Taylor & Francis
Taylor & Francis Group

Gut Microbes

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/kgmi20

Exclusive breastfeeding is associated with the gut microbiome maturation in infants according to delivery mode

Nathalia F. Naspolini, Paulo A. Schüroff, Pedro A. R. Vanzele, Davi Pereira-Santos, Tamires Amabili Valim, Kevin S. Bonham, André Fujita, Maria Rita Passos-Bueno, Patricia C. B. Beltrão-Braga, André C. P. L. F. Carvalho, Vanja Klepac-Ceraj, Guilherme V. Polanczyk, Alline C. Campos & Carla R. Taddei

To cite this article: Nathalia F. Naspolini, Paulo A. Schüroff, Pedro A. R. Vanzele, Davi Pereira-Santos, Tamires Amabili Valim, Kevin S. Bonham, André Fujita, Maria Rita Passos-Bueno, Patricia C. B. Beltrão-Braga, André C. P. L. F. Carvalho, Vanja Klepac-Ceraj, Guilherme V. Polanczyk, Alline C. Campos & Carla R. Taddei (2025) Exclusive breastfeeding is associated with the gut microbiome maturation in infants according to delivery mode, Gut Microbes, 17:1, 2493900, DOI: 10.1080/19490976.2025.2493900

To link to this article: https://doi.org/10.1080/19490976.2025.2493900

GUT MICROBES
2025, VOL. 17, NO. 1, 2493900
https://doi.org/10.1080/19490976.2025.2493900

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

OPEN ACCESS Check for updates

Exclusive breastfeeding is associated with the gut microbiome maturation in infants according to delivery mode

Nathalia F. Naspolini[a*], Paulo A. Schüroff[b*], Pedro A. R. Vanzele[c*], Davi Pereira-Santos[d,e], Tamires Amabili Valim[f], Kevin S. Bonham[g], André Fujita[h,i], Maria Rita Passos-Bueno[j], Patricia C. B. Beltrão-Braga[a,k], André C. P. L. F. Carvalho[d], Vanja Klepac-Ceraj[g], Guilherme V. Polanczyk[l], Alline C. Campos[f], and Carla R. Taddei[a,m]

[a]Institute of Biomedical Sciences, University of Sao Paulo, Sao Paulo, Brazil; [b]School of Arts, Sciences and Humanity, University of Sao Paulo, Sao Paulo, Brazil; [c]Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of Sao Paulo, Sao Paulo, Brazil; [d]Department of Applied Mathematics and Statistics, Institute of Mathematics and Computer Sciences, University of Sao Paulo, Sao Carlos, Brazil; [e]Departamento Acadêmico de Computação, Universidade Tecnológica Federal do Paraná (UTFPR), Câmpus Medianeira, Medianeira, Brazil; [f]Department of Pharmacology, Ribeirao Preto Medical School, University of Sao Paulo, Sao Paulo, Brazil; [g]Department of Biological Sciences, Wellesley College, Wellesley, MA, USA; [h]Division of Network AI Statistics, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan; [i]Department of Computer Science, Institute of Mathematics and Statistics, University of Sao Paulo, Sao Paulo, Brazil; [j]Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of Sao Paulo, Sao Paulo, Brazil; [k]Laboratory of Disease Modeling, Institut Pasteur de Sao Paulo, Sao Paulo, Brazil; [l]Department of Psychiatry, Faculdade de Medicina FMUSP, University of Sao Paulo, Sao Paulo, Brazil; [m]Division of Clinical Laboratory, University Hospital - University of Sao Paulo, Sao Paulo, Brazil

ABSTRACT

Exclusive breastfeeding (EBF) plays a crucial role in infant gut microbiome assembly and development. However, few studies have investigated the effects of EBF in restoring a perturbed microbiome. In this study, we applied whole metagenomic sequencing to assess the gut microbiome assembly in 525 Brazilian infants from 3 to 9 months of age of the Germina Cohort, demonstrating the early determinants of microbial taxonomy and function modulation. Our analysis shows that EBF alters the relative abundance of genes related to the microbiome taxonomy and function, with effects varying by delivery mode. EBF alters the pattern of carbohydrates, lipid metabolism, and cell structure pathways depending on the delivery mode. The microbiome age is closer to chronological infant age in EBF than in non-EBF infants, meaning a lower microbiome maturation index (MMI). Using a complementary machine learning approach, we show that *Escherichia coli*, *Ruminococcus gnavus*, and *Clostridium neonatale*, as well as vitamin K and o-antigen pathways contribute strongly to EBF prediction. Moreover, EBF influences the microbiome maturation in early life, toward a microbiome age more similar to the chronological infant's age.

ARTICLE HISTORY
Received 23 October 2024
Revised 29 January 2025
Accepted 8 April 2025

KEYWORDS
Early-life gut microbiome; microbiome functional pathways; exclusive breastfeeding

Background

Early assembly and development of the gut microbiome is a complex process and is influenced by maternal factors,[1] preterm delivery,[2] delivery mode,[3] breastfeeding, complementary feeding,[4] and antibiotic use,[5] among others. Previous studies focused on the effect of breastfeeding in driving infant microbial composition, the so-called milk-oriented microbiota,[6,7] however, the mechanisms behind exclusive breastfeeding (EBF) in shaping a perturbed microbiome, for example, as a result of Cesarean delivery (C-section), have received less attention.

The microbial community and function of EBF infants differ from those in mixed-fed infants, regardless of delivery mode and infant age.[8,9] Breastfeeding and delivery mode are essential drivers of gut microbiota composition.[10] Furthermore, breastfeeding has the potential to minimize the effect of C-section delivery on the infant's gut microbiome.[8,11] Nevertheless, the current knowledge about this theme is limited to the characterization of microbial taxa and lacks information regarding microbial gene function and trajectories.[4,12] One of the protective effects of

CONTACT Carla R. Taddei crtaddei@usp.br Av. Prof Lineu Prestes, 1374 - Prédio Didáticas Laboratório de Microbioma Humano Departamento de Microbiologia, Instituto de Ciências Biomédicas (ICB-II), Universidade de São Paulo, Sao Paulo 05508-000, Brazil

*These authors contributed equally to this work and shared first authorship.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/19490976.2025.2493900.

breastfeeding on the microbiome of C-section-delivered infants is mediated by α1–2 fucosylated human milk oligosaccharides, and mothers that produce more of these sugars in their breast milk are designated "secretors".[1] C-section-delivered infants born to secretor mothers exhibit higher abundances of *Bifidobacterium breve* and *B. bifidum* and lower abundances of *Enterococcus lactis* and *E. faecalis*. While maternal secretor status did not affect the microbiota composition in vaginally delivered infants, it was hypothesized that maternal secretory capacity might be an essential factor, especially among infants born by C-section.[1] A similar study in Brazil supported those findings and showed an increased abundance of *Akkermansia* only in C-section-delivered infants of maternal secretors.[13]

Considerable efforts have focused on studying infants' gut microbiomes and their early determinants, such as delivery mode or breastfeeding.[14] A study by Bäckhed et al. demonstrated that the cessation of breastfeeding leads to a shift toward a more adult-like microbiota composition.[15] Recently, Shenhav et al. identified microbiome colonization patterns which, in conjunction with milk components, were associated with breastfeeding and predictive of asthma.[7] Selma-Royo et al. reported infant gut microbiome differences according to birth settings (hospital- or home-born), microbiome vertical transmission and the influence of breastfeeding on *B. longum* colonization and microbial functional diversity.[16] Other available studies are restricted to small sample sizes and/or limited methodologies to investigate microbial mechanisms, such as 16S rRNA amplicon sequencing.[4,12,17] The role of exclusive breastfeeding in shaping microbial assembly remains to be fully elucidated, requiring confirmation across diverse human populations and further investigation into the underlying mechanisms.

Evidence of the effects of exclusive breastfeeding on the assembly of the infant gut microbiome may provide insights into how exclusive breastfeeding influences microbiome changes in perturbed microbiome colonization, as observed in C-section-delivered infants,[3,18] toward a microbiome more similar to that of vaginally delivered infants. In this article, we evaluate the early determinants of the succession and establishment of the fecal microbiome from 3 to 9 months of age in a large cohort of 525 infants enrolled in the German Cohort (https://www.projetogermina.com.br/).[19] Employing metagenomic sequencing and machine learning approach, we suggest that EBF is associated with microbial composition, function, and trajectory during the first year of life according to delivery mode.

Methods

Study design and population

The Germina Project[19] is an ongoing prospective cohort designed to assess mother-infant dyads across the first 1,000 days of life, from 3 months to 36 months of age. Families with 3-month-old infants in the metropolitan area of São Paulo (a city in the Southeast of Brazil) were recruited with the following eligibility criteria: maternal age (20-−45 years old); infant age (3–4 months old); gestational age at birth (>37 weeks); infant birth weight >2,000 g; no substance abuse during pregnancy; no history of psychosis or bipolar disorder; no childbirth complications (e.g. perinatal asphyxia, shoulder dystocia, excessive bleeding); infant not previously diagnosed with genetic syndromes or auditory/visual deficiencies; and availability for in-person appointments/assessments.

Sample collection and processing

Fecal samples were collected at 3 and 5–9 months of age from a diaper during follow-up visits by trained personnel using sterile collector tubes (Sarstedt, Germany) and stored at −20°C for a few hours until −80°C storage. When the above procedures were not possible, mothers were instructed to collect the stool samples at home and store them at −20°C for a few hours until the Germina specialized service could pick them up for subsequent storage at −80°C freezers. Samples were collected between December 2021 and May 2023. Genomic DNA was extracted from stool samples using the ZymoBIOMICS™ DNA Miniprep Kit (Zymo Research, USA), according to the manufacturer's instructions. The ZymoBIOMICS® Microbial Community Standard (ZMCS, Catalog D6300, Zymo Research) was used as an extraction and

sequencing positive control and was included in each sequencing run. Negative controls were not included in this study because fecal samples have a very high microbial load, which overwhelms minor contaminants.

Metagenome library preparation and sequencing

Libraries were prepared using the genomic DNA (100 ng) and the Illumina DNA Prep kit (Illumina) quantified by qPCR using the KAPA Library Quantification Kit (Roche). Paired-end reads (2 × 150 bp) were sequenced in a P3 flow cell on the Illumina NextSeq 2000, under the technological services and expertise of the NGS Soluções Genômicas (Piracicaba, São Paulo-Brazil).

Bioinformatics

Bioinformatic tools from bioBakery workflows[20] were used for meta-omics data analyses. The pipelines include: i) KneadData (v.0.10.0) was used for quality control of metagenomic sequencing data using default parameters. Human reads were filtered by aligning to the *Homo sapiens* Bowtie 2 hg37 database; ii) MetaPhlAn (v3.1.0, database "mpa_v31_CHOCOPhlAn_201901") was employed for bacterial taxonomic composition; and iii) HUMAnN (v3.6) was applied to generate pathway and gene abundance annotation using the MetaCyc and UniRef90 databases, respectively, with a subsequent normalization to relative abundance at community level. We also used the HUMAnN (v3.6) outputs at species level to assess the contribution of each species to pathway abundance, measured in reads per kilobase (RPKs). Then, we normalized the species contribution, considering the sum of the total community for each pathway to be 100% and the species stratification was made against that total.

Statistical analysis

Alpha diversity was measured using the Shannon index and Chao1. The Bray-Curtis dissimilarity was calculated to obtain pairwise beta diversity. Both diversity measures were calculated using the *Phyloseq* R package.[21] Differences in the Bray-Curtis matrix were detected using Permutational Multivariate Analysis of Variance (PERMANOVA) with the *vegan* R package,[22] tested with 999 permutations, and visualized with principal coordinates analysis (PCoA). This analysis was used to test significant differences in several variables, including maternal age, maternal BMI pregestational, infant antibiotic use, and birth weight. Analysis of variance (ANOVA) was used to compare alpha diversity and the relative abundance of species and pathways between groups at each time point, with Tukey's 'Honest Significant Difference' (HSD) post-hoc correction, a significance threshold of p-value ≤0.05.

For the subsequent analysis, a threshold prevalence filtering of 5% for species and functional pathways and an abundance filter of 0.1% for species and 3×10^{-7}% for pathways were applied based on previous publication.[23] The *Maaslin2* R package[24] was employed to detect overrepresented/depleted bacterial features related to EBF and delivery mode within each time point. All p-values were corrected for false discovery rate (FDR) at a significance threshold of q-value <0.10. Further, we described the unique and shared altered features in those infant groups. Finally, the top ten species contributing to each of the significantly altered carbohydrate pathways from Maaslin2 results were described for each delivery mode. In the analysis of the 5–7 and 8–9 month time points, the early introduction of solid foods (before 6 months of age), formula feeding and infant age (in days) were included as covariates. All analyses were conducted in R version 4.3.0.

Machine learning model

We aimed to model taxonomic and functional microbiome features as predictors of EBF using a Random Forest classifier with a hundred trees. The rationale for selecting Random Forest is its robustness against overfitting and its ability to handle high-dimensional data,[25,26] making it particularly well-suited for the complex and heterogeneous nature of microbiome data. The model targeted two labels, EBF = 0 and EBF = 1, enabling us to capture the binary outcome of interest effectively. We stratified infants by time points and delivery modes and further categorized them into functional pathways and species, resulting in 12

distinct models. This stratification was crucial for capturing potential variations in microbiome features influenced by these factors,[27] enabling a more detailed understanding of their predictive capabilities.

The number of trees in the Random Forest was set to the maximum feasible value within our computational resources and time constraints, ensuring an adequate number of trees were generated for robust model performance. Preliminary tests revealed that inducing fewer than 100 trees already achieved stability in classical error measures, with no further reduction in error as additional trees were added. This decision was crucial, as both the permutation significance test[28] and the calculation of SHAP values[29] are computationally demanding processes. The permutation test, which involved 10,000 evaluations on shuffled target values, was used to compute an empirical p-value against the null hypothesis (H0) that the model cannot find any dependency between the microbiome features and the target. This provided a rigorous method to assess the statistical significance of our model's findings. SHAP values were employed to determine the marginal contribution of each microbiome feature to predict EBF = 1 for each infant.[30]

The Leave-One-Out Cross-Validation (LOOCV) scheme[31] was selected for model validation to maximize the use of available data, ensuring that each training set – comprising all infants' data except the one used for prediction – was as comprehensive as possible. This approach was particularly important given the small sample sizes, which ranged from 57 to 361 depending on breastfeeding type and delivery mode.

We kept the default "number of random paths" setting (25) in the SHAP library to balance computational efficiency with the reliability of variable attributions.[30] Results from models with a p-value <0.05 for balanced accuracy were considered significant, ensuring that only robust predictors were reported. To enhance interpretability, we introduced a target-oriented SHAP value (SHAP*), defined as the SHAP value multiplied by −1 if it contributed against the target (i.e., EBF = 0 or EBF = 1). Predictors with a SHAP* p-value ≥0.01 were excluded from the resulting figure to streamline the visual presentation.

For trajectory modeling, we used the Microbiome Toolbox software, a web-based collection of tools and methods for microbiome data analysis, to study longitudinal microbiome progression in infant cohorts.[32] The MMI (Microbiome Maturation Index) was defined using the approach of Subramanian et al.[33] This method employed a machine learning approach to analyze 16S rRNA sequencing data, using fecal samples from infants with consistent healthy growth to define a benchmark for healthy gut microbiome maturation. Briefly, the predictive accuracy of the trajectory model was evaluated using a range of machine learning techniques, including Random Forest, Gradient Boosting, and a custom-developed deep learning model tailored to the dataset. This multi-model approach facilitated the selection of the most suitable algorithm for each dataset, ensuring reliable predictions and highlighting the key factors influencing microbiome maturation within infant cohorts. All infant samples and feature columns for taxonomy or functional pathways were used to train the Random Forest regression model to generate the reference trajectory. The infants were categorized into four groups based on their mode of delivery (vaginal or cesarean) and feeding practices (exclusive breastfeeding or non-exclusive breastfeeding), and the trajectories of gut microbiome maturation were analyzed across these groups.

Results

Study population characteristics

The Germina cohort enrolled 560 infants at the first time point (T1 = 3 months), and 532 attended the second appointment (T2 = 5–9 months). A total of 552 fecal samples were collected at the first and 495 at the second appointment, of which 525 (93%) and 441 (89%) underwent WGS, respectively (Supplementary Figure S1). Infants with WGS data available were then categorized as exclusive breastfeed (EBF, n = 362) and non-exclusive breastfeed (non-EBF, n = 161) if whether they were EBF for at least 3 months or not (the time of first sample collection), and according to delivery mode within each EBF category (vaginally delivered-EBF, n = 188; C-section delivered-EBF, n =

174; vaginally delivered-non-EBF, n = 58; C-section delivered-non-EBF, n = 103). Since our goal was to estimate the effect of exclusive breastfeeding, any infant not included in this group was classified as non-EBF, as nearly all infants (95%) received breast milk at some point. Additionally, 71.2% of infants exclusively breastfed at 3 months continued to be exclusively breastfed until 6 months. Therefore, EBF status at 3 months was used to analyze subsequent microbiome timepoints. Maternal and infant characteristics are described in Table 1. C-section delivered-EBF infants had higher birth weights compared to vaginally delivered-EBF (p = 0.017). EBF infants had lower formula feeding and lower early introduction of solid food compared to non-EBF infants, despite delivery mode (p < 0.05). Vaginally delivered-EBF infants had slightly younger mothers compared to C-section delivered-non-EBF infants (p = 0.04). Pre-pregnancy body mass index (BMI) was lower in vaginal birth within the EBF category (p = 0.021) and lower comparing vaginally delivered-EBF with the C-section delivered-non-EBF category (p < 0.001). However, birth weight, maternal age, and pregestational BMI did not have a significant effect on microbiome composition at 3 months of age (Supplementary Figure S2). Antibiotic use within 3-months prior to the sample collection (n = 2; 0.4%) did not significantly impact the gut bacterial composition (PERMANOVA; $R^2 = 0.00149$; p = 0.826). No significant differences were observed for the other variables between the infant groups, such as gestational age, infant sex, maternal ethnicity, maternal education, and total family income. Also, we did not observe significant differences between the number of quality-filtered read pairs present for the breastfeeding and delivery mode subgroups across the time points analyzed (All p > 0.05; Supplementary Figure S3).

EBF changes the microbial composition and function according to the delivery mode and infant's age

Due to the variability in infant ages during the second time point, the microbiome composition was stratified in two primary clusters: 5–7 months, and 8–9 months (PERMANOVA; $R^2 = 0.022$; $p < 0.001$; Supplementary Figure S4). Subsequent analyses were conducted based on these age groups. Delivery mode had a lower impact on taxonomic (PERMANOVA; $R^2_{taxa} = 0.028, 0.013$; all $p = 0.001$) and functional β-diversity ($R^2_{function} = 0.069, 0.020$; all $p < 0.01$) in

Table 1. Maternal and infants' characteristics. Germina study – Sao Paulo, Brazil, 2021–2022.

	EBF		non-EBF		
	Vaginal (N = 188)	C-section (N = 174)	Vaginal (N = 58)	C-section (N = 103)	p-value
Gestational age, weeks (sd)	39.18 (1.21)	39.07 (1.33)	39.13 (0.99)	38.81 (0.99)	0.095
Infant sex					0.317
female	111 (59.0%)	94 (54.0%)	38 (65.5%)	54 (52.4%)	
male	77 (41.0%)	80 (46.0%)	20 (34.5%)	49 (47.6%)	
Birth weight, kg (sd)[a]	3.21 (0.34)	3.33 (0.37)	3.26 (0.44)	3.32 (0.39)	0.042
Formula feeding at 6 months of age					<0.001
yes	34 (8.7%)	50 (12.8%)	44 (11.3%)	71 (18.2%)	
no	106 (27.2%)	75 (19.2%)	2 (0.5%)	8 (2.1%)	
Introduction of solid foods before 6 months of age					<0.001
yes	3 (0.7%)	11 (2.7%)	23 (5.7%)	36 (8.9%)	
no	143 (35.2%)	122 (30.0%)	25 (6.2%)	43 (10.6%)	
Maternal ethnicity					0.687
white	121 (64.4%)	57 (32.8%)	23 (40.4%)	33 (32.0%)	
non-white	67 (35.6%)	117 (67.2%)	34 (59.6%)	70 (68.0%)	
Maternal age, years (sd)[b]	32.96 (5.20)	34.27 (5.31)	32.42 (5.56)	34.59 (5.12)	0.006
Maternal BMI pregestation[c]	24.44 (4.88)	25.98 (5.05)	25.10 (4.76)	26.98 (5.84)	< 0.001
Maternal educational attainment					0.079
> college degree	145 (76.7%)	146 (83.9%)	40 (69.0%)	78 (75.7%)	
< college degree	44 (23.3%)	28 (16.1%)	18 (31.0%)	25 (24.3%)	
Total family income (US dollar)	2722.63 (7724.13)	2410.68 (2600.78)	1656.34 (1461.72)	2166.49 (2959.78)	0.528

Comparisons between groups were conducted using analysis of variance (for numerical variables) or chi-square tests (for categorical variables).
[a]Difference related to C-section delivered-EBF compared to vaginally delivered-EBF infant group (p=0.017).
[b]Difference related to Vaginally delivered-EBF compared to C-section delivered-non-EBF infant group (p=0.04).
[c]Difference related to vaginal birth within the EBF category (p=0.021) and vaginally delivered-EBF compared to C-section delivered-non-EBF infant group (p<0.001).

EBF compared to non-EBF infants ($R^2_{taxa} = 0.032$, 0.032 and $R^2_{function} = 0.073$, 0.028; all $p < 0.03$) at 3–4 and 5–7 months of age, respectively. At 8–9 months of age, delivery mode did not affect the taxonomic and functional β-diversity of EBF infants ($R^2_{taxa} = 0.013$ and $R^2_{function} = 0.013$; all $p > 0.05$). In contrast, taxonomic β-diversity remained different in non-EBF infants ($R^2_{taxa} = 0.037$, $p < 0.05$), although functional β-diversity showed no significant differences ($R^2_{function} = 0.039$, $p > 0.05$), after adjusting for introduction of solid foods before 6 months of age and formula feeding (Figure 1a,b).

At 3–4 months of age, EBF infants had lower taxonomic α-diversity (Chao1 index), compared to non-EBF infants, despite delivery mode ($p < 0.001$). By 5–7 months, α-diversity was increased in vaginally delivered non-EBF infants compared to vaginally delivered EBF infants ($p = 0.01$). However, at 8–9 months, α-diversity did not differ by infant groups (Figure 2a ; Supplementary Figure S5). By 5–7 months, α-diversity was significantly higher in vaginally delivered non-EBF infants compared to vaginally delivered EBF infants ($p = 0.01$). At 8–9 months, α-diversity did not differ significantly among the infant groups (Figure 2a; Supplementary Figure S5).

E. coli and *Bifidobacterium longum* were the dominant species up to 7 months of age, regardless of the breastfeeding practices and delivery mode. At earlier time points, *Bacteroides vulgatus, Bacteroides fragilis, Bifidobacterium breve, Bifidobacterium bifidum,* and *Ruminococcus gnavus* were consistently among the most abundant species alongside *C. neonatale* at 3 months and *Veillonella parvula* from 5 to 7 months. By 8–9 months, *B. longum* and *B. fragilis* remained the most abundant species, followed by *B. vulgatus*, *B. breve, R. gnavus, E. coli*, and *Prevotella copri* (Figure 2b).

At 3–4 months of age, *E. coli* was significantly increased in EBF infants, regardless of delivery mode. *Bacteroides* spp. (*B. fragilis*, *B. vulgatus*, *B. dorei*, *B. thetaiotaomicron*, and *B. uniformis*) were significantly increased in vaginally delivered infants, with some species enriched in EBF and others in non-EBF infants. *R. gnavus* and *Klebsiella pneumoniae* were increased in C-section-delivered infants during the same period. In addition, *C. neonatale* was significantly increased in C-section-delivered EBF infants. By 5–7 months of age, these patterns persisted but were less pronounced (Figure 2b). At 8–9 months, a distinct profile of species relative abundances

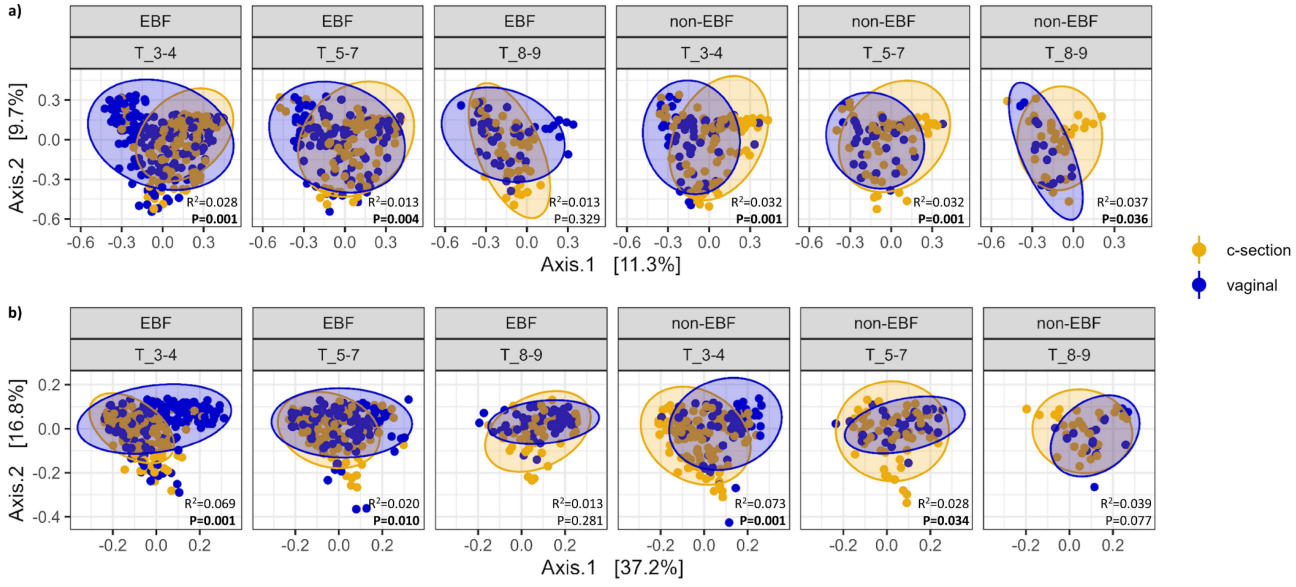

Figure 1. Compositional and functional profiles of the gut microbiome in 966 samples from infants aged 3–9 months. PCoA of Bray-Curtis distances illustrating a) microbial species; b) functional pathways. The analysis is colored by delivery mode and further stratified by breastfeeding status (EBF = exclusive breastfeeding; non-EBF = non-exclusive breastfeeding) and time points. Statistical significance was assessed using PERMANOVA (two-sided).

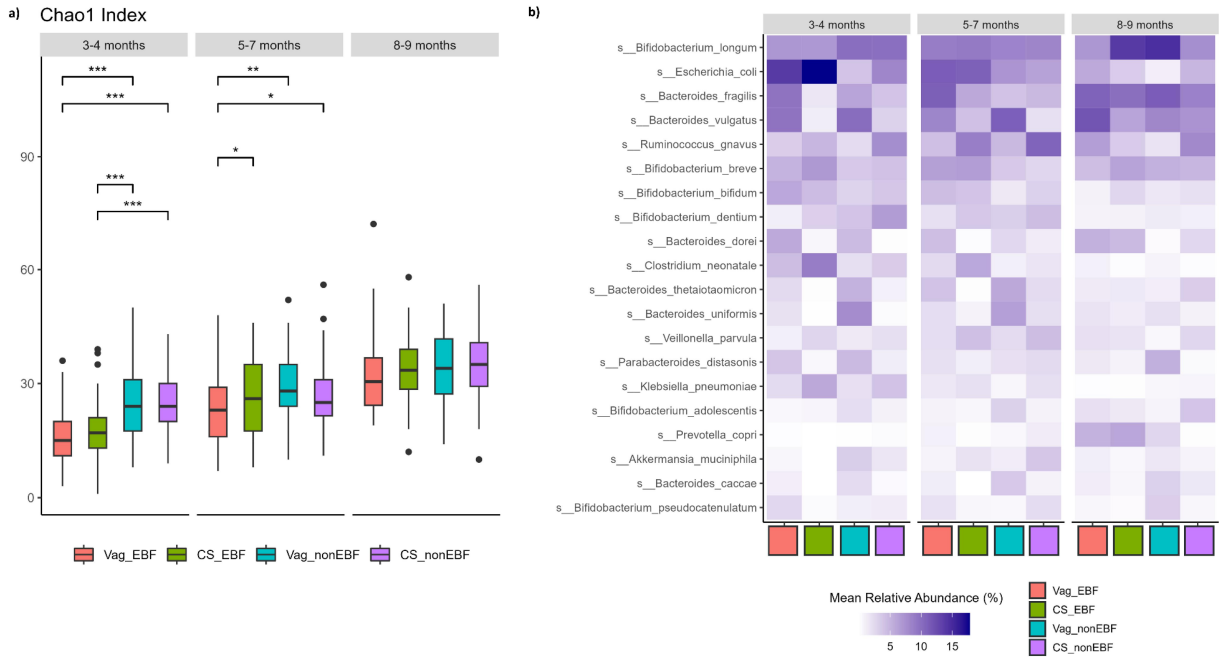

Figure 2. Alpha diversity and taxonomic composition of the gut microbiome in 966 samples from infants aged 3–9 months. a) Chao1 index (alpha diversity) and b) Heatmap of the relative abundance of the top 20 bacterial species, stratified by time points and colored according to infant groups. Statistical comparisons between groups were performed using ANOVA, followed by Tukey's honest significant difference post-hoc test for pairwise comparisons. Significance levels are denoted as ***$p \leq 0.001$; **$p \leq 0.01$; *$p \leq 0.05$. Infant groups: Vag_EBF = vaginally delivered-EBF; CS_EBF = C-section delivered-EBF; Vag_nonEBF = vaginally delivered-non-EBF; CS_nonEBF = C-section delivered-non-EBF.

was observed, with *R. gnavus* remaining significantly increased only in C-section-delivered non-EBF infants. Full statistical details, including The Tukey's HSD post-hoc corrections, are available in Supplementary File 1.

EBF is associated with a lower relative abundance of several species linked to delivery mode, and this is reflected in microbial functional profile

We used *Maaslin2* to perform multivariable linear modeling to assess the relationship between microbial species or functions and delivery mode at each time point. At 3–4 months of age, EBF was associated with alterations in 29 species (all q-value <0.10) in both vaginally and C-section-delivered infants. Of those, 14 species were shared between the delivery mode groups (Figure 3a). While the direction of the associations was consistent across both groups, the effect of EBF was greater in vaginally delivered infants for species such as *Parabacteroides distasonis* ($\beta_{vag} = -2.54$; $\beta_{c\text{-}sec} = -0.84$), *Collinsella aerofaciens* ($\beta_{vag} = -1.41$; $\beta_{c\text{-}sec} = -0.62$), *Clostridioides difficile* ($\beta_{vag} = -1.69$; $\beta_{c\text{-}sec} = -0.93$), and *B. vulgatus* ($\beta_{vag} = -1.67$; $\beta_{c\text{-}sec} = -0.79$). Conversely, *Clostridium innocuum* ($\beta_{vag} = -0.85$; $\beta_{c\text{-}sec} = -1.80$) and *Veillonella atypica* ($\beta_{vag} = -1.94$; $\beta_{c\text{-}sec} = -2.39$) exhibited stronger depletion in C-section delivered-EBF infants.

At 3–4 months, EBF was associated with distinct alterations in species depending on delivery mode. In C-section delivered-EBF infants, decreased relative abundances were observed for *R. gnavus* ($\beta = -2.63$), *Klebsiella michiganensis* ($\beta = -1.25$), *Erysipelatoclostridium ramosum* ($\beta = -1.37$), *Bifidobacterium dentium* ($\beta = -1.55$), and *Streptococcus* species (*S. parasanguinis*, $\beta = -2.11$; *S. mitis*, $\beta = -1.72$; *S. salivarius*, $\beta = -1.53$; *S. vestibularis*, $\beta = -0.83$) (all q-value <0.10), among others. In vaginally delivered infants, EBF was associated with lower relative abundance of *Akkermansia muciniphila* ($\beta = -2.17$), *Parabacteroides merdae* ($\beta = -1.30$), and *Bacteroides* species (*B. uniformis*, $\beta = -2.58$; *B. caccae*, $\beta = -1.95$; *B. stercoris*, $\beta = -0.93$; *B. xylanisolvens*, $\beta = -0.58$), among others. In contrast, *Staphylococcus aureus* ($\beta = 0.99$) and *E. coli* ($\beta = 1.95$) were increased in this group (Figure 3a).

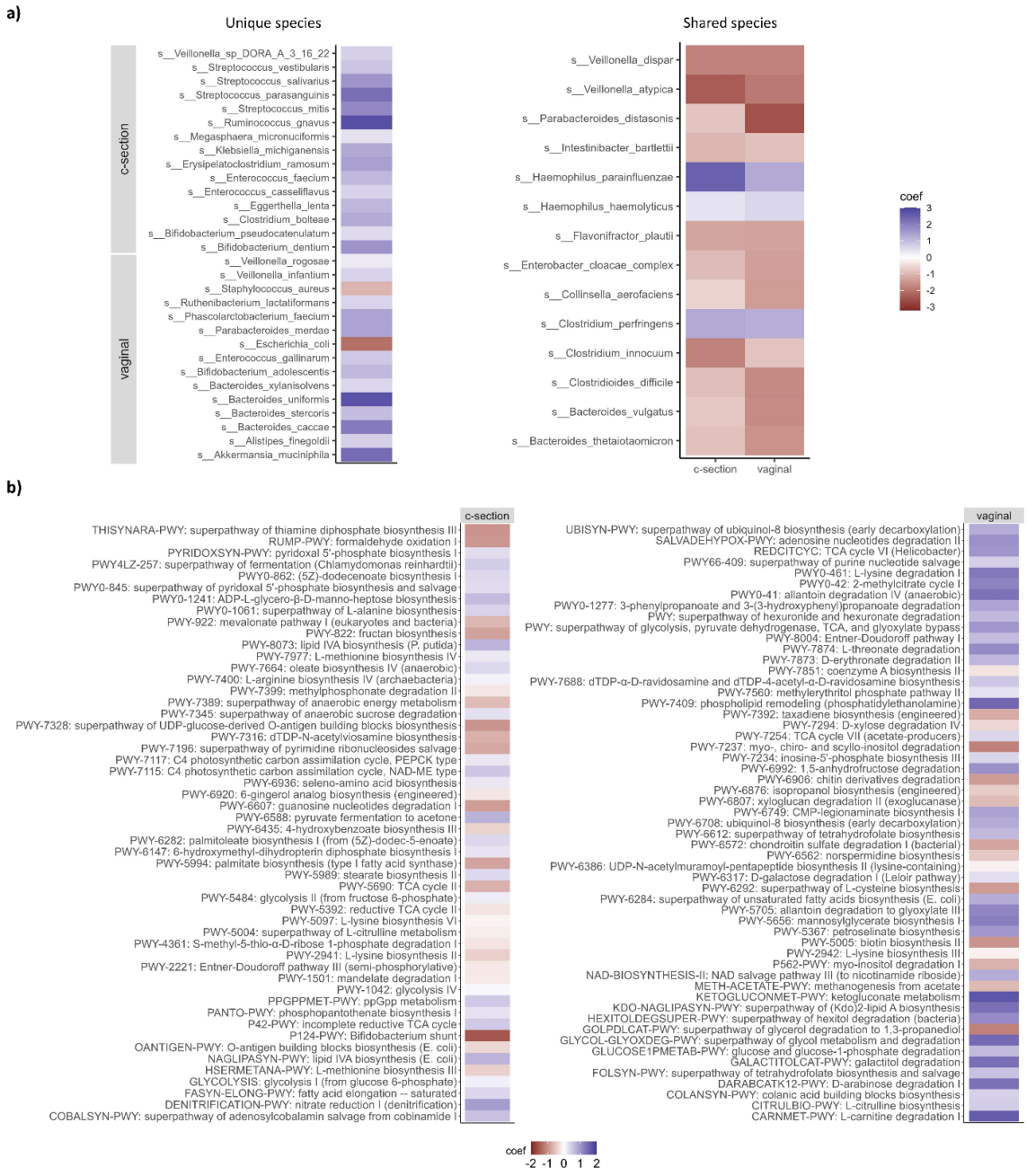

Figure 3. Effect of exclusive breastfeeding on the taxonomic and functional profiles of the gut microbiome in vaginally and C-section delivered infants. a) differentially abundant species at 3–4 months and; b) differentially abundant functional pathways at 3–4 months. The plots depict unique and shared species or pathways for each delivery mode, with colors representing the coefficient value of the significant associations with EBF ($q < 0.10$). Unique features are specific to either vaginal or C-section delivery, while shared features are common to both delivery modes.

At later timepoints, EBF did not alter the microbial composition regardless of delivery mode.

At 3–4 months, EBF was associated with significant changes in 147 functional pathways ($q < 0.10$). Among these, 55 pathways were significantly altered in vaginally delivered infants, and 52 pathways in C-section-delivered infants, with 107 unique pathways visualized in Figure 3b. In vaginally delivered infants, EBF mostly affected degradation pathways, increasing the breakdown of carbohydrates (D-galactose, D-arabinose, anhydrofructose, and glucose), amine, and polyamine (such as carnitine and allantoin), carboxyl/carboxylate compounds, including intermediates of the tricarboxylic acid (TCA) cycle. Furthermore, EBF was associated with an increase in the biosynthesis

of various molecules, including keto sugar, colanic acid, phospholipid, cofactors, carriers, and vitamins such as nicotinamide adenine dinucleotide-NAD salvage, ubiquinol, and tetrahydrofolate, as well as TCA cycle intermediates. In contrast, the biosynthesis of amino acids and biotin, along with the degradation of carbohydrates (xylose, xyloglucan, chitin, chondroitin) and inositol, was decreased. In infants delivered via C-section, EBF led to an increase in the biosynthesis of fatty acids and lipids, cofactors and carriers, cell structure components such as lipopolysaccharides, glycolysis, and fermentation pathways. EBF also affected the biosynthesis and degradation of amino acids and carbohydrates, with some pathways showing increases while others showed decreases. By 5–7 months and 8–9 months of age, EBF no longer significantly impacted microbial function.

At 3-4 months of age, the effect of EBF on carbohydrate-related microbial function was attributed to distinct microbial species. In C-section-delivered infants, species such as *B. dentium*, *B. longum*, *E. coli*, *B. fragilis*, *C. neonatale*, *K. pneumoniae*, *Streptococcus salivarius*, and *Lactobacillus gasseri* predominantly contributed to pathways that exhibited reduced activity. These pathways include the superpathway of UDP-glucose-derived O-antigen building blocks biosynthesis, fructan biosynthesis, dTDP-N-acetylviosamine biosynthesis, *E. coli* O-antigen building blocks biosynthesis, the *Bifidobacterium* shunt, and the Entner-Doudoroff pathway III. Conversely, pathways that showed increased activity in C-section-delivered infants – such as ADP-L-glycero-β-D-manno-heptose biosynthesis, superpathway of anaerobic sucrose degradation, glycolysis II, glycolysis I, and glycolysis IV – were predominantly attributed to *E. coli*, *B. fragilis*, *K. pneumoniae*, *K. variicola*, *C. neonatale*, and *V. parvula*. Notably, the enzymes encoded by the genes within the OANTIGEN-PWY pathway, attributed to *B. dentium* and *B. longum*, were limited to the synthesis of sugars such as dTDP-β-L-rhamnose and UDP-N-acetyl-D-glucosamine (Figure 4a).

In vaginally delivered infants at 3–4 months of age, *B. thetaiotaomicron* and *K. pneumoniae* were the primary species contributing to the decreased abundance of the chondroitin sulfate degradation I and D-xylose degradation IV pathways. However, it was not possible to assign specific species to other pathways with decreased activity in EBF infants, such as chitin derivatives degradation and xyloglucan degradation II. In contrast, several pathways were increased in vaginally delivered infants, including colanic acid building blocks biosynthesis, CMP-legionaminate biosynthesis I, dTDP-α-D-ravidosamine, and dTDP-4-acetyl-α-D-ravidosamine biosynthesis, Entner-Doudoroff pathway I, glucose and glucose-1-phosphate degradation, D-arabinose degradation I, 1,5-anhydrofructose degradation, and D-galactose degradation I. The major contributors to these pathways were *E. coli*, *K. pneumoniae*, *K. oxytoca*, *B. longum*, *C. freundii*, *K. variicola*, *B. fragilis*, and *P. distasonis* (Figure 4b).

The machine learning approach reveals a distinct microbial composition and function predictive of EBF across time

The top-ranked features in the machine learning model predicting EBF were selected based on the significance criteria explained in Methods (Supplementary Figure S6). Additionally, we selected significant features according to the target-oriented SHAP (SHapley Additive exPlanation) values (Supplementary Figure S7). Features relative abundance and SHAP values of each infant are shown in Figure 5. Overall, models using species for EBF prediction had balanced accuracy ranging from 0.74 to 0.57, with the lower end corresponding to the vaginal group at later time points. At 3–4 months, a higher relative abundance of *E. coli* and *C. perfringens* strongly contributed to EBF prediction, while a higher relative abundance of *Veillonella* spp. and *Parabacteroides distancionis* led to a lower contribution to EBF prediction regardless of delivery mode. In the C-section group, a higher relative abundance of *R. gnavus* led to a lower contribution to EBF prediction. Among vaginally delivered infants, higher relative abundance of *B. adolecentis*, and *Bacteorides* spp. led to a lower contribution to EBF prediction. A higher relative abundance of *C. neonatale* had a high positive contribution in C-sections, but the opposite was observed in vaginally delivered infants. Those effects were less pronounced at 5–7 months and persisted longer in the C-section group (Figure 5a).

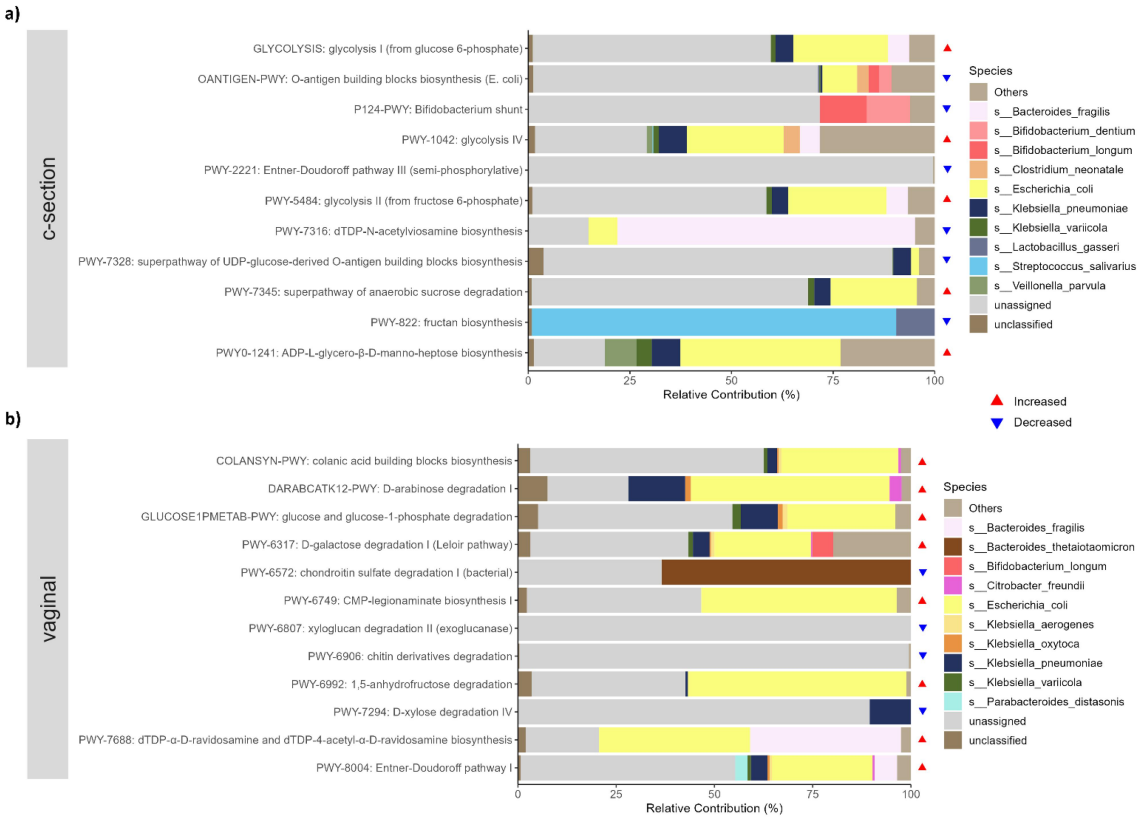

Figure 4. Top ten species attributed to carbohydrate pathways significantly associated with EBF at 3–4 months of age. a) C-section delivered infants and b) vaginally delivered infants. The plots display the top ten species attributed to carbohydrate metabolism pathways that show a significant association with exclusive breastfeeding. Blue and red triangles indicate pathways that increase and decrease in abundance at 3–4 months of age, respectively.

Models using microbial pathways had balanced accuracy ranging from 0.61 to 0.64. At the 3–4 time point, a higher relative abundance of menaquinol biosynthesis and a lower relative abundance of ketogluconate metabolism had a positive contribution to EBF prediction in the vaginal group. In contrast, the opposite was observed in C-section-delivered infants. Other pathways, including gluconeogenesis III, anaerobic energy metabolism, and L-histidine degradation III, when in high relative abundance, lead to lower prediction of EBF regardless of infants' delivery mode (Figure 5b).

EBF determines the taxonomic microbiome trajectory, while the functional trajectory is influenced by both EBF and delivery mode

Using the Microbiome Toolbox software, infants with the same breastfeeding status but differing delivery modes exhibited taxonomic microbiome trajectories with non-significant slopes (k) and intersections (n). This finding indicates that EBF significantly influenced both the starting point (n) and the overall trajectory (k) of microbiome development, regardless of delivery mode (all $p < 0.05$; Figure 6a). Furthermore, non-EBF infants consistently displayed higher microbial taxonomic Microbiome Maturation Index (MMI) compared to EBF infants, who showed the strongest correlation between MMI and chronological age. Notably, EBF also influenced the speed of maturation in C-section infants, who began with lower MMI but achieved the highest MMI levels by the end of the trajectory (Figure 6a).

Regarding the microbial functional trajectory, all infant groups begin with distinct trajectories, as evidenced by significant differences in intercepts (all $p < 0.05$; Figure 6b). This indicates that both EBF and delivery mode play a role in shaping the microbial functional trajectory. Among EBF infants, those born via C-section had the lowest

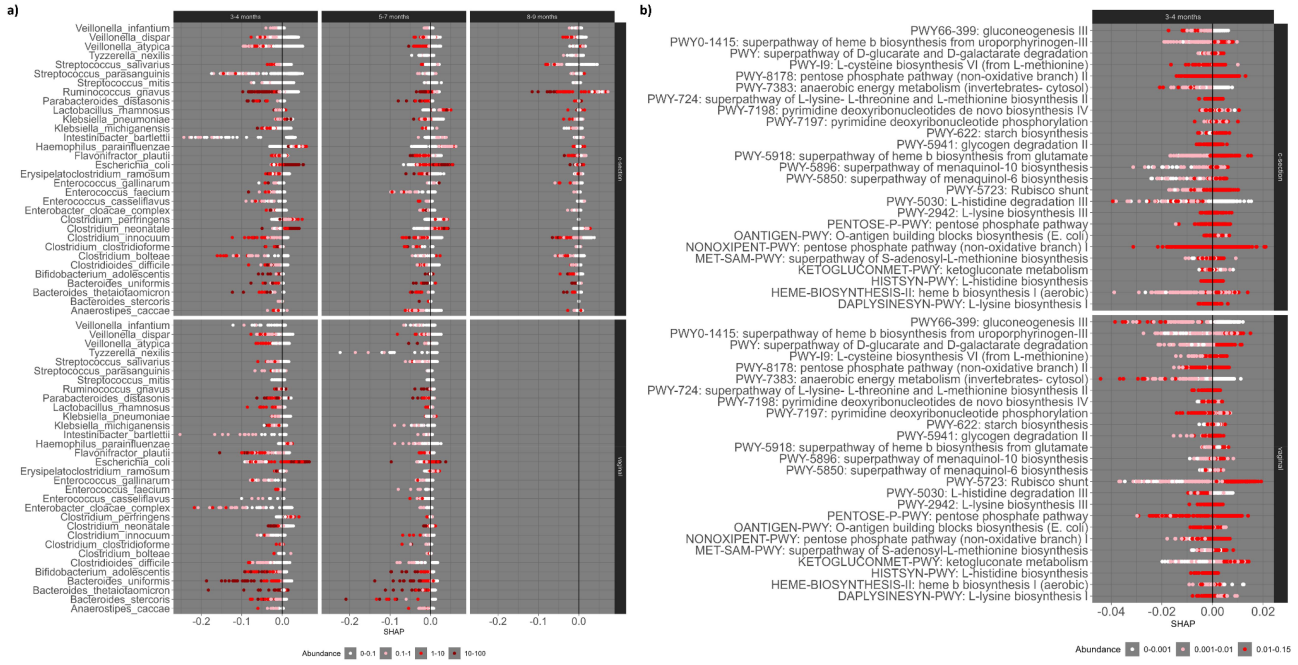

Figure 5. Contribution of each EBF predictors in a random forest classifier. a) top-ranked species predicting EBF at each time point and delivery mode. b) top-ranked functional pathways predicting EBF at each time point and delivery mode. The X-axis represents SHAP values, and the Y-axis lists species/functional pathways. Each dot represents an infant, with colors indicating the relative abundance of the species/pathways. Only statistically significant models ($p \leq 0.05$) and features ($p \leq 0.01$) are included for clarity and conciseness. Thus, functional pathways at time points 5–9 months are excluded from the figure.

functional MMI, while vaginally delivered non-EBF infants exhibited the highest index. Notably, differences in the rate of maturation were observed: C-section-delivered EBF infants had an accelerated maturation speed at 5 months of age, aligning with the trajectories of other groups. At 9 months, functional MMI were comparable across infant groups, though C-section-delivered non-EBF infants had a slightly higher functional MMI (Figure 6b).

Discussion

In this study, we explored the early determinants of the taxonomic and functional microbiome trajectory in a large cohort of healthy infants. Our findings reveal that EBF supports a non-accelerated maturation of the gut microbiome up to 5 months, irrespective of the delivery mode. EBF is associated with a reduced relative abundance of species commonly linked to delivery mode. Specifically, in infants delivered via C-section, EBF is associated with lower abundance of *Streptococcus* spp., *Klebsiella* and *R. gnavus*, while in vaginally delivered infants, it is associated with reduced levels of *Bacteroides* spp., *Veillonella* spp., and *Parabacteroides merdae*. Additionally, EBF contributes to a functional adaptation of the gut microbiome according to delivery mode.

Delivery mode significantly impacts the microbial community structure in infants at 3–4 months of age. However, this effect is less pronounced in exclusively breastfed infants, highlighting the potential of EBF to mitigate the impact of delivery mode on the initial colonization and establishment of the intestinal microbiome. A comparable effect was observed on microbial structure at 5–7 months but not at 8–9 months. This effect at 5–7 months is likely due to the continuation of exclusive breastfeeding among most infants who were exclusively breastfed at 3 months, persisting until 6 months of age. In contrast, in infants who were not exclusively breastfed, the effect of delivery mode on the microbial structure remained stronger and persisted until 8–9 months. These findings underscore the critical role of EBF in shaping gut microbial development.

Although the initial composition of gut microbiome is influenced by delivery mode, breastfeeding, in combination with environmental factors,

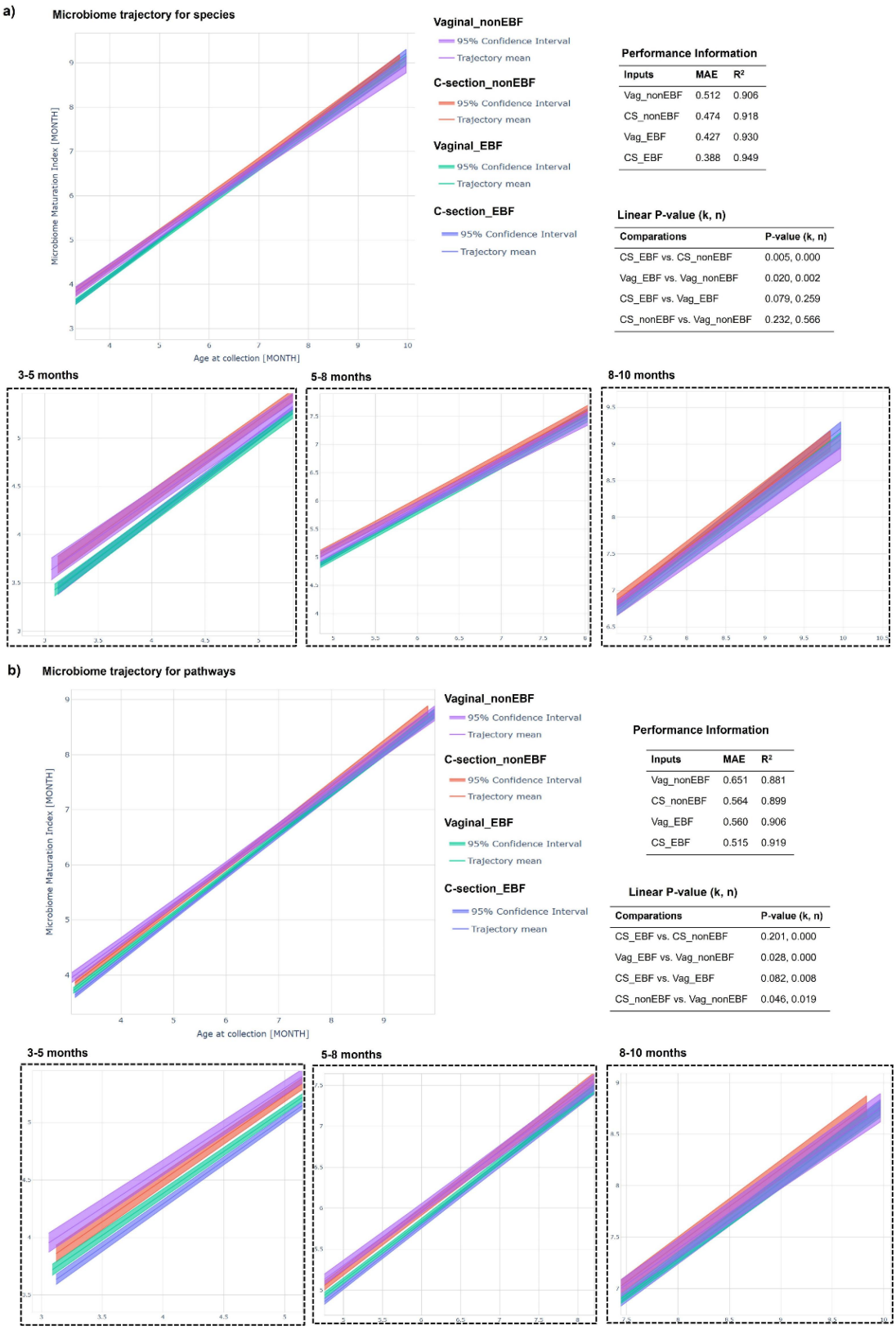

Figure 6. Compositional and functional microbiome trajectories by infant groups. Microbiome trajectory mean and 95% confidence interval of a) species and; b) functional pathways are shown for infants grouped by delivery mode and EBF status. The X-axis represents infants' chronological age at sample collection, and the Y-axis represents the microbiome maturation index. The zoomed-in portion of the plot emphasizes differences in trajectories at each time point. The tables summarize performance metrics, including mean absolute error (MAE), R^2 values, and linear p-values (k, n), for comparisons between infant groups. Infant groups: Vag_EBF = vaginally delivered-EBF; CS_EBF = C-section delivered-EBF; Vag_nonEBF = vaginally delivered-non-EBF; CS_nonEBF = C-section delivered-non-EBF.

plays a crucial role in shaping the infant microbiome during the first year.[16] In our study, EBF appears to support a more gradual maturation of the infant microbiome, maintaining lower taxonomic diversity compared to non-EBF infants. Shenhav et al. reported that a paced microbiome colonization could facilitate a healthy respiratory development.[7] Non-breastfeeding at 3 months contributes to the premature acquisition of species, such as *Ruminococcus gnavus*, which has been correlated with an increased risk of asthma and allergic diseases, when not acquired at the ideal timing.[7,34] Our data shows that at 3 months, *R. gnavus* has a higher abundance in C-section non-EBF infants compared to all other infant groups, while it showed a decrease in C-section delivered EBF infants, although this was not statistically significant. This finding suggests that EBF may attenuate the early acquisition of *R. gnavus*. Conversely, a later colonization of this species, along with other mucin-degrading species such as *Akkermansia muciniphila*, is associated with a decreased risk of allergic diseases.[35] These species ferment non-digestible galacto-oligosaccharides, from non-human-milk dietary sources, and host mucin, into short-chain fatty acids (SCFAs), enabling them to colonize the infant gut microbiome once non-human milk food sources are introduced.[36,37]

B. longum and *E. coli* were the dominant bacterial species across all infants and time points, regardless of breastfeeding practices or delivery mode. This observation is consistent with previous studies that demonstrated maternal-to-infant transmission, with these species detected in maternal and infant stool as well as human milk samples.[16,38] In our cohort, other *Bifidobacterium* species, such as *B. breve*, *B. bifidum*, and *B. dentium*, were also highly abundant across all infants, appearing to be supported by both partial or EBF. Feehily et al. reported *B. breve* exhibiting the highest proportion of isolates from human milk, while *B. dentium* was less frequent and predominantly detected in infant stool.[38]

E. coli has been described as an important member of the microbiome in Brazilian infants[13,39–41] and other developing countries.[42] While *E. coli* is recognized for its role in vitamin K production,[43] a high relative abundance has been associated with dysbiotic conditions in developed countries.[44,45] However, in developing countries, *E. coli* has been proposed as a crucial factor for the establishment of the infant microbiome, as previously hypothesized in Brazilian cohorts.[41] Our study reinforces these findings using WGS approach in a large cohort of healthy infants. Notably, *E. coli* exhibited a significant increase in vaginally delivered EBF infants, along with an increase in the vitamin K biosynthesis pathway and high importance in predicting EBF (superpathway of demethylmenaquinol-9 biosynthesis). These results suggest that the role *E. coli* extends beyond geographical variation, highlighting its contribution to vitamin K biosynthesis and its association with exclusive breastfeeding practices.

Our data also point to a reduction in *Bacteroides* spp. colonization in vaginally delivered infants at 3–4 and 5–7 months of age, suggesting a sustained influence of exclusive breastfeeding, whether practiced during the first 3 months or continued up to 6 months. Colonization by *Bacteroides* is widely recognized as being associated with vaginal delivery,[3,16] with persistence from the first week postpartum.[3] For example, Mitchell et al. reported that *B. thetaiotaomicron*, *B. uniformis*, *B. xylanisolvens*, and *B. caccae* were more abundant in vaginally delivered neonates,[3] the same species we observed to be decreased in EBF infants. This discrepancy may stem from differences in feeding practices, possibly because their study included a high proportion of formula-fed infants, limiting their ability to assess the impact of breastfeeding on *Bacteroides* persistence. We suggest that the reduced relative abundance of *Bacteroides* in EBF infants could be linked to the cleavage of HMO by *B. longum*. *Bacteroides* species are also important consumers of HMO. However, their feeding mechanism involves external degradation of the oligosaccharides structure, whereas *Bifidobacterium* internalizes HMO for cleavage.[46] This internalization strategy provides a competitive advantage by sequestering sugar monomers, thereby limiting the growth of competing microbial strains in the colon.[6]

Our study highlights distinct patterns of carbohydrate metabolism dependent on the mode of delivery. In vaginally delivered infants, EBF was linked to alterations in carbohydrate metabolism

pathways, including increased degradation of human milk sugars such as D-galactose, D-arabinose, anhydrofructose, and glucose.[6] Additionally, pathways related to the biosynthesis of building block sugars (e.g., keto sugar and colanic acid) were elevated, while the degradation of xylose, xyloglucan, chitin, and chondroitin were decreased. Conversely, C-section delivered infants who were exclusively breastfed showed carbohydrate metabolism pathways associated with microbial energy production and maintenance.[47] These differences in functional profiles suggest that the gut microbiome of C-section-delivered infants may adopt alternative nutrient utilization strategies. This aligns with the concept of nutritional energy harvesting or wasting microbiome, previously proposed to discuss the microbiome adaptation supporting infant growth under conditions such as malnutrition.[48]

Several factors may account for the observed differences in carbohydrate metabolism: i) Variations in microbiome composition, and consequently its metabolic machinery influencing energy acquisition strategies; ii) EBF may modulate microbiome function by promoting more efficient nutrient utilization from the diet, prompting infants growth; iii) HMOs in breast milk may reduce the adhesion of pathobionts[49] as supported by our findings that O-antigen biosynthesis was inversely associated with EBF; iv) A reduction in the relative abundance of *Bacteroides* spp., associated with EBF, may also contribute to these differences. Many species within this genus employs a strategy of externally degrading HMOs to release sugar monomers, which can cross-feed pathogenic bacteria or pathobionts.[6]

EBF is also associated with differences in lipid metabolism between C-section and vaginally delivered infants, indicating nuanced impacts on fatty acid biosynthesis. For instance, C-section-delivered infants exhibited increased biosynthesis of several saturated and mono-unsaturated fatty acids, whereas vaginally delivered infants were less impacted, presenting increased pathways of petroselinate and other unsaturated fatty acids. Wilson et al. report a similar profile of altered functional pathways in pediatric populations with asthma,[50] mirroring our finding in C-section-delivered infants. In addition, EBF was linked to alterations in amino acids and polyamine metabolism. Among C-section-delivered infants, several pathways of the homocysteine-methionine cycle, seleno-amino acid metabolism, and methyl group production were altered, which is notable given the role of methyl-donating vitamins in gut inflammation.[51] In contrast, vaginally delivered infants exhibited reduced norspermidine biosynthesis, a compound implicated in the formation and dissolution of biofilms of exopolysaccharides.[52] This finding suggests that EBF may influence bacterial quorum-sensing specifically in vaginally delivered infants.

EBF emerges as a critical determinant of the microbiome trajectory, shaping both species diversity and maturation rate. The lower taxonomic diversity observed in EBF infants aligns with the findings from a meta-analysis conducted across different populations.[53] Additionally, the speed of gut maturation in EBF infants was consistently slower compared to non-EBF infants, whose accelerated gut maturation led to an earlier increase in taxonomic diversity. The extended duration of EBF in our cohort further supports this finding. A review study by Iozzo and Sanguinetti described that breastfeeding duration delays microbiota's maturation toward an adult-type microbiome.[54] After 5 months of age, this effect gradually decreases, probably due to the introduction of solid foods, which alters the dietary composition by increasing the intake of carbohydrates, fat, fiber, and protein[55]; thereby, increasing the microbiome diversity. Our findings are consistent with this transition, as the effect of EBF on species and pathways abundances diminishes at later time points when formula feeding and solid food introduction are accounted for the multivariate model. The dietary shift associated with solid food introduction alters the metabolic activity of gut bacteria, reducing the abundance of genes involved in breast milk sugars degradation. Thus, the gut microbiota undergoes functional maturation, adapting to the newly available energy sources and acquiring the capacity to degrade complex sugars and other nutrients from solid foods.[15]

Overall, our cohort demonstrated homogeneity in terms of breastfeeding practices and delivery mode. However, pre-gestational maternal BMI and maternal age were statistically different between the exclusive and non-EBF groups.

Several factors, including maternal characteristics, can influence the composition of the milk microbiome, as highlighted by Notarbartolo et al.,[56] thereby indirectly affecting the infants gut microbiome. While maternal BMI is recognized as a critical factor influencing milk microbiome composition, recent evidence from Cortés-Macías et al. suggests that breastfeeding practices play a more significant role in shaping the composition and diversity of the breast milk microbiota.[57] Indeed, exclusively breastfeeding stimulates direct contact between infants' mouths and the mother's breasts, which serves as a vital route for microbial exchange.[56] Notably, our results were not confounded by factors such as pregestational BMI, birth weight, and mother's age. Therefore, we attribute the observed impact on the infant gut microbiome primarily to breastfeeding practices and their influence on the mothers' milk microbiome, underscoring the pivotal role of EBF in shaping the infants' gut microbiome.

To the extent of our knowledge, this study adopts a sounder approach to establishing feature importance through machine learning when compared to the available literature. For instance, we avoided setting an arbitrary number of relevant features according to a model internal overall measure of importance.[23] Specifically, we focused on evaluating feature importance only in models that demonstrated the ability to make statistically significant predictions, ensuring the reliability and relevance of the identified features. Additionally, we set a p-value-based criteria to streamline the visual presentation of features. This approach relied on a target-oriented SHAP value specifically designed for our analysis, focusing on each infant individually rather than the entire sample. This approach utilized a target-oriented SHAP value specifically tailored for our analysis, positive values were assigned to features that contributed to correct predictions. Such a scheme enabled us to objectively remove features with questionable relevance due to low statistical support – provided the selected subset was not interpreted as a statistically significant result on its on; otherwise, a post hoc analysis would be needed. Lastly, our study strengthens the value of machine learning-assisted microbiome analysis, as the findings from traditional statistical modeling were in accordance with those derived from machine learning approaches.

Our results, however, have several limitations: (i) the age range of the second appointment was not precisely defined, leading to non-linear trajectory time points; (ii) the microbiome trajectory was modeled using only two time points, which may miss subtle variances and dynamic changes in microbiome composition and function over time; (iii) fewer than 5% of the cohort were never breastfed, which limits our ability to use a no-breastfeeding group as a comparison. Consequently, our non-EBF group comprises infants with mixed feeding (formula and breast milk) and, in some cases, those who were ever breastfed. As a result, we were unable to directly compare the effects of breastfeeding versus never breastfeeding, although the impact of formula feeding was accounted for in our analysis; (iv) information on antibiotic use was collected through a general question about medication use, which may have led to an underestimation of this variable; (v) while genes encoding enzymes, metabolites, and other elements involved in microbial pathways were sequenced, microbial metabolites were not measured to confirm the activity of metabolic processes. The strengths of this study include the use of a large sample size analyzed through shotgun metagenomics, enabling the identification of differential effects of exclusive breastfeeding on microbiome signatures based on delivery mode, as well as revealing alterations in microbial species and functional pathways. A notable strength is the characterization of microbiome trajectories using spline and linear regression statistical analysis, which allowed us to test universality across different infant groups. Furthermore, the integration of a robust machine learning approach enhanced the reliability and depth of our findings. Lastly, the inclusion of a novel Brazilian cohort contributes valuable data from a population underrepresented in microbiome research, thereby enhancing the study's relevance and potential generalizability.

Conclusions

Our findings provide valuable insights into the functional pathways influenced by EBF in the infants' gut microbiome and demonstrate how EBF may mitigate the impact of delivery mode on microbiome establishment during the first year of

life. This study has the potential to guide preventive intervention strategies and inform policy changes and educational practices aimed at promoting exclusive breastfeeding, particularly in C-section delivered infants. Our results reinforce existing knowledge of bacterial succession in the early-life gut microbiome, highlighting the role of EBF in fostering a gradual bacterial colonization. By showing that the reduced diversity in the gut microbiome is associated with EBF, regardless of delivery mode, our study suggests that this slower microbiome maturation may play a role in immune system training during the first 1000 days of life.[58]

Exclusive breastfeeding is associated with changes in the microbiome composition and function based on the mode of delivery, influencing microbiome development in distinct ways for C-section and vaginally delivered infants. Moreover, EBF plays a critical role in driving microbiome maturation in early life, attenuating the negative impacts of C-section delivery. Additionally, our findings highlight the importance of exclusive breastfeeding in shaping pathways related to carbohydrate and lipid metabolism, as well as cell structure. The microbiome trajectory in EBF infants closely aligns microbiome age with the infant's chronological age. Given the demonstrated benefits of EBF in early life, our results suggest that its influence on microbiome maturation supports a non-accelerated development of microbiome composition and function, as reflected in the data presented.

Acknowledgments

We would like to thank Justin O'Sullivan and Robert Britton for mentoring assistance and analysis development in Leap's Microbiome Working Group.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by funding from the Wellcome Leap 1kD Program (The First 1000 Days; https://wellcomeleap.org/1kd/.), a multi-million-dollar, multi-year funding.

ORCID

Paulo A. Schüroff http://orcid.org/0000-0002-7808-7818
Pedro A. R. Vanzele http://orcid.org/0000-0002-0305-8262
Davi Pereira-Santos http://orcid.org/0000-0002-3766-3137
Tamires Amabili Valim http://orcid.org/0000-0002-2645-126X
Kevin S. Bonham http://orcid.org/0000-0003-3200-7533
André Fujita http://orcid.org/0000-0002-7756-7051
Maria Rita Passos-Bueno http://orcid.org/0000-0002-9248-3008
Patricia C. B. Beltrão-Braga http://orcid.org/0000-0003-0004-1091
André C. P. L. F. Carvalho http://orcid.org/0000-0002-4765-6459
Vanja Klepac-Ceraj http://orcid.org/0000-0001-5387-5706
Guilherme V. Polanczyk http://orcid.org/0000-0003-2311-3289
Alline C. Campos http://orcid.org/0000-0003-4258-3198
Carla R. Taddei http://orcid.org/0000-0003-0869-0651

Author contributions

C.R.T., A.C.C., and G.V.P. conceptualized and designed the study. G.V.P, M.R.P.B., C.R.T., A.C.C., P.B.B.B., A.F., and A. C.P.L.F.C. coordinated the Germina cohort. Germina Consortium: collected data and biospecimen. N.F.N., P.A.S., and T.A.V. processed biospecimen. N.F.N., P.A.S., and P.A.R. V. conducted bioinformatics and D.P. machine learning analysis. C.R.T., N.F.N., P.A.S., and P.A.R.V. analyzed and interpreted the data. C.R.T. N.F.N, P.A.S., and P.A.R.V. wrote the original manuscript draft. K.S.B. and V.K. supervised and verified the data. All authors reviewed the manuscript.

Availability of data and material

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials. All raw sequencing data generated for this study are deposited in the NCBI sequence read archive (SRA) and is available at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1072081. The code for the implementation of the machine learning analysis is available at https://github.com/davips/germina/blob/main/experiments/microbiome/breastfeeding-LOO—2023-11-11.py.. The metadata associated with this study will be available from the authors upon reasonable request and with permission of the Germina project.

Ethics approval and consent to participate

Ethics approvals were obtained from the Ethics Committee for the Analysis of Research Projects (CAPPESq) and the National Council of Ethics in Research (ref. number CAAE 49,671,221.20000.0068). Following the Declaration of Helsinki, all caregivers provide written informed consent

before completing any study measure. Cases of developmental, health, and mental health issues are being referred to specialized health services.

References

1. Korpela K, Salonen A, Hickman B, Kunz C, Sprenger N, Kukkonen K, Savilahti E, Kuitunen M, de Vos WM. Fucosylated oligosaccharides in mother's milk alleviate the effects of caesarean birth on infant gut microbiota. Sci Rep. 2018;8(1). doi: 10.1038/s41598-018-32037-6.
2. Healy DB, Ryan CA, Ross RP, Stanton C, Dempsey EM. Clinical implications of preterm infant gut microbiome development. Nat Microbiol. 2022;7(1):22–33. doi: 10.1038/s41564-021-01025-4.
3. Mitchell CM, Mazzoni C, Hogstrom L, Bryant A, Bergerat A, Cher A, Pochan S, Herman P, Carrigan M, Sharp K, et al. Delivery mode affects stability of early infant gut microbiota. Cell Reports Med. 2020;1(9):1. doi: 10.1016/j.xcrm.2020.100156.
4. Ma J, Li Z, Zhang W, Zhang C, Zhang Y, Mei H, Zhuo N, Wang H, Wang L, Wu D. Comparison of gut microbiota in exclusively breast-fed and formula-fed babies: a study of 91 term infants. Sci Rep. 2020;10(1):10. doi: 10.1038/s41598-020-72635-x.
5. Gasparrini AJ, Wang B, Sun X, Kennedy EA, Hernandez-Leyva A, Ndao IM, Tarr PI, Warner BB, Dantas G. Persistent metagenomic signatures of early-life hospitalization and antibiotic treatment in the infant gut microbiota and resistome. Nat Microbiol. 2019;4(12):2285–2297. doi: 10.1038/s41564-019-0550-2.
6. Charbonneau MR, Blanton LV, Digiulio DB, Relman DA, Lebrilla CB, Mills DA, Gordon JI. A microbial perspective of human developmental biology. Nature. 2016;535(7610):48–55. doi: 10.1038/nature18845.
7. Shenhav L, Fehr K, Reyna ME, Petersen C, Dai DLY, Dai R, Breton V, Rossi L, Smieja M, Simons E, et al. Microbial colonization programs are structured by breastfeeding and guide healthy respiratory development. Cell. 2024;187(19):5431–5452.e20. doi: 10.1016/j.cell.2024.07.022.
8. Liu Y, Qin S, Song Y, Feng Y, Lv N, Xue Y, Liu F, Wang S, Zhu B, Ma J, et al. The perturbation of infant gut microbiota caused by cesarean delivery is partially restored by exclusive breastfeeding. Front Microbiol. 2019;10:10. doi: 10.3389/fmicb.2019.00598.
9. Guo C, Zhou Q, Li M, Zhou L, Xu L, Zhang Y, Li D, Wang Y, Dai W, Li S, et al. Breastfeeding restored the gut microbiota in caesarean section infants and lowered the infection risk in early life. BMC Pediatrics. 2020;20(1):20. doi: 10.1186/s12887-020-02433-x.
10. Princisval L, Rebelo F, Williams BL, Coimbra AC, Crovesy L, Ferreira AL, Kac G. 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. 2022;80(1):113–127. doi: 10.1093/nutrit/nuab008.
11. Akagawa S, Tsuji S, Onuma C, Akagawa Y, Yamaguchi T, Yamagishi M, Yamanouchi S, Kimata T, Sekiya S, Ohashi A, et al. Effect of delivery mode and nutrition on gut microbiota in neonates. Ann Nutr Metab. 2019;74(2):132–139. doi: 10.1159/000496427.
12. Raspini B, Vacca M, Porri D, De Giuseppe R, Calabrese FM, Chieppa M, Liso M, Cerbo RM, Civardi E, Garofoli F, et al. Early life microbiota colonization at six months of age: a transitional time point. Front Cell Infect Microbiol. 2021;11:11. doi: 10.3389/fcimb.2021.590202.
13. Tonon KM, Morais TB, Taddei CR, Araújo-Filho HB, Acfv A, Miranda A, de Morais MB, Staley C. Gut microbiota comparison of vaginally and cesarean born infants exclusively breastfed by mothers secreting α1–2 fucosylated oligosaccharides in breast milk. PLOS ONE. 2021;16(2):e0246839. doi: 10.1371/journal.pone.0246839.
14. Madan JC, Hoen AG, Lundgren SN, Farzan SF, Cottingham KL, Morrison HG, Sogin ML, Li H, Moore JH, Karagas MR. Association of cesarean delivery and formula supplementation with the intestinal microbiome of 6-week-old infants. JAMA Pediatrics. 2016;170(3):212–219. doi: 10.1001/jamapediatrics.2015.3732.
15. Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, Li Y, Xia Y, Xie H, Zhong H, et al. Dynamics and stabilization of the human gut microbiome during the first year of life. Cell Host Microbe. 2015;17(5):690–703. doi: 10.1016/j.chom.2015.04.004.
16. Selma-Royo M, Dubois L, Manara S, Armanini F, Cabrera-Rubio R, Valles-Colomer M, González S, Parra-Llorca A, Escuriet R, Bode L, et al. Birthmode and environment-dependent microbiota transmission dynamics are complemented by breastfeeding during the first year. Cell Host Microbe. 2024;32(6):996–1010.e4. doi: 10.1016/j.chom.2024.05.005.
17. Raspini B, Porri D, De Giuseppe R, Chieppa M, Liso M, Cerbo RM, Civardi E, Garofoli F, Monti MC, Vacca M, et al. Prenatal and postnatal determinants in shaping offspring's microbiome in the first 1000 days: study protocol and preliminary results at one month of life. Ital J Pediatr. 2020;46(1). doi: 10.1186/s13052-020-0794-8.
18. Reyman M, van Houten MA, van Baarle D, Bosch AATM, Man WH, Chu MLJN, Arp K, Watson RL, Sanders EAM, Fuentes S, et al. Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. Nat Commun. 2019;10(1):10. doi: 10.1038/s41467-019-13014-7.
19. Fatori D, Shephard E, Naspolini NF, Guzman GC, Wang JYT, Tótolo P, Mafra AL, Isaias C, dos

Santos DP, Russo FB, et al. Identifying biomarkers and trajectories of executive functions and language development in the first 3 years of life: design, methods and initial findings of the Germina cohort study [internet]. OSF Preprints. 2024 [cited 2024 June 12]; osf.io/ed4fb.
20. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, Mailyan A, Manghi P, Scholz M, Thomas AM, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife. 2021;10:10. doi: 10.7554/eLife.65088.
21. McMurdie PJ, Holmes S, Watson M. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLOS ONE. 2013;8(4):e61217. doi: 10.1371/journal.pone.0061217.
22. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Solymos P, Stevens MHH, Szoecs E, et al. Vegan: community ecology package [internet]. 2022; https://github.com/vegandevs/vegan.
23. Robertson RC, Edens TJ, Carr L, Mutasa K, Gough EK, Evans C, Geum HM, Baharmand I, Gill SK, Ntozini R, et al. The gut microbiome and early-life growth in a population with high prevalence of stunting. Nat Commun. 2023;14(1):654. doi: 10.1038/s41467-023-36135-6.
24. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, Tickle TL, Weingart G, Ren B, Schwager EH, et al. Multivariable association discovery in population-scale meta-omics studies. PLOS Comput Biol. 2021;1–27. doi:10.1371/journal.pcbi.1009442.
25. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32. doi: 10.1023/A:1010933404324.
26. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. Scikit-learn: machine learning in Python. J Mach Learn Res [Internet]. 2011 [cited 2024 Jan 18]. 12:2825–2830. http://scikit-learn.sourceforge.net.
27. Bokulich NA, Chung J, Battaglia T, Henderson N, Jay M, Li H, Lieber A, Wu F, Perez-Perez GI, Chen Y, et al. Antibiotics, birth mode, and diet shape microbiome maturation during early life. Sci Transl Med. 2016;8(343). doi: 10.1126/scitranslmed.aad7121.
28. Ojala M, Garriga GC. Permutation tests for studying Classifier performance. J Mach Learn Res. 2010;11:1833–1863.
29. Biecek PD. Explainers for complex predictive models in R. J Mach Learn Res [Internet] 2018 [cited 2024 Jan 18]. 19:1–5. http://jmlr.org/papers/v19/18-416.html.
30. Lundberg S, Lee S-I. A unified approach to interpreting model predictions [internet]. In: Proceedings of the 31st international conference on neural information processing systems. 2017. p. 4768–4777. http://arxiv.org/abs/1705.07874 .
31. Wong TT. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognit. 2015;48(9):2839–2846. doi: 10.1016/j.patcog.2015.03.009.
32. Banjac J, Sprenger N, Dogra SK, Marschall T. Microbiome toolbox: methodological approaches to derive and visualize microbiome trajectories. Bioinformatics. 2023;39(1):39. doi: 10.1093/bioinformatics/btac781.
33. Subramanian S, Huq S, Yatsunenko T, Haque R, Mahfuz M, Alam MA, Benezra A, Destefano J, Meier MF, Muegge BD, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature. 2014;510(7505):417–421. doi: 10.1038/nature13421.
34. Chua H-H, Chou H-C, Tung Y-L, Chiang B-L, Liao C-C, Liu H-H, Ni Y-H. Intestinal dysbiosis featuring abundance of ruminococcus gnavus associates with allergic diseases in infants. Gastroenterology. 2018;154(1):154–167. doi: 10.1053/j.gastro.2017.09.006.
35. Lee M-J, Kang M-J, Lee S-Y, Lee E, Kim K, Won S, Suh DI, Kim KW, Sheen YH, Ahn K, et al. Perturbations of gut microbiome genes in infants with atopic dermatitis according to feeding type. J Allergy Clin Immunol Pract. 2018;141(4):1310–1319. doi: 10.1016/j.jaci.2017.11.045.
36. Crost EH, Coletto E, Bell A, Juge N. Ruminococcus gnavus: friend or foe for human health. FEMS Microbiol Rev. 2023;47(2):47. doi: 10.1093/femsre/fuad014.
37. Fallani M, Amarri S, Uusijarvi A, Adam R, Khanna S, Aguilera M, Gil A, Vieites JM, Norin E, Young D, et al. Determinants of the human infant intestinal microbiota after the introduction of first complementary foods in infant samples from five European centres. Microbiol (NY). 2011;157(5):1385–1392. doi: 10.1099/mic.0.042143-0.
38. Feehily C, O'Neill IJ, Walsh CJ, Moore RL, Killeen SL, Geraghty AA, Lawton EM, Byrne D, Sanchez-Gallardo R, Nori SRC, et al. Detailed mapping of bifidobacterium strain transmission from mother to infant via a dual culture-based and metagenomic approach. Nat Commun. 2023;14(1):14. doi: 10.1038/s41467-023-38694-0.
39. Naspolini NF, Meyer A, Moreira JC, Sun H, Froes-Asmus CIR, Dominguez-Bello MG. Environmental pollutant exposure associated with altered early-life gut microbiome: results from a birth cohort study. Environ Res. 2022;205:112545. doi: 10.1016/j.envres.2021.112545.
40. Brandt K, Taddei CR, Takagi EH, Oliveira FF, Duarte RTD, Irino I, Martinez MB, Carneiro-Sampaio M. Establishment of the bacterial fecal community during the first month of life in Brazilian newborns. Clinics. 2012;67(2):113–123. doi: 10.6061/clinics/2012(02)05.
41. Taddei CR, Oliveira FF, Duarte RTD, Talarico ST, Takagi EH, Ramos Carvalho GF II, Brandt K, Martinez MB, Martinez MB. High abundance of Escherichia during the establishment of fecal

microbiota in Brazilian children. Microb Ecol. 2014;67(3):624–634. doi: 10.1007/s00248-014-0381-x.
42. Fontana F, Mancabelli L, Lugli GA, Taracchini C, Alessandri G, Longhi G, Anzalone R, Viappiani A, Famo R, Brognan M, et al. Investigating the infant gut microbiota in developing countries: worldwide metagenomic meta-analysis involving infants living in suburban areas of Côte d'Ivoire. Environ Microbiol Rep. 2021;13(5):626–636. doi: 10.1111/1758-2229.12960.
43. Ramotar K, Conly JM, Chubb H, Louie TJ. Production of menaquinones by intestinal anaerobes. J Infect Dis. 1984;150(2):213–218. doi: 10.1093/infdis/150.2.213.
44. Hoffman LR, Pope CE, Hayden HS, Heltshe S, Levy R, McNamara S, Jacobs MA, Rohmer L, Radey M, Ramsey BW, et al. Escherichia coli dysbiosis correlates with gastrointestinal dysfunction in children with cystic fibrosis. Clin Infect Dis. 2014;58(3):396–399. doi: 10.1093/cid/cit715.
45. Zeng MY, Inohara N, Nuñez G. Mechanisms of inflammation-driven bacterial dysbiosis in the gut. Mucosal Immunol. 2017;10(1):18–26. doi: 10.1038/mi.2016.75.
46. Turroni F, Milani C, Duranti S, Mahony J, van Sinderen D, Ventura M. Glycan utilization and cross-feeding activities by bifidobacteria. Trends Microbiol. 2018;26(4):339–350. doi: 10.1016/j.tim.2017.10.001.
47. van Bodegom P. Microbial maintenance: a critical review on its quantification. Microb Ecol. 2007;53(4):513–523. doi: 10.1007/s00248-006-9049-5.
48. Robertson RC, Manges AR, Finlay BB, Prendergast AJ. The human microbiome and child growth – first 1000 days and beyond. Trends Microbiol. 2019;27(2):131–147. doi: 10.1016/j.tim.2018.09.008.
49. Bode L. Human milk oligosaccharides: every baby needs a sugar mama. Glycobiology. 2012;22(9):1147–1162. doi: 10.1093/glycob/cws074.
50. Wilson NG, Hernandez-Leyva A, Schwartz DJ, Bacharier LB, Kau AL. The gut metagenome harbors metabolic and antibiotic resistance signatures of moderate-to-severe asthma. FEMS Microbes. 2024;5. doi: 10.1093/femsmc/xtae010.
51. Vaccaro JA, Naser SA. The role of methyl donors of the methionine cycle in gastrointestinal infection and inflammation. Healthcare (Switz). 2022;10(1):10. doi: 10.3390/healthcare10010061.
52. Ou M, Ling J. Norspermidine changes the basic structure of S. mutans biofilm. Mol Med Rep. 2017;15(1):210–220. doi: 10.3892/mmr.2016.5979.
53. Ho NT, Li F, Lee-Sarwar KA, Tun HM, Brown BP, Pannaraj PS, Bender JM, Azad MB, Thompson AL, Weiss ST, et al. Meta-analysis of effects of exclusive breastfeeding on infant gut microbiota across populations. Nat Commun. 2018;9(1). doi: 10.1038/s41467-018-06473-x.
54. Iozzo P, Sanguinetti E. Early dietary patterns and microbiota development: still a way to go from descriptive interactions to health-relevant solutions. Front Nutr. 2018;5. doi: 10.3389/fnut.2018.00005.
55. Johnson CL, Versalovic J. The human microbiome and its potential importance to pediatrics. Pediatrics. 2012;129(5):950–960. doi: 10.1542/peds.2011-2736.
56. Notarbartolo V, Giuffre M, Montante C, Corsello G, Carta M. Composition of human breast milk microbiota and its role in Children's health. Pediatr Gastroenterol Hepatol Nutr. 2022;25(3):194–210. doi: 10.5223/pghn.2022.25.3.194.
57. Cortés-Macías E, Selma-Royo M, Martínez-Costa C, Collado MC. Breastfeeding practices influence the breast milk microbiota depending on pre-gestational maternal bmi and weight gain over pregnancy. Nutrients. 2021;13(5):1518. doi: 10.3390/nu13051518.
58. Donald K, Finlay BB. Early-life interactions between the microbiota and immune system: impact on immune system development and atopic disease. Nat Rev Immunol. 2023;23(11):735–748. doi: 10.1038/s41577-023-00874-w.