



OPEN ACCESS

EDITED BY

Wenjing Zhao,
Shenzhen Campus, China

REVIEWED BY

Anita Kozyrskyj,
University of Alberta, Canada
David C. B. Taras,
Boehringer Ingelheim, Germany

*CORRESPONDENCE

Anne Salonen
anne.salonen@helsinki.fi

SPECIALTY SECTION

This article was submitted to
Microorganisms in Vertebrate
Digestive Systems,
a section of the journal
Frontiers in Microbiology

RECEIVED 26 May 2022

ACCEPTED 07 September 2022

PUBLISHED 06 October 2022

CITATION

Matharu D, Ponsero AJ, Dikareva E,
Korpela K, Kolho K-L, de Vos WM and
Salonen A (2022) *Bacteroides*
abundance drives birth mode
dependent infant gut microbiota
developmental trajectories.
Front. Microbiol. 13:953475.
doi: 10.3389/fmicb.2022.953475

COPYRIGHT

© 2022 Matharu, Ponsero, Dikareva,
Korpela, Kolho, de Vos and Salonen.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Bacteroides abundance drives birth mode dependent infant gut microbiota developmental trajectories

Dollwin Matharu¹, Alise J. Ponsero^{1,2}, Evgenia Dikareva¹,
Katri Korpela¹, Kaija-Leena Kolho^{1,3,4}, Willem M. de Vos^{1,5} and
Anne Salonen^{1*}

¹Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland, ²Department of Biosystems Engineering and BIO5 Institute, University of Arizona, Tucson, AZ, United States, ³Children's Hospital, Pediatric Research Center, University of Helsinki and HUS, Helsinki, Finland, ⁴Faculty of Medicine and Health Technology, University of Tampere, Tampere, Finland, ⁵Laboratory of Microbiology, Wageningen University, Wageningen, Netherlands

Background and aims: Birth mode and other early life factors affect a newborn's microbial colonization with potential long-term health effects. Individual variations in early life gut microbiota development, especially their effects on the functional repertoire of microbiota, are still poorly characterized. This study aims to provide new insights into the gut microbiome developmental trajectories during the first year of life.

Methods: Our study comprised 78 term infants sampled at 3 weeks, 3 months, 6 months, and 12 months ($n = 280$ total samples), and their mothers were sampled in late pregnancy ($n = 50$). Fecal DNA was subjected to shotgun metagenomic sequencing. Infant samples were studied for taxonomic and functional maturation, and maternal microbiota was used as a reference. Hierarchical clustering on taxonomic profiles was used to identify the main microbiota developmental trajectories in the infants, and their associations with perinatal and postnatal factors were assessed.

Results: In line with previous studies, infant microbiota composition showed increased alpha diversity and decreased beta diversity by age, converging toward an adult-like profile. However, we did not observe an increase in functional alpha diversity, which was stable and comparable with the mother samples throughout all the sampling points. Using a *de novo* clustering approach, two main infant microbiota clusters driven by *Bacteroidaceae* and *Clostridiaceae* emerged at each time point. The clusters were associated with birth mode and their functions differed mainly in terms of biosynthetic and carbohydrate degradation pathways, some of which consistently differed between the clusters for all the time points. The longitudinal analysis indicated three main microbiota developmental trajectories, with the majority of the infants retaining their characteristic cluster until 1 year. As many as 40% of vaginally delivered infants were grouped with infants delivered by C-section due to their clear and persistent depletion in *Bacteroides*. Intrapartum antibiotics, any perinatal or postnatal factors, maternal microbiota composition, or other maternal factors did not explain the depletion in *Bacteroides* in the subset of vaginally born infants.

Conclusion: Our study provides an enhanced understanding of the compositional and functional early life gut microbiota trajectories, opening avenues for investigating elusive causes that influence non-typical microbiota development.

KEYWORDS

***Bacteroides*, shotgun metagenomics, infant gut microbiota, functional maturation, hierarchical clustering, vaginal delivery, cesarean section**

Introduction

The acquisition and development of early life gut microbiota have been linked to health outcomes during infancy and later life (Avershina et al., 2021; Sarkar et al., 2021). The establishment of gut microbiota is affected by several maternal factors, perinatal and postnatal exposures such as mode of birth, breastfeeding, infections, antibiotics, hosts genetics, and living environment (Penders et al., 2006; Tamburini et al., 2016; Korpela and de Vos, 2018). Delivery mode is a crucial factor especially in the first months of life (Arboleya et al., 2018; Kumbhare et al., 2019; Korpela, 2021). As infants develop, their microbiota matures through a phase characterized by high dynamics and low diversity in early infancy (from birth to 6 months) and reaches a more stable and diverse microbial community during early childhood (3–5 years old) (Milani et al., 2017; Korpela and de Vos, 2018; Korpela, 2021). Importantly, the maturation of infant gut microbiota is affected by the timing of weaning and breastfeeding cessation, as these events contribute to shifting of the microbial community composition toward an adult-like profile (Bäckhed et al., 2015).

Several authors have reported the critical impact of delivery mode on early-life microbiota acquisition. Typically, Cesarean-section delivered (CSD) infants show a strong depletion in *Bacteroides* that typically resolves between 6 and 12 months of age in relation to weaning (Shao et al., 2019; Yang et al., 2019; Korpela, 2021; Princival et al., 2021). Some studies suggest that birth mode-related impact on early life microbiota may take 3–5 years to normalize (Roswall et al., 2021). On the other hand, vaginally delivered (VD) infants' gut microbiota is typically dominated by *Bifidobacterium* and *Bacteroides spp.*; both genera are known to be central in digestion of human milk oligosaccharides (Marcobal et al., 2010). The CS birth-related scarcity of *Bacteroides spp.* has been linked to increased risk

of intestinal colonization by opportunistic pathogens (Reyman et al., 2019; Shao et al., 2019) and altered immunostimulatory potential in infants (Wampach et al., 2018). Overall, *Bacteroides spp.* have been suggested to play a pivotal role in infant health because of their role in modulating and training the immune system (Vatanen et al., 2016), and higher ratio of *Enterobacteriaceae* to *Bacteroidaceae* in CSD infants has been implicated in increased risk of development of atopy and *Clostridioides difficile* infection in late infancy (Vu et al., 2021). Some species of *Bacteroides* are involved in cross-feeding nutrients with other gut microbes, producing antimicrobial toxins and competing against pathogens (Wexler, 2007; Zafar and Saier, 2021). The lack of exposure to mother's gut microbiota at birth explains the delayed establishment of *Bacteroides spp.* in CSD babies (Korpela et al., 2020; Korpela, 2021). However, some VD infants are also essentially devoid of *Bacteroides spp.* (Yassour et al., 2016; Shao et al., 2019; Wilson et al., 2021). Characterization of individual variations influencing the abundance of *Bacteroides spp.* in the infant gut will contribute to the enhanced understanding of the physiological role of this key early life colonizer.

Most studies to date have described the dynamics of infant gut microbiota development through 16S rRNA gene amplicon sequencing approaches. When combined with unsupervised clustering approaches, these studies have provided important insights into temporal, ethnicity related, and age-specific development of microbiota trajectories as well as the stability, inferred functions, and maturation rates of community types (Stewart et al., 2018; Borewicz et al., 2019; Roswall et al., 2021; Tun et al., 2021; Xiao et al., 2021). Since amplicon-based studies restrict exploring the functional capacities of microbial communities, much less is known about the functional aspects of infant microbiota development. A recent Canadian study based on the well-characterized CHILD birth cohort used clustering approaches and conducted metagenomic imputation to provide novel insights into the potential role of infant's gut microbiota-mediated sphingolipid metabolism in food sensitization, linking *Bacteroides*-depleted microbiota clusters to higher odds of developing atopic sensitization in a host genotype-dependent manner (Tun et al., 2021). A recent meta-analysis of multiethnic infant microbiota enterotypes also characterized the functional aspects of microbiota maturation during the first years and

Abbreviations: VD, Vaginally delivered; CSD, Cesarean-section delivered; CS, Cesarean section; HMOs, Human milk oligosaccharides; IAP, Intrapartum antibiotic prophylaxis; HELMi, Health and Early Life Microbiota; RBB, Repeated bead beating; FDR, False discovery rate; PERMANOVA, Permutational multivariate analysis of variance; ANOVA, Analysis of variance; PCoA, Principal coordinate analysis; qPCR, Quantitative polymerase chain reaction.

showed that *Bacteroides*- and *Prevotella*-dominated infant enterotypes share enriched metabolic pathways, e.g., glycolysis and starch degradation (Xiao et al., 2021). Earlier metagenome studies have provided important insights into the temporal development of gut microbiota mainly in VD infants (Bäckhed et al., 2015; Stewart et al., 2018; Busi et al., 2021). However, the anticipated impact of delivery mode on the functional capacities of infant gut microbiota is still not fully understood. A recent metagenome study reported that the characteristic signature of low *Bacteroides* abundance in CSD infants led to the underrepresentation of several biosynthetic pathways at the age of 3 months (Wilson et al., 2021). Our study leverages the shotgun metagenomic sequencing approach to explore the dynamics of infant gut microbiota development. This study comprises samples spanning 3 weeks to 1 year of age in a group of 78 term infants from the well-characterized Finnish Health and Early Life Microbiota (HELMi) birth cohort (Korpela et al., 2019). Using a clustering approach, we defined three main microbiota development trajectories differing in both taxonomic and functional compositions. While the trajectories were strongly dependent on delivery mode, we identified a group of VD infants presenting depletion in *Bacteroides* spp. comparable to that in CSD infants. We further characterized this unusual microbiota composition for VD infants and investigated the underlying factors as well its impacts on the functional capabilities of the microbiota.

Materials and methods

Sample collection and sequencing

The HELMi birth cohort study ($N = 1,055$) is a prospective follow-up study on early life microbiota and health (Korpela et al., 2019). Healthy infants born on gestational weeks 37–42 without known congenital defects and exceeding the birth weight of 2.5 kg were included in the cohort. For this study, 90 infants representative of key early exposures ($N = 25$ born by CS, $N = 24$ VD with intrapartum antibiotics (IAP), and $N = 41$ VD without IAP) were selected from the HELMi cohort. The infants were selected from the broader HELMi cohort based on the number of samples available and their birth mode. Infants who had a long-term non-allergic disease reported by the age of 2 years were excluded. All CS deliveries involved IAP, and among the CSD infants, 14 were born by planned CS and 11 were by emergency CS. No restrictions in terms of maternal health status or maternal antibiotic treatment during the pregnancy period were applied in selecting the infants for this study.

Fecal samples were collected at the age of 3 weeks and 3, 6, and 12 months, and their mothers' samples were collected within 2 weeks prior to childbirth and were subjected to shotgun metagenome sequencing. The study was approved by the ethical committee of the Hospital District of Helsinki and

Uusimaa and performed in accordance with the principles of the Helsinki Declaration. The parents signed informed consent for enrolment. They collected maternal and infant fecal samples and stored them at home in freezers at -20°C . The samples were transported frozen to the study center within 6 months of collection and were stored in freezers at -80°C upon arrival. Information on early life exposures was collected through online questionnaires and hospital records. The infants were generally healthy, although two were diagnosed with an allergic disease by the age of 1 year.

Fecal DNA was extracted from fecal material (0.026 to 2.46 g, median = 0.266 g) by suspension in 1 ml of sterile ice-cold PBS, and 250 μl of the fecal suspension was combined with 340 μl of RBB lysis buffer (500 mM NaCl, 50 mM Tris-HCl (pH 8), 50 mM EDTA, and 4% SDS) in a bead-beating tube from Ambion Magmax™ Total Nucleic Acid Isolation Kit (Life Technologies, Carlsbad, CA, United States). After repeat bead-beating, 200 μl of the supernatant was used for DNA extraction with a KingFisher™ Flex automated purification system (Thermo Fisher Scientific, Waltham, MA, United States) using MagMAX™ Pathogen High Vol. DNA was quantified using the Quanti-iT™ Pico Green dsDNA Assay (Invitrogen, San Diego, CA, United States). Sequencing libraries were prepared according to the Nextera DNA Flex Library Prep Reference Guide (v07) (Illumina, San Diego, CA, United States), with the exception that the reaction volumes were scaled down to $\frac{1}{4}$ of the protocol volumes. Sequencing was performed with an Illumina NovaSeq system using S4 flow cells with a lane divider (Illumina, San Diego, CA, United States) at the sequencing laboratory of the Institute for Molecular Medicine Finland (FIMM) Technology Center, University of Helsinki. Each pool was sequenced in a single lane. The read length for the paired-end run was 2×151 bp.

Quality control and human read filtering

Quality filtering and removal of human sequences were performed on Fastq files from NovaSeq using fastqc v0.11.9 and trimGalore v0.6.6 with default parameters. Quality-filtered sequences were screened to remove human sequences using bowtie2 v2.4.2 against a non-redundant version of the Genome Reference Consortium Human Build 38, patch release 7 (available at https://genome-index.s3.amazonaws.com/bt/GRCh38_noalt_as.zip). After quality control and human read filtering, metagenomes at the early time points (3 and 12 weeks) containing <10 million paired-end reads were discarded based on the observed drop in species richness in shallow sequenced samples. For the later time points (6 and 12 months) and mother samples, metagenomes with <20 million paired-end reads were discarded. Infants with <3 samples remaining were excluded from the analysis. Overall, 12 infants ($N = 2$ born by CS, $N = 9$ VD delivery, and $N = 1$ VD with IAP) were excluded because

of low sequencing depth or insufficient number of samples. A total of 78 infant samples were kept for further analysis.

Taxonomic and functional profiling

Taxonomic profiling of the metagenomic samples was performed using Kraken2 (Wood et al., 2019) and Bracken (Lu et al., 2017). Briefly, Kraken2 v2.1.1 was run on the paired read using the HumGut database (Hiseni et al., 2021), and Bracken v2.6.1 was run on Kraken2 outputs. Functional profiling was realized on the quality-controlled and human-filtered reads using HumaNn 3.0 (Beghini et al., 2021). The translated search was carried out using the UniRef90 database and hits were mapped to their corresponding Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthogroups (KOs). Gene family counts were normalized as count per million reads before the analysis. Gene counts were mapped to MetaCyc pathway definitions, and MinPath was used to calculate pathway abundances.

Alpha diversity measure (Shannon index), richness (Chao1 index), and beta diversity distances between samples (Bray-Curtis) for the taxonomic and functional profiles were estimated with the R packages *vegan* (Dixon, 2003) and *microbiome* (Lahti and Shetty, 2012–2019). The stability of bacterial composition during growth was explored by calculating the presence-absence Jaccard index for successive time points for each child. The Jaccard index was computed on the relative abundances at the genus level using the package *vegan*.

kmer-composition profiling

Simka v1.5.3 (Benoit et al., 2016) was run on metagenomic datasets with the following parameters: `-abundance-min 2 -max-reads [MINCOUNT] -simple-dist`, where [MINCOUNT] is the smallest sequence count across the analyzed samples. A Bray-Curtis distance computed on the content in the kmer of metagenomes was obtained and used for further analysis.

Hierarchical clustering of the samples based on taxonomic composition

The compositional analyses were performed using R (version 3.4.4). First, the taxonomy data were aggregated at the family level. Low-abundance families (<1% relative abundance and <0.5% prevalence) were filtered out, and zero values were replaced by an estimate of 1. Then, the dataset was normalized by taking the centered log ratio (CLR), and a distance matrix was calculated from the transformed data using the Euclidean distance (Aitchison

distance) (Aitchison et al., 2000; Tsilimigras and Fodor, 2016). Hierarchical clustering was carried out separately for each time point with the function *hclust* and with the Wald.D2 method. To determine the appropriate number of clusters, the silhouette coefficient was calculated and the number of clusters allowing the highest silhouette coefficient was selected. The significance of the obtained clusters was confirmed by PERMANOVA at the family level using the Aitchinson distance. The association of the sample clusters to exposures such as intrapartum and postnatal antibiotics was assessed by Fisher's exact test and was not found to be significant at any time point.

Defining VD high- and low-*Bacteroides* groups

The VD infants were grouped as “*Bacteroides*-present VD” and “*Bacteroides*-depleted VD” according to the relative abundance of the *Bacteroides* genus in the samples collected at 3 weeks. A cut-off of 1% relative abundance at the genus-level was used to define the groups. The 1% cut-off was chosen as all the CS infants had a relative abundance of *Bacteroides* below this threshold. The association of the infant groups with IAP and postnatal antibiotic exposure was assessed by Fisher's exact test and was found to be non-significant.

MaAsLin2 differential abundance analysis

To assess the associations between infant microbiota clusters at each time point and MetaCyc pathway abundances, pathways with at least 0.0001% relative abundance in at least 10% of the samples were selected. The associations between pathways and clusters were determined with generalized linear and mixed models, using Multivariate Association with Linear Models (MaAsLin2) (Mallick et al., 2021). Any pathway association with an FDR-adjusted *p*-value < 0.1 was considered statistically significant.

Associations between maternal microbiota composition and the grouping of VD infants according to their relative abundance of *Bacteroides* were assessed using the Negative Binomial model from the MaAsLin2 package (-m NEGBIN parameter). Taxonomic profiles were aggregated at the species level, low abundance hits were filtered out (detection below 0.001% and prevalence below 1%), and counts were normalized by TMM (Trimmed Mean by M-Values). The results obtained through negative binomial models were confirmed using generalized linear models from the MaAsLin2 package. Any species association with an FDR-adjusted *p*-value < 0.1 was considered statistically significant.

Association between infant microbiota and environmental variables

A directed acyclic graph (DAG) was built detailing causal relationships affecting the association between perinatal and postnatal exposures and infant gut microbiota development (Supplementary Figure 1). Nodes represent the exposure of interest, and each node is interrelated by directional arrows that represent theoretical associations based on the researchers' assessment of previous literature and confirmed by association testing in the HELMi cohort. From this representation, we assessed that birth mode is a potential confounder of the effect of perinatal variables (intrapartum antibiotics, gestational age, time of first skin-to-skin contact, and number of previous deliveries) and breastfeeding variables (length of breastfeeding time, length of exclusive breastfeeding, and age of first solids). Tests of associations between these variables and the infant gut microbiota were therefore controlled for birth mode effects.

For testing the association between infant microbiota trajectories and environmental variables (Results Section Taxonomic and functional features of the birth mode-driven infant microbiota clusters) as well as between the VD *Bacteroides* infant group and environmental variables (Results Section CS-like *Bacteroides* depletion in early life microbiota can be observed in a subset of VD infants), a large number of tests were carried out in an attempt to identify any variables that may be significantly associated to a group of samples. An FDR correction was applied to account for the increased risk of Type I error in multiple testing settings, and FDR adjusted p -values <0.1 were considered statistically significant.

A *post-hoc* sensitivity analysis was conducted with the G*Power software version 3.1.9.2 to retrospectively examine the observed power of the study. The effect size required for a power level of 0.8 was calculated with the number of samples used at the significance level ($p = 0.05$) for the Wilcoxon–Mann–Whitney test.

Other statistical analysis

The statistical analysis was conducted using R (version 3.4.4) with the packages *vegan* and *microbiome* (Dixon, 2003; Lahti and Shetty, 2012–2019). Alluvial diagrams were constructed with the R package *ggalluvial* (Brunson, 2020).

For univariate data comparison, a statistically significant difference was evaluated by the unpaired Wilcoxon test when applied to numeric variables or the Fisher test when comparing categorical groups, and p -values <0.05 were considered statistically significant. Additionally, when a pair-wise Wilcoxon test was carried out on more than two groups of samples, an FDR correction was applied, and FDR adjusted p -values <0.05 were considered statistically significant.

Results

Cohort overview

The present study included 78 infants from the HELMi cohort [NCT03996304 (Korpela et al., 2019)], born in Finland at hospitals at term, and followed up for 1 year. A total of 55 infants were born *via* VD. All the 23 CS deliveries and 23 VDs involved IAP. Baseline characteristics and antibiotic exposures over the first year of life for all 78 infants, stratified by mode of delivery and IAP exposure, are summarized in Supplementary Table 1. Perinatal and background variables were evenly distributed over the three birth groups except for two environmental exposures (siblings and pets at home) and antibiotic exposures during the first year of life (Supplementary Table 1). All the children were breastfed during the first months of life. Most of the children were breastfed for more than 9 months, and only 8 were breastfed for 6 months or less. The introduction of solid food took place between 18 and 31 weeks of age. Seventeen (22%) children received at least one antibiotic treatment during their first year of life, and some received multiple courses (30 courses in total for all the children) (Supplementary Table 1). Infants' gut microbiota was assessed by shotgun metagenomic sequencing of fecal samples collected at 3 weeks ($n = 75$) and 3 months ($n = 77$), 6 months ($n = 71$), and 1 year ($n=57$). Each infant was sampled a minimum of 3 times, with 60% of the infants ($n = 46$) having all 4 samples. For 50 infants, metagenomic data were also available from their mother's stool samples taken during the last 2 weeks before delivery.

Microbiota maturation over the first year of life

After quality filtering and removing human sequences, the metagenomes had an average of 39 million paired-end reads per sample. These high-quality, human-filtered reads were subjected to taxonomic annotation with Kraken2 (Wood et al., 2019) using the HumGut database (Hiseni et al., 2021). The taxonomic assignment obtained with Kraken2 was recalculated using Braken (Lu et al., 2017). This method has been previously shown to achieve high taxonomic precision for gut taxonomic profiling (Tamames et al., 2019; Ye et al., 2019; Allnutt et al., 2021). On average, 91% of the total reads at the phylum level and 77% at the family level were successfully annotated. In the infant samples, the most abundant bacterial families were *Enterobacteriaceae* (27% global average relative abundance), *Bifidobacteriaceae* (22%), *Bacteroidaceae* (15%), *Lachnospiraceae* (7%), *Veillonellaceae* (4%), *Clostridiaceae* (4%), *Oscillospiraceae* (4%), and *Streptococcaceae* (1%). Other families represented in total 14% of the sequences (Supplementary Figure 2). The infants' microbiota maturation was characterized by a global decrease in interindividual beta diversity and convergence

toward an adult-like profile (Figure 1A) and had a progressive increase in *Lachnospiraceae* and *Oscillospiraceae* during the first 12 months (Supplementary Figure 2). However, at 1 year, the infants' gut microbiota composition was still significantly different from that of the mothers' microbiota (PERMANOVA on genus-level Aitchinson distance, $p = 0.001$; 999 permutations). The alpha diversity (Shannon diversity index) did not significantly increase between 3 weeks and 3 months but significantly increased between 3 and 6 months (Wilcoxon $p < 0.05$) (Figure 1B), coinciding with the age when solid foods were introduced to the children's diet (median = 22 weeks in the cohort). Significant differences in alpha diversity were also observed between 6 and 12 months, and between 12 months and adult microbiota. Similarly, species richness (Chao1 index) increased during infancy, especially between 6 and 12 months, but was significantly higher in the adult samples than in the 12-month samples (Wilcoxon $p < 0.05$) (Figure 1B). The results were similar after rarefying the samples to the same sequencing depth (data not shown) (Bäckhed et al., 2015).

We also analyzed gut microbiota maturation using an annotation-free approach that computes distances between microbial communities directly based on the number and abundance of shared words of length k (k-mers) (Benoit et al., 2016). Using this approach, we observed a global maturation of the infant samples toward an adult-like sequence composition (Supplementary Figure 3A) and a significant decrease in the inter-individual beta diversity during growth. However, no interindividual beta diversity decrease was observed between the 12-month samples and the mother samples (Supplemental Figure 3B). This could be explained by the increased heterogeneity in subspecies diversity in the adult samples.

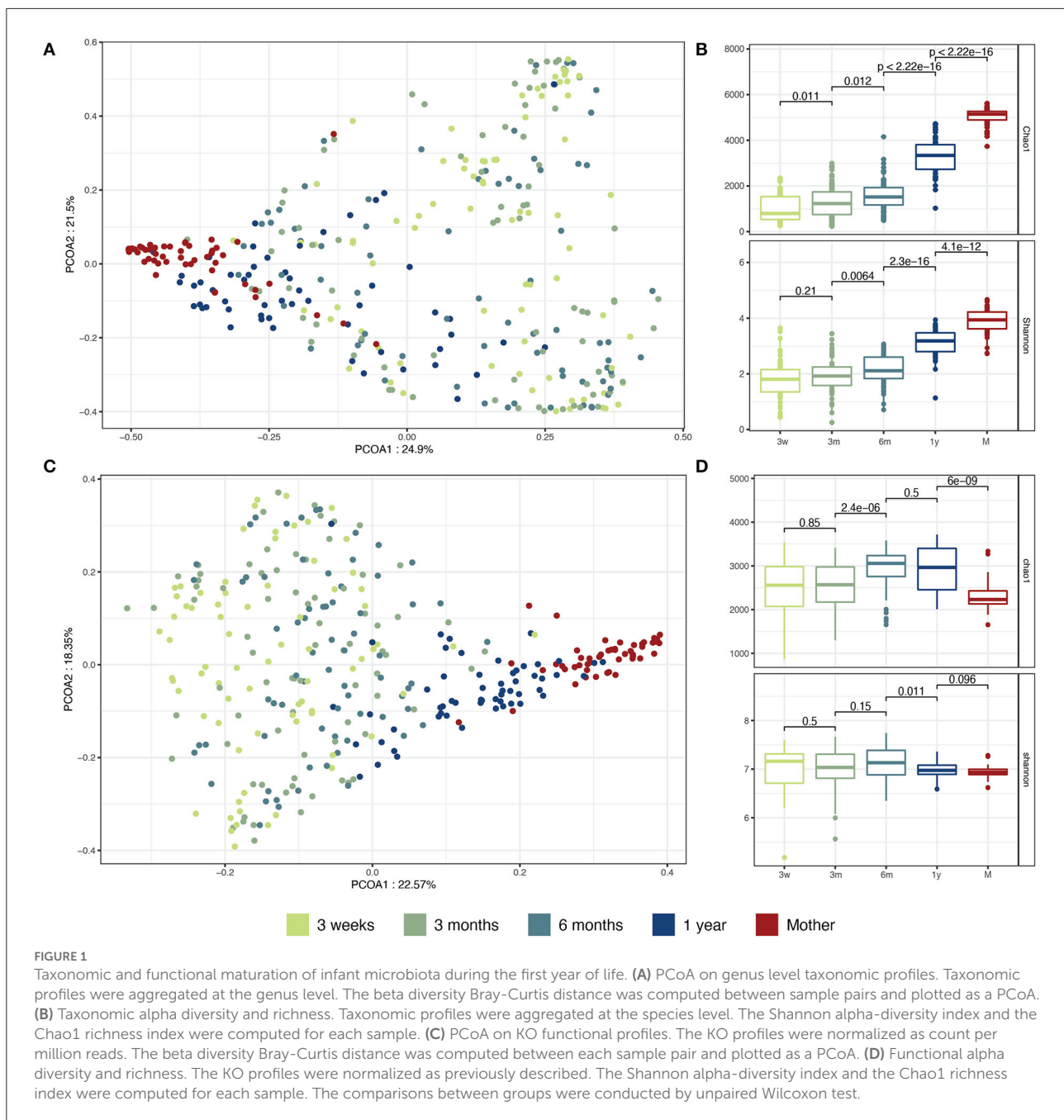
Next, we explored how the maturation of the microbiota affected the microbial metabolic and functional pathways. KEGG Orthogroup (KO) counts were obtained for all the samples and mapped to the main KEGG cellular metabolic pathways. KOs involved in amino acid metabolism, carbohydrate metabolism, and energy metabolism were found to be the most abundant at all infant ages and in the mother samples (Supplementary Figure 4B). Similar to the taxonomic maturation, we observed functional maturation of the infant microbiota, with the infant samples slowly converging toward an adult-like composition during the first year of life (Figure 1C). The functional alpha diversity (Shannon diversity index on KO counts) was globally stable during the first year of life and comparable between the infant and mother samples. On the other hand, the gene richness (Chao1 index on KO counts) significantly increased in infants between 3 and 6 months but was surprisingly found to be lower in the mother samples (Figure 1D). To control for potential biases in the functional annotation as an explanation for the decreasing trend of gene richness, we looked at the proportion of sequences left out of functional annotation. Indeed, the proportion of sequences

unmapped to a gene family increased significantly between 6 and 12 months in the infants and was highest in the mother samples (Supplementary Figure 4A), suggesting increased complexity of the microbiome in late infancy and further in adulthood.

Taxonomic and functional features of the birth mode-driven infant microbiota clusters

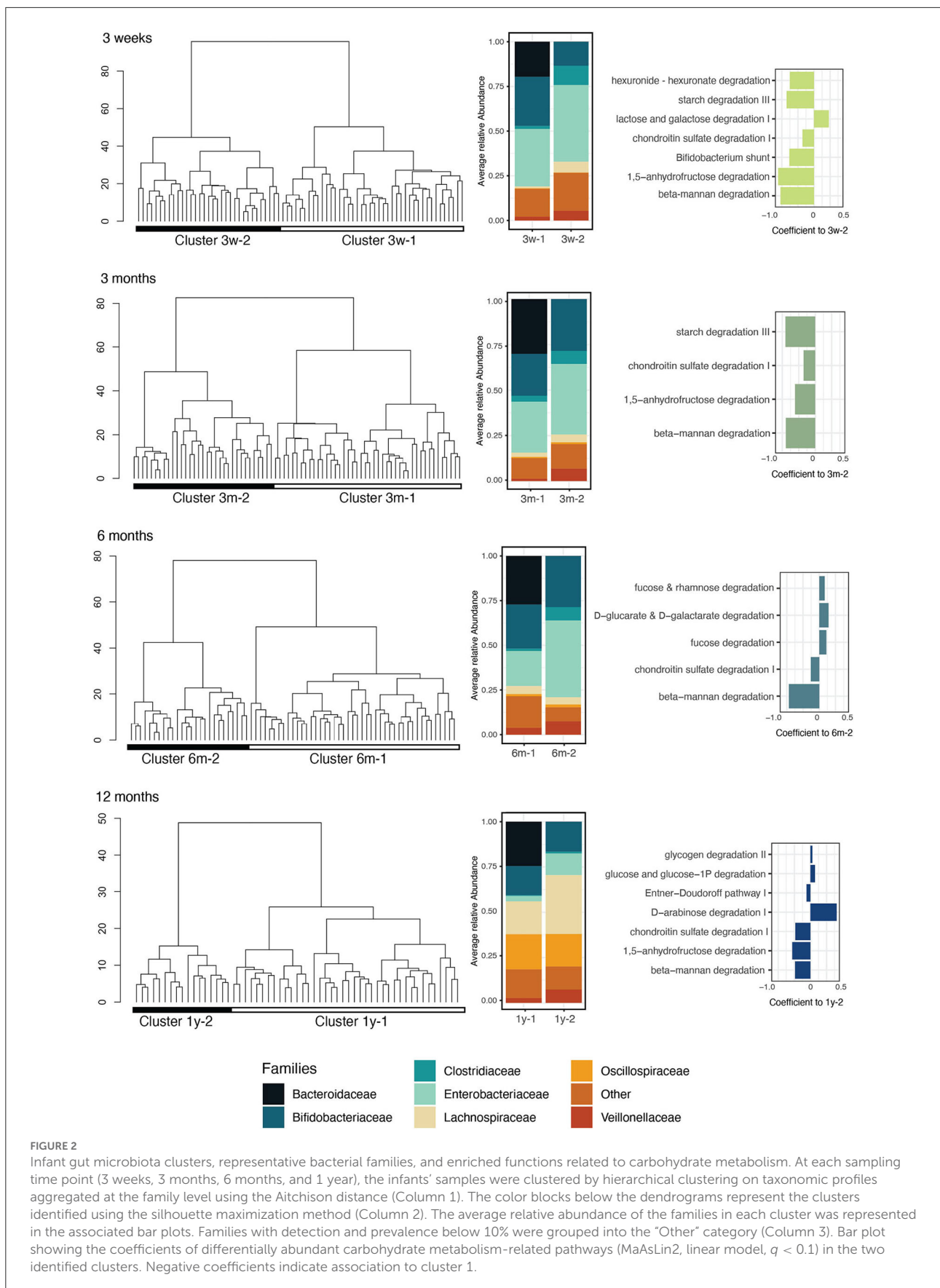
To identify the main microbiota establishment trajectories, we clustered the infant samples on their taxonomic composition at the family level for each sampling time point. We performed hierarchical clustering of the samples using the ward linkage on the Aitchinson distance. The bacterial communities clustered into two groups at each sampling time point (Figure 2). The clusters were confirmed by PERMANOVA and showed that the composition of the samples was significantly different between clusters at the family level at each time point (PERMANOVA on family-level Aitchinson distance, $p = 0.001$; 999 permutations). However, the alpha diversity was not significantly different between the clusters, and their taxonomic richness was significantly distinct between the clusters only in the 3-week and 12-month samples (Wilcoxon $p < 0.05$) (Supplementary Figure 5), suggesting that the identified clusters differed in terms of composition but not in terms of taxonomic complexity. The PCA bi-plot for the clusters revealed that the clustering was driven at all time points by the abundance of the *Bacteroidaceae*, *Tannerallaceae*, *Rikenellaceae*, and *Ordoribacteriaceae* families in the first cluster, and the presence of *Clostridiaceae* in the second (Supplementary Figure 6). The cluster membership was strongly associated with the birth mode at all time points (Fisher's exact test $p < 0.05$) but was not associated with IAP exposure (Fisher's exact test $p > 0.05$ between VD and VD + IAP infants) or postnatal antibiotic exposure (Fisher's exact test $p > 0.05$).

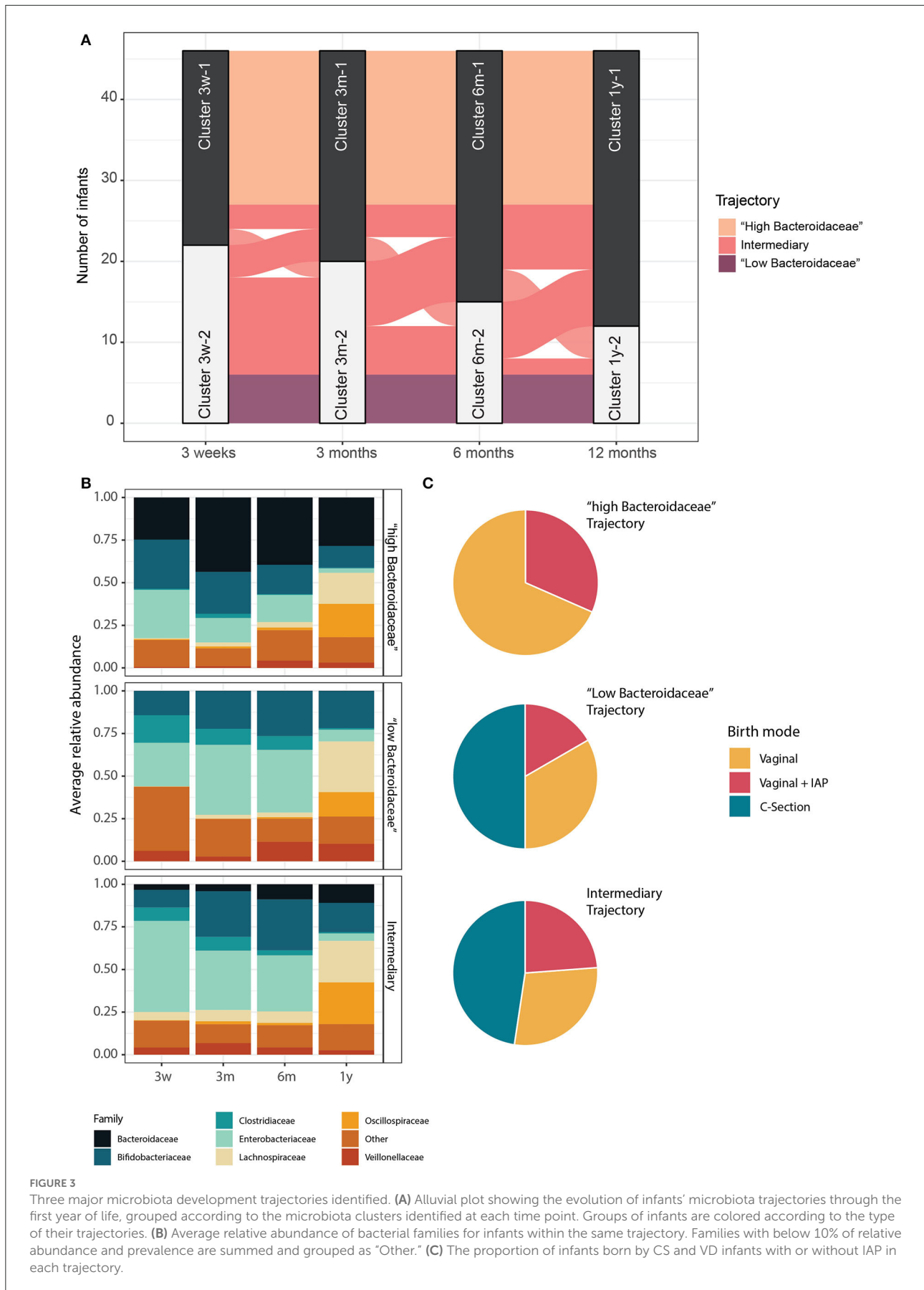
We next compared the identified clusters in terms of their functional composition. At all the time points, the clusters had a significantly different functional composition (PERMANOVA on KO Bray-Curtis distance, $p = 0.001$; 999 permutations). The KO alpha diversity (Shannon index) was not significantly distinct between the clusters at any sampling time point; however, the KO richness (Chao1 index) was found to be significantly lower in cluster 2 at 3 weeks (cluster 3w-2) (Supplementary Figure 5). To provide an overview of the variability patterns of microbiota functions between the clusters, we mapped the gene families to 401 prokaryotic MetaCyc pathways and identified differentially abundant pathways between the clusters for each time point. At 3 weeks, 49 pathways were identified as significantly differentially abundant in the two clusters; at 3 months, 55 pathways; at 6 months,

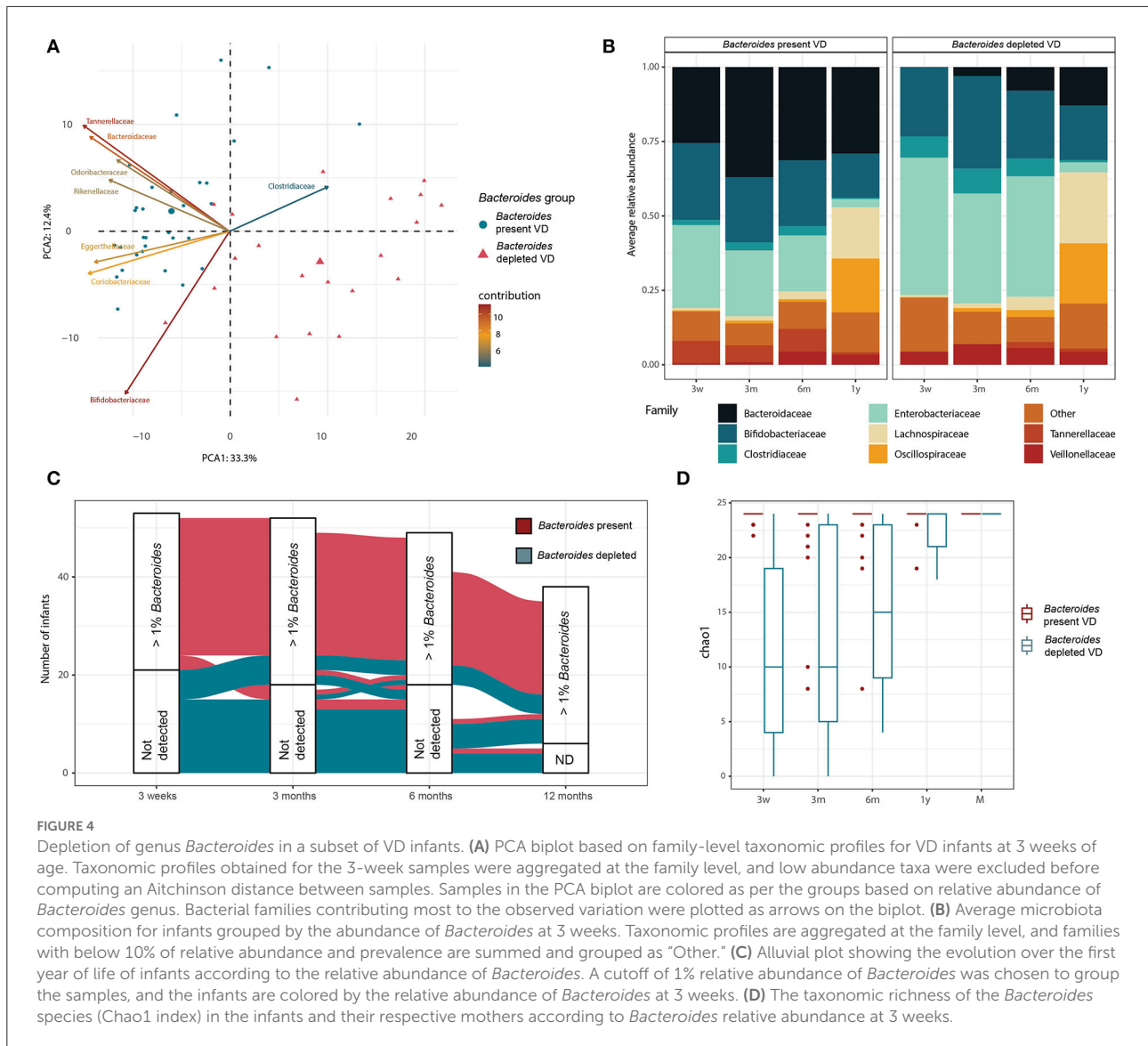


73 pathways and at 12 months, 120 pathways were found significantly differentially abundant between the clusters (linear model, $q < 0.1$). Across all the time points, in total, 68 unique pathways were enriched in clusters 3w-2, 3m-2, 6m-2, and 1y-2, in particular pathways involved in nucleotide biosynthesis (18% of the differentially abundant pathways), carbohydrate biosynthesis (13%), and cofactor biosynthesis (10%). On the other hand, 124 pathways were enriched in clusters 3w-1, 3m-1, 6m-1, and 1y-1, in particular pathways involved in amino acid

biosynthesis (15%), cofactor biosynthesis (14%), and carboxylate degradation (8%) (Supplementary File 1). Pathways involved in carbohydrate degradation were consistently differentially abundant between the two clusters at all the time points and were mostly associated with *Bacteroidaceae*-driven cluster 1, especially in the early time points (Figure 2). Interestingly, pathways involved in the degradation of beta-mannan, a plant-based dietary polysaccharide, were consistently enriched in this cluster throughout the first year and already before







weaning age (Figure 2). The complete list of identified pathways and their differences between the clusters are listed in Supplementary File 1.

We next investigated the temporal stability of the identified clusters in individual infants. For this, we focused on a subset of 46 infants for which all four samples were available. Interestingly, 54% of infants belonging to one of the two clusters, either *Bacteroidaceae*- or *Clostridiaceae*-driven cluster at 3 weeks, stayed in the same cluster throughout all the time points, suggesting that the initial microbiota composition after birth has a critical effect on future microbiota developmental trajectory until 1 year (Figure 3A). Consequently, two stable microbiota development trajectories and an unstable trajectory emerged. One microbiota developmental trajectory was followed by 19 (41%) infants, all VD, and was characterized by

a high relative abundance of the family *Bacteroidaceae* (mean > 25% at all time points, IQR = 36%). Another microbiota trajectory was followed by 6 (13%) infants, with 50% CS infants and 50% VD infants, and was characterized by a virtual lack of *Bacteroidaceae* at all the time points (mean < 0.05% at all time points, IQR = 2%) and higher relative abundance of *Clostridiaceae*. Finally, the temporally instable intermediary trajectory was followed by 21 (46%) infants, with 48% CS infants and 52% VD infants (Figures 3B,C). The relative abundance of *Bacteroidaceae* was significantly different between all the trajectories, and *Clostridiaceae* was significantly different in the second and third trajectories compared to the first one (Pair-wise Wilcoxon, FDR-adjusted, $p < 0.05$). Interestingly, the abundance of *Bifidobacteriaceae*, *Enterobacteriaceae*, *Lachnospiraceae*, *Oscillospiraceae*, and *Veillonellaceae* did not

differ consistently between the trajectories (Pair-wise Wilcoxon, FDR-adjusted, $p > 0.05$). Additionally, the alpha diversity (Shannon diversity index) was similar for all the trajectories at all the time points (Pair-wise Wilcoxon, FDR-adjusted, $p > 0.05$, on Shannon diversity index). We compared the trajectories in terms of stability of bacterial composition by calculating the presence-absence Jaccard index of the genus-level composition for successive time points for each child. As expected, the “intermediary” trajectory presented a less stable composition over time, with successive communities for each infant being more dissimilar than for the first and second trajectory (pair-wise Wilcoxon, FDR-adjusted, $p < 0.05$ on presence/absence Jaccard distance). The trajectories were strongly associated with delivery mode (Fisher test, FDR-adjusted, $p < 0.1$) but not with IAP exposure (Fisher test, FDR-adjusted, $p > 0.1$ between VD and VD + IAP infants). The trajectories were also not significantly associated with the number of antibiotic courses received by the infants during the first year of life (Fisher’s exact test, FDR-adjusted, $p > 0.1$). Additionally, starting age of solid food and breastfeeding habits (exclusive breastfeeding and breastfeeding period length) were not significantly different between the three trajectories (Fisher’s exact test, FDR-adjusted, $p > 0.1$). All the comparisons between trajectories are available in [Supplementary File 5](#).

CS-like *Bacteroides* depletion in early life microbiota can be observed in a subset of VD infants

We further investigated the abundance and dynamics of the genus *Bacteroides* that was driving the clustering and trajectories. At 3 weeks, the microbiota of all the CSD infants contained a low relative abundance of the genus *Bacteroides* (mean = 0.2% relative abundance, IQR = 0.2%). At 3 weeks, we had 53 VD infants as opposed to 55, since samples from two infants at this time point were not available. Strikingly, roughly 40% ($n = 21$) of the VD infants ($n = 10$ in group VD with IAP, $n = 11$ in group VD without IAP) presented an equivalently low relative abundance (below 1%) of the genus *Bacteroides* at 3 weeks. To explore this unusual microbiota composition for the VD infants, we defined the two groups of VD infants according to the relative abundance of the genus *Bacteroides* at the first sampling point, i.e., at the age of 3 weeks. A group of 21 infants, called hereafter “*Bacteroides*-depleted VD” had a <1% relative abundance of *Bacteroides* in their microbiota. The rest of the VD infants ($n = 32$) were called “*Bacteroides*-present VD.” The “*Bacteroides* VD” groups were consistent with the trajectories defined previously, with 90% of the “*Bacteroides*-present VD” infants following the first trajectory and 92% of the “*Bacteroides*-depleted VD” infants following the second or intermediary trajectory ([Supplementary File 6](#)). The global

bacterial composition of the two groups was significantly distinct (PERMANOVA on family-level Aitchinson distance, $p = 0.001$; 999 permutations) at 3 weeks ([Figure 4A](#)). The two groups were also distinct in terms of their functional composition (PERMANOVA on KO Bray-Curtis distance, $p = 0.001$; 999 permutations). Altogether, 62 MetaCyc prokaryotic pathways were significantly differentially abundant in the two VD infant groups at 3 weeks (linear model, $q < 0.1$), in a large part, overlapping with the pathways identified as differentially abundant in clusters 3w-1 and 3w-2 ([Supplementary File 2](#)).

When comparing the microbiota developmental trajectories of the two groups of VD infants, the “*Bacteroides*-depleted VD” infants showed a gradual acquisition of *Bacteroides* during the first year of life ([Figure 4B](#)). Yet, while strongly depleted in *Bacteroides* at 3 weeks, 75% of the “*Bacteroides*-depleted VD” infants had a *Bacteroides* relative abundance above 1% at 12 months (average 13%). On the other hand, only three infants from the “*Bacteroides*-present group” showed a loss in *Bacteroides* relative abundance of <1% at any time point, while the remaining infants for that group had a *Bacteroides* relative abundance above 1% at all the sampling time points ([Figure 4C](#)). Globally, the richness of the *Bacteroides* species increased during growth, while the “*Bacteroides*-depleted VD” infant group had significantly lower *Bacteroides* richness at all the time points (Wilcoxon $p < 0.05$) ([Figure 4D](#)). A total of 24 species of *Bacteroides* were detected in the dataset, with *B. fragilis*, *B. dorei*, and *B. vulgatus* being the most abundant. However, no *Bacteroides* species were found to be present in one only infant group, indicating differences in the abundances of *Bacteroides* species but not in their prevalence ([Supplementary Figure 7](#)). Taken together, these results suggest that the *Bacteroides* depletion observed in the “*Bacteroides*-depleted VD” infant group at 3 weeks is general rather than a lack of specific species of *Bacteroides*, and has a long-term effect that impacts the microbiota development of the infants during the first year of life.

To explore the potential reasons for the depletion of *Bacteroides* in VD infants, we compared the prenatal and perinatal variables between the “*Bacteroides*-present VD” and “*Bacteroides*-depleted VD” infants. We first explored birth-related differences between the two groups, such as potential IAP exposure, intensive care, the time between water breakage and delivery, gestational age, and birth weight. None of these variables were significantly associated with the two infant groups. Next, we explored maternal variables such as the mother’s age, maternal health during pregnancy, exposure to antibiotics or probiotics during pregnancy, number of prior pregnancies, and pre-pregnancy BMI. Once more, none of these variables were significantly associated with the two infant groups. We also tested against group membership the antibiotics received by the children in the first 3 weeks of life and breastfeeding habits; but once again, no significant association could be found. A complete list of the variables and their

definition, and the result of the statistical comparison are available in [Supplementary File 3](#).

We next compared the composition of the maternal gut microbiota in the samples collected prior to delivery for the two VD groups ([Supplementary Figure 8](#)), and in particular in the *Bacteroides* species ([Supplementary Figure 7](#)). Overall, no differences were observed between the microbiota composition of the mothers of “*Bacteroides*-present VD” and “*Bacteroides*-depleted VD” infants at the family level (PERMANOVA on family-level Aitchinson distance, $p = 0.26$; 999 permutations), and we could not identify any significantly differentially abundant taxa between the two groups of mothers using MaAsLin2 at the family and species levels (negative binomial, $q > 0.1$). In order to rule out any effect of technical variations, we explored the association between the two groups of infants with the length of sample storage, the person performing DNA extraction, DNA extraction yield, and sequencing depth. Interestingly, DNA yield (ng of DNA per gram of feces) was significantly higher in the “*Bacteroides*-present VD” than in the “*Bacteroides*-depleted VD” (FDR-adjusted $p < 0.1$, Wilcoxon test). Similarly, the DNA yield obtained from 3 weeks samples of CSD infants was significantly lower (median 1,762 ng/g of feces) compared to the “*Bacteroides*-present VD” (median 5,495 ng/g of feces) but was similar to the “*Bacteroides*-depleted VD” samples (median 1,575 ng/g of feces, FDR-adjusted $p < 0.1$ Wilcoxon test). Importantly, the lack of significant associations between birth and maternal variables and the “*Bacteroides*-depleted VD” group could be due to our limited cohort sampling size. Indeed, our study was powered to only reliably detect large effect sizes ($d = 0.82$ for the power of 0.8 by the Wilcoxon test).

Discussion

Birth mode is a major factor influencing the development of infant gut microbiota and has anticipated effects on infants' immune systems and long-term health ([Mueller et al., 2015](#); [Korpela, 2021](#)). More specifically, CSD infants have a characteristically reduced colonization and abundance of *Bacteroides* spp. ([Dominguez-Bello et al., 2010](#); [Bokulich et al., 2016](#); [Wampach et al., 2018](#); [Guittar et al., 2019](#); [Reyman et al., 2019](#); [Busi et al., 2021](#)). In this study, we leveraged the shotgun metagenomics approach on 78 infants from a well-characterized longitudinal birth cohort followed up from 3 weeks to 1 year of age. Our results recapitulate the previous findings that show a strong influence of birth mode on infant microbiota acquisition. Apart from comparing the infant microbiota composition and function based on birth groups, we zoomed into data-driven infant microbiota clusters and trajectories. Our results show a strong depletion of genus *Bacteroides* in 40% of the VD infants. This study expands our understanding of the impact of various early life factors on the colonization and dynamics of *Bacteroides* spp. in infants.

As previously reported in several longitudinal infant cohorts, we observed the maturation of the infant gut microbiota over the first year of life in terms of taxonomic composition and diversity, with an increase in taxonomic alpha diversity and richness especially observed around the age of weaning ([Bäckhed et al., 2015](#); [Beller et al., 2021](#)). Consistent with previous studies ([Yatsunenکو et al., 2012](#); [Bäckhed et al., 2015](#)), the compositional microbiota maturation during the first year of life resulted in reduced beta diversity and global convergence in the gene content of the infant gut microbiota toward an adult-like composition. At the age of 1 year, the taxonomic composition of the infant gut microbiota was still distinct from that of the adult gut microbiota. Interestingly and contrary to the taxonomic maturation, the functional microbiota maturation was not characterized by an increase in gene family richness. This result also observed in previous studies ([Yatsunenکو et al., 2012](#); [Wang et al., 2021](#)), is likely due to a smaller fraction of sequences with assignable KEGG annotations, in particular in the adult samples, as a result of increased complexity. We did not observe any significant differences in the abundance of broad functional categories between the infant and mother samples or in the infant samples across the different time points. As reported previously ([Wang et al., 2021](#)), unlike taxonomic signatures, there is a high similarity between the microbial metabolic functions between infants and adults, underlining the importance of conserved core microbial functions.

We used a clustering approach to identify the main microbiota community type characteristics for each sampling time point. We applied hierarchical clustering on the Aitchinson distance, which allows us to take into account the compositionality of microbiome data and reduces effects due to sequencing depth differences between samples ([Galloway-Pena et al., 2017](#)). This method allowed us to identify two main infant microbiota clusters at each time point, driven in part by the relative abundance of the *Bacteroidaceae* family. Using cut-offs different from those in our study, Eck et al. reported comparable patterns, in particular describing two distinct infant gut microbiota settler types based on a cut-off of 30% of relative abundance of *Bacteroidetes* ([Eck et al., 2020](#)). A large meta-analysis covering > 10,000 microbiota samples from 17 countries and derived from children sampled between birth and 3 years identified four robust infant enterotypes typified by *Firmicutes*, *Bifidobacterium*, *Bacteroides*, and *Prevotella* ([Xiao et al., 2021](#)). While age was the strongest predictor of enterotype and all enterotypes included children from both birth modes, interestingly, the *Bacteroides*-dominated enterotype was detected in Northern European countries such as Finland, Norway, and Estonia in most sampling months. Hence, our results on Finnish infants may not be generalized across populations. In any case, many previous studies using the clustering approach have also reported the importance of *Bacteroidaceae* in infant or child gut microbiota datasets and its associations to geography, breastfeeding duration, butyrate

synthesis, birth mode, human milk oligosaccharide (HMO) metabolizing capacity, orthogonal or collateral relationship with clusters driven by other taxa, and developmental stages (Yatsunenkov et al., 2012; Stewart et al., 2018; Zhong et al., 2019; Berger et al., 2020; Casaburi et al., 2021; Roswall et al., 2021). Importantly, we observed the infant microbiota clusters to be associated with differences in several functional pathways, including pathways implicated in carbohydrate metabolism. In particular, degradation pathways of two glycans, β -mannans, and chondroitin sulfate were significantly associated with the *Bacteroidaceae*-rich cluster at all the time points. Chondroitin sulfate is a glycosaminoglycan prevalent in human milk and is associated with anti-inflammatory properties and protective effects in neonates (Knowles et al., 2021); it can be degraded by *Bacteroides* species (Shang et al., 2016). β -Mannans are complex plant-based sugars widespread in the human diet and can be degraded and used by *Bacteroides* species apart from some *Firmicutes* species (la Rosa et al., 2019). Overall, the results suggest that depletion in *Bacteroidaceae* contributes to a significant alteration in the ability of an infant's microbiota to metabolize complex sugars (Marcobal et al., 2011; Bäckhed et al., 2015; Casaburi et al., 2021). Our results expand on previous reports that have addressed the functional differences of infant gut microbiota in relation to CS birth and reduced abundance of *Bacteroides*. A previous study has reported that during the first 3 months of life, CSD infants have underrepresented biosynthetic pathways and that the vast majority of which (14/20) could be assigned to *Bacteroides* spp. (Wilson et al., 2021). A recent meta-analysis on infant metagenomes reported functions such as starch degradation, glycolysis, and queuosine biosynthesis to be enriched in a *Bacteroides*-dominated community type, which comprised 90% of VD infants (Xiao et al., 2021). In a recent randomized controlled trial, functional differences between CSD and VD infants were investigated on days 7 and 27, and the results indicated that 133 and 663 functional genes differed between the two birth groups, respectively (T. Dierikx et al., 2022). Wang et al. have also reported differences between metabolic functions in relation to birth mode, with functions such as vitamin, sugar, and cell wall biosynthesis higher in VD infants (Wang et al., 2021).

Strikingly, considering the infants that had samples available at all the time points, 54% of the infants clustering together at 3 weeks clustered together throughout all the sampling time points. This indicates the lasting influence of the pioneering community assemblance and aligns with previous reports about deterministic and largely predictable transition patterns between infant gut microbiota community types (Stewart et al., 2018; Xiao et al., 2021). The fact that more than half of the infants retained their characteristic high or low abundance of *Bacteroidaceae* allowed us to define three infant microbiota trajectories strongly associated with the birth mode. As expected, the "high *Bacteroidaceae* trajectory" was only observed in the VD infants, while the "low *Bacteroidaceae* trajectory" and

"intermediary trajectory," characterized by complete or milder depletion in *Bacteroidaceae* until 1 year of age, respectively, were observed not only in the CS born infants but also in the VD infants. While this phenomenon has not been well-discussed in the scientific literature, we are not the first to report it. Our results confirm previous observations that *Bacteroides* depletion is not only present in CS-born infants but can also be observed in 20–49% of VD infants (Yassour et al., 2016; Shao et al., 2019; Wilson et al., 2021). We further explored the plausible causes of *Bacteroides* depletion in VD infants. The depletion was prolonged and was only partially resolved after 1 year. While many studies (e.g., Shao et al., 2019) and a systematic review (T. H. Dierikx et al., 2020) have documented that intrapartum antibiotic exposures lead to depletion of *Bacteroides* in VD infants, we did not confirm such an association in our cohort. Similarly, Stearns et al. (2017) did not find IAP to decrease the relative abundance of *Bacteroides* among VD infants at 12 weeks, and a clinical trial where no VD infants were exposed to IAP found that 20% of their VD samples had a low *Bacteroides* signature (Wilson et al., 2021). Also, Yassour et al. observed a similar *Bacteroides* depletion profile in seven Finnish VD infants; however, with such a small subset of infants, they were unable to search for significant correlations with any clinical variables such as IAP (Yassour et al., 2016). Our results from the same cohort using absolute abundance estimates from 16S rRNA gene amplicon data (Jokela et al., In Press), as well as earlier targeted qPCR studies, identified no statistically significant differences in the absolute abundance of the *Bacteroides fragilis* group in VD infants in relation to IAP exposure (Aloisio et al., 2014; Corvaglia et al., 2016). In this study, the incidental courses of postnatal antibiotics did not explain the microbiota clustering or trajectories at any time point.

Despite extensive exposure and other metadata available for the study infants, we were not able to associate the differences in the relative abundance of *Bacteroides* in VD infants to any prenatal variable such as the mother's pre-pregnancy BMI or parity, or to perinatal and postnatal factors such as gestational age, stay in intensive care, time between water break and delivery, and infant weight at birth. Importantly, all the infants in our study were partially or exclusively breastfed until 3 months, with the majority of infants breastfed until 6 months, and no association was found between breastfeeding habits and the *Bacteroides* depletion observed. Similarly, an earlier study found *Bacteroides* depletion within the first 3 months of life not to be related to feeding mode or antibiotic exposure after birth (Wilson et al., 2021). We also comprehensively explored the possible effect of technical variations such as time of storage of samples before processing, extraction date, and sequencing run. Interestingly, we found DNA extraction yield per gram of feces to be significantly reduced in the *Bacteroides*-depleted VD infants. While the relative abundance of this genus is known to be sensitive to sample storage and DNA extraction methods (Salonen et al., 2010; Bahl et al., 2012; Rinninella et al., 2019;

Zhang et al., 2021), in our study, all the samples were frozen without delays with stabilization buffers and were extracted using the same DNA extraction protocol. Also, the sequencing libraries were normalized for the input DNA. Hence, unlike across different studies, we believe that within this study, the link between DNA yield and community composition is a biological effect rather than a technical effect. As in the fecal samples, the majority of the extracted DNA is from bacteria (Qin et al., 2010; Li et al., 2014; Sender et al., 2016; Matijašić et al., 2020), and our findings suggest that infants with *Bacteroides*-depleted community type may have a lower absolute abundance of bacteria in their stools. While outside of the scope of this study, this hypothesis could be explored by qPCR, potentially in conjunction with sequencing, to convert the relative abundance of microbiota profiles into estimates of absolute abundances (Jian et al., 2020).

Several studies have documented the prominent transmission of *Bacteroides* spp. from mothers to infants (Ferretti et al., 2018; Korpela et al., 2018, 2020; Wampach et al., 2018; Shao et al., 2019; Mitchell et al., 2020). Our team has also proven experimentally that maternal fecal microbiota transplant immediately after birth restores the levels of *Bacteroides* in CSD infants (Korpela et al., 2020). Strain-level metagenome studies have shown that maternal strains are more likely to persist than other bacteria and that the retention of *Bacteroides* is higher in the infant's gut compared to other genera, terming them as "persisters" (Ferretti et al., 2018; Korpela et al., 2018; Lou et al., 2021; Wang et al., 2021). To investigate a possible role of the mother-infant transmission of *Bacteroides* in the present or depleted-*Bacteroides* groups, we studied the effect of maternal fecal microbiota composition. Fecal samples were collected from the mothers in the 2 weeks preceding delivery, and no differences in terms of fecal microbiota composition, either overall or within the composition or abundance of *Bacteroides* spp., could be detected between the two groups of mothers.

Conclusion

Our study explored the gut microbiota developmental trajectories in healthy infants from 3 weeks to 1 year of age. Our results confirm previous reports indicating depletion of *Bacteroides* in CSD infants but also in a significant proportion of VD infants. Despite screening of extensive metadata, the cause of the depletion in VD infants still remains unresolved and would require larger infant cohorts where particularly the effect of different breastfeeding patterns could be investigated.

Data availability statement

The sequence data that support the findings of this study are available under the BioProject ID: PRJEB52774. The names

of the repository/repositories and accession number(s) can be found in the article/[Supplementary material](#).

Ethics statement

The studies involving human participants were reviewed and approved by the Ethical Committee of the Hospital District of Helsinki and Uusimaa. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

AS, KK, K-LK, and WMdV designed and established the cohort. AS and WMdV managed the cohort. ED performed the sample collection, management, and DNA extraction. DM, AP, and AS planned and carried out the computational experiments, analyzed the data, and wrote the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

Funding

This study was supported by grants from Tekes 329/31/2015 (WMdV), European Union's Horizon 2020 Research and Innovation Program H2020 MSCA (Sweet Crosstalk) project under grant agreement No. 814102 (AS), Academy of Finland 1325103 (AS), and 339172 (AP), Mary and Georg C. Ehrnrooth Foundation (AS), and Päivikki and Sakari Sohlberg foundation (AS). DM acknowledges the funding for Ph.D. received through European Union's H2020-MSCA-ITN-2018 Sweet Crosstalk project under grant agreement No. 814102.

Acknowledgments

The study nurses, especially Jaana Valkeapää, along with Heli Suomalainen, Anna Mantere, Eevi Heitto, and Janica Bergström are thanked for the recruitment and sample and data collection. Jaana Valkeapää and Susanna Gardemeister are acknowledged for extracting the hospital charts on the use of intrapartum antibiotics. We thank the Finnish IT Center for Science for providing the computational resources for this project. We thank Roosa Jokela, Emilia Lahtinen, and the other group members for their helpful discussions. We thank Tinja Kanerva for the sample and laboratory management along with Jessica Manngård, Prabin Koirala, and Sofia Rybas. PikkuJätti medical centers at Tapiola, Itäkeskus, and Myyrmäki are acknowledged for offering the premises for study visits and the personnel of

Women's Hospital at Jorvi, Kättilöopisto, and Naistenklinikka for their flexibility and support in the recruitment. We acknowledge Wellworks Ltd. for the development and maintenance of the interactive HELMi questionnaire system and database. We are truly grateful to the participating families for their efforts and commitment that made this study possible.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Aitchison, J., Barceló-Vidal, C., Martín-Fernández, J. A., and Pawłowsky-Glahn, V. (2000). Logratio Analysis and Compositional Distance. *Math. Geol.* 32, 271–275. doi: 10.1023/A:1007529726302
- Allnut, T. R., Roth-Schulze, A. J., and Harrison, L. C. (2021). Expanding the taxonomic range in the fecal metagenome. *BMC Bioinformatics* 22, 312. doi: 10.1186/s12859-021-04212-6
- Aloisio, I., Mazzola, G., Corvaglia, L. T., Tonti, G., Faldella, G., Biavati, B., et al. (2014). Influence of intrapartum antibiotic prophylaxis against group B Streptococcus on the early newborn gut composition and evaluation of the anti-Streptococcus activity of Bifidobacterium strains. *Appl. Microbiol. Biotechnol.* 98, 6051–6060. doi: 10.1007/s00253-014-5712-9
- Arbolea, S., Suárez, M., Fernández, N., Mantecón, L., Solís, G., Gueimonde, M., et al. (2018). C-section and the neonatal gut microbiome acquisition: consequences for future health. *Ann. Nutr. Metab.* 73, 17–23. doi: 10.1159/000490843
- Avershina, E., Solís Sánchez, G., Griselda Binetti, A., Li, Z., Zhang, C., Li, L., et al. (2021). The effects of delivery mode on the gut microbiota and health: state of art. *Front. Microbiol.* 12, 724449. doi: 10.3389/fmicb.2021.724449
- Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., et al. (2015). Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe* 17, 690–703. doi: 10.1016/j.chom.2015.04.004
- Bahl, M. I., Bergström, A., and Licht, T. R. (2012). Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis. *FEMS Microbiol. Lett.* 329, 193–197. doi: 10.1111/j.1574-6968.2012.02523.x
- Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., et al. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife* 10, e65088. doi: 10.7554/eLife.65088.sa2
- Beller, L., Deboutte, W., Falony, G., Vieira-Silva, S., Tito, R. Y., Valles-Colomer, M., et al. (2021). Successional stages in infant gut microbiota maturation. *MBio* 12, e0185721. doi: 10.1128/mbio.01857-21
- Benoit, G., Peterlongo, P., Mariadassou, M., Drezon, E., Schbath, S., Lavenier, D., et al. (2016). Multiple comparative metagenomics using multiset *k*-mer counting. *PeerJ. Comput. Sci.* 2, e94. doi: 10.7717/peerj-cs.94
- Berger, B., Porta, N., Foata, F., Grathwohl, D., Delley, M., Moine, D., et al. (2020). Linking human milk oligosaccharides, infant fecal community types, and later risk to require antibiotics. *MBio* 11, e03196-19. doi: 10.1128/mBio.03196-19
- Bokulich, N. A., Chung, J., Battaglia, T., Henderson, N., Jay, M., Li, H., et al. (2016). Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* 8, 343ra82. doi: 10.1126/scitranslmed.aad7121
- Borewicz, K., Gu, F., Saccenti, E., Arts, I. C., Penders, J., Thijs, C., et al. (2019). Correlating infant fecal microbiota composition and human milk oligosaccharide consumption by microbiota of 1-month-old breastfed infants 1801214 (1 of 13). *Mol. Nutr. Food Res.* 63, 1801214. doi: 10.1002/mnfr.201801214
- Brunson, J. (2020). ggalluvial: layered grammar for alluvial plots. *J. Open Source Softw.* 5, 2017. doi: 10.21105/joss.02017
- Busi, S. B., de Nies, L., Habier, J., Wampach, L., Fritz, J., v., et al. (2021). Persistence of birth mode-dependent effects on gut microbiome composition, immune system stimulation and antimicrobial resistance during the first year of life. *ISME Commun.* 1, 8. doi: 10.1038/s43705-021-00003-5
- Casaburi, G., Duar, R. M., Brown, H., Mitchell, R. D., Kazi, S., Chew, S., et al. (2021). Metagenomic insights of the infant microbiome community structure and function across multiple sites in the United States. *Sci. Rep.* 11, 1472. doi: 10.1038/s41598-020-80583-9
- Corvaglia, L., Tonti, G., Martini, S., Aceti, A., Mazzola, G., Aloisio, I., et al. (2016). Influence of intrapartum antibiotic prophylaxis for group B Streptococcus on Gut microbiota in the first month of life. *J. Pediatr. Gastroenterol. Nutr.* 62, 304–308. doi: 10.1097/MPG.0000000000000928
- Dierikx, T., Berkhout, D., Eck, A., Tims, S., van Limbergen, J., Visser, D., et al. (2022). Influence of timing of maternal antibiotic administration during caesarean section on infant microbial colonisation: A randomised controlled trial. *Gut* 71, 1803–1811. doi: 10.1136/gutjnl-2021-324767
- Dierikx, T. H., Visser, D. H., Benninga, M. A., van Kaam, A. H. L. C., de Boer, N. K. H., de Vries, R., et al. (2020). The influence of prenatal and intrapartum antibiotics on intestinal microbiota colonisation in infants: a systematic review. *J. Infect.* 81, 190–204. doi: 10.1016/j.jinf.2020.05.002
- Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Veget. Sci.* 14, 927–930. doi: 10.1111/j.1654-1103.2003.tb02228.x
- Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., et al. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Nat. Acad. Sci. U. S. A.* 107, 11971–11975. doi: 10.1073/pnas.1002601107
- Eck, A., Rutten, N. B. M. M., Singendonk, M. M. J., Rijkers Id, G. T., Savelkoul, P. H. M., Meijssen, C. B., et al. (2020). Neonatal microbiota development and the effect of early life antibiotics are determined by two distinct settler types. *PLoS ONE* 15, e022813. doi: 10.1371/journal.pone.0228133
- Ferretti, P., Pasolli, E., Tett, A., Asnicar, F., Gorfer, V., Fedi, S., et al. (2018). Mother-to-infant microbial transmission from different body sites shapes the developing infant gut microbiome. *Cell Host Microbe* 24, 133–145.e5. doi: 10.1016/j.chom.2018.06.005
- Galloway-Pena, J., Stearns, J., Gloor, G. B., Macklaim, J. M., Pawłowsky-Glahn, V., and Egozcue, J. J. (2017). Microbiome datasets are compositional: and this is not optional. *Front. Microbiol.* 8, 2224. doi: 10.3389/fmicb.2017.02224

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.953475/full#supplementary-material>

- Guittar, J., Shade, A., and Litchman, E. (2019). Trait-based community assembly and succession of the infant gut microbiome. *Nat. Commun.* 10, 512. doi: 10.1038/s41467-019-08377-w
- Hiseni, P., Rudi, K., Wilson, R. C., Hegge, F. T., and Snipen, L. (2021). HumGut: a comprehensive human gut prokaryotic genomes collection filtered by metagenome data. *Microbiome* 9, 165. doi: 10.1186/s40168-021-01114-w
- Jian, C., Luukkonen, P., Yki-Jä Rvinen, H., Salonen, A., and Korpela, K. (2020). Quantitative PCR provides a simple and accessible method for quantitative microbiota profiling. *PLoS ONE* 15, e0227285. doi: 10.1371/journal.pone.0227285
- Jokela, R., Korpela, K., Jian, C., Dikareva, E., Nikkonen, A., Saisto, T., et al. (In Press). Quantitative insights into effects of intrapartum antibiotics and birth mode on infant gut microbiota in relation to well-being during the first year of life. *Gut Microbes*. doi: 10.1101/2021.11.01.21265735
- Knowles, T. A., Hosfield, B. D., Pecoraro, A. R., Li, H., Shelley, W. C., and Markel, T. A. (2021). It's all in the milk: chondroitin sulfate as potential preventative therapy for necrotizing enterocolitis. *Pediatr. Res.* 89, 1373–1379. doi: 10.1038/s41390-020-01125-7
- Korpela, K. (2021). Impact of delivery mode on infant gut microbiota. *Rev. Article Ann. Nutr. Metab.* 77, 11–19. doi: 10.1159/000518498
- Korpela, K., Costea, P., Coelho, L. P., Kandels-Lewis, S., Willemsen, G., Boomsma, D. I., et al. (2018). Selective maternal seeding and environment shape the human gut microbiome. *Genome Res.* 28, 561–568. doi: 10.1101/gr.233940.117
- Korpela, K., and de Vos, W. M. (2018). Early life colonization of the human gut: microbes matter everywhere. *Curr. Opin. Microbiol.* 44, 70–78. doi: 10.1016/j.mib.2018.06.003
- Korpela, K., Dikareva, E., Hanski, E., Kolho, K.-L., de Vos, W. M., and Salonen, A. (2021). Cohort profile: Finnish health and early life microbiota (HELMi) longitudinal birth cohort. *BMJ Open* 9, e028500. doi: 10.1136/bmjopen-2018-028500
- Korpela, K., Helve, O., Kolho, K.-L., Saisto, T., Skogberg, K., Dikareva, E., et al. (2020). Maternal fecal microbiota transplantation in cesarean-born infants rapidly restores normal gut microbial development: a proof-of-concept study. *Cell* 183, 324–334.e5. doi: 10.1016/j.cell.2020.08.047
- Kumbhare, S., v. Patangia, D. V., v. Patil, R. H., Shouche, Y. S., and Patil, N. P. (2019). Factors influencing the gut microbiome in children: from infancy to childhood. *J. Biosci.* 44, 49. doi: 10.1007/s12038-019-9860-z
- la Rosa, S. L., Leth, M. L., Michalak, L., Hansen, M. E., Pudlo, N. A., Glowacki, R., et al. (2019). The human gut Firmicute *Roseburia intestinalis* is a primary degrader of dietary β -mannans. *Nat. Commun.* 10, 905. doi: 10.1038/s41467-019-08812-y
- Lahti, L., and Shetty, S. A. (2012–2019). *Microbiome R Package: Tools for Microbiome Analysis in R*. Available online at: <https://github.com/microbiome/microbiome>
- Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., et al. (2014). An integrated catalog of reference genes in the human gut microbiome. *Nat. Biotechnol.* 32, 834–841. doi: 10.1038/nbt.2942
- Lou, Y. C., Olm, M. R., Diamond, S., Crits-Christoph, A., Firek, B. A., Baker, R., et al. (2021). Infant gut strain persistence is associated with maternal origin, phylogeny, and traits including surface adhesion and iron acquisition. *Cell Rep. Med.* 2, 100393. doi: 10.1016/j.xcrm.2021.100393
- Lu, J., Breitwieser, F. P., Thielen, P., and Salzberg, S. L. (2017). Bracken: estimating species abundance in metagenomics data. *PeerJ. Comput. Sci.* 3, e104. doi: 10.7717/peerj-cs.104
- Mallick, H., Rahnavard, A., McIver, L. J., Ma, S., Zhang, Y., Nguyen, L. H., et al. (2021). Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput. Biol.* 17, e1009442. doi: 10.1371/journal.pcbi.1009442
- Marcobal, A., Barboza, M., Froehlich, J. W., Block, D. E., German, J. B., Lebrilla, C. B., et al. (2010). Consumption of human milk oligosaccharides by gut-related microbes. *J. Agric. Food Chem.* 58, 5334–5340. doi: 10.1021/jf9044205
- Marcobal, A., Barboza, M., Sonnenburg, E. D., Pudlo, N., Martens, E. C., Desai, P., et al. (2011). Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways. *Cell Host Microbe* 10, 507–514. doi: 10.1016/j.chom.2011.10.007
- Matijašić, M., Meštrović, T., Paljetak, H. C., Perić, M., Barešić, A., and Verbanac, D. (2020). Gut microbiota beyond bacteria-mycobiome, virome, archaeome, and eukaryotic parasites in IBD. *Int. J. Mol. Sci.* 21, 2668. doi: 10.3390/ijms21082668
- Milani, C., Duranti, S., Bottacini, F., Casey, E., Turrioni, F., Mahony, J., et al. (2017). The first microbial colonizers of the human gut: composition, activities, and health implications of the infant gut microbiota. *Microbiol. Mol. Biol. Rev.* 81, e00036-17. doi: 10.1128/MMBR.00036-17
- Mitchell, C. M., Mazzoni, C., Hogstrom, L., Bryant, A., Bergerat, A., Cher, A., et al. (2020). Delivery mode affects stability of early infant gut microbiota. *Cell Rep. Med.* 1, 100156. doi: 10.1016/j.xcrm.2020.10.0156
- Mueller, N. T., Bakacs, E., Combellick, J., Grigoryan, Z., and Dominguez-Bello, M. G. (2015). The infant microbiome development: mom matters HHS Public Access. *Trends Mol. Med.* 21, 109–117. doi: 10.1016/j.molmed.2014.12.002
- Penders, J., Thijs, C., Vink, C., Stelma, F. F., Snijders, B., Kummeling, I., et al. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* 118, 511–521. doi: 10.1542/peds.2005-2824
- Princival, L., Rebelo, F., Williams, B. L., Coimbra, A. C., Crovesy, L., Ferreira, A. L., et al. (2021). Association between the mode of delivery and infant gut microbiota composition up to 6 months of age: a systematic literature review considering the role of breastfeeding. *Nutr. Rev.* 80, 113–127. doi: 10.1093/nutrit/nuab008
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: 10.1038/nature08821
- Reyman, M., van Houten, M. A., van Baarle, D., Bosch, A. A. T. M., Man, W. H., Chu, M. L. J. N., et al. (2019). Author correction: impact of delivery mode-associated gut microbiota composition on health in the first year of life. *Nat. Commun.* 10, 5352. doi: 10.1038/s41467-019-13373-1
- Rinninella, E., Raoul, P., Cintoni, M., Franceschi, F., Miggiano, G., Gasbarrini, A., et al. (2019). What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases. *Microorganisms* 7, 14. doi: 10.3390/microorganisms7010014
- Roswall, J., Olsson, L. M., Kovatcheva-Datchary, P., Nilsson, S., Tremaroli, V., Simon, M. C., et al. (2021). Developmental trajectory of the healthy human gut microbiota during the first 5 years of life. *Cell Host Microbe* 29, 765–776.e3. doi: 10.1016/j.chom.2021.02.021
- Salonen, A., Nikkilä, J., Jalanka-Tuovinen, J., Immonen, O., Rajilić-Stojanović, M., Kekkonen, R. A., et al. (2010). Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *J. Microbiol. Methods* 81, 127–134. doi: 10.1016/j.mimet.2010.02.007
- Sarkar, A., Yoo, J. Y., Dutra, S. V. O., Morgan, K. H., and Groer, M. (2021). The association between early-life gut microbiota and long-term health and diseases. *J. Clin. Med.* 10, 1–24. doi: 10.3390/jcm10030459
- Sender, R., Fuchs, S., and Milo, R. (2016). Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.* 14, e1002533. doi: 10.1371/journal.pbio.1002533
- Shang, Q., Yin, Y., Zhu, L., Li, G., Yu, G., and Wang, X. (2016). Degradation of chondroitin sulfate by the gut microbiota of Chinese individuals. *Int. J. Biol. Macromol.* 86, 112–118. doi: 10.1016/j.ijbiomac.2016.01.055
- Shao, Y., Forster, S. C., Tsaliki, E., Vervier, K., Strang, A., Simpson, N., et al. (2019). Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* 574, 117–121. doi: 10.1038/s41586-019-1560-1
- Stearns, J. C., Simioni, J., Gunn, E., McDonald, H., Holloway, A. C., Thabane, L., et al. (2017). Intrapartum antibiotics for GBS prophylaxis alter colonization patterns in the early infant gut microbiome of low risk infants. *Sci. Rep.* 7, 16527. doi: 10.1038/s41598-017-16606-9
- Stewart, C. J., Ajami, N. J., O'Brien, J. L., Hutchinson, D. S., Smith, D. P., Wong, M. C., et al. (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* 562, 583–588. doi: 10.1038/s41586-018-0617-x
- Tamames, J., Cobo-Simón, M., and Puente-Sánchez, F. (2019). Assessing the performance of different approaches for functional and taxonomic annotation of metagenomes. *BMC Genomics* 20, 960. doi: 10.1186/s12864-019-6289-6
- Tamburini, S., Shen, N., Chih Wu, H., and Clemente, J. C. (2016). The microbiome in early life: implications for health outcomes. *Nat. Med.* 22, 713–722. doi: 10.1038/nm.4142
- Tsilimigras, M. C. B., and Fodor, A. A. (2016). Compositional data analysis of the microbiome: fundamentals, tools, and challenges. *Ann. Epidemiol.* 26, 330–335. doi: 10.1016/j.annepidem.2016.03.002
- Tun, H. M., Peng, Y., Chen, B., Konya, T. B., Morales-Lizcano, N., adia., P., et al. (2021). Ethnicity associations with food sensitization are mediated by gut microbiota development in the first year of life. *Gastroenterology* 161, 94–106. doi: 10.1053/j.gastro.2021.03.016
- Vatanen, T., Kostic, A. D., d'Hennezel, E., Siljander, H., Franzosa, E. A., Yassour, M., et al. (2016). Variation in microbiome LPS immunogenicity contributes to autoimmunity in humans. *Cell* 165, 842–853. doi: 10.1016/j.cell.2016.04.007
- Vu, K., Lou, W., Tun, H. M., Konya, T. B., Morales-Lizcano, N., Chari, R. S., et al. (2021). From birth to overweight and atopic disease: multiple and common

pathways of the infant gut microbiome. *Gastroenterology* 160, 128–144.e10. doi: 10.1053/j.gastro.2020.08.053

Wampach, L., Heintz-Buschart, A., Fritz, J., v., Ramiro-Garcia, J., Habier, J., et al. (2018). Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential. *Nat. Commun.* 9, 5091. doi: 10.1038/s41467-018-07631-x

Wang, S., Zeng, S., Egan, M., Cherry, P., Strain, C., Morais, E., et al. (2021). Metagenomic analysis of mother-infant gut microbiome reveals global distinct and shared microbial signatures. *Gut Microbes* 13. doi: 10.1080/19490976.2021.1911571

Wexler, H. M. (2007). Bacteroides: the good, the bad, and the nitty-gritty. *Clin. Microbiol. Rev.* 20, 593–621. doi: 10.1128/CMR.00008-07

Wilson, B. C., Butler, É. M., Grigg, C. P., Derraik, J. G. B., Chiavaroli, V., Walker, N., et al. (2021). Oral administration of maternal vaginal microbes at birth to restore gut microbiome development in infants born by caesarean section: a pilot randomised placebo-controlled trial. *EBioMed.* 69, 103443. doi: 10.1016/j.ebiom.2021.103443

Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biol.* 20, 257. doi: 10.1186/s13059-019-1891-0

Xiao, L., Wang, J., Zheng, J., Li, X., and Zhao, F. (2021). Deterministic transition of enterotypes shapes the infant gut microbiome at an early age. *Genome Biol.* 22, 243. doi: 10.1186/s13059-021-02463-3

Yang, R., Gao, R., Cui, S., Zhong, H., Zhang, X., Chen, Y., et al. (2019). Dynamic signatures of gut microbiota and influences of delivery and feeding

modes during the first 6 months of life. *Physiol. Genomics* 51, 368–378. doi: 10.1152/physiolgenomics.00026.2019

Yassour, M., Vatanen, T., Siljander, H., Hämäläinen, A.-M., Härkönen, T., Ryhänen, S. J., et al. (2016). Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. *Science Transl. Med.* 8, 343ra81. doi: 10.1126/scitranslmed.aad0917

Yatsunencko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., et al. (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 222–227. doi: 10.1038/nature11053

Ye, S. H., Siddle, K. J., Park, D. J., and Sabeti, P. C. (2019). Benchmarking metagenomics tools for taxonomic classification. *Cell* 178, 779–794. doi: 10.1016/j.cell.2019.07.010

Zafar, H., and Saier, M. H. (2021). Gut Bacteroides species in health and disease. *Gut Microbes* 13(, 1–20. doi: 10.1080/19490976.2020.1848158

Zhang, B., Brock, M., Arana, C., Dende, C., van Oers, N. S., Hooper, L., et al. (2021). Impact of bead-beating intensity on the genus- and species-level characterization of the gut microbiome using amplicon and complete 16S rRNA gene sequencing. *Front. Cell. Infect. Microbiol.* 11, 678522. doi: 10.3389/fcimb.2021.678522

Zhong, H., Penders, J., Shi, Z., Ren, H., Cai, K., Fang, C., et al. (2019). Impact of early events and lifestyle on the gut microbiota and metabolic phenotypes in young school-age children. *Microbiome* 7, 2. doi: 10.1186/s40168-018-0608-z