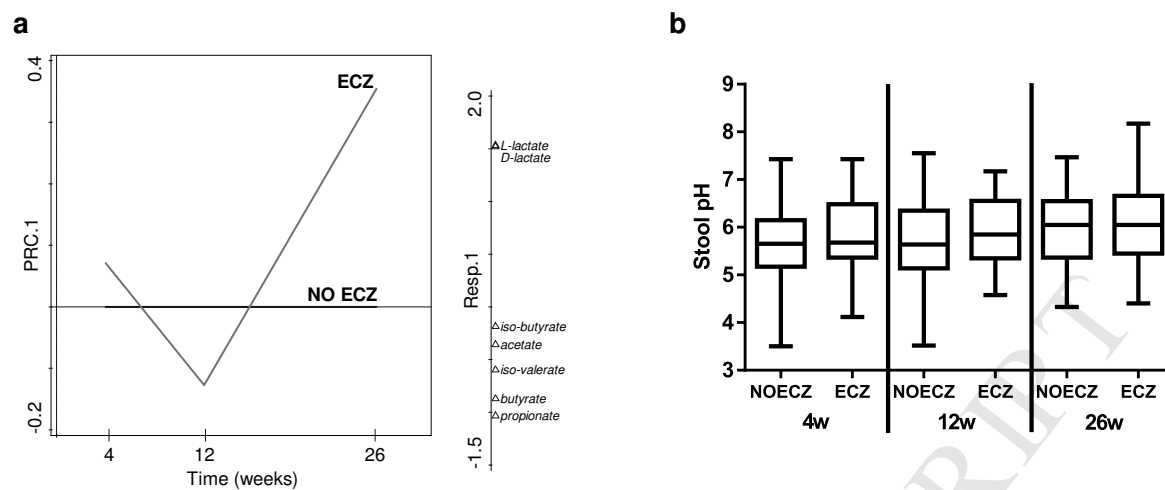
a**b****c****d**

a

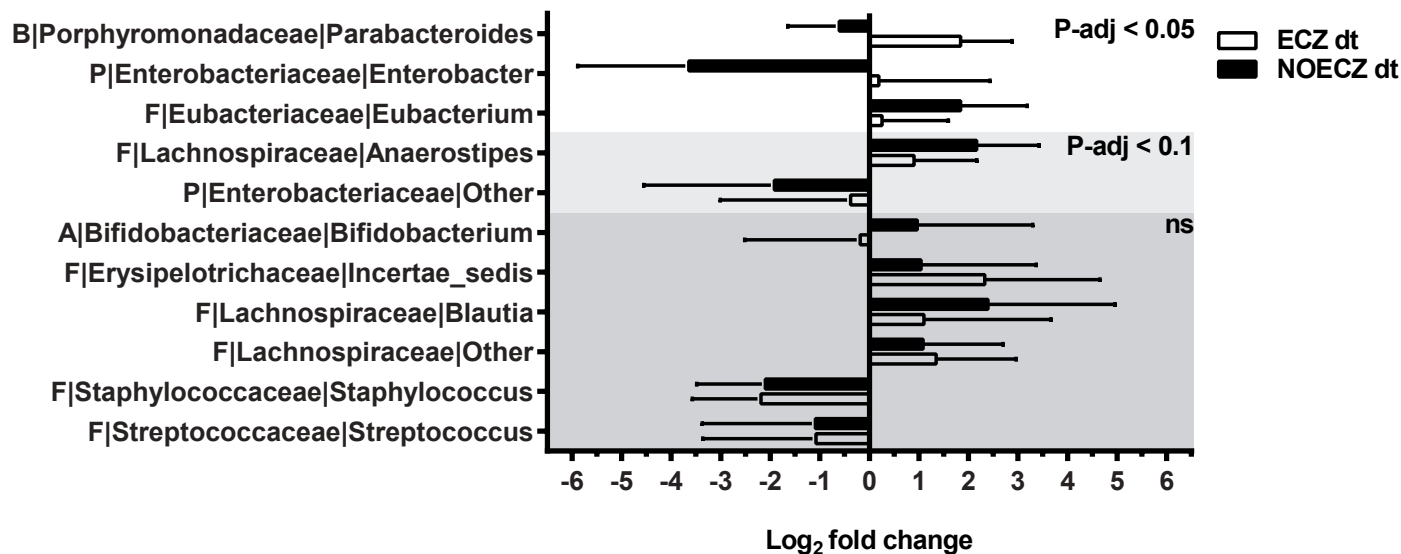


b

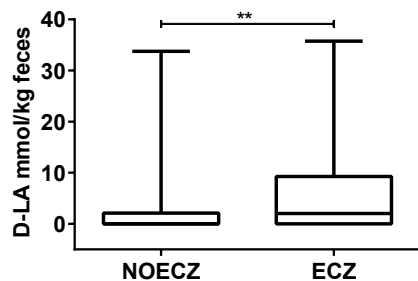




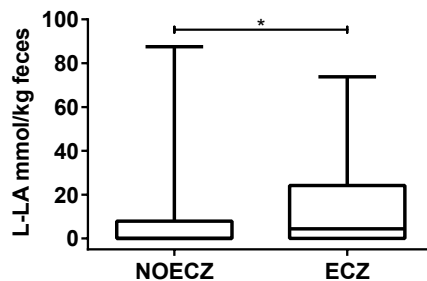
a



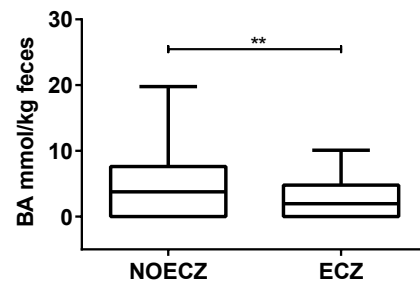
b



c



d



Supplementary figures and tables.**Table S1:** Sample size and sequence details for sequences passing quality filters

summarized per feeding group. Avg = Average, seq. = sequence, SD = standard deviation.

Feeding Group	Age	Count of subjects	Sum of reads	Avg seq. depth	SD of seq. depth
CONTROL	4w	55	326 169	5930	3300
	26w	57	345 935	6069	2959
pHF-OS	4w	49	305 530	6235	4459
	26w	51	299 308	5869	4113
BREASTFED	4w	26	166 838	6417	3950
	26w	30	220 892	7363	4014
Total		268	1 664 672	6211	3767

Table S2: Summary of gut microbial compositions of breastfed infants (B), pHF-OS (A) and control (C) at 4 and 26 weeks of age. Relative abundances and (standard deviations) are given in percentages for taxa present in >15% of the samples and average relative abundances >0.1% at family level, else designated “Other”. Phylum names are indicated as; A|. = Actinobacteria; B|. = Bacteroidetes; F|. = Firmicutes; P|. = Proteobacteria; V|. = Verrucomicrobia.

Taxa	4 weeks			26 weeks			Grand Total
Phylum Order Family	C	A	B	C	A	B	
Total Actinobacteria	55.28 (32.66)%	61.83 (31.56)%	69.66 (35.33)%	63.01 (28.43)%	73.25 (27.08)%	73.12 (28.76)%	64.92 (31.21)%
A Bifidobacteriales Bifidobacteriaceae	51.39 (32.25)%	60.18 (31.18)%	68.95 (35.13)%	59.99 (27.59)%	71.02 (26.73)%	72.38 (28.52)%	62.6 (30.98)%
A Coriobacteriales Coriobacteriaceae	3.56 (9.53)%	1.19 (3.14)%	0.34 (1.01)%	2.99 (6.92)%	2.2 (4.13)%	0.65 (1.63)%	2.12 (6.01)%
A Propionibacteriales Propionibacteriaceae	0.28 (0.86)%	0.4 (1.24)%	0.26 (1.01)%	0.01 (0.04)%	0.01 (0.05)%	0 (0)%	0.15 (0.74)%
A Other	0.06 (0.09)%	0.06 (0.27)%	0.11 (0.23)%	0.02 (0.04)%	0.02 (0.05)%	0.08 (0.21)%	0.05 (0.16)%
Total Bacteroidetes	1.88 (8.01)%	1.15 (3.65)%	3.06 (5.45)%	0.33 (1.42)%	1.53 (4.49)%	1.94 (5.82)%	1.49 (5.26)%
B Bacteroidales Bacteroidaceae	1.77 (7.9)%	1 (3.17)%	2.45 (5.21)%	0.31 (1.39)%	1.17 (3.48)%	1.44 (3.5)%	1.25 (4.71)%
B Bacteroidales Porphyromonadaceae	0.10 (0.63)%	0.14 (0.83)%	0.22 (0.73)%	0.01 (0.04)%	0.31 (1.58)%	0.48 (2.45)%	0.18 (1.2)%
B Other	0.01 (0.02)%	0.01 (0.03)%	0.39 (1.9)%	0.01 (0.04)%	0.05 (0.24)%	0.02 (0.04)%	0.06 (0.63)%
Total Firmicutes	31.09 (24.75)%	23.34 (21.26)%	8.63 (8.59)%	32.04 (24.86)%	21.85 (22.06)%	21.16 (25.58)%	24.81 (23.72)%
F Bacillales Staphylococcaceae	0.52 (1.44)%	0.44 (1.17)%	0.72 (1.57)%	0.03 (0.06)%	0.03 (0.14)%	0.02 (0.06)%	0.27 (1.00)%
F Lactobacillales Enterococcaceae	8.69 (13.3)%	1.19 (2.53)%	0.6 (2.02)%	4.84 (8.26)%	4.11 (5.92)%	2.38 (4.33)%	4.18 (8.36)%
F Lactobacillales Lactobacillaceae	2.54 (7.67)%	2.82 (5.67)%	0.78 (1.69)%	2.01 (5.27)%	1.05 (2.91)%	1.31 (2.94)%	1.87 (5.22)%
F Lactobacillales Streptococcaceae	12.68 (16.91)%	12.42 (16.31)%	3.05 (4.92)%	8.85 (16.43)%	5.34 (8.99)%	3.94 (10.11)%	8.44 (14.37)%
F Clostridiales Clostridiaceae	0.29 (1.21)%	0.05 (0.15)%	0.11 (0.51)%	0.17 (0.46)%	0.04 (0.08)%	0.07 (0.19)%	0.13 (0.64)%
F Clostridiales Eubacteriaceae	0.06 (0.25)%	0 (0.01)%	0 (0.01)%	0.34 (1.38)%	0.12 (0.38)%	0.08 (0.28)%	0.12 (0.69)%
F Clostridiales Lachnospiraceae	2.74 (8.59)%	5.7 (15.79)%	2.59 (5.52)%	13.56 (22.46)%	7.87 (15.07)%	8.5 (20.01)%	7.17 (16.55)%
F Clostridiales Peptostreptococcaceae	0.08 (0.29)%	0.00 (0.01)%	0.00 (0.00)%	0.14 (0.35)%	0.08 (0.18)%	0.20 (0.66)%	0.09 (0.33)%
F Clostridiales Ruminococcaceae	0.01 (0.05)%	0.00 (0.00)%	0.01 (0.03)%	0.15 (0.96)%	0.34 (2.14)%	0.13 (0.56)%	0.11 (1.05)%
F Clostridiales Veillonellaceae	0.87 (2.23)%	0.05 (0.14)%	0.11 (0.29)%	0.06 (0.19)%	0.18 (0.53)%	0.41 (1.33)%	0.30 (1.2)%
F Erysipelotrichales Erysipelotrichaceae	2.51 (9.88)%	0.63 (1.73)%	0.63 (2.1)%	1.83 (3.63)%	2.65 (7.78)%	4.08 (10.63)%	2.06 (7.07)%
F Other	0.07 (0.18)%	0.03 (0.09)%	0.04 (0.05)%	0.05 (0.05)%	0.03 (0.07)%	0.05 (0.12)%	0.05 (0.11)%
Total Proteobacteria	11.72 (19.76)%	13.67 (22)%	18.64 (30.75)%	4.07 (13.77)%	3.13 (8.88)%	3.78 (11.22)%	8.61 (19.03)%
P Enterobacteriales Enterobacteriaceae	11.54 (19.24)%	13.61 (21.96)%	18.5 (30.7)%	4 (13.74)%	3.11 (8.86)%	3.53 (11.08)%	8.5 (18.87)%
P Other	0.18 (0.83)%	0.06 (0.12)%	0.14 (0.32)%	0.07 (0.33)%	0.02 (0.05)%	0.25 (0.83)%	0.11 (0.52)%
V Verrucomicrobiales Verrucomicrobiaceae	0.01 (0.09)%	0.01 (0.07)%	0.00 (0.00)%	0.54 (2.74)%	0.23 (0.88)%	0.00 (0.00)%	0.16 (1.35)%
Other	0.02 (0.06)%	0.01 (0.02)%	0.01 (0.02)%	0.01 (0.02)%	0.01 (0.02)%	0.01 (0.01)%	0.01 (0.03)%

Table S3: Results of forward selection based on canonical correspondence analysis (CCA) of fecal taxonomic compositions. The sample covariates shown were included based on the percentage explained of the total taxonomic variation (Explains %). Inclusion of sample covariates was stopped to prevent overfitting of the ordination model leading to the identification of the 9 factors (and variables) listed in the table. False discovery rate (FDR) method for P-value correction was used to identify the major covariates with an adjusted P-value ($P\text{-adj}$) < 0.05 (in bold) and shown alongside with the uncorrected P-value (P).

Sample covariates (variables)	Explains %	P	P-adj
Age (4 weeks, 26 weeks)	5.2	0.002	0.01733
Ethnicity (asian, caucasian, other)	3.7	0.002	0.0104
Feedinggroup (control, pHF-OS, breastfed)	2.6	0.002	0.00867
Siblings (yes, no)	1.6	0.002	0.00867
Antibiotics before 26 weeks (yes, no)	1.2	0.022	0.08089
Birth weight (continuous data)	0.6	0.028	0.104
Gender (male, female)	0.8	0.206	0.44
Weaning before 18 weeks (yes, no)	0.8	0.226	0.41947
Parental allergic history (mother, father, both)	1	0.266	0.43225

Figure S1: Barplots (mean with 95% confidence interval) summarizing the microbial intestinal richness **(a)** and phylogenetic diversity **(b)** at 4 and 26 weeks of age in infants developing eczema (ECZ) compared to infants not developing eczema (NOECZ). Comparisons were done at a sequencing depth of 1636 reads per sample. Statistics were performed by a one-way ANOVA with Bonferroni's multiple comparison correction (not significantly different).

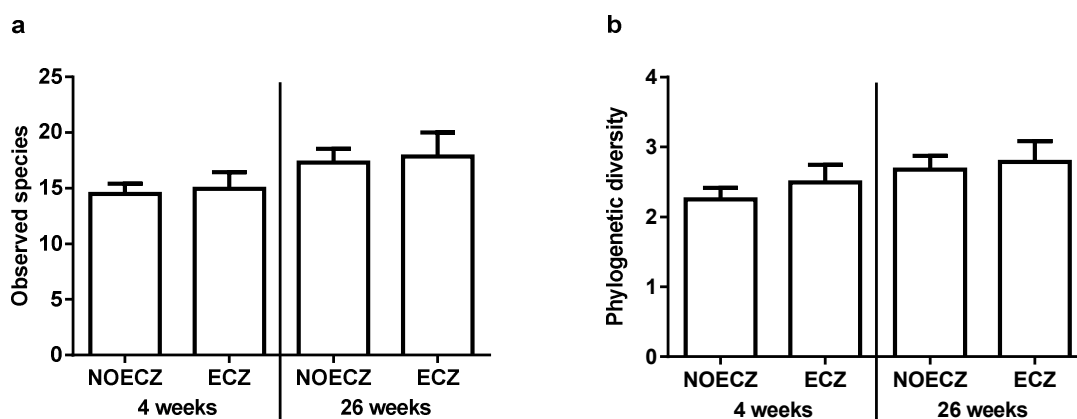
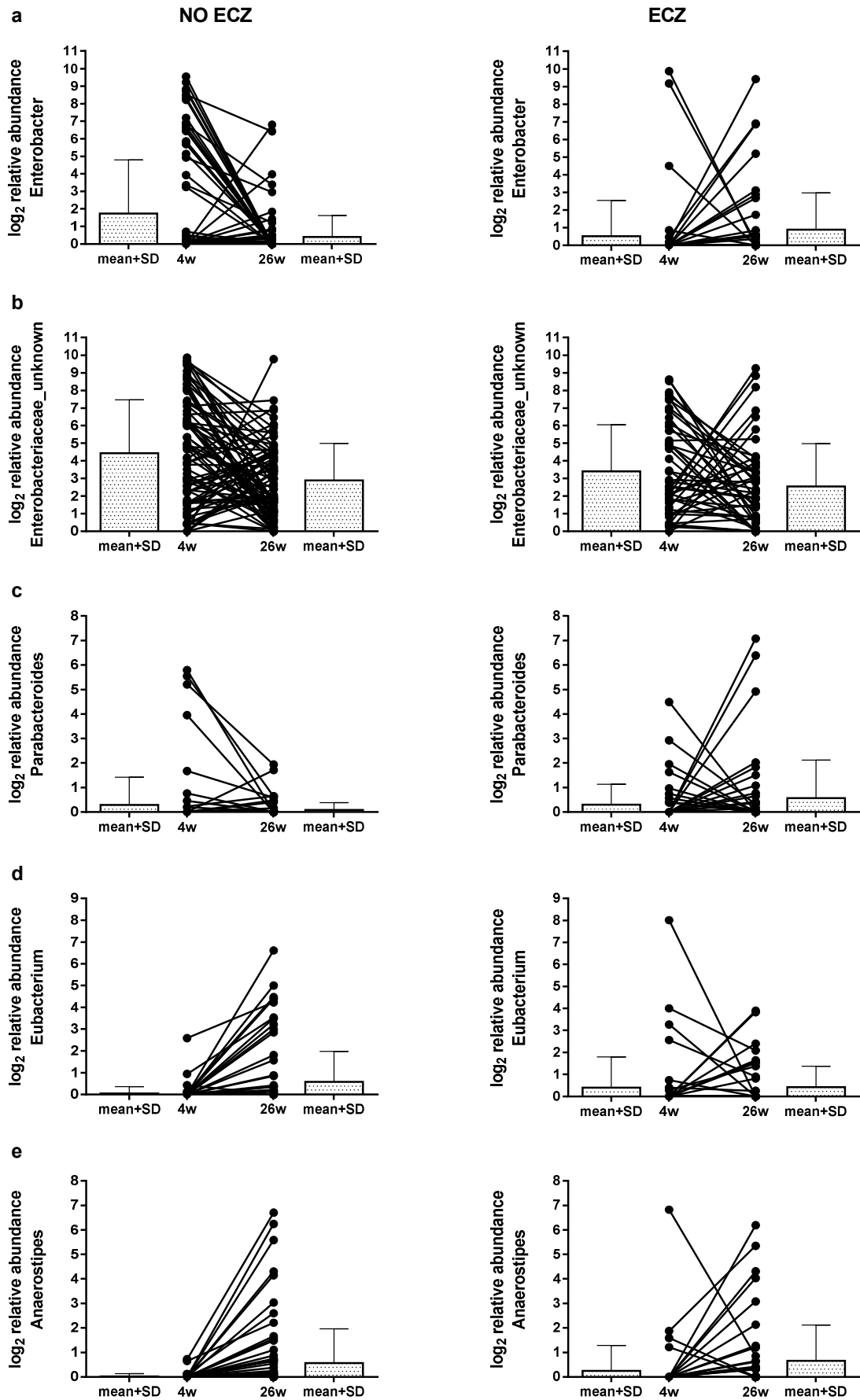
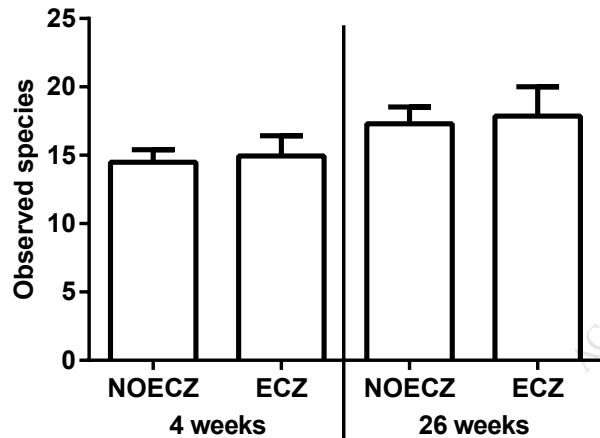


Figure S2: Taxa identified with MetagenomeSeq that show differential relative abundances over time comparing healthy infants with infants developing eczema ($P\text{-adj}<0.1$). Levels for sample-pairs are shown as log-transformed data, as well as means and standard deviations at 4 and 26 weeks, are shown for *Enterobacter* **(a)**, an unassigned genus of Enterobacteriaceae **(b)**, *Parabacteroides* **(c)**, *Eubacterium* **(d)**, and *Anaerostipes* **(e)**, respectively. The left panels show data for infants not developing eczema (NO ECZ, $n=82$ sample-pairs) and the right panels for infants that did develop eczema (ECZ, $n=48$ sample-pairs), respectively.



a**b**