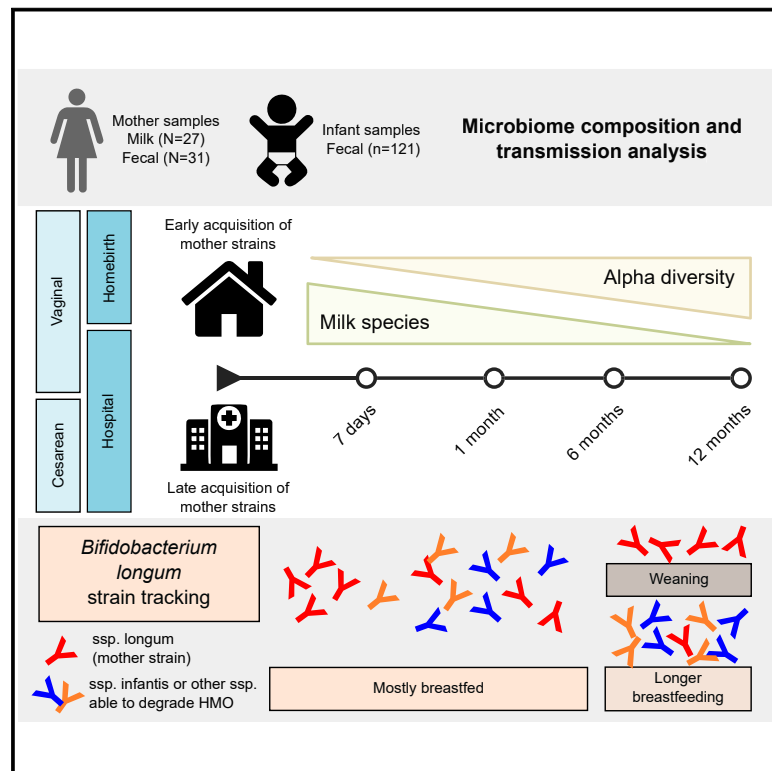


Cell Host & Microbe

Birthmode and environment-dependent microbiota transmission dynamics are complemented by breastfeeding during the first year

Graphical abstract



Authors

Marta Selma-Royo, Léonard Dubois, Serena Manara, ..., Cecilia Martínez-Costa, Nicola Segata, Maria Carmen Collado

Correspondence

nicola.segata@unitn.it (N.S.), mcolam@iata.csic.es (M.C.C.)

In brief

Selma-Royo et al. report that delivery mode and place of birth impact gut microbiota transmission, while breastfeeding practices influence the replacement patterns of *B. longum* subspecies. Overall, these factors collectively shape vertical transmission and influence infant gut colonization in early life.

Highlights

- Hospital versus at-home delivery impacts infant microbiota transmission
- Human milk is a key modulator of infant microbiota during the first year
- Species transmissibility varies across delivery modes and places, except Bifidobacteria
- Breastfeeding duration impacts *B. longum* strain retention and functional diversity



Article

Birthmode and environment-dependent microbiota transmission dynamics are complemented by breastfeeding during the first year

Marta Selma-Royo,^{1,2,11} Léonard Dubois,^{2,11} Serena Manara,^{2,11} Federica Armanini,² Raúl Cabrera-Rubio,¹ Mireia Valles-Colomer,^{2,3} Sonia González,^{4,5} Anna Parra-Llorca,⁶ Ramon Escuriet,⁷ Lars Bode,⁸ Cecilia Martínez-Costa,^{9,10} Nicola Segata,^{2,*} and María Carmen Collado^{1,12,*}

¹Institute of Agrochemistry and Food Technology-Spanish National Research Council (IATA-CSIC), Paterna, Valencia, Spain

²Department of Cellular, Computational and Integrative Biology, University of Trento, Trento, Italy

³MELIS Department, Universitat Pompeu Fabra, Barcelona, Spain

⁴Department of Functional Biology, University of Oviedo, Oviedo, Spain

⁵Diet Microbiota and Health Group, Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, Spain

⁶Health Research Institute La Fe, Neonatal Research Group, Division of Neonatology, Valencia, Spain

⁷Gerencia de Procesos Integrales de Salud. Area Asistencial, Servicio Catalán de la Salud, Generalitat de Catalunya, Centre for Research in Health and Economics, Universitat Pompeu Fabra, Barcelona, Spain

⁸Department of Pediatrics, Larsson-Rosenquist Foundation Mother-Milk-Infant Center of Research Excellence (LRF MOMI CORE), Human Milk Institute (HMI), University of California, San Diego, La Jolla, CA, USA

⁹Department of Pediatrics, Hospital Clínico Universitario, University of Valencia, Spain

¹⁰Nutrition Research Group of INCLIVA, Valencia, Spain

¹¹These authors contributed equally

¹²Lead contact

*Correspondence: nicola.segata@unitn.it (N.S.), mcolam@iata.csic.es (M.C.C.)

<https://doi.org/10.1016/j.chom.2024.05.005>

SUMMARY

The composition and maturation of the early-life microbiota are modulated by a number of perinatal factors, whose interplay in relation to microbial vertical transmission remains inadequately elucidated. Using recent strain-tracking methodologies, we analyzed mother-to-infant microbiota transmission in two different birth environments: hospital-born (vaginal/cesarean) and home-born (vaginal) infants and their mothers. While delivery mode primarily explains initial compositional differences, place of birth impacts transmission timing—being early in homebirths and delayed in cesarean deliveries. Transmission patterns vary greatly across species and birth groups, yet certain species, like *Bifidobacterium longum*, are consistently vertically transmitted regardless of delivery setting. Strain-level analysis of *B. longum* highlights relevant and consistent subspecies replacement patterns mainly explained by breastfeeding practices, which drive changes in human milk oligosaccharide (HMO) degrading capabilities. Our findings highlight how delivery setting, breastfeeding duration, and other lifestyle preferences collectively shape vertical transmission, impacting infant gut colonization during early life.

INTRODUCTION

Early-life microbial colonization plays a key role in the establishment of the human microbiota,^{1–3} which performs several essential functions.^{4,5} Initial seeding of the infant gut is set off by maternal transmission,^{6–8} and during this process, microbes establish a complex interaction with the developing systems, especially the intestinal epithelium⁹ and the immune system.^{10,11} Even though this process is not fully understood, some perinatal factors such as cesarean delivery, formula feeding, or antibiotic use could cause shifts in children's microbial colonization^{12–15}; this impact has been linked to non-communicable diseases including obesity^{16,17} and atopy,^{18–20} highlighting the impor-

tance and long-term influence of the microbe-host interplay during early life.

While a limited number of causality studies showed normalization of the microbiota in cesarean-delivered infants, notably by maternal fecal microbiota transfer,^{21,22} most of the available studies have assessed the impact of these factors solely on the infant microbiota composition,⁸ without considering their effect on vertical microbial transmission and its potential relevance for infant health. Previous analysis using strain-resolved metagenomics revealed that strains from the maternal gut microbiota were ecologically better adapted to colonizing the neonatal gut than strains from other sources,⁶ highlighting the role of the maternal microbiota as the main microbial seeding source for the newborn.^{6,7,23}



Table 1. Clinical and anthropometric characteristics of the participants

	HB (n = 10)	VAG (n = 13)	CS (n = 11)	p value
Maternal characteristics				
Maternal age	35 (4)	33 (7)	35 (2)	0.325
Pre-gestational BMI (kg/m ²)	22.71 ± 5.13	21.86 ± 2.64	22.95 ± 2.27	0.721
Gestational weight gain (kg)	13.9 ± 4.41	12.02 ± 3.86	14.27 ± 4.19	0.366
ATB during pregnancy (% pos. cases)	0 (0)	2 (15.4)	4 (36.4)	0.099
IAP (% pos. cases)	0 (0) ^a	0 (0) ^a	11 (100) ^b	<0.001
Infant data				
Gestational age (weeks)	40 (1.75) ^b	39 (1) ^a	40 (1.5) ^b	0.010
Sex (female/male)	7/3 (70/30)	9/4 (69/31)	3/8 (27/73)	0.079
Siblings (% pos. cases)	4 (40)	5 (38.5)	4 (36.4)	0.999
Pets	6 (60)	7 (53.9)	2 ^c (20)	0.108
Breastfeeding duration (months)	12 (0) ^a	6 (8) ^b	12 (5.5) ^{a,b}	0.021
Exclusive breastfeeding duration (months)	5.75 (0.5)	2 (4.75)	5 (2)	0.020
Breastfeeding type during the first 4 months of life^d				
Exclusive breastfed	9	6	9	0.167
Partially breastfed	1	5	1	
No breastfed	0	2	1	
Infant anthropometric data				
Birth weight (g)	3,489 ± 543	3,252 ± 444	3,608 ± 521	0.220
Birth height (cm)	51.2 ± 1.92	49.92 ± 1.53	50.68 ± 2.08	0.258
BMI Z score at birth	-0.16 ± 1.24	-0.32 ± 0.97	0.39 ± 1.01	0.261
BMI Z score at 1 month	-0.18 ± 0.86	-0.57 ± 0.84	-0.57 ± 0.84	0.464
BMI Z score at 6 months	-0.24 ± 0.84	-0.31 ± 0.66	0.22 ± 1.07	0.299
BMI Z score at 12 months	-0.32 ± 1.53	0.17 ± 0.95	0.53 ± 1.12	0.288

Continuous variables are presented as mean ± standard deviation and median (IQR, expressed as quartile q3–q1) for the normally and non-normally distributed variables, respectively. One-way ANOVA test was used to assess differences in normally distributed variables, while Kruskal-Wallis test was performed for those with non-normal distribution (Dunn's test was performed as a post hoc test). Categorical variables are shown as the number of positive cases (percentage within the group). Differences in these variables were assessed by Fisher's exact test. IAP, intrapartum antibiotics; BMI, body mass index; ATB, antibiotics; HB, homebirth; VAG, vaginal delivery at the hospital; CS, cesarean delivery at the hospital.

^aFor each variable with significant differences between the delivery group, different letters (a-b) represent significant differences in the corresponding post hoc test.

^bFor each variable with significant differences between the delivery group, different letters (a-b) represent significant differences in the corresponding post hoc test.

^cMissing data (n = 1).

^dClassification according to breastfeeding practices was performed considering the first 4 months of life when most of the infants change from exclusive to partially breastfeeding.

Despite the importance of the maternal microbiota on neonatal seeding, little is known about how maternal lifestyle and associated factors could shape bacterial transmission to the neonate.

In this study, we analyzed the impact of place and mode of delivery along with feeding mode on infant gut colonization and vertical microbiota transmission during the first year of life. In previous studies, we²⁴ and others^{25,26} already reported the effect of place and mode of delivery on the neonatal microbiota profile up to 1 month of life. However, these initial analyses were performed using the 16S rRNA gene amplicon sequencing approach, which lacks the strain-level resolution required to assess strain sharing and therefore vertical transmission. In this study, we show that mode and place of delivery affect microbiota composition and transmission only during the first few months of life. Most importantly, we show that many species were consistently transmitted regardless of delivery mode and

place, including key infant gut microbiota species like *Bifidobacterium* spp. For the highly prevalent and transmitted *Bifidobacterium longum*, we moreover highlighted how feeding habits can strongly impact its strain retention and replacement dynamics, therefore supporting more in-depth investigations on the role of breastfeeding in microbiota seeding and establishment.

RESULTS AND DISCUSSION

A cohort to study the influence of perinatal factors on infant gut microbiota development

To investigate the impact of place (hospital vs. home) and mode of delivery (vaginal vs. cesarean) on the development of the infant gut microbiota during the first year of life, we included a total of 34 mother-infant pairs (Table 1) and followed them during this period, collecting fecal (longitudinal study) and human milk

samples (Table S1). Families were classified according to the place and mode of delivery into three groups as follows: mothers who decided in advance to give birth at home (homebirth [HB], $n = 10$); mothers who had a vaginal delivery at the hospital (VAG, $n = 13$); and mothers who had a cesarean delivery at the hospital (CS, $n = 11$). Participants were enrolled in the frame of the “maternal microbes” (MAMI) cohort,²⁷ which aims to uncover the effect of perinatal factors on infant gut colonization and health outcomes. Fecal microbiota from the infant was assessed through metagenomic sequencing at 1 week (7 days), 1 month, 6 months, and 12 months after birth. Maternal fecal and milk samples collected 1 month postpartum were also metagenomically sequenced to explore the contribution of these microbial sources in vertical mother-infant bacterial transmission. For all infant and maternal fecal ($n = 161$) and milk ($n = 27$, only 6 were kept based on sequencing depth >1 million reads) samples (Figure S1A), strain-resolved computational tools were used to track the vertical transmission of microbial strains and their retention in the infant’s microbiota over time.^{28,29} DNA was extracted and sequenced, producing a median of 31 million reads (IQR = [20–43] million) per fecal sample after quality filtering (see STAR Methods; Table S1). With human milk, even though the DNA concentration was enough for sequencing, most of the samples were removed from the final analysis due to the low number of microbial reads (in some cases, <1%; Table S1). Thus, only six samples yielded a number of microbial reads above the threshold (>1 million reads for milk samples) and were retained for further analysis.

Mothers included in the study had a normal body mass index (BMI; normal weight in the [18.5–24.9] range according to World Health Organization) with a total gestational weight gain within the recommended range.³⁰ All the infants were delivered at term, and most of them (31/34, 91%) were at least partially breastfed during the first 6 months of life (Table 1; Figure S1B). Pairs were selected from the MAMI cohort to have homogeneous breastfeeding rates at 1 week across the three different groups, even though children born at home showed longer exclusive breastfeeding and higher rates of at least partial breastfeeding at 12 months of age (Table 1; Figure S1B; Table S1). Indeed, hospital-delivered infants were less frequently exclusively breastfed during the first month of life (Fisher’s exact test, odds ratio [OR] = 0, p adj. = $8.89\text{e-}3$; Figure S1B; Table S1) and less frequently at least partially breastfed from 6 to 12 months (Fisher’s exact test, OR = 0, p adj. = $3.61\text{e-}2$; Figure S1B; Table S1). All and only cesarean delivery pairs were exposed to intrapartum antibiotics (IAPs; Table 1).

Hospital versus at-home delivery impacts microbiota composition of MAMI infants

To assess the temporal evolution of the early-life gut microbiota composition, we analyzed the infant fecal samples collected during the first year of life. As reported in previous studies,³¹ the infant gut microbiota steadily acquires new species and progressively resembles an adult-like composition, a pattern already evident at 12 months (Figure 1A; Table S2). In the first month and especially during the first week, the neonatal microbiota was dominated by the presence of *Escherichia coli* and *Bifidobacterium* sp., such as *B. breve* and *B. longum* (highest 90th percentile of abundance at 1 week: *E. coli* and *B. breve* in cesar-

ean delivery, *B. breve* and *B. longum* in HB, *E. coli* and *B. longum* in vaginal group; Figure 1A; Tables S2 and S5). These species are considered pioneers of the infant microbiota³¹ since they are acquired before the first week but were also among the most abundant during two time points (highest 90th percentile of abundance, at 1 month: *B. breve* and *B. longum* in all groups, at 6 months: *B. breve* and *B. longum* in cesarean and *B. longum* and *B. pseudocatenulatum* in HB and vaginal groups; Figure 1A; Tables S2 and S5). At 1 year, the infant gut microbiota had shifted to a more adult-like profile, closer to the one observed in the maternal 1-month postpartum gut microbiota (Figures 1A and 1B; Table S2), with the Jaccard distance between infant and mother being lower at 12 months compared with the first month of life (post hoc Dunn test all p adj. < 0.05; Figure S1B; Table S2). The new species appearing at 6 and 12 months are known for their capacity to ferment different carbohydrates^{32–34} including *Faecalibacterium prausnitzii* and several species from *Blautia*, *Eubacterium*, and *Roseburia* genera. These changes in microbiota composition were overall characterized by an increasing Shannon diversity over infant age, especially at 12 months (Kruskal-Wallis test $\chi^2 = 39.74$, post hoc Dunn tests between 12 months and other time points: all p adj. < 0.05; Figure 1C; Table S2), which however remains lower than in mothers (post hoc Dunn test p adj. = $5.89\text{e-}3$; Figure 1C; Table S2). We hypothesize that these changes are probably due to weaning and the introduction of solid food (exclusively breastfed infants: at 6 months = 11/32 [34%], at 12 months = 0/29 [0%]; Figure S1B; Table S1). We specifically tested the differences in prevalence of species before and after introduction of solid food (as the end of exclusive breastfeeding). We found 21 species having differential prevalence (Fisher’s exact test p adj. < 0.05; Table S2), but only 2 were more prevalent during exclusive breastfeeding (OR > 1): *Staphylococcus hominis* and *Staphylococcus epidermidis*, likely due to skin contact between mother and infant during breastfeeding.

The maturation of the overall gut microbiota to resemble the maternal one is further supported by temporal shifts in the difference between mother and infant overall microbiota composition (i.e., beta diversity). Indeed, even if all time points’ microbiota composition of mothers and infants was different (PERMANOVA adj. $R^2 = 14.41\%$, $p = 1\text{e-}3$; Figure 1B; Table S2; STAR Methods), this difference was the highest at 1 week and decreased over the infant’s first year of life (adj. $R^2 = 22.81\%$, 19.52%, 17%, and 13.86% and all p adj. = $1\text{e-}3$ at 1 week, 1 month, 6 months, and 12 months, respectively; Table S2). While age conveys a significant portion of the infant microbiota variability among the covariates, the delivery condition also had a significant impact, with place of delivery (home vs. hospital, cesarean section excluded) being more important (stepwise PERMANOVA infant age adj. $R^2 = 8.03\%$, delivery group adj. $R^2 = 0.93\%$ beyond infant age, $p = 2\text{e-}3$) than mode of delivery (cesarean vs. vaginal at the hospital; stepwise PERMANOVA infant age adj. $R^2 = 9.88\%$, $p = 2\text{e-}3$ and delivery adj. $R^2 = 0.25\%$ beyond infant age, $p = 4.4\text{e-}2$). Over the first year of life, when considering individual time points, the place of delivery had a more persisting impact on the variability of the infant microbiota composition (PERMANOVA adj. $R^2 = 2.21\%$, 2.28%, and 1.83% and all p adj. = $2\text{e-}2$ at 1 week, 1 month, and 12 months, respectively), compared with mode of delivery

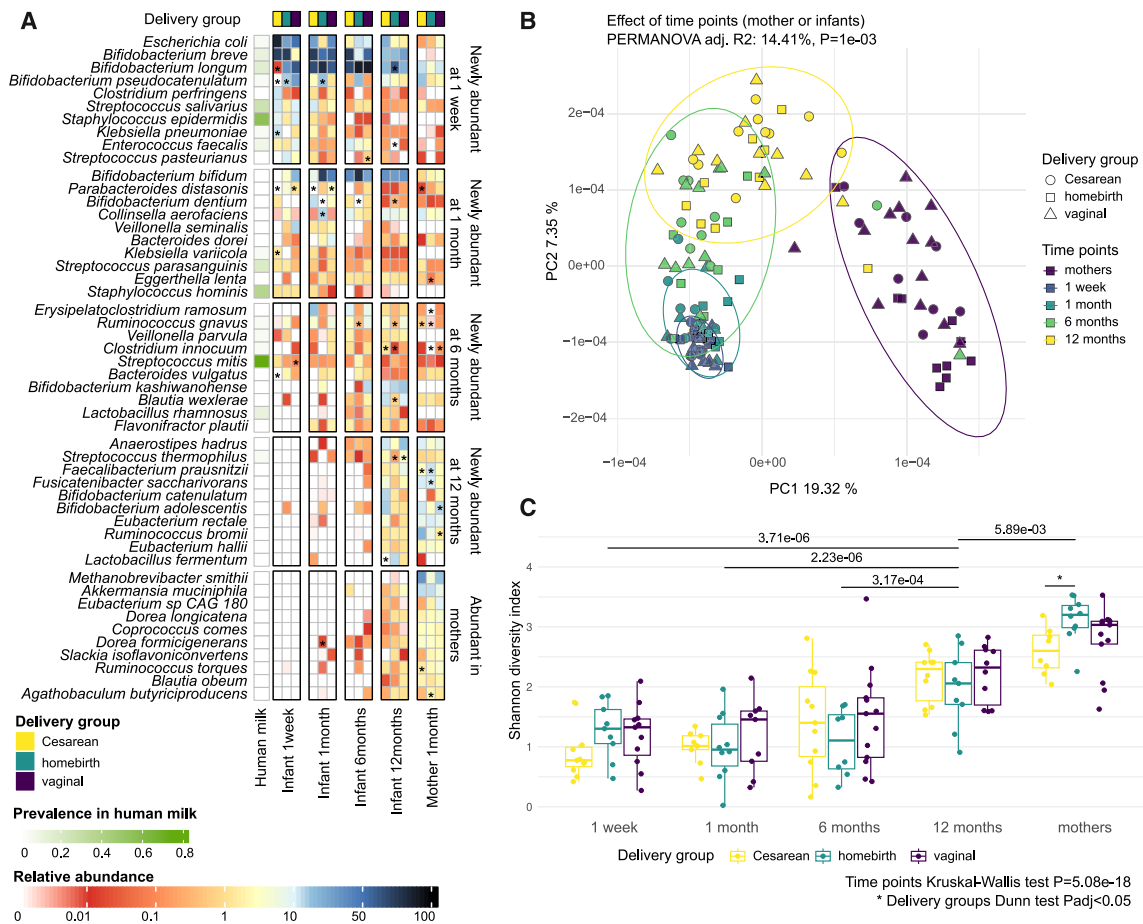


Figure 1. Mother and infant microbiota composition according to time point and place and mode of delivery

(A) Microbiota taxonomic profiles at each collection time point (1 week and 1, 6, and 12 months). The microbiota taxonomic composition of maternal samples at 1 month is shown in the last column of the figure. At each delivery-mode time point pair, the species' relative abundance is summarized by its 90th percentile across all samples. At each time point, only the 10 most abundant species at that time point, which are not included in the previous time points, are displayed as "newly abundant." Newly abundant species are not newly acquired at that specific time point and might be present at lower abundances in previous ones. The prevalence of microbial species in the 27 human milk samples is reported on the left of the main heatmap. Extended heatmap available in Figure S2C. Complete species relative abundance table available in Tables S2 (Tab1 and Tab4) and Tab5.

(B) Ordination (PCoA) of time effect on the infant gut microbiota composition. Maternal samples are also included in the plot. The symbol shape indicates the mode/place of delivery. (All PERMANOVA results in Table S2).

(C) Alpha diversity of gut microbiota expressed as Shannon diversity index over time points (Kruskal-Wallis test $\chi^2 = 87.24$, $p = 5.08e-18$). Post hoc Dunn test p adj. is shown. Only mothers have different Shannon diversity between homebirth and cesarean delivery groups (Kruskal-Wallis test $\chi^2 = 7.07$, $p = 2.91e-2$; post hoc Dunn test p adj. = $1.44e-2$; Table S2).

(only significant at 1 week: PERMANOVA $R^2 = 2.36\%$, p adj. = $5e-2$). The infant sex did not influence the composition at any time point (all PERMANOVA p adj. > 0.05). In addition, it was never selected in stepwise models beyond delivery mode (all p adj. > 0.05).

Accordingly, while in the initial time points (≤ 1 month of life), we observed differences in species composition of the infant gut according to the mode of delivery (CS w.r.t. vaginal delivery groups; Figure 1A; Table S2); in the following time points the most important differences were found by place of delivery: HB w.r.t. hospital deliveries. While no species reached statistical significance (p adj. < 0.05; possibly due to the high number of tests), several of them were significant before adjustment ($p < 0.05$). Then, 1 week after birth, CS neonates showed a lower relative abundance of *Bifidobacterium longum* (Wilcoxon rank-

sum test $r = 0.379$, $p = 3.40e-2$) and *Bacteroides vulgatus* ($r = 0.373$, $p = 3.72e-2$) and an enrichment in *Klebsiella pneumoniae* ($r = 0.480$, $p = 7.22e-3$) and *K. varicola* ($r = 0.402$, $p = 2.446e-6$), while at 12 months, children born at home harbored a microbiota richer in *Bifidobacterium longum* ($r = 0.403$, $p = 3e-2$) and *Enterococcus faecalis* ($r = 0.432$, $p = 2.15e-2$) and diminished in *Ruminococcus gnavus* ($r = 0.551$, $p = 2.04e-3$), *Clostridium innocuum* ($r = 0.566$, $p = 2.48e-3$), or *Streptococcus thermophilus* ($r = 0.373$, $p = 4.70e-2$), compared with the other two groups (all results in Table S2). Related to this, we found an effect of HB on the infant microbiota that is detectable also after 6 months of life when the microbial profile of the infant becomes more similar to the maternal one. We hypothesized that the shifts in the microbiota composition of infants delivered at home could be related to the differences in maternal lifestyle

and environment, which are reflected in divergences in their microbiota composition with respect to hospital-delivering mothers (Figure S2). In fact, differences in Shannon diversity were also observed across delivery groups in mothers exclusively (Kruskal-Wallis test $\chi^2 = 7.07$, $p = 2.91e-2$; Table S2), with mothers from the HB group having increased diversity with respect to the cesarean group (post hoc Dunn test p adj. = $1.43e-2$; Table S2). Considering that mother samples were collected 1 month after birth, this could be due to intrapartum antibiotics that were used exclusively in the cesarean births and could impact mother microbiota diversity, compared with non-exposed mothers. However, some epidemiologic registers previously described that mothers who planned a home birth showed specific population-based differences, some of which related to their lifestyle.^{35,36} We explored diet as a potential factor modulating the maternal microbiota from HB mothers; however, no differences were observed in the main nutritional categories according to delivery groups (Table S1). Then, other maternal-related factors not surveyed in this study, such as lifestyle, could also be involved in the observed differences in maternal microbiota composition.

Human milk as a key modulator of microbiota seeding dynamics during the first year

Among factors surveyed here, the decision to give birth at home was previously associated with longer breastfeeding duration,³⁷ as we observed also in our study population (Fisher's exact test, OR = 0, p adj. = $3.61e-2$; Table S1). While this association is probably non-causal, the implication of HB-correlated factors on the reported observations needs to be considered.

In order to survey the role of breastfeeding in shaping microbiota assembly in different birth groups, we investigated the impact of human milk as a potential source of infant microbiota species by analyzing the microbial composition of the 27 shotgun metagenomics samples from human milk (10/27 samples from mothers who vaginally delivered at the hospital, 8/27 C-section, 9/27 from HB). Due to the limited number of reads in some samples (median = 0.24 million reads; min = 53, 133; max = 31 million; Table S1), we focused on the sole presence or absence of species and not on their relative abundance. Interestingly, species present in human milk samples are also among those most abundant in the infant gut microbiota during the first 6 months, such as *Bifidobacterium longum*, *Staphylococcus epidermidis*, and *Enterococcus faecalis* (Figures 1A and S1C; STAR Methods). Besides this, the species found in milk samples represented a decreasing fraction of the species richness in infant gut. At 1 week, 43.06% of the species present in the infant gut were species found in human milk (IQR = [35.61%–46.15%]). Over the first year of life, this fraction steadily diminished, as fewer and fewer species found in human milk were found in the infant's gut (median 24.04%, 15.13%, and 8.13%, respectively, at 1 month, 6 months, and 12 months; Kruskal-Wallis $\chi^2 = 37.2$, $p = 4.17e-8$; Figures S1D and S1E; Table S2). Despite the limited sample size and low read count discarding many samples in the strain analysis, we observed a number of strain-sharing events between milk and gut samples of the mother ($N = 1$) and between milk and gut samples of the infant ($N = 4$). These include *Enterococcus faecalis* and *Escherichia coli* (Table S3).

Our results suggest, through metagenomics, that human milk could represent one of the sources of bacterial species seeding the infant microbiota during the first months of life, expanding on previous studies conducted using other sequencing approaches and culture-dependent methods.^{38,39} This needs to be further confirmed in an analysis that specifically addresses the transmission between human milk microbes and infant gut, using strain-resolution metagenomics techniques and probably bigger cohorts.

Species transmissibility varies across delivery modes and places, except for Bifidobacteria

To better understand the effect of mode and place of delivery on infant gut microbiota acquisition, we applied strain-resolved mapping-based metagenomic tools^{28,29} to test what fraction of the bacterial strains found in the gut of infants were shared with their respective mothers, thus representing possible vertical transmission events (see STAR Methods; Table S3). To this aim, we calculated two estimators that characterize the maternal vertical transmission, named “strain-sharing rate” and “infant acquisition rate” (see STAR Methods). The strain-sharing rate is defined as the fraction of the species that are present in both mother and infant, which are shared at the strain level. The infant acquisition rate reflects the percentage of all the infant gut strains that are shared with the mother. These two metrics give different information about vertical transmission: the former highlights how likely it is for species found in both mother and infant to be shared at the strain level (i.e., could be transmitted from the mother), whereas the latter highlights the overall impact of the mother as a source of seeding strains in the infant's gut by showing how much of the infant's gut microbiota could have a maternal origin. We finally defined “species transmissibility” as the fraction of families in which a mother-infant sharing event is detected for a given species. It highlights species frequently involved in strain-sharing events.

Overall, the strain-sharing rate was higher at early time points and later decreased (Kruskal-Wallis $\chi^2 = 10.04$, $p = 1.82e-2$; post hoc Dunn tests between 1 week and 12 months p adj. = $2.5e-2$, between 1 and 12 months p adj. = $3e-2$; Figure 2A; Table S3), probably as a consequence of increased diversity of their microbiotas with the introduction of solid food. During the first year of life, we observed no significant change in the fraction of the infant microbiota that was shared with the mother. Infant acquisition rate (i.e., the percentage of infant gut strains found also in the mother) was indeed stable over time (Kruskal-Wallis test $\chi^2 = 4.12$, $p = 0.25$; Figure 2A; Table S3) and similar across the different groups (Kruskal-Wallis test $\chi^2 = 2.88$, $p = 0.24$; Figure 2B; Table S3).

By looking at the species that were most shared at the strain level between mother and infant, we exposed interesting patterns linked to place and mode of delivery (Figure 2C). Some species (16.13%, 10/62) were indeed vertically shared across all delivery groups, as in the case of four *Bifidobacterium* species (*B. longum*, *B. pseudocatenulatum*, *B. bifidum*, and *B. adolescentis*) whose transmissibility was consistently higher than 33% (Figure 2C). Interestingly, *B. longum* was transmitted in more than 50% of the families where it was present as a species regardless of place or mode of delivery, and it was extremely prevalent in vaginally delivered infants soon after birth

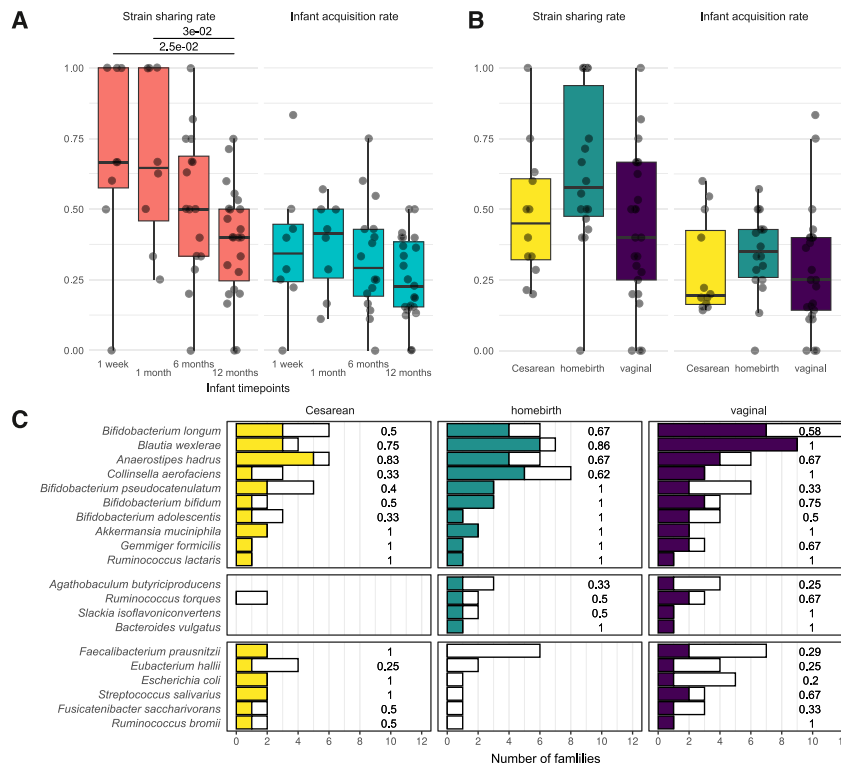


Figure 2. Overall strain acquisition and strain-sharing rates in the mother-infant cohort

(A) Infant-mother strain sharing and baby acquisition rates by time points.

(B) Infant-mother strain-sharing and baby acquisition rates by delivery groups (across all infant points). For (A) and (B), strain-sharing rate is defined as the number of shared species' strains divided by the number of species found in both infant and mother samples. Baby acquisition rate is the number of shared species' strains divided by the number of species detected in the infant gut microbiota. Only strain-sharing rate shows significant differences over time (Kruskal-Wallis $\chi^2 = 10.04$, $p = 1.82e-2$), significant post hoc Dunn tests results are shown (Table S3).

(C) Mother-infant transmissibility of the most transmitted species according to delivery groups. Transmissibility is computed as the number of families (length of the barplot) in which a species is shared at the strain level between mother and infant (colored part of the barplot) divided by the number of times the species is found in both microbiotas (blank part of the barplot). Here, results are organized into three groups to show species transmitted in all delivery modes/places, only in vaginal delivery, and only in hospital delivery. Complete plot is available in Figure S4.

(prevalence at 1 week: CS 9%, VAG 55%, and HB 67%; Table S2). An even more striking observation regards *Blautia wexlerae* (Figure 2C), which was shared at the strain level in >75% of the families where it was present as a species, regardless of place/mode of delivery, but reached 100% transmissibility in VAG. As for *B. longum*, *B. wexlerae* was highly prevalent, with peaks in 12-month-old infants and in mothers (prevalence at 12 months: CS 60%, VAG 90%, and HB 66.7%; prevalence in mothers: CS 87.5%, VAG 66.7%, and HB 100%; Table S2).

The potential role of these maternally transmitted species in the infant intestinal and immune system development is still to be fully elucidated, despite *Bifidobacterium* being one of the most studied genera in the field of early-life microbiota, consistently found among the most vertically transmitted^{40–43} with probable implications for infant development.^{11,44} Fewer studies instead focused on *Blautia wexlerae*, which was previously found to be depleted in obese children and those with insulin resistance, while showing an anti-inflammatory effect in peripheral blood mononuclear cell cultures *in vitro*, suggesting a potential interaction with the immune system in early life.^{45,46} In contrast to our results on *B. wexlerae*, Nayfach et al. found that it was rarely transmitted by the mother and usually acquired from other sources.⁴⁷ Thus, the relevance of these species in vertical transmission needs to be further explored.

We moreover observed that some species were transmitted only in vaginal deliveries, regardless of the place where delivery occurred. This is the case for *Agathobaculum butyriciproducens*, *Ruminococcus torques*, *Slackia isoflavoniconvertens*, and *Bacteroides vulgatus* (Figure 2C). Another interesting group of species (including *F. prausnitzii*, *E. coli*, and *Streptococcus salivarius*) was present but never shared between mother and infant

in the HB delivery groups, highlighting a possible effect of hospital environment on their transmission. Interestingly, species linked with vaginal delivery were totally absent or very lowly prevalent in cesarean-delivered babies. In agreement with our results, Mitchell et al. found several *Bacteroides* spp. enriched in vaginally delivered infants, compared with those born by cesarean delivery, with *Bacteroides vulgatus* being the species with the most identified transmission events in vaginal delivery.¹² Similar results were described by Wampach et al. in a different cohort⁴² where they also identified several functional pathways with immune-stimulatory potential enriched in vertically transmitted strains in vaginal deliveries.

On the contrary, species associated with hospital delivery were similarly prevalent in HB-delivered babies, although not shared with their mother (Figure 2C). This suggests that these species are not acquired from the hospital environment but that in the clinical setting, these species are more often acquired from the mother, whereas at home, they are acquired from other sources (median transmissibility 50% vs. 0%) probably because of the less sanitized environment.

Most of the maternal contribution to infant gut microbiota occurs earlier in HB neonates

Taking a deeper look at the transmission dynamics, we identified differences in the first acquisition time across species and delivery groups. We observed that during the first month of life, vaginally delivered infants acquire a higher number of new species from their mothers, irrespective of place of delivery (until 1 month of age, CS vs. vaginal delivery groups, Fisher's exact test OR = 3.44, p adj. = 2.53e-2; Figure 3A; Table S3). However, in the following months (6–12 months), the infants delivered at home

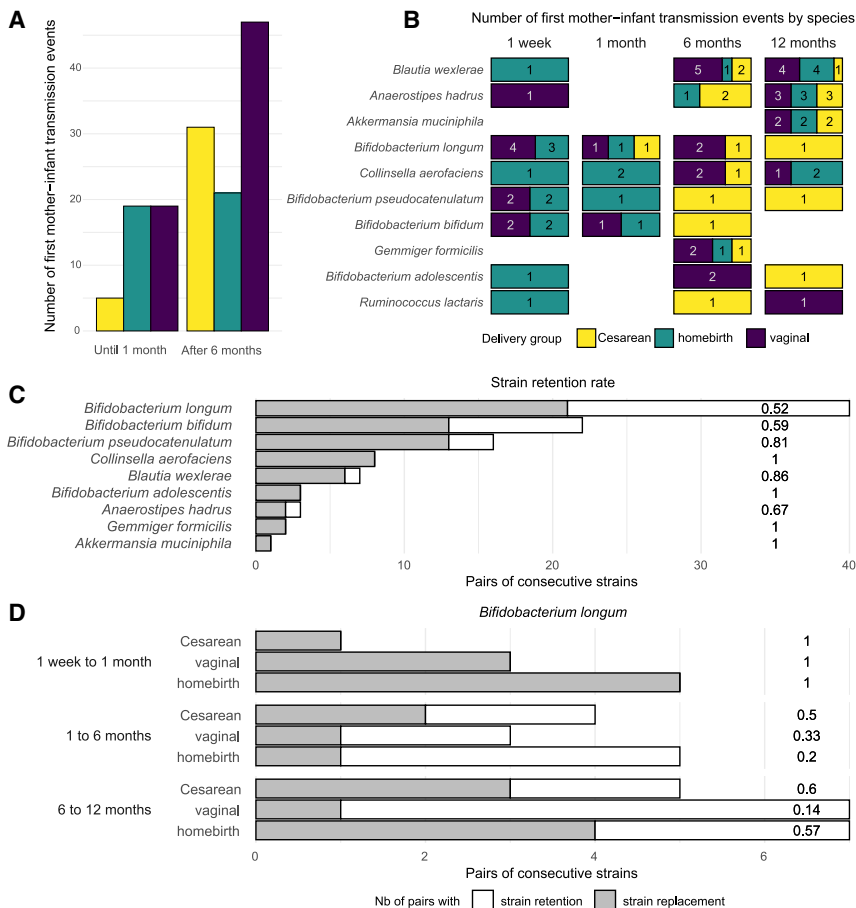


Figure 3. Time course of the transmission events occurrence and strain retention rates by species

(A) Overall counts of the first transmission event across all species in delivery groups until the first month of life (1 week and 1 month time points) and after 6 months (6- and 12-month time points). The cesarean group shows a delayed acquisition, with more strains being first acquired after 6 months than in the first month of life (Fisher’s exact test OR = 3.44, p adj. = $2.53e-2$) while the homebirth group has anticipated acquisition with less first acquisitions, compared with the hospital delivery groups after 6 months (Fisher’s exact test OR = 0.34, p adj. = $2.53e-2$; Table S3).

(B) Time of first transmission of the maternal strain in each family for those species transmitted in all delivery groups during the first year of life. Complete plot available at Figure S5.

(C) Retention rate is defined as the number of strains persisting over two time points divided by the number of consecutive strains detected and typed with StrainPhlAn. Only species with transmission events detected in every delivery group are shown in the plot.

(D) *Bifidobacterium longum* retention rate by delivery group and time point.

share fewer newly transmitted species than the hospital-delivered ones (at 6 months and after, HB vs. hospital delivery groups, Fisher’s exact test OR = 0.34, p adj. = $2.53e-2$). When comparing the first month of life with the 6- to 12-month period, cesarean-delivered infants received 6.2 more strains for the first time after 6 months, vaginal delivered at the hospital 2.47 more, and homebirth infants 1.11 times more strains (Table S3). This suggests that while CS limits early-life strain sharing, home-born neonates acquire most of the maternal strains earlier than the other groups and that at later time points, strain transmission from the mother could be altered by different place-dependent or lifestyle factors. We hypothesized that one of these factors could be the longer breastfeeding duration in the HB group, which could hinder the acquisition of new adult-like species in replacement of those specialized in human milk sugar metabolism. Indeed, hospital-delivered infants were less frequently exclusively breastfed during the first month of life (Figure S1B; Table S1).

Among others, species like *B. longum*, *Bifidobacterium pseudocatenulatum*, and *Bifidobacterium bifidum* were first acquired earlier in life in vaginally delivered newborns regardless of the place of delivery (Figures 3B and S4), in agreement with previous works.^{48,49} Interestingly, these early-acquired species in vaginally delivered infants were also among those that showed higher rates of strain replacement (i.e., the portion of consecutive time points in the same infant presenting different strains; see STAR Methods) during the first year of life (Figure 3C). The mech-

anisms that may drive this increased pace of early-strain replacement and its effect on infant microbiota development still need to be elucidated, and they may include the massive changes in diet typically happening during the first year.

Conversely, other species like *Akkermansia muciniphila*, *Anaerostipes hadrus*, and *Gemmiger formicilis*, which were first acquired later in time, were shared at the strain level with the mother at similar rates regardless of mode and place of delivery. These species are typical components of the adult microbiota, and they are known to be acquired later in life,⁵⁰ probably due to their specialization to use different carbon sources than those coming from human milk.⁵¹

Taken together, these observations support the species-specific variability in transmissibility and strain retention, highlighting that strain transmission dynamics might be driven and supported by specific events, such as the delivery by cesarean that causes a delay in strain acquisition,⁵² the birth at home, or the change in diet during weaning.

Breastfeeding duration impacts strain retention and functional diversity of *Bifidobacterium longum*

To better investigate the hypothesis that strain retention dynamics might be related to changes in infants’ diet, we explored in depth the transmissibility and strain replacement patterns of *B. longum*. This species is revealed as one of the earliest colonizers of the MAMI infants’ gut microbiota (Table S2), and it is also among the ones most vertically transmitted (Figure 2C). It is moreover enriched in children born at home (Table S2), potentially due to prolonged breastfeeding duration in these children (Table S4).

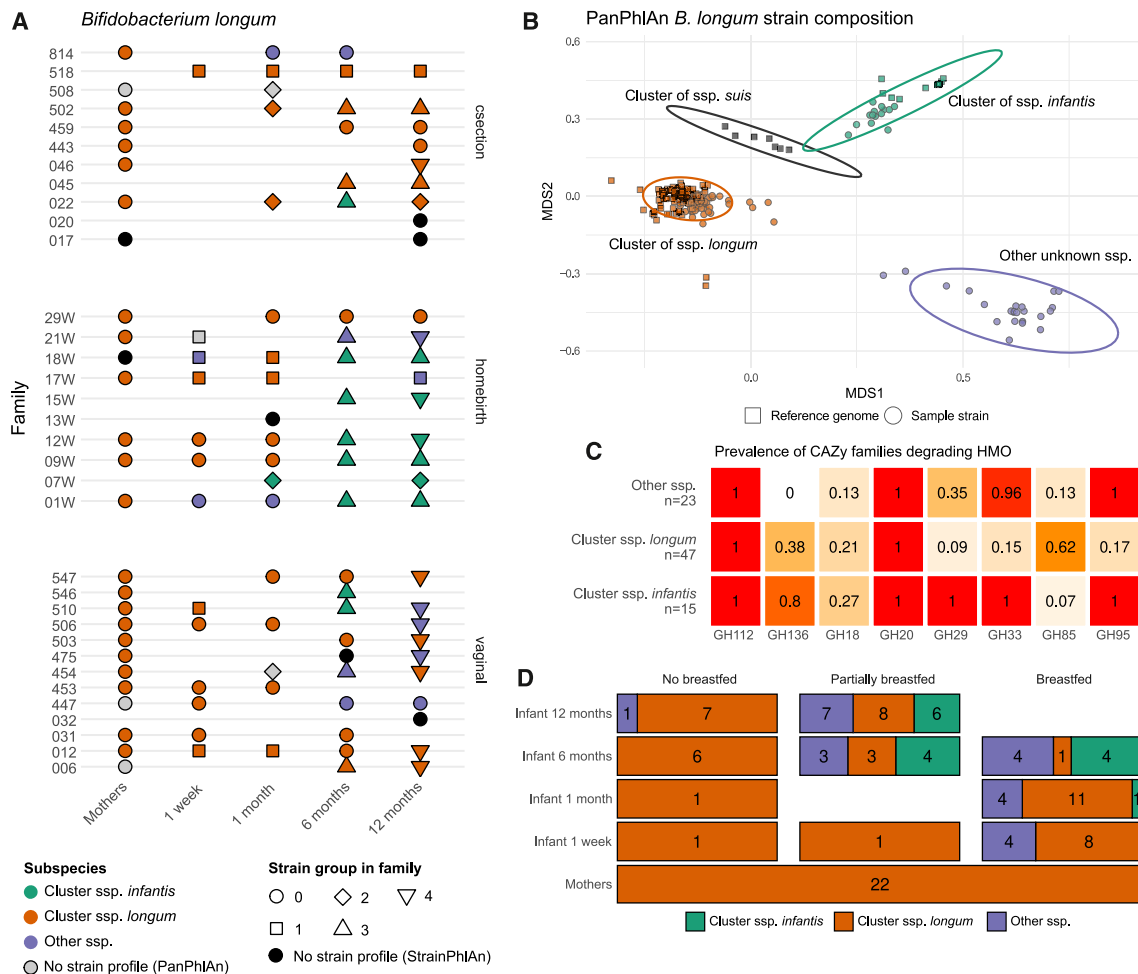


Figure 4. *Bifidobacterium longum* retention and change in functional diversity over time

(A) Each type of dot represents one strain in a given family. The absence of a dot means that the species is not detected by MetaPhlAn at that time point. A black round dot represents a sample that has been discarded by StrainPhlAn (species present but strain not typeable). A gray dot represents a sample that has been discarded by PanPhlAn. Other colors indicate subspecies of *B. longum* predicted with the PanPhlAn analysis of the panel. Shapes indicate strain groups within a family; similar shapes thus indicate transmission or persistence over time.

(B) Multidimensional scaling (MDS) scaling of the *B. longum* PanPhlAn profile (binary matrix of each strain presence-absence of each gene in the pangenome), highlighting 3 distinct clusters including reference genomes of *B. longum* ssp. *longum* and ssp. *infantis*. One cluster lacking reference genomes is labeled as “other ssp”.

(C) Prevalence of the CAZy families involved in HMO degradation across the 3 subspecies (Table S4).

(D) Relative prevalence of each subspecies over time, according to breastfeeding status.

Indeed, *B. longum* is the species showing the lowest retention rate overall (52%; Figure 3C), with the strain acquired in the first week always conserved until the first month (Figure 3D), followed by a drop in the retention rate between 1 and 6 months, with more than half of the strains being replaced (average retention rate = 34%). Strikingly, this retention rate rose back between 6 and 12 months but only in the CS and HB delivery groups, in which more than half of the strains were conserved (Figure 3D). With strain replacements occurring mainly at 6 months (Figures 3C and 4A), the recommended weaning age⁵³ when most of the infants of the cohort are no longer exclusively breastfed (exclusively breastfed infants: at 6 months = 11/32, at 12 months = 0/29; Figure S1B), we hypothesized that different strains of *B. longum* might have different capacities of adaptations to the

changing diet. Indeed, as the replacement is never observed in the very first month of life, nor showed a straightforward association with the mode of delivery, it suggests that the difference in strain retention/replacement dynamics could be due to the different breastfeeding practices and duration across delivery groups (Figure S1B), as we already observed for the delayed first acquisition of new species in infants breastfed for longer.

To explore this hypothesis, we profiled the *B. longum* strain composition with Pangenome-based Phylogenomic Analysis (PanPhlAn 3),²⁹ a tool that exploits species-specific pangenomic information to expose genes’ presence-absence patterns (Table S4). This pangenome analysis highlighted four clusters of *B. longum* strains that we further characterized based on their closest reference genomes (Figure 4B; see

STAR Methods). Two clusters were identified as *B. longum* ssp. *longum* and *B. longum* ssp. *infantis*, while another one contained only reference genomes closely related to *B. longum* ssp. *suillum/suis*, which was not detected in our population. A fourth cluster contained only gut strains from our cohort but no available reference genome and was thus labeled as “other subspecies.”

We also labeled genes from the PanPhlAn matrix with CAZy families⁵⁴ by using the dbCAN2 database (see STAR Methods) and focused on genes involved in carbohydrate degradation (glycosyl hydrolases [GHs]; Table S4). PanPhlAn analyses defined a core and an accessory GH repertoire. Interestingly, this last variable part contains most of the enzymes needed for human milk oligosaccharide (HMO) degradation (GH18, GH20, GH29, GH33, GH85, GH95, GH112, and GH136⁵⁵), supporting the hypothesis that the *B. longum* strain replacement is related to weaning and changes in carbon sources availability.

To investigate the catabolic capabilities of these subspecies, we focused on the CAZy families involved in HMO degradation, including GH18 (endo- β -N-acetylglucosaminidase), GH29 (α -L-fucosidase), GH136 (lacto-N-biosidase), GH33 (2,3-2,6- α -sialidase), and GH95 (α -1,2-L-fucosidase), among others.⁵⁵ We found that some specific genes with α -L-fucosidase and sialidase activity (GH29, GH95, and GH33, respectively) were more prevalent in the ssp. *infantis*, compared with the ssp. *longum* (Fisher’s exact test OR = 11.32, 5.73 and 6.53 and p adj. = 2.22e-4, 2.03e-3, and 2.03e-3, respectively; Figure 4C; Table S4). In addition, analysis of subspecies dynamics along with CAZy gene presence indicates that replacement of a ssp. *longum* strain by a ssp. *infantis* strain is significantly associated with the gain of GH29, GH33, and GH136 genes, compared with retention of ssp. *longum* strains (Fisher’s exact tests, OR = 0; p adj. = 1.5e-3, 5.2e-4, and 2.1e-2, respectively; Table S4). The acquisition of these CAZy plays an essential part in HMO degradation as fucosidase and sialidase activities (expressed by GH29 and GH33) remove L-fucose (Fuc) and N-acetylneuraminic acid (Neu5Ac) decorations from HMOs, allowing them to be then degraded by lacto-N-biosidase enzymes expressed by GH20 and GH136 families.⁵⁶ Consequently, ssp. *infantis* seems to outperform ssp. *longum* in HMO degradation.

Moreover, we observed two clear trends by linking these subspecies’ presence with the time of collection and breastfeeding practices (Figure 4D): the first maternal strain characterized was always ssp. *longum*, which became less prevalent in the infant’s gut over time and especially at 6 months, when we saw the greatest rate of strain replacement in favor of the ssp. *infantis* (Figure 4D). However, this frequent replacement was observed only in infants that were still at least partially breastfed (Fisher’s exact tests of ssp. *longum* vs. others ssp. at 6 months, OR = 0, p adj. = 2.37e-3 and at 12 months, OR = 0.09, p adj. = 3.52e-2; Figure 4D).

Overall, the *B. longum* subspecies had a very different CAZy repertoire, with strains of *B. longum* ssp. *longum* harboring more genes coding for GH not involved in HMO degradation and likely more useful in the case of a mixed diet (Kruskal-Wallis test $\chi^2 = 66.3$, $p = 4.1e-5$; Dunn test p adj. = 4.3e-10 and 6.47e-10 with ssp. *infantis* and the other ssp., respectively; Figure 5A). In addition, we used maternal secretor status ($n = 20$; STAR Methods; Figure 5B) and HMO metabolomic profiles of the

corresponding human milk samples. The presence of *B. longum* ssp. *longum* was clearly associated with non-secretor mothers (Fisher’s exact test of ssp. *longum* vs. others ssp. OR = 0.12, $p = 1e-3$; Figure 5C). It highlights that the presence of enzymes GH29 and GH95 in the genomes of *B. longum* subspecies provides a clear advantage over the other subspecies when fucosylated HMOs are present in the gut environment.

Regarding this, other studies have reported variations in HMO composition associated with lactation period.^{57,58} However, a recent study that assessed the *B. longum* subspecies dynamics did not find an association between HMO composition, neither with specific *B. longum* subspecies nor strain replacement.⁵⁹ Further studies with longitudinal sample collection of both infant fecal and milk samples could explore this interesting issue to have a complete overview of how human milk and its components could support the strain replacement in the infant microbiota.

These results support the idea that the microbiota of infants with longer exclusive or partial breastfeeding duration (Figure S1B) harbors more strains specialized in human milk-derived energy sources, potentially delaying the acquisition of new maternal species.

Overall, these results suggest that the strain replacement of *B. longum* occurring over time in the infant’s gut is highly dependent on the breastfeeding duration and that while the strain is replaced, it is often switching from *B. longum* ssp. *longum* (maternal strain) to either *B. longum* ssp. *infantis* or one of the other subspecies groups because of their better ability to degrade HMOs. Vatanen et al. also found the replacement of *Bifidobacterium longum* subspecies related to the transition from exclusive to partial breastfeeding in a population of Bangladeshi children.⁶⁰ Interestingly, they identified a specific group of *B. longum* strains that would be specialized to metabolize both human milk-derived sugars and complex carbohydrates from their mixed diet. They hypothesized that this subgroup would be adapted to each mixed diet and could be different between geographical locations. In fact, we also describe the switch between *Bifidobacterium* subspecies triggered by the transitional feeding period when the introduction of solid food is complemented with human milk. Similar results regarding the appearance of *B. longum* ssp. *infantis* late in breastfeeding were also recently reported by Ennis et al.⁵⁹ in a population of 21 children, supporting the observation of the present study.

In addition to these findings, our analysis also highlighted some technical limitations common to most metagenomic studies. More specifically, PanPhlAn profiling of *B. longum* strains often raised a warning of putative multi-strain detection. In that case, only the gene composition of the dominant strain is assessed. The same happened with the StrainPhlAn analysis, where only the dominant strain is characterized. This is a limitation common to most software for strain detection from metagenomic samples.⁶¹ However, recent advances paved the way for multi-strain detection.^{62–64} We showed that the initial *B. longum* strain is replaced over time, but further analysis is however needed to confirm whether or not the new strain is acquired *de novo* or if a change in the diet will disrupt the fitness of the dominant and non-dominant strain, leading to a shift in their relative abundance that would be detected as strain replacement. Another limitation of the study is the reduced sample size that prevents the inclusion of other host-related factors in the

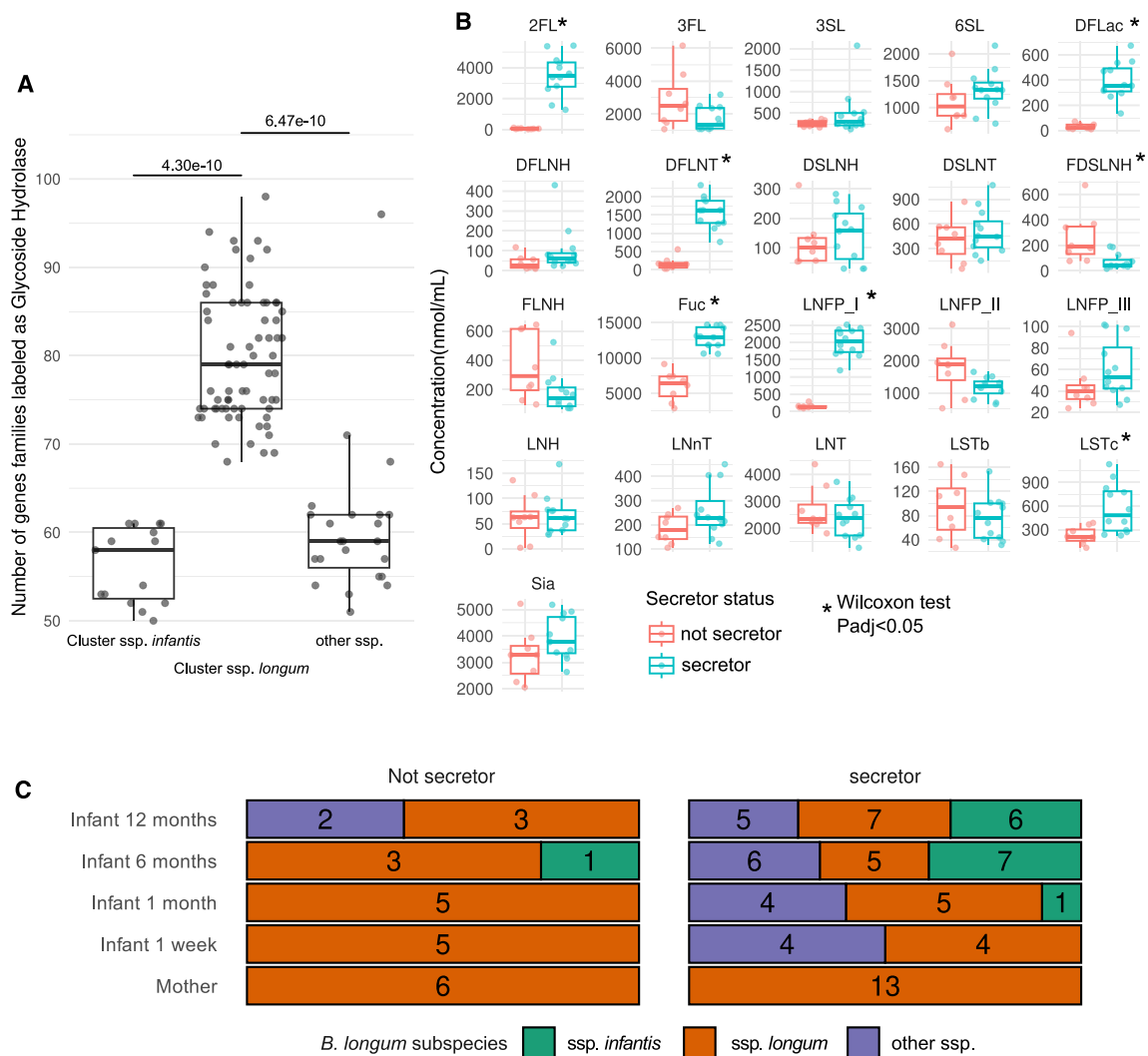


Figure 5. *B. longum* subspecies, human milk HMO profile, and secretor status

(A) Comparison of GH richness between *B. longum* subspecies. GHs related to HMO degradation (GH18, GH20, GH29, GH33, GH85, GH95, GH112, and GH136) are not counted. Kruskal-Wallis $\chi^2 = 66.3$, $p = 4.1 \times 10^{-5}$. Only significant post hoc Dunn tests are shown (p adj. < 0.05).

(B) Human milk samples' HMO composition. Only samples for which secretor status is known are used ($n = 20$). * indicates a significant difference in concentration of the HMO between secretors and non-secretors.

(C) Prevalence count of *B. longum* subspecies according to secretor status. *B. longum* ssp. *longum* is more frequent in infants from non-secretor mothers (Fisher's exact test OR = 0.12, $p = 1 \times 10^{-3}$).

statistical models, which could be a source of bias in our analysis. Apart from that, the gap in the sample collection between 1 and 6 months of life prevents us from deciphering strain dynamics during this period. Despite these intrinsic limitations, the use of strain-level resolution metagenomic sequencing enabled us to reach some interesting observations regarding the effect of these perinatal factors on the vertical microbiome transmission. Further analysis, using bigger cohorts, could confirm the hypothesis raised by the present study and decipher its implications for infant development.

Conclusions

While delivery mode is recognized as one of the main perinatal factors affecting early-life microbial colonization and more spe-

cifically maternal microbiota transmission, the effect of other factors, including place of birth and breastfeeding duration, is less well characterized. In the present study, we evaluated the impact of both place and mode of delivery and breastfeeding duration on the maturation of the infant microbiota during the first year of life, an impact that is detectable even months after birth not only in terms of gut microbiota composition but also of strain transmission dynamics.

By analyzing the early-life microbiota composition and dynamics over time in a cohort of 34 mother-infant pairs, we highlighted that at-home vs. hospital delivery has a significant impact that is visible also at 6 months, when differences linked with mode of delivery have vanished. Additionally, despite some species being constitutively transmitted, like *Bifidobacterium* spp.,

mode and place of delivery affected not only the transmissibility of a number of gut species but also the timing of acquisition and their retention in later months. Indeed, our results also revealed that home-born neonates acquire most of the maternal strains earlier than the others and that cesarean-delivered infants get most of them after 6 months. These findings would suggest that maternal strain transmission is affected by mode of delivery earlier in life while place of delivery shapes transmission at later months, an observation that would be difficult to explain if not taking into consideration other factors correlated with the maternal choice of giving birth at home. Indeed, we hypothesized and showed that a longer breastfeeding period that characterizes the home-born infants in the MAMI cohort may be one of the factors explaining the effect of place of delivery on the microbiota dynamics of infants older than 6 months. Our hypothesis is further supported by previous results based on amplicon sequencing and culture-based approaches proposing breastfeeding and its duration as a key modulating factor of maternal microbial transmission after birth.^{38,39}

On a more subtle level, breastfeeding duration was also correlated with the retention and the functional diversity of *Bifidobacterium longum*, one of the most consistently vertically transmitted species in all groups and especially prevalent in infants born at home. Indeed, in our cohort, the first acquired strain is often replaced depending on the feeding habits of the infant and the resulting changes in nutrients availability. As we showed here, *B. longum* strain replacement is also reflected in changes in the functional repertoire linked with HMOs' degradation and their prolonged availability in infants that are at least partially breastfed at 6 or 12 months, as in the case of the home-born infants in our cohort. These findings call for a more comprehensive analysis of strain-level dynamics in the light of more perinatal factors, as the widely studied mode of delivery may represent only one of many facets affecting early-life microbiota assembly.

Although the role of cesarean delivery in the delayed colonization of the neonatal microbiota has already been extensively studied,^{8,48,65,66} the few available studies that included HB pairs and reported differences in neonatal microbial colonization due to hospital birth were performed by amplicon sequencing^{26,48,67} without assessing species-level differences, vertical transmission, or strain replacement patterns. Moreover, the same is true for studies that focused on the role of human milk in shaping the infant gut microbiota, because technical challenges like a high proportion of human DNA hinder the possibility of getting good quality metagenomes, as we experienced for many human milk samples in this and previous studies.^{6,68} Here, we obtained high-quality metagenomes for six human milk samples, which allowed us to shed light on the role and importance of breastfeeding in the early stages of microbiota assembly, showing that while most milk species are found in the neonate's gut early on, their prevalence tends to diminish over time. Despite the limited number of samples, strain-sharing analysis showed that some of these milk strains were indeed present in the infant gut as well. While the metagenomes from human milk samples are enriched in oral and skin bacteria, this is not a new observation,⁶⁹ and we cannot consider them as contamination. This study aims at surveying the transmission events between the mother and the infant and which species, independently on the maternal source, could be potentially transferred to the neonate.

Further studies with a larger sample size are needed to elucidate the stochasticity of these observations and their biological significance.

Overall, our results highlight the importance of considering not only the mode and place of delivery but also key correlated factors, such as breastfeeding duration and possibly other lifestyle preferences, when assessing infant microbiota assembly and seeding to reach a better understanding of the infant bacterial colonization and its potential long-term implications for infant development.

CONSORTIA

MAMI team investigators: M. Carmen Collado, Cecilia Martínez Costa, Máximo Vento, Isabel Úbeda-Sansano, Anna Parra-Llorca, Ramón Escuriet, Izaskun García-Mantrana, Marta Selma-Royo, Raúl Cabrera-Rubio, Christine Bäuerl, María Cernada, Elena Crehuá-Gaudiza, Jacobo Martínez-Santamaria, Erika Cortés-Macías, Anna Samarra, Julian Beltrán, Andrea Ahi-cart, Bibiana Bertúa-Ríos, Alba Peretó-Moll, Irene Rausell, Jose Luis Tortajada, Ignacio Sorribes-Monrabal, Luis C. Blesa-Baviera, Amelia Peris, and the midwives Amparo Rodríguez, María Dolores Soler, Teresa Gonzalo del Moral, Pepi Domínguez-Cano, Marga Franch i Ferrer; Concha Delgado and Adela Atero.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead contact
 - Materials availability
 - Data and code availability
- **STUDY PARTICIPANT DETAILS**
 - Mother-infant cohort described in this study
- **METHOD DETAILS**
 - Sample collection, DNA extraction, and sequencing
 - Human Milk Oligosaccharide (HMO) Profile and Secretor Status
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Sequences pre-processing
 - MetaPhlAn 3 profiling
 - Microbial Presence in human milk
 - Phylogenetic trees building with StrainPhlAn 3
 - Strain-sharing events identification
 - Strain-sharing rate, infant acquisition rate, species transmissibility, and retention rate
 - PanPhlAn strain composition profiling
 - Labelling CAZy analysis
 - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.chom.2024.05.005>.

ACKNOWLEDGMENTS

We would like to thank all the families who were involved in the study as well as the whole MAMI team (consortia section), which includes neonatologists, pediatricians, midwives, nurses, research scientists, and computer/laboratory technicians. We would like to thank the Biobank (Biobanco para la Investigación Biomédica y en Salud Pública de la Comunidad Valenciana, IBSP-CV) for their work in processing the biological samples. The MAMI

team would like to acknowledge the support from H2020-ERC Starting Grant (MAMI-639226 project) and from Spanish Ministry of Science and Innovation (MCIN) research grant (ref. PID2022-139475OB-I00). R.C.-R. thanks Generalitat Valenciana for the grant Plan GenT project (CDEIGENT 2020). M.C.C., M.-S.R., and R.C.-R. would also like to acknowledge the award of the Spanish Government MCIN/AEI to the Institute of Agrochemistry and Food Technology (IATA-CSIC) as Centre of Excellence Severo Ochoa (CEX2021-001189-S MCIN/AEI/10.13039/501100011033). This work was also supported by the European Research Council (ERC-STG project MetaPG-716575 and ERC-CoG microTOUCH-101045015) to N.S., by the European H2020 programme (ONCOBIOME-825410 project, MASTER-818368 project, and IHMCSA-964590) to N.S., by the National Cancer Institute of the National Institutes of Health (1U01CA230551) to N.S., and by EMBO ALTF 593–2020 to M.V.-C.

AUTHOR CONTRIBUTIONS

Conceptualization, M.C.C. and N.S.; software, L.D.; formal analysis, M.S.-R., L.D., S.M., F.A., S.G., and L.B.; investigation, M.S.-R., L.D., and S.M.; resources, A.P.-L., R.E., C.M.-C., M.C.C., and N.S.; data curation, M.S.-R., L.D., S.M., and F.A.; writing – original draft, M.S.-R., L.D., and S.M.; writing – review & editing, M.S.-R., L.D., S.M., R.C.-R., M.V.-C., M.C.C., and N.S.; visualization, L.D.; supervision, M.C.C. and N.S.; project administration, M.S.-R., S.M., M.C.C., and N.S.; funding acquisition, M.C.C. and N.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 9, 2024

Revised: April 24, 2024

Accepted: May 10, 2024

Published: June 12, 2024

REFERENCES

- Vatanen, T., Kostic, A.D., d’Hennezel, E., Siljander, H., Franzosa, E.A., Yassour, M., Kolde, R., Vlamakis, H., Arthur, T.D., Hämäläinen, A.-M., et al. (2016). Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. *Cell* 165, 842–853. <https://doi.org/10.1016/j.cell.2016.04.007>.
- Uzan-Yulzari, A., Turta, O., Belogolovski, A., Ziv, O., Kunz, C., Perschbacher, S., Neuman, H., Pasolli, E., Oz, A., Ben-Amram, H., et al. (2021). Neonatal antibiotic exposure impairs child growth during the first six years of life by perturbing intestinal microbial colonization. *Nat. Commun.* 12, 443. <https://doi.org/10.1038/s41467-020-20495-4>.
- Xiao, L., and Zhao, F. (2023). Microbial transmission, colonisation and succession: from pregnancy to infancy. *Gut* 72, 772–786. <https://doi.org/10.1136/gutjnl-2022-328970>.
- Oliphant, K., and Allen-Vercoe, E. (2019). Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. *Microbiome* 7, 91. <https://doi.org/10.1186/s40168-019-0704-8>.
- Liang, L., Liu, L., Zhou, W., Yang, C., Mai, G., Li, H., and Chen, Y. (2022). Gut microbiota-derived butyrate regulates gut mucus barrier repair by activating the macrophage/WNT/ERK signaling pathway. *Clin. Sci. (Lond)* 136, 291–307. <https://doi.org/10.1042/CS20210778>.
- Ferretti, P., Pasolli, E., Tett, A., Asnicar, F., Gorfer, V., Fedi, S., Armanini, F., Truong, D.T., Manara, S., Zolfo, M., 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. <https://doi.org/10.1016/j.chom.2018.06.005>.
- Yassour, M., Jason, E., Hogstrom, L.J., Arthur, T.D., Tripathi, S., Siljander, H., Selvenius, J., Oikarinen, S., Hyöty, H., Virtanen, S.M., et al. (2018). Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. *Cell Host Microbe* 24, 146–154.e4. <https://doi.org/10.1016/j.chom.2018.06.007>.
- Shao, Y., Forster, S.C., Tsaliki, E., Vervier, K., Strang, A., Simpson, N., Kumar, N., Stares, M.D., Rodger, A., Brocklehurst, P., et al. (2019). Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* 574, 117–121. <https://doi.org/10.1038/s41586-019-1560-1>.
- Hill, D.R., Huang, S., Nagy, M.S., Yadagiri, V.K., Fields, C., Mukherjee, D., Bons, B., Dedhia, P.H., Chin, A.M., Tsai, Y.-H., et al. (2017). Bacterial colonization stimulates a complex physiological response in the immature human intestinal epithelium. *eLife* 6, e29132. <https://doi.org/10.7554/eLife.29132>.
- Ignacio, A., Shah, K., Bernier-Latmani, J., Köller, Y., Coakley, G., Moyat, M., Hamelin, R., Armand, F., Wong, N.C., Ramay, H., et al. (2022). Small intestinal resident eosinophils maintain gut homeostasis following microbial colonization. *Immunity* 55, 1250–1267.e12. <https://doi.org/10.1016/j.immuni.2022.05.014>.
- Henrick, B.M., Rodriguez, L., Lakshmikanth, T., Pou, C., Henckel, E., Arzoomand, A., Olin, A., Wang, J., Mikes, J., Tan, Z., et al. (2021). Bifidobacteria-mediated immune system imprinting early in life. *Cell* 184, 3884–3898.e11. <https://doi.org/10.1016/j.cell.2021.05.030>.
- Mitchell, C.M., Mazzoni, C., Hogstrom, L., Bryant, A., Bergerat, A., Cher, A., Pochan, S., Herman, P., Carrigan, M., Sharp, K., et al. (2020). Delivery Mode Affects Stability of Early Infant Gut Microbiota. *Cell Rep. Med.* 1, 100156. <https://doi.org/10.1016/j.xcrm.2020.100156>.
- Davis, E.C., Castagna, V.P., Sela, D.A., Hillard, M.A., Lindberg, S., Mantis, N.J., Seppo, A.E., and Järvinen, K.M. (2022). Gut microbiome and breast-feeding: Implications for early immune development. *J. Allergy Clin. Immunol.* 150, 523–534. <https://doi.org/10.1016/j.jaci.2022.07.014>.
- Roswall, J., Olsson, L.M., Kovatcheva-Datchary, P., Nilsson, S., Tremaroli, V., Simon, M.-C., Kiilerich, P., Akrami, R., Krämer, M., Uhlén, M., 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. <https://doi.org/10.1016/j.chom.2021.02.021>.
- Healy, D.B., Ryan, C.A., Ross, R.P., Stanton, C., and Dempsey, E.M. (2022). Clinical implications of preterm infant gut microbiome development. *Nat. Microbiol.* 7, 22–33. <https://doi.org/10.1038/s41564-021-01025-4>.
- Park, S.Y., Ustulin, M., Park, S., Han, K.-D., Kim, J.Y., Shin, D.W., and Rhee, S.Y. (2021). Systemic Antibiotics and Obesity: Analyses from a Population-Based Cohort. *J. Clin. Med.* 10, 2601. <https://doi.org/10.3390/jcm10122601>.
- Martín-Calvo, N., Martínez-González, M.Á., Segura, G., Chavarro, J.E., Carlos, S., and Gea, A. (2020). Caesarean delivery is associated with higher risk of overweight in the offspring: within-family analysis in the SUN cohort. *J. Epidemiol. Community Health* 74, 586–591. <https://doi.org/10.1136/jech-2019-213724>.
- Ślabuszczyńska-Józwiak, A., Szymański, J.K., Ciebiera, M., Sarecka-Hujar, B., and Jakiel, G. (2020). Pediatrics Consequences of Caesarean Section—A Systematic Review and Meta-Analysis. *Int. J. Environ. Res. Public Health* 17, 8031. <https://doi.org/10.3390/ijerph17218031>.
- Levin, M.E., Botha, M., Basera, W., Facey-Thomas, H.E., Gaunt, B., Gray, C.L., Kiragu, W., Ramjith, J., Watkins, A., and Genuneit, J. (2020). Environmental factors associated with allergy in urban and rural children from the South African Food Allergy (SAFFA) cohort. *J. Allergy Clin. Immunol.* 145, 415–426. <https://doi.org/10.1016/j.jaci.2019.07.048>.
- Zhong, Y., Zhang, Y., Wang, Y., and Huang, R. (2021). Maternal antibiotic exposure during pregnancy and the risk of allergic diseases in childhood: A meta-analysis. *Pediatr. Allergy Immunol.* 32, 445–456. <https://doi.org/10.1111/pai.13411>.
- Korpela, K., Helve, O., Kolho, K.-L., Saisto, T., Skogberg, K., Dikareva, E., Stefanovic, V., Salonen, A., Andersson, S., and de Vos, W.M. (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. <https://doi.org/10.1016/j.cell.2020.08.047>.

22. Korpela, K., and de Vos, W.M. (2022). Infant gut microbiota restoration: state of the art. *Gut Microbes* *14*, 2118811. <https://doi.org/10.1080/19490976.2022.2118811>.
23. Korpela, K., Costea, P., Coelho, L.P., Kandels-Lewis, S., Willemsen, G., Boomsma, D.I., Segata, N., and Bork, P. (2018). Selective maternal seeding and environment shape the human gut microbiome. *Genome Res.* *28*, 561–568. <https://doi.org/10.1101/gr.233940.117>.
24. Selma-Royo, M., Calatayud Arroyo, M., García-Mantrana, I., Parra-Llorca, A., Escuriet, R., Martínez-Costa, C., and Collado, M.C. (2020). Perinatal environment shapes microbiota colonization and infant growth: impact on host response and intestinal function. *Microbiome* *8*, 167. <https://doi.org/10.1186/s40168-020-00940-8>.
25. Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y., Xie, H., Zhong, H., et al. (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* *17*, 690–703. <https://doi.org/10.1016/j.chom.2015.04.004>.
26. Combellick, J.L., Shin, H., Shin, D., Cai, Y., Hagan, H., Lacher, C., Lin, D.L., McCauley, K., Lynch, S.V., and Dominguez-Bello, M.G. (2018). Differences in the fecal microbiota of neonates born at home or in the hospital. *Sci. Rep.* *8*, 15660. <https://doi.org/10.1038/s41598-018-33995-7>.
27. García-Mantrana, I., Alcántara, C., Selma-Royo, M., Boix-Amorós, A., Dzidic, M., Gimeno-Alcañiz, J., Úbeda-Sansano, I., Sorribes-Monrabal, I., Escuriet, R., Gil-Raga, F., et al. (2019). MAMI: a birth cohort focused on maternal-infant microbiota during early life. *BMC Pediatr.* *19*, 140. <https://doi.org/10.1186/s12887-019-1502-y>.
28. Valles-Colomer, M., Blanco-Míguez, A., Manghi, P., Asnicar, F., Dubois, L., Golzato, D., Armanini, F., Cumbo, F., Huang, K.D., Manara, S., et al. (2023). The person-to-person transmission landscape of the gut and oral microbiomes. *Nature* *614*, 125–135. <https://doi.org/10.1038/s41586-022-05620-1>.
29. Beghini, F., McIver, L.J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Thomas, A.M., et al. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* *10*, e65088. <https://doi.org/10.7554/eLife.65088>.
30. National Research Council; Institute of Medicine; Board on Children, Youth, and Families; Food and Nutrition Board; Committee to Reexamine IOM Pregnancy Weight Guidelines (2010). *Weight Gain During Pregnancy: Reexamining the Guidelines* (National Academies Press).
31. Stewart, C.J., Ajami, N.J., O'Brien, J.L., Hutchinson, D.S., Smith, D.P., Wong, M.C., Ross, M.C., Lloyd, R.E., Doddapaneni, H., Metcalf, G.A., et al. (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* *562*, 583–588. <https://doi.org/10.1038/s41586-018-0617-x>.
32. Rios-Covian, D., Gueimonde, M., Duncan, S.H., Flint, H.J., and de los Reyes-Gavilan, C.G. (2015). Enhanced butyrate formation by cross-feeding between *Faecalibacterium prausnitzii* and *Bifidobacterium adolescentis*. *FEMS Microbiol. Lett.* *362*, fnv176. <https://doi.org/10.1093/femsle/fnv176>.
33. Cockburn, D.W., Suh, C., Medina, K.P., Duvall, R.M., Wawrzak, Z., Henrissat, B., and Koropatkin, N.M. (2018). Novel carbohydrate binding modules in the surface anchored α -amylase of *Eubacterium rectale* provide a molecular rationale for the range of starches used by this organism in the human gut. *Mol. Microbiol.* *107*, 249–264. <https://doi.org/10.1111/mmi.13881>.
34. Liu, X., Mao, B., Gu, J., Wu, J., Cui, S., Wang, G., Zhao, J., Zhang, H., and Chen, W. (2021). *Blautia*-a new functional genus with potential probiotic properties? *Gut Microbes* *13*, 1–21.
35. Hildingsson, I.M., Lindgren, H.E., Haglund, B., and Rådestad, I.J. (2006). Characteristics of women giving birth at home in Sweden: a national register study. *Am. J. Obstet. Gynecol.* *195*, 1366–1372. <https://doi.org/10.1016/j.ajog.2006.03.050>.
36. Anthony, S., Buitendijk, S.E., Offerhaus, P.M., Dommelen, P., and Pal-de Bruin, K.M. (2005). Maternal factors and the probability of a planned home birth. *BJOG* *112*, 748–753. <https://doi.org/10.1111/j.1471-0528.2004.00520.x>.
37. Quigley, C., Taut, C., Zigman, T., Gallagher, L., Campbell, H., and Zgaga, L. (2016). Association between home birth and breast feeding outcomes: a cross-sectional study in 28 125 mother-infant pairs from Ireland and the UK. *BMJ Open* *6*, e010551. <https://doi.org/10.1136/bmjopen-2015-010551>.
38. Jost, T., Lacroix, C., Braegger, C.P., Rochat, F., and Chassard, C. (2014). Vertical mother-neonate transfer of maternal gut bacteria via breastfeeding. *Environ. Microbiol.* *16*, 2891–2904. <https://doi.org/10.1111/1462-2920.12238>.
39. Laursen, M.F., Pekmez, C.T., Larsson, M.W., Lind, M.V., Yonemitsu, C., Larnkjær, A., Mølgaard, C., Bode, L., Dragsted, L.O., Michaelsen, K.F., et al. (2021). Maternal milk microbiota and oligosaccharides contribute to the infant gut microbiota assembly. *ISME Commun.* *1*, 21. <https://doi.org/10.1038/s43705-021-00021-3>.
40. Makino, H., Kushiro, A., Ishikawa, E., Muylaert, D., Kubota, H., Sakai, T., Oishi, K., Martin, R., Ben Amor, K., Oozeer, R., et al. (2011). Transmission of intestinal *Bifidobacterium longum* subsp. *longum* strains from mother to infant, determined by multilocus sequencing typing and amplified fragment length polymorphism. *Appl. Environ. Microbiol.* *77*, 6788–6793. <https://doi.org/10.1128/AEM.05346-11>.
41. Makino, H., Kushiro, A., Ishikawa, E., Kubota, H., Gawad, A., Sakai, T., Oishi, K., Martin, R., Ben-Amor, K., Knol, J., et al. (2013). Mother-to-infant transmission of intestinal bifidobacterial strains has an impact on the early development of vaginally delivered infant's microbiota. *PLoS One* *8*, e78331. <https://doi.org/10.1371/journal.pone.0078331>.
42. Wampach, L., Heintz-Buschart, A., Fritz, J.V., Ramiro-Garcia, J., Habier, J., Herold, M., Narayanasamy, S., Kaysen, A., Hogan, A.H., Bindl, L., et al. (2018). Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential. *Nat. Commun.* *9*, 5091. <https://doi.org/10.1038/s41467-018-07631-x>.
43. Li, W., Tapiainen, T., Brinkac, L., Lorenzi, H.A., Moncera, K., Tejesvi, M.V., Salo, J., and Nelson, K.E. (2021). Vertical Transmission of Gut Microbiome and Antimicrobial Resistance Genes in Infants Exposed to Antibiotics at Birth. *J. Infect. Dis.* *224*, 1236–1246. <https://doi.org/10.1093/infdis/jiaa155>.
44. Meng, D., Zhu, W., Ganguli, K., Shi, H.N., and Walker, W.A. (2016). Anti-inflammatory effects of *Bifidobacterium longum* subsp. *infantis* secretions on fetal human enterocytes are mediated by TLR-4 receptors. *Am. J. Physiol. Gastrointest. Liver Physiol.* *311*, G744–G753. <https://doi.org/10.1152/ajpgi.00090.2016>.
45. Hosomi, K., Saito, M., Park, J., Murakami, H., Shibata, N., Ando, M., Nagatake, T., Konishi, K., Ohno, H., Tanisawa, K., et al. (2022). Oral administration of *Blautia wexlerae* ameliorates obesity and type 2 diabetes via metabolic remodeling of the gut microbiota. *Nat. Commun.* *13*, 4477. <https://doi.org/10.1038/s41467-022-32015-7>.
46. Benítez-Páez, A., Gómez Del Pugar, E.M., López-Almela, I., Moya-Pérez, Á., Codoñer-Franch, P., and Sanz, Y. (2020). Depletion of *Blautia* Species in the Microbiota of Obese Children Relates to Intestinal Inflammation and Metabolic Phenotype Worsening. *mSystems* *5*, e00857-19. <https://doi.org/10.1128/mSystems.00857-19>.
47. Nayfach, S., Rodriguez-Mueller, B., Garud, N., and Pollard, K.S. (2016). An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Res.* *26*, 1612–1625. <https://doi.org/10.1101/gr.201863.115>.
48. Hesla, H.M., Stenius, F., Jäderlund, L., Nelson, R., Engstrand, L., Alm, J., and Dicksved, J. (2014). Impact of lifestyle on the gut microbiota of healthy infants and their mothers—the ALADDIN birth cohort. *FEMS Microbiol. Ecol.* *90*, 791–801. <https://doi.org/10.1111/1574-6941.12434>.
49. Rutayisire, E., Huang, K., Liu, Y., and Tao, F. (2016). The mode of delivery affects the diversity and colonization pattern of the gut microbiota during the first year of infants' life: a systematic review. *BMC Gastroenterol.* *16*, 86. <https://doi.org/10.1186/s12876-016-0498-0>.

50. Guo, M., Miao, M., Wang, Y., Duan, M., Yang, F., Chen, Y., Yuan, W., and Zheng, H. (2020). Developmental differences in the intestinal microbiota of Chinese 1-year-old infants and 4-year-old children. *Sci. Rep.* *10*, 19470. <https://doi.org/10.1038/s41598-020-76591-4>.
51. Kostopoulos, I., Elzinga, J., Ottman, N., Klievink, J.T., Blijenberg, B., Aalvink, S., Boeren, S., Mank, M., Knol, J., de Vos, W.M., et al. (2020). *Akkermansia muciniphila* uses human milk oligosaccharides to thrive in the early life conditions in vitro. *Sci. Rep.* *10*, 14330. <https://doi.org/10.1038/s41598-020-71113-8>.
52. Jakobsson, H.E., Abrahamsson, T.R., Jenmalm, M.C., Harris, K., Quince, C., Jernberg, C., Björkstén, B., Engstrand, L., and Andersson, A.F. (2014). Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section. *Gut* *63*, 559–566. <https://doi.org/10.1136/gutjnl-2012-303249>.
53. World Health Organization (2003). *Global Strategy for Infant and Young Child Feeding* (World Health Organization and UNICEF).
54. Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* *50*, D571–D577. <https://doi.org/10.1093/nar/gkab1045>.
55. Ioannou, A., Knol, J., and Belzer, C. (2021). Microbial Glycoside Hydrolases in the First Year of Life: An Analysis Review on Their Presence and Importance in Infant Gut. *Front. Microbiol.* *12*, 631282. <https://doi.org/10.3389/fmicb.2021.631282>.
56. Arzamasov, A.A., and Osterman, A.L. (2022). Milk glycan metabolism by intestinal bifidobacteria: insights from comparative genomics. *Crit. Rev. Biochem. Mol. Biol.* *57*, 562–584. <https://doi.org/10.1080/10409238.2023.2182272>.
57. Samuel, T.M., Binia, A., de Castro, C.A., Thakkar, S.K., Billeaud, C., Agosti, M., Al-Jashi, I., Costeira, M.J., Marchini, G., Martínez-Costa, C., et al. (2019). Impact of maternal characteristics on human milk oligosaccharide composition over the first 4 months of lactation in a cohort of healthy European mothers. *Sci. Rep.* *9*, 11767. <https://doi.org/10.1038/s41598-019-48337-4>.
58. Azad, M.B., Robertson, B., Atakora, F., Becker, A.B., Subbarao, P., Moraes, T.J., Mandhane, P.J., Turvey, S.E., Lefebvre, D.L., Sears, M.R., et al. (2018). Human Milk Oligosaccharide Concentrations Are Associated with Multiple Fixed and Modifiable Maternal Characteristics, Environmental Factors, and Feeding Practices. *J. Nutr.* *148*, 1733–1742. <https://doi.org/10.1093/jn/nxy175>.
59. Ennis, D., Shmorak, S., Jantscher-Krenn, E., and Yassour, M. (2024). Longitudinal quantification of *Bifidobacterium longum* subsp. *infantis* reveals late colonization in the infant gut independent of maternal milk HMO composition. *Nat. Commun.* *15*, 894. <https://doi.org/10.1038/s41467-024-45209-y>.
60. Vatanen, T., Ang, Q.Y., Siegwald, L., Sarker, S.A., Le Roy, C.I., Duboux, S., Delannoy-Bruno, O., Ngom-Bru, C., Boulangé, C.L., Stražar, M., et al. (2022). A distinct clade of *Bifidobacterium longum* in the gut of Bangladeshi children thrives during weaning. *Cell* *185*, 4280–4297.e12. <https://doi.org/10.1016/j.cell.2022.10.011>.
61. Olm, M.R., Crits-Christoph, A., Bouma-Gregson, K., Firek, B.A., Morowitz, M.J., and Banfield, J.F. (2021). inStrain profiles population microdiversity from metagenomic data and sensitively detects shared microbial strains. *Nat. Biotechnol.* *39*, 727–736. <https://doi.org/10.1038/s41587-020-00797-0>.
62. Da Silva, K., Pons, N., Berland, M., Plaza Oñate, F.P., Almeida, M., and Peterlongo, P. (2021). StrainFLAIR: strain-level profiling of metagenomic samples using variation graphs. *PeerJ* *9*, e11884. <https://doi.org/10.7717/peerj.11884>.
63. Podlesny, D., Arze, C., Dörner, E., Verma, S., Dutta, S., Walter, J., and Fricke, W.F. (2022). Metagenomic strain detection with SameStr: identification of a persisting core gut microbiota transferable by fecal transplantation. *Microbiome* *10*, 53. <https://doi.org/10.1186/s40168-022-01251-w>.
64. Luo, C., Knight, R., Siljander, H., Knip, M., Xavier, R.J., and Gevers, D. (2015). ConStrains identifies microbial strains in metagenomic datasets. *Nat. Biotechnol.* *33*, 1045–1052. <https://doi.org/10.1038/nbt.3319>.
65. Reyman, M., van Houten, M.A., van Baarle, D., Bosch, A.A.T.M., Man, W.H., Chu, M.L.J.N., Arp, K., Watson, R.L., Sanders, E.A.M., Fuentes, S., et al. (2019). Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. *Nat. Commun.* *10*, 4997. <https://doi.org/10.1038/s41467-019-13014-7>.
66. Mäklin, T., Thorpe, H.A., Pöntinen, A.K., Gladstone, R.A., Shao, Y., Pesonen, M., McNally, A., Johnsen, P.J., Samuelson, Ø., Lawley, T.D., et al. (2022). Strong pathogen competition in neonatal gut colonisation. *Nat. Commun.* *13*, 7417. <https://doi.org/10.1038/s41467-022-35178-5>.
67. Penders, J., Thijs, C., Vink, C., Stelma, F.F., Snijders, B., Kummeling, I., van den Brandt, P.A., and Stobberingh, E.E. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* *118*, 511–521. <https://doi.org/10.1542/peds.2005-2824>.
68. Asnicar, F., Manara, S., Zolfo, M., Truong, D.T., Scholz, M., Armanini, F., Ferretti, P., Gorfer, V., Pedrotti, A., Tett, A., et al. (2017). Studying Vertical Microbiome Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling. *mSystems* *2*, e00164-16. <https://doi.org/10.1128/mSystems.00164-16>.
69. Kordy, K., Gauvin, T., Mwangi, M., Li, F., Cerini, C., Lee, D.J., Adisetiyo, H., Woodward, C., Pannaraj, P.S., Tobin, N.H., et al. (2020). Contributions to human breast milk microbiome and enteromammary transfer of *Bifidobacterium breve*. *PLoS One* *15*, e0219633. <https://doi.org/10.1371/journal.pone.0219633>.
70. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* *9*, 357–359. <https://doi.org/10.1038/nmeth.1923>.
71. Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* *40*, W445–W451. <https://doi.org/10.1093/nar/gks479>.
72. Buchfink, B., Reuter, K., and Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* *18*, 366–368. <https://doi.org/10.1038/s41592-021-01101-x>.
73. Thiele, C. (2020). cutpointr: Determine and Evaluate Optimal Cutpoints in Binary Classification Tasks. R package cutpointr version 1.0.32 (CRAN). <https://cran.r-project.org/web/packages/cutpointr/index.html>.
74. Kassambara, A. (2023). rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package rstatix version 0.7.2 (CRAN). <https://cran.r-project.org/web/packages/rstatix/index.html>.
75. Dinno, A. (2017). dunn.test: Dunn's Test of Multiple Comparisons Using Rank Sums (CRAN). <https://cran.r-project.org/web/packages/dunn.test/index.html>.
76. Ahlmann-Eltze, C., and Patil, I. (2021). ggsignif: R Package for Displaying Significance Brackets for “ggplot2”. Preprint at PsyArXiv. <https://doi.org/10.31234/osf.io/7awm6>.
77. Wilke, C.O. (2020). cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2”. R package cowplot version 1.1.1 (CRAN). <https://cran.r-project.org/web/packages/cowplot/index.html>.
78. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* *32*, 2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>.
79. Gu, Z. (2022). Complex heatmap visualization. *iMeta* *1*, e43. <https://doi.org/10.1002/imt2.43>.
80. Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., et al. (2022). vegan: Community Ecology Package (CRAN). <http://CRAN.R-project.org/package=vegan>.
81. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Comprehensive R Archive Network (CRAN)). <https://cran.r-project.org/web/packages/ggplot2/index.html>.
82. García-Mantrana, I., Selma-Royo, M., González, S., Parra-Llorca, A., Martínez-Costa, C., and Collado, M.C. (2020). Distinct maternal

- microbiota clusters are associated with diet during pregnancy: impact on neonatal microbiota and infant growth during the first 18 months of life. *Gut Microbes* 11, 962–978. <https://doi.org/10.1080/19490976.2020.1730294>.
83. Cervera, P., Farran, A., and Zamora-Ros, R. (2004). Tablas de composición de alimentos del CESNID: Taules de composició d'aliments del CESNID. *Rev. Esp. Salud Publ.* 78, 10–1590.
 84. Marlett, J.A., and Cheung, T.F. (1997). Database and quick methods of assessing typical dietary fiber intakes using data for 228 commonly consumed foods. , 1151; quiz 1149–1150. *J. Am. Diet. Assoc.* 97, 1139–1148. [https://doi.org/10.1016/S0002-8223\(97\)00275-7](https://doi.org/10.1016/S0002-8223(97)00275-7).
 85. Selma-Royo, M., González, S., Gueimonde, M., Chang, M., Fürst, A., Martínez-Costa, C., Bode, L., and Collado, M.C. (2022). Maternal Diet Is Associated with Human Milk Oligosaccharide Profile. *Mol. Nutr. Food Res.* 66, e2200058. <https://doi.org/10.1002/mnfr.202200058>.
 86. Marionneau, S., Airaud, F., Bovin, N.V., Le Pendu, J., and Ruvoën-Clouet, N. (2005). Influence of the combined ABO, FUT2, and FUT3 polymorphism on susceptibility to Norwalk virus attachment. *J. Infect. Dis.* 192, 1071–1077. <https://doi.org/10.1086/432546>.
 87. Truong, D.T., Tett, A., Pasolli, E., Huttenhower, C., and Segata, N. (2017). Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res.* 27, 626–638. <https://doi.org/10.1101/gr.216242.116>.
 88. Scholz, M., Ward, D.V., Pasolli, E., Tolio, T., Zolfo, M., Asnicar, F., Truong, D.T., Tett, A., Morrow, A.L., and Segata, N. (2016). Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nat. Methods* 13, 435–438. <https://doi.org/10.1038/nmeth.3802>.
 89. Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P.K., Xu, Y., and Yin, Y. (2018). dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 46, W95–W101. <https://doi.org/10.1093/nar/gky418>.
 90. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. <https://doi.org/10.1038/nmeth.3176>.
 91. R Development Core Team (2022). R: A language and environment for statistical computing (R Foundation for Statistical Computing).

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical Commercial Assays		
Master-Pure DNA extraction kit	Epicentre, Madison, USA	Catalog No. 12888-50
DNA Purification Kit	Macherey-Nagel, Duren, Germany	Catalog No. MC89010
NexteraXT DNA Library Preparation Kit	Illumina, California, USA	FC-131-1096
Deposited Data		
A newly sequenced cohort from Spain: MAMI cohort	This study	ENA: PRJEB74322
Software and Algorithms		
Trim Galore v0.6.6	https://github.com/FelixKrueger/TrimGalore	https://github.com/FelixKrueger/TrimGalore
Bowtie2 v2.3.4.3	Langmead and Salzberg ⁷⁰	https://github.com/BenLangmead/bowtie2
MetaPhlAn v3.0.7 + database v30_CHOCOPHiAn_201901	Beghini et al. ²⁹	https://github.com/biobakery/MetaPhlAn
StrainPhlAn 3	Beghini et al. ²⁹	https://github.com/biobakery/MetaPhlAn
PanPhlAn 3 + <i>Bifidobacterium longum</i> Pangenome	Beghini et al. ²⁹	https://github.com/biobakery/MetaPhlAn
CAZy database CAZyDB_07312020	Yin et al. ⁷¹	https://bcbl.unl.edu/dbCAN2/download/
DIAMOND v2.0.8	Buchfink et al. ⁷²	https://github.com/bbuchfink/diamond
R package cutpointr v1.1.2	Thiele ⁷³	https://github.com/Thie1e/cutpointr
R package rstatix v0.7.2	Kassambara ⁷⁴	https://github.com/kassambara/rstatix
R package dunn.test v1.6.5	Dinno et al. ⁷⁵	https://github.com/cran/dunn.test
R package ggsignif v0.6.4	Ahlmann-Eltze and Patil ⁷⁶	https://github.com/const-ae/ggsignif
R package cowplot v1.1.1	Wilke ⁷⁷	https://github.com/wilkelab/cowplot
R package ComplexHeatmap 2.12.1	Gu et al. ^{78,79}	https://github.com/jokergoo/ComplexHeatmap
R package vegan 2.6-4	Oksanen et al. ⁸⁰	https://github.com/vegandevs/vegan
R package ggplot2 3.4.1	Wickham ⁸¹	https://github.com/tidyverse/ggplot2

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Maria Carmen Collado (mcolam@iata.csic.es)

Materials availability

This study did not generate new unique reagents.

Data and code availability

Raw sequences of shotgun metagenomics samples are available on the European Nucleotide archive (ENA) under the accession number PRJEB74322.

STUDY PARTICIPANT DETAILS

Mother-infant cohort described in this study

A subset of mother-infant pairs (n=34) from the MAMI cohort was selected to be included in the analysis based on the sample availability and the homogeneity of the clinical/anthropometric characteristics such as the breastfeeding practices. MAMI is a longitudinal prospective birth cohort that followed mother-infant pairs from Spain with the aim of exploring the impact of perinatal factors on the maternal and infant microbiota and how this interaction could affect the intestinal, metabolic and immune system development in the infant.²⁷ Mothers

were enrolled during pregnancy and/or during the first week postpartum between 2015–2016 and inclusion criteria included: mothers older than 18 years old with healthy pregnancies without signs of chorioamnionitis or health complications during the gestational period. Clinical and anthropometric data were collected at delivery and in the following time points. Neonates born at hospital remained in the hospital for two days. Infant length and weight were registered at birth, 1, 6, and 12 months and the data were used to calculate BMI Z-scores. These anthropometric measures that are standardised by age and sex were electronically computed using WHO Anthro software (<https://www.who.int/childgrowth/software/en/>). We selected mother-infant pairs that were not subjected to intrapartum antibiotics (IAP), except for mothers that delivered by Caesarean section that were all treated with IAP as a standard procedure.

Dietary data regarding the consumption during pregnancy was collected using a 140-items food-frequency questionnaire as described in Garcia-Mantrana et al.⁸² The energy and daily intake of macro- and micronutrients were extracted by using the nutrient Food Composition Tables developed by the Centro de Enseñanza Superior de Nutrición Humana y Dietética (CESNID)⁸³ that were completed by the Marlett food composition table⁸⁴ for specific fibre types.

Among the longitudinal sampling performed in the MAMI cohort, samples from 7 days and 1, 6 and 12 months were included in this analysis. One faecal and one milk sample were collected 1-month postpartum from each mother. As the aim of this study was to explore the evolution of the infant microbiota during the first year of life, we collected samples before and after weaning to consider the impact of the introduction of solid food on microbiota development. Similarly, we avoided the first-pass faeces sample since previous studies have shown that it would be mainly composed of environmental and transient species that do not significantly contribute to the seeding of infant microbiota.⁷ We selected mother-infant pairs from the three delivery groups enrolled in the MAMI cohort already explored in a previous work using 16S rRNA gene sequencing,²⁴ including infants born at the hospital by both vaginal (n=13) and caesarean (n=11) delivery and those born at home by vaginal delivery (n=10).

Ethical approval for the study was obtained from the Ethics Committee for Clinical Research of the healthcare centres involved in the recruitment including the Hospital La Fe, Hospital Clínico Universitario, Parc de Salut MAR, and CSIC (Consejo Superior de Investigaciones Científicas) [ClinicalTrials.gov NCT03552939].

METHOD DETAILS

Sample collection, DNA extraction, and sequencing

Infant faecal samples were collected at home by their parents who were previously trained by clinical personnel in the health care centres where they were enrolled. Briefly, faecal samples were deposited in provided sterile containers and immediately kept at -20°C, before the final storage at -80°C until further analysis. Similarly, milk samples were collected by the mother following the provided instructions. Breast skin was cleaned with 0.5% chlorhexidine solution and the first drops were discarded. Then, human milk was collected with a sterile pumper in sterile bottles to normalise milk collection.

DNA was isolated from the faecal sample (50–100 mg) and from human milk sample (2.0 mL) using the Master-Pure DNA extraction kit (Epicentre, Madison, WI, USA) following the recommended protocol with some additional steps as described in Selma-Royo et al.²⁴ Briefly, a pre-treatment with lysozyme and mutanolysin was included as well as a cell disruption step with 3- μ m diameter glass beads in a FastPrep 24–5G Homogenizer (MP Biomedicals). After extraction, DNA Purification Kit (Macherey-Nagel, Duren, Germany) was used for DNA purification. For the human milk samples (2 ml), an additional preliminary step of centrifugation at 4000g for 20 min was carried out to separate fat and cells from whey before the DNA extraction which was initiated from the pellet fraction.

Sequencing libraries were constructed using the NexteraXT DNA Library Preparation Kit (Illumina) and sequenced on the Illumina HiSeq2500 platform with a target depth of 5Gb/sample.

Human Milk Oligosaccharide (HMO) Profile and Secretor Status

A subset of the available HMO profiling data, obtained by quantitative HPLC, from the MAMI cohort described elsewhere⁸⁵ were used in this study. In brief, the procedure included 20–50 μ L aliquots of each human milk samples that were used to determine the absolute concentrations for each of the annotated HMO and corrected for internal standard. (Oligosaccharide detection limit: \approx 20 pmol, dynamic range between 20 and 5000 pmol; milk samples were diluted accordingly). The identified HMOs in the samples were: 2'-fucosyllactose (2'FL), 3-fucosyllactose (3FL), 3'-sialyllactose (3'SL), 6'-sialyllactose (6'SL), difucosyllactose (DFLac), difucosyl-lacto-N-hexaose (DFLNH), difucosyl-lacto-N-tetraose (DFLNT), disialyl-lacto-N-hexaose (DSLNH), disialyl-lacto-N-tetraose (DSLNT), fucosyl-disialyl-lacto-N-hexaose (FDSLNH), fucosyl-lacto-N-hexaose (FLNH), lacto-N-fucopentaose (LNFP)I, LNFPII, LNFPIII, lactose-N-hexaose (LNH), lacto-N-neotetraose (LNNT), lacto-N-tetraose (LNT), sialyl-lacto-N-tetraose b (LSTb), and sialyl-lacto-N-tetraose c (LSTc). The HMO Simpson's diversity and evenness indexes were also calculated based on relative HMO abundances.

The Secretor status was determined taking into account the abundance of α (1,2)-Fuc containing structures, mainly the most abundant one: 2'-fucosyllactose (2'FL). We have also checked the secretor/non-secretor phenotype assignment by genotyping the FUT2 gene from the human milk total DNA by PCR-random fragment length polymorphisms (RFLPs) as described previously.⁸⁶

QUANTIFICATION AND STATISTICAL ANALYSIS

Sequences pre-processing

Following the pipeline described in <https://github.com/SegataLab/preprocessing>, sequences were filtered to remove the low-quality shotgun metagenomic reads (quality score <Q20), fragmented short reads (<75 bp), and reads with >2 ambiguous nucleotides

(removed with Trim Galore (v0.6.6)). Next, contaminant and host DNA (phiX 174 Illumina spike-in and human-associated reads (hg19 genome release)) was identified with Bowtie2 (v2.3.4.3)⁷⁰ using the `-sensitive-local` parameter. The remaining high-quality reads were sorted and split to create standard forward, reverse and unpaired reads output files for each metagenome.

MetaPhlAn 3 profiling

Samples taxonomic profiles were estimated using the MetaPhlAn 3.0.7 (database v30_CHOCOPhlan_201901) software with default parameters.²⁹ Profiles generated from samples with less than 1 million reads were discarded. Thus, the relative abundance of taxa was used to assess the abundance and presence (relative abundance $\geq 0.05\%$) of species. In order to identify potential targets for strain-tracking, we retained only species co-detected in at least one mother-infant sample pair.

Microbial Presence in human milk

Using the MetaPhlAn 3 profiles, species were categorised as “found in human milk” if their relative abundance was higher than 0.05% in at least 3 of the human milk samples ($n=27$). This corresponds to an overall prevalence $> 11\%$ and prevents adding too much noise in the set of breast milk associated species.

Phylogenetic trees building with StrainPhlAn 3

Species detected with MetaPhlAn 3 in at least one mother-infant pair irrespective of time point and body site were analysed using StrainPhlAn 3^{29,87} and phylogenetic trees were built. Here we tried to improve the robustness of the analysis by including as many samples as possible in the trees. Samples were divided into two groups (`-samples` and `-secondary_samples` input parameters) depending on whether the relative abundance of the targeted species was ≥ 0.05 or between 0.05 and 0. To keep more marker genes for the analysis we lowered the `-marker_in_n_samples` from default 80% to 50% so that markers detected in at least half of the samples will be kept. On the other hand, we adapted the number of marker genes that should be detected in a sample in order for it to be kept (parameters `-sample_with_n_markers` and `-secondary_sample_with_n_markers`) to 20 (default value) if the species has at least 50 markers and 10 otherwise.

Strain-sharing events identification

The strain-tracking analysis was done according to the method described in Valles-Colomer et al.²⁸ The set of pairwise distances between samples was extracted for each StrainPhlAn tree. We then used metadata to divide it into 3 groups: samples coming from related individuals, unrelated individuals, or the same individual (same person, multiple time points). The aim is to define for each species a threshold of the genetic distance below which strains will be considered similar. When possible we computed this threshold as the distance maximising the Youden’s index in the same individual VS unrelated individuals classification (R package `cutpointr` v1.1.2 with `metric="youden"`). When this index was impossible to compute, we used the 3rd percentile of the distribution of unrelated individuals’ strain phylogenetic distance.

Strain-sharing rate, infant acquisition rate, species transmissibility, and retention rate

From the raw counts of species sharing events (detected in both samples with MetaPhlAn 3) and strain-sharing events (both strains profiled with StrainPhlAn 3 and distance below the cutoff), we defined the strain-sharing and infant acquisition rates and the species transmissibility.

The strain-sharing rate is defined as the fraction of the species that are present in both mother and infant that are shared at the strain level. The infant acquisition rate reflects the percentage of all the infant gut strains that are shared with the mother. These two rates were only defined for the sample pairs having at least 3 species in common profiled at the strain level (present in the phylogenetic tree generated by StrainPhlAn 3).

The species transmissibility was defined as the fraction of families in which a mother-infant sharing event is detected for a given species at any time point, so multiple strain-sharing events in the same family would count only once.

Finally, retention rate is computed as the number of strains persisting over two consecutive time points in infants divided by the number of consecutive time points with the species present.

In addition, we define the first acquisition in a given family as the earliest infant time point at which a given species was detected as shared with the mother.

PanPhlAn strain composition profiling

The gene composition of the dominant bacterial strain was assessed by running PanPhlAn 3^{29,88} on each sample using the default mapping parameters and the very sensitive profiling parameters from the PanPhlAn wiki of the Github repository (<https://github.com/SegataLab/panphlan>, `panphlan_profiling.py` script with the `-min_coverage 1 -left_max 1.70 -right_min 0.30` arguments). The *Bifidobacterium longum* pangenome used for the PanPhlAn analysis was obtained from the database provided through the `panphlan_download_pangenome.py` script.

Labelling CAZY analysis

Carbohydrate Active Enzymes (CAZy) in the gene families of *Bifidobacterium longum* pangenome were labelled using the dbCAN2 software.⁸⁹ A CAZY pre-annotated sequence database (CAZYDB_07312020) was downloaded from the dbCAN2 website

(<https://bcb.unl.edu/dbCAN2/download/>) and used DIAMOND (v2.0.8, using default parameters).^{72,90} We classified hits on the E-value threshold of $1e-102$ as recommended in Zhang et al.⁸⁹ We focused on the GH families that have been reported to be in HMO metabolism⁵⁵: GH18, GH20, GH29, GH33, GH85, GH95, GH112 and GH136.

Statistical analysis

The significance of statistical tests used in this paper was defined as $P_{adj} < 0.05$. Correction for multiple testing followed the Benjamini-Hochberg procedure (`p.adjust()` function of the R package *stats* 4.2.1,⁹¹ parameter *method=BH*). All tests are two-sided unless specified otherwise in the main text. Differences between two groups were assessed with Wilcoxon rank-sum tests. For more than two groups, the Kruskal–Wallis test with post hoc Dunn tests was used. The PERMANOVA analysis was performed on the Aitchison distance (function `vegdist(method = "aitchison", pseudocount = 1e-20)` of the *vegan* R package v2.6-4) by using the `capscale()` function for dbRDA followed by `anova.cca()` and `RsquaredAdj()` of the same package in order to get a P value and the adjusted R². The analysis of multivariate models used with the function `ordiR2step(direction = "forward")` of the same package. The analysis and the graphical representations were performed using the R software for Statistical Computing (version 4+) using the packages: *vegan* 2.6-4,⁸⁰ *rstatix* 0.7.2,⁷⁴ *dunn.test* 1.6.5,⁷⁵ *ggplot2* 3.4.1,⁸¹ *ggsignif* 0.6.4,⁷⁶ *cowplot* 1.1.1,⁷⁷ *ComplexHeatmap* 2.12.1.^{78,79}