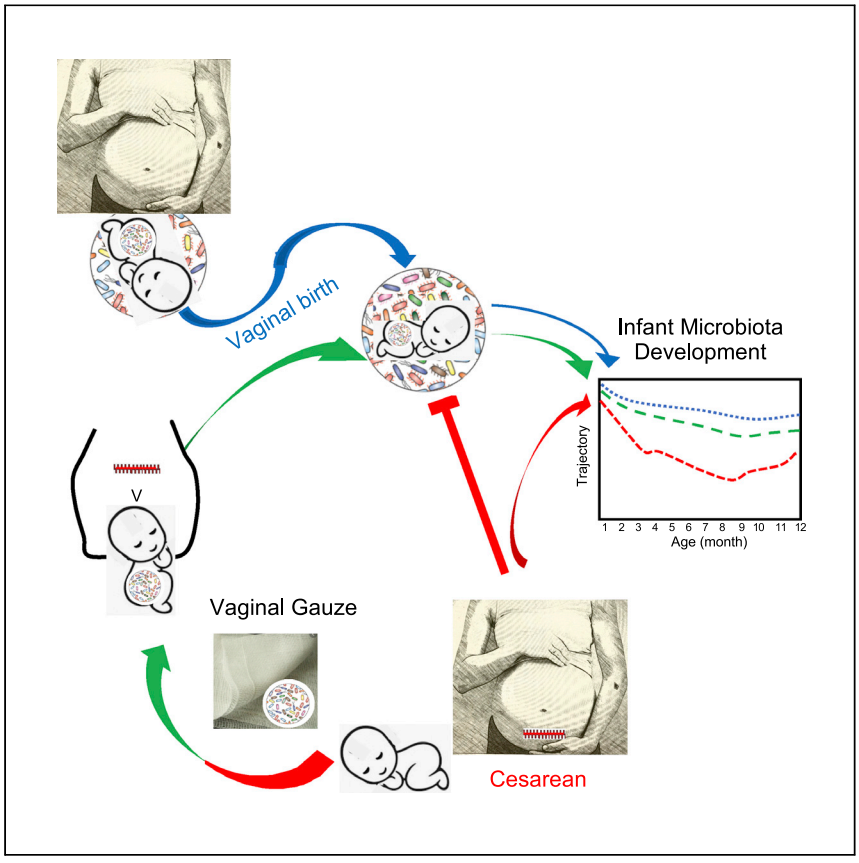
Med

CellPress

Clinical and Translational Article

# Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding




Cesarean section alters maturation of the infant microbiota. Song et al. show that restoring maternal microbes immediately after birth in Cesarean section-born infants naturalizes the developmental trajectory of their microbiota. Perinatally, maternal vaginal fluids appear to be pluripotent to provide pioneer bacterial colonizers for the newborn's body sites.

Se Jin Song, Jincheng Wang, Cameron Martino, ..., Jean F. Ruiz-Calderon, Rob Knight, Maria Gloria Dominguez-Bello

mg.dominguez-bello@rutgers.edu

## Highlights

Vaginal seeding of cesarean section-born babies naturalizes their microbiota

Bacteria from multiple body sites compose the perinatal maternal vaginal microbiome

Bacteria typical in vaginal birth engraft different sites of cesarean section-born babies

**Translation to Humans**

Song et al., Med *2*, 1–14
August 13, 2021 
https://doi.org/10.1016/j.medj.2021.05.003

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003

Med

CellPress

Clinical and Translational Article

# Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding

Se Jin Song,[1,2,23] Jincheng Wang,[3,23] Cameron Martino,[1,2,4,23] Lingjing Jiang,[5] Wesley K. Thompson,[5] Liat Shenhav,[6] Daniel McDonald,[1] Clarisse Marotz,[1] Paul R. Harris,[7] Caroll D. Hernandez,[7] Nora Henderson,[8] Elizabeth Ackley,[10] Deanna Nardella,[11] Charles Gillihan,[8] Valentina Montacuti,[8] William Schweizer,[9] Melanie Jay,[8] Joan Combellick,[12] Haipeng Sun,[3] Izaskun Garcia-Mantrana,[13] Fernando Gil Raga,[14] Maria Carmen Collado,[13] Juana I. Rivera-Viñas,[15] Maribel Campos-Rivera,[16] Jean F. Ruiz-Calderon,[17] Rob Knight,[1,2,18,19,24] and Maria Gloria Dominguez-Bello[3,20,21,22,24,25,*]

## SUMMARY

**Background:** Early microbiota perturbations are associated with disorders that involve immunological underpinnings. Cesarean section (CS)-born babies show altered microbiota development in relation to babies born vaginally. Here we present the first statistically powered longitudinal study to determine the effect of restoring exposure to maternal vaginal fluids after CS birth.

**Methods:** Using 16S rRNA gene sequencing, we followed the microbial trajectories of multiple body sites in 177 babies over the first year of life; 98 were born vaginally, and 79 were born by CS, of whom 30 were swabbed with a maternal vaginal gauze right after birth.

**Findings:** Compositional tensor factorization analysis confirmed that microbiota trajectories of exposed CS-born babies aligned more closely with that of vaginally born babies. Interestingly, the majority of amplicon sequence variants from maternal vaginal microbiomes on the day of birth were shared with other maternal sites, in contrast to non-pregnant women from the Human Microbiome Project (HMP) study.

**Conclusions:** The results of this observational study prompt urgent randomized clinical trials to test whether microbial restoration reduces the increased disease risk associated with CS birth and the underlying mechanisms. It also provides evidence of the pluripotential nature of maternal vaginal fluids to provide pioneer bacterial colonizers for the newborn body sites. This is the first study showing long-term naturalization of the microbiota of CS-born infants by restoring microbial exposure at birth.

**Funding:** C&D, Emch Fund, CIFAR, Chilean CONICYT and SOCHIPE, Norwegian Institute of Public Health, Emerald Foundation, NIH, National Institute of Justice, Janssen.

## INTRODUCTION

Over the past few decades, we have learned much about the multitude of ways in which microbiotas affect development of their hosts. Studies using model organisms show that fetal development can be modulated by microbial products from the pregnant mother's microbiota and that early colonization is critical for immune system development.[1,2]

### Context and significance

Cesarean section birth alters the infant microbiota and is associated with increased risk of immune and metabolic disorders. We restored the natural microbial exposure of babies to their mothers' birth canal fluids right after Cesarean section birth and found that, perinatally, the maternal birth canal contains very high proportions of bacteria typical of other body sites and that engraftment of these maternal bacteria normalizes microbiota development in different infant sites. In the context of risks and benefits of Cesarean section procedures, normalizing the infant microbiota from birth might mitigate the collateral effects of missing colonization by important early bacteria and reduce the increased risk to immune and metabolic diseases associated with cesarean section birth.

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



**Table 1. Key characteristics of families included in this study for analysis**

| | | Vaginal | CS | CS-seeded |
|---|---|---|---|---|
| Number of families | | 97 | 49 | 28 |
| Number of babies | total | 98 | 49 | 30 |
| | USA | 62 | 23 | 16 |
| | Spain | 7 | 5 | 8 |
| | Chile | 26 | 18 | 6 |
| | Bolivia | 3 | 3 | 0 |
| Baby female sex (%) | | 52.0 | 38.8 | 46.7 |
| Mean baby follow-up, months (standard deviation) | | 6.2 (5.1) | 8.8 (4.7) | 7.1(5.2) |
| Baby antibiotics use within the study period (after week 1) | | | | |
| Any (%) | | 18.4 | 20.4 | 16.7 |
| 1 dose (%) | | 7.1 | 14.3 | 3.3 |
| >1 doses (%) | | 11.3 | 6.1 | 13.4 |
| Breastfeeding dominant within first 4 months (%)[a] | | 75.5 | 69.4 | 53.3 |
| Mother tested positive for group B *Streptococcus* (%) | | 8.9 | 11.1 | 0 |
| Use of perinatal antibiotics in mother (%)[b] | | 13.3 | 95.6 | 95.8 |
| Any use of antibiotics during pregnancy (%) | | 20 | 26.7 | 16.7 |

[a]Breastfeeding dominant is defined as mothers who reported breastfeeding exclusively or more than 50% breastfeeding in at least 60% of follow-up visits.

[b]Use of perinatal antibiotics is part of the standard of care for women undergoing CS birth.

Natural transmission and colonization of maternal microbes is impaired by delivery via cesarean section (CS).[3–7] Furthermore, CS birth is associated with reduced levels of various cytokines and their receptors,[8] increased risk of opportunistic neonatal infections,[6] immune diseases,[9,10] and obesity.[11,12] These associations have been shown to be causal in mouse models for conditions such as obesity[13,14] and immune disorders.[15,16] Neuroendocrine abnormalities, including cognitive and behavioral disorders, have also been associated with early microbiome perturbations.[17,18] Understanding the contribution of microbiotas to healthy development remains a crucial challenge to address the current epidemic of immune and metabolic diseases in urban societies.

Although used without medical indication in many countries, CS delivery is often medically necessary and a life-saving procedure, and, thus, restoration may be one solution to help reduce the risk of associated disorders related to the microbiome. Two proof-of-concept studies have demonstrated the principle of engraftment of maternal bacteria on CS-born babies after deliberate microbial exposure: the first one used maternal vaginal gauze as a source,[19] and the second pilot study used maternal feces.[20] Here we present the first large observational study of the long-term effect of maternal vaginal seeding after CS delivery to restore microbial development during the first year of life.

## RESULTS

### Vaginal seeding of CS-born infants

A total of 177 infants born to 174 mothers were studied (Figure S1A), of whom 101 were born in the United States, 50 in Chile, 6 in Bolivia, and 20 in Spain (Table 1). 98 infants were born vaginally, and 79 were delivered by CS, of whom 30 who complied with the inclusion criteria (STAR Methods), were swabbed with a maternal vaginal gauze at birth (vaginal seeding).[19] Microbiota development was followed during

[1]Department of Pediatrics, School of Medicine, University of California, San Diego, La Jolla, CA 92093, USA

[2]Center for Microbiome Innovation, University of California, San Diego, La Jolla, CA 92093, USA

[3]Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ 08901, USA

[4]Bioinformatics and Systems Biology Program, University of California, San Diego, La Jolla, CA 92093, USA

[5]Division of Biostatistics, University of California, San Diego, La Jolla 92093, CA, USA

[6]Department of Computer Science, University of California, Los Angeles, Los Angeles, CA 90095, USA

[7]Department of Infectious Diseases and Pediatric Immunology, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile

[8]Departments of Medicine and Population Health, New York University Grossman School of Medicine, New York University, New York, NY 10016, USA

[9]Department of Obstetrics and Gynecology, New York University Grossman School of Medicine, New York University, New York, NY 10016, USA

[10]Yale New Haven Hospital, New Haven, CT 06510, USA

[11]Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

[12]Yale University School of Nursing, VA Connecticut Healthcare System, West Haven, CT 06516, USA

[13]Department of Biotechnology, Institute of Agrochemistry and Food Technology-Spanish National Research Council (IATA-CSIC), 46980 Paterna, Spain

[14]Department of Obstetrics and Gynecology, Manises Hospital, 46940 Manises, Spain

[15]Department of Obstetrics and Gynecology, Medical Science Campus, University of Puerto Rico, San Juan, PR 00925, USA

[16]Center for Community Outreach for Health Across the Lifespan, Medical Sciences Campus, University of Puerto Rico, San Juan, PR 00925, USA

[17]Medical Science Campus, University of Puerto Rico, San Juan, PR 00925, USA

[18]Department of Computer Science & Engineering, Jacobs School of Engineering, University of California, San Diego, La Jolla, CA 92093, USA

[19]Department of Bioengineering, Jacobs School of Engineering, University of California, San Diego, La Jolla 92093, CA, USA

[20]Department of Anthropology, Rutgers University, New Brunswick, NJ 08901, USA

[21]New Jersey Institute for Food, Nutrition and Health, Rutgers University, New Brunswick, NJ 08901, USA

[22]Humans and the Microbiome Program, Canadian Institute for Advanced Research, Toronto, ON M5G 1M1, Canada

[23]These authors contributed equally

[24]Senior author

[25]Lead contact

*Correspondence: mg.dominguez-bello@rutgers.edu

https://doi.org/10.1016/j.medj.2021.05.003

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003





the first year of life. A total of 8,104 samples from stool, mouth, and skin of infants and their mothers were obtained, with additional nasal and vaginal samples from mothers (Figures S1A–S1C). None of the seeded infants had any complications, and all children developed normally during the 12 months of the study.

### Vaginal seeding partly normalizes microbiome trajectories in CS-delivered infants

Across the different body sites, the samples yielded a good overall sequencing depth (mean depth of 63,035 paired-end reads per sample) with a low probability of sample contamination, as indicated by a survey of negative controls (Figure S1D). Analysis of the vaginal gauze stored in the vagina for 1 h before the CS procedure with which the neonates were swabbed showed that ~76% of bacterial amplicon sequence variants (ASVs; STAR Methods) contained in maternal vaginal swabs were also present in the gauze (Figure S2A).

Some studies have reported decreased alpha diversity in CS-born versus vaginally born infants.[21] Others have reported no differences by birth mode.[4,5] Using a linear mixed-effects model, we found inconsistent results depending on the body site and alpha diversity metric (Methods S1). One possibility for this inconsistency is that the dynamic nature of the developing microbiome can be highly non-linear, and data collected longitudinally often vary in frequency and timing across individuals. To account for these potential irregularities, we applied a novel method called Bayesian sparse functional principal-component analysis (SFPCA)[22] to estimate individual trajectories (STAR Methods). Using SFPCA, we found that alpha diversity trajectories did not differ among birth modes when measured as Shannon diversity (Figure S3) or when accounting for phylogenetic relatedness (SFPCA on Faith's Phylogenetic Diversity -PD-, data not shown).

However, significant birth group differences were found in beta diversity when using an unsupervised dimensionality reduction method called compositional tensor factorization (CTF).[23] CTF accounts for repeated measurements, allowing comparisons of beta diversity over time ("trajectory") while accounting for the sparse compositional nature of next-generation microbiome sequencing data.[24,25] The trajectory of gut microbiota development in CS-born infants diverged from that of vaginally born infants through the entire first year of life (Figure 1). These results are consistent with findings from previous studies that used more traditional analysis approaches.[4,5] CTF also detected measurable differences in microbial development of the mouth (Figure 2) and skin (Figure 3), underscoring the importance of birth mode in affecting multiple microbial niches during human development.

Seeding CS-born infants led to a developmental trajectory that more closely resembled that of vaginally born infants, most prominently in feces (Figures 1, S4A, and S4B) and skin (Figures 3, S4A, and S4B); this trend held when considering only the 101 babies born in the United States (Figure S4C), but other countries lacked sufficient sample size for individual analysis. Furthermore, a stepwise redundancy analysis based on the first three principal components of CTF ordination[26] confirmed that birth mode significantly contributed to differences in microbial community structures in the gut and on skin but not in the mouth, with effect sizes of 0.17 ($R^2$) in fecal samples and 0.09 in skin samples (Methods S2). Analyzing these data using more conventional tools for comparing beta diversity that do not account for inter-individual variation in repeated-measure studies (Methods S3 and S4), evaluated through PERMANOVA (on unweighted UniFrac distance) or RDA (on principal-coordinate analysis [PCoA] PCs), reveals individuals as the primary driver of variation, as

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



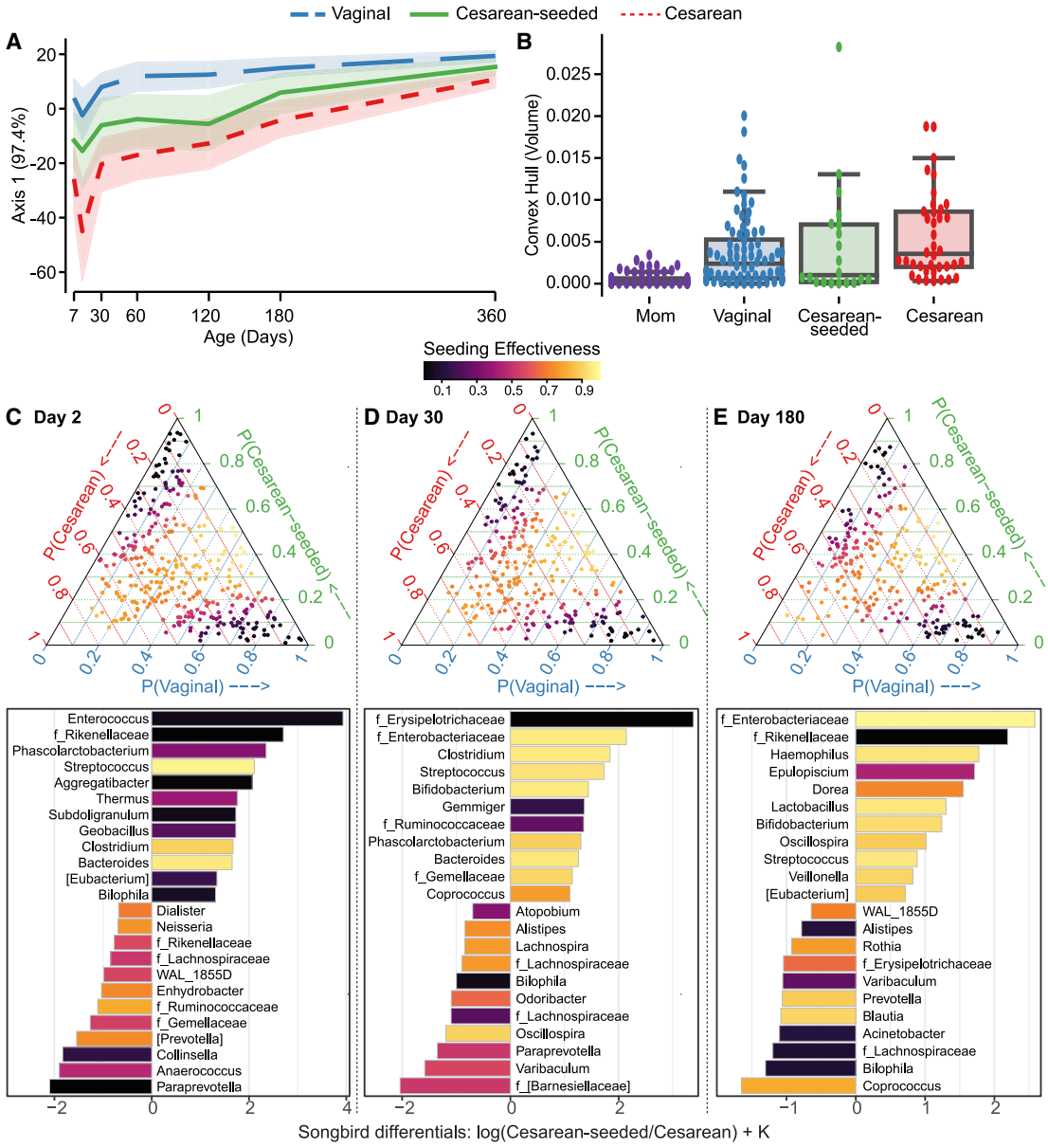



**Figure 1. Fecal microbiota development during the first year of life in babies is discordant with birth mode/exposure**

(A) Compositional tensor factorization (CTF) first principal component (y axis) of infant samples over age in days (x axis).

(B) Convex hull volume (y axis, median and interquantile range) on the first three principal coordinates (unweighted UniFrac distances) in mothers (purple) and infants by birth mode or exposure (x axis). CS-born infants show the highest volumes and vaginally born the lowest, with CS-seeded babies showing intermediate volumes; all pairwise comparisons are significant using Mann-Whitney test with Bonferroni corrections at 0.05 level (Methods S3).

(C–E) Songbird differentials for days 2, 30, and 180 after birth; ternary plots of the inverse additive log-ratio transform (inverse ALR) of Songbird differentials give the estimated probability of a microbe being observed for CS (left axes, red), vaginal (bottom axes, blue), or CS seeded (right axes, green). The color of the dots represents the seeding effectiveness, with yellow indicating effectively seeded/suppressed and black indicating not effectively seeded. Below each triangle, bar plots of the top and bottom 20% Songbird differentials are summarized at genus level between CS-seeded and CS-born babies; a positive value indicates higher association with the CS-seeded group, and a negative value indicates higher association with CS. Bars are colored by the ASVs' seeding effectiveness. The majority of discordant taxa overrepresented in the CS-seeded group over the CS group are shown as yellow-orange, indicating ASVs seeded effectively in the CS-seeded group, and these are observed at all ages.

See also Figure S4 and Methods S2, S3, S4, S5, S6, and S7.

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



CellPress

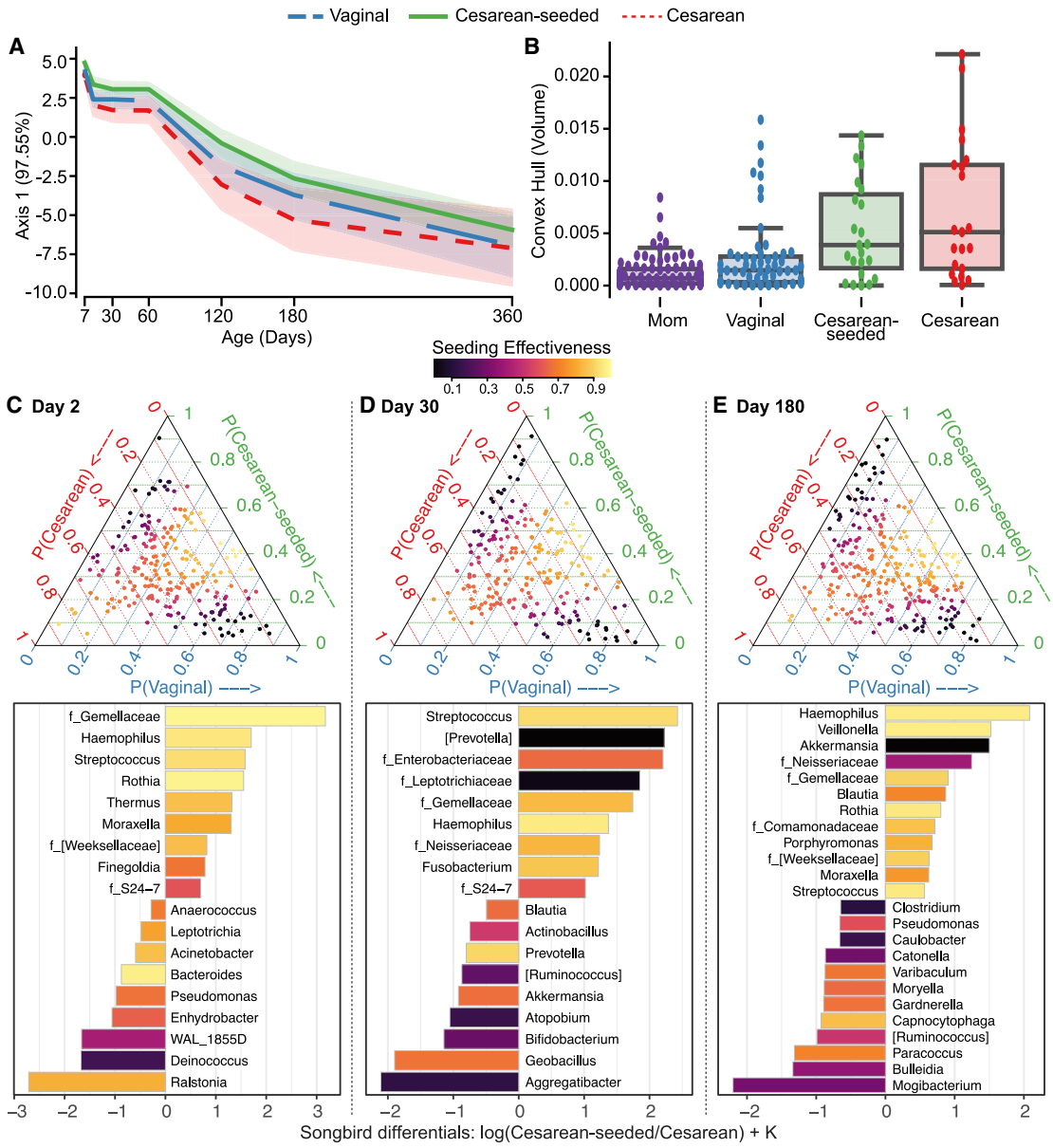



**Figure 2. Oral microbiota development during the first year of life in babies is discordant with birth mode/exposure**

(A) CTF first principal component (y axis, median and interquantile range) of infant samples over age in days (x axis).

(B) Convex hull volume (y axis) on the first three principal coordinates (unweighted UniFrac distances) in mothers (purple) and infants by birth mode or exposure (x axis). CS-born infants show the highest volumes and vaginally born infants the lowest, with CS-seeded babies showing intermediate volumes; all pairwise comparisons are significant using Mann-Whitney test with Bonferroni corrections at 0.05 level (Table S3).

(C–E) Songbird differentials for days 2, 30, and 180 after birth; ternary plots of the inverse ALR of Songbird differentials give the estimated probability of a microbe being observed for CS (left axes, red), vaginal (bottom axes, blue), or CS seeded (right axes, green). The color of the dots represents the seeding effectiveness, with yellow indicating effectively seeded/suppressed and black indicating not effectively seeded. Below each triangle, bar plots of the top and bottom 20% Songbird differentials are summarized at genus level between CS-seeded and CS-born babies; a positive value indicates higher association with the CS-seeded group, and a negative value indicates higher association with CS. Bars are colored by the ASVs' seeding effectiveness. The majority of discordant taxa overrepresented in the CS-seeded group over the CS group are yellow-orange, indicating ASVs seeded effectively in the CS-seeded group, and these are observed at all ages.

See also Figure S4 and Methods S2, S3, S4, S5, S6, and S7.

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



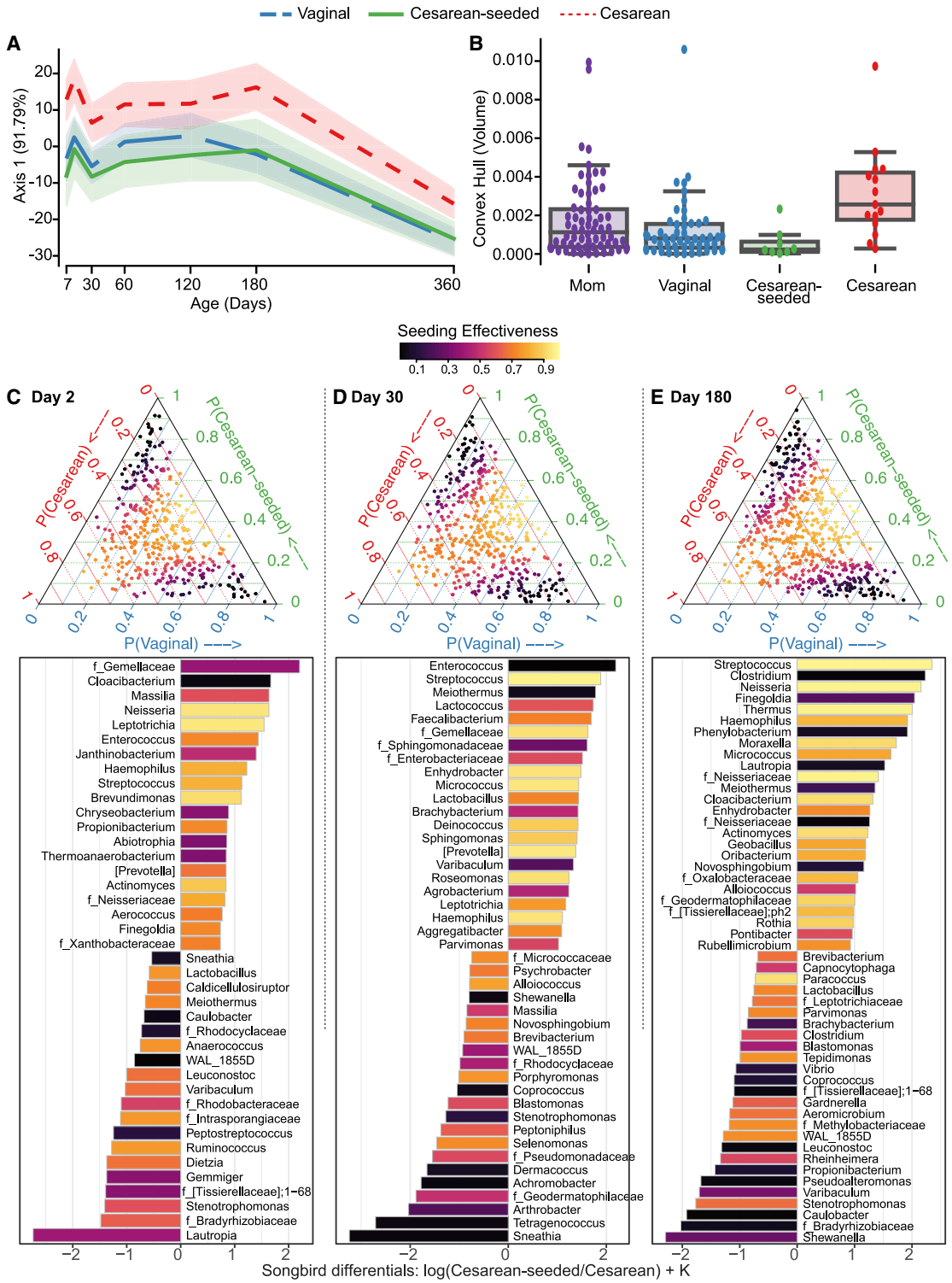



**Figure 3. Skin microbiota development during the first year of life in babies is discordant with birth mode/exposure**

(A) CTF first principal component (y axis, median and interquantile range)) of infant samples over age in days (x axis).

(B) Convex hull volume (y axis) on the first three principal coordinates (unweighted UniFrac distances) in mothers (purple) and infants by birth mode or exposure (x axis). CS-born infants show highest volumes and vaginally born the lowest, with CS-seeded babies showing intermediate volumes; all but one pairwise comparison are significant using Mann-Whitney test with Bonferroni corrections at 0.05 level (Table S3).

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003





expected (PERMANOVA F-statistic = 5.45, $p \leq 0.001$; RDA adjusted $R^2 = 0.113$; Methods S5). High interindividual variation obscured the ability to detect differences because of more muted factors, such as birth mode, using these methods. These findings reveal that birth mode affects the development of microbial communities and that this effect may be undetected upon analysis with traditional bioinformatics tools.

Differences in microbial composition stability have been used to differentiate phenotypes in longitudinal studies.[27,28] Accordingly, we next compared variability across samples over time within a given individual. To leverage the dense sampling design, we calculated the volume of the shape determined by an individual's samples in the first 3 principal coordinates of unweighted UniFrac space using a convex hull analysis (STAR Methods). As expected, the average variability of the microbiome over an infant's first year of life was much greater than the variability in the mother's microbiome (Figures 1B, 2B, and 3B). CS-born infants had significantly greater microbial variability than vaginally born infants, and the variability of seeded infants was intermediate (Figures 1B, 2B, and 3B; Methods S6). This finding held true for fecal, oral, and skin samples, suggesting that vaginal seeding may also help stabilize microbiome development. This trend can also be observed using data within the first 6 months (Methods S7). Possible confounders, such as antibiotic consumption (which was similar between baby groups; Table 1), were discarded; in the CS-born and restored babies, stepwise RDA did not recognize antibiotic consumption as a factor altering seeding efficiency. These results indicate that vaginal seeding resulted in partial recovery of the microbiome in CS-delivered infants.

### Bacterial taxa associated with effective seeding

To determine whether specific microbial taxonomies were being seeded well or whether the overall seeding across all microbes was partial, we first identified which taxa were most associated with vaginal birth compared with CS birth using Songbird[25] and then calculated a seeding effectiveness score for those taxa (STAR Methods; 0 indicates poor seeding, and 1 indicates effective seeding or effectively suppressed). Effectively seeded microbes are those shared by vaginally and CS-seeded infants. Effectively suppressed microbes are those highly associated only with unseeded CS infants, indicating that seeding excludes that microbe. Many taxa highly associated with CS-seeded infants had a seeding effectiveness score of greater than 0.8, indicating that the vaginal seeding method was able to establish microbes missing in CS-born babies (Figures 1C–1E, 2C–2E, and 3C–3E). Notably, in the infant gut, ASVs from common gut-associated genera, such as *Bacteroides*, *Streptococcus*, and *Clostridium*, were identified to be enriched in CS-seeded infants and have high seeding effectiveness scores at early time points (Figures 1C–1E; Methods S8 and S9). Especially of note, *Bacteroides* was identified consistently as being associated with vaginal seeding (Methods S10) using other algorithms, such as ANCOM (Methods S11), MaAsLin2 (Methods S12), and LEfSe (Methods S13). In the mouth, bacteria with high seeding effectiveness scores included ASVs from Gemellaceae, *Haemophilus*, and *Streptococcus* (Figures 2C–2E). In the skin, taxa

(C–E) Songbird differentials for days 2, 30, and 180 after birth; ternary plots of the inverse ALR of Songbird differentials give the estimated probability of a microbe being observed in CS (left axes, red), vaginal (bottom axes, blue), or CS-seeded (right axes, green). The color of the dots represents the seeding effectiveness, with yellow indicating effectively seeded/suppressed and black indicating not effectively seeded. Below each triangle, bar plots of the top and bottom 20% Songbird differentials are summarized at genus level between CS-seeded and CS-born babies; a positive value indicates higher association with the CS-seeded group, and a negative value indicates higher association with CS. Bars are colored by the ASVs' seeding effectiveness. The majority of discordant taxa overrepresented in the CS-seeded group over the CS group are yellow-orange, indicating ASVs seeded effectively in the CS-seeded group, and these are observed at all ages.
See also Figure S4 and Methods S2, S3, S4, S5, S6, and S7.

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



included ASVs from *Streptococcus*, *Neisseria*, *Thermus*, and Neisseriaceae (Figure 3c-e). However, across all three body sites, most of the taxa associated with CS had a moderate to low seeding effectiveness score, indicating that this method was not effective at attenuating the presence of microbes typically depleted in vaginally born babies.

### Maternal sites contribute to the infant microbiota

To determine which body sites from the mother were most likely to have the highest contributions to shaping the infant microbiome, we also used the source-tracking tool fast expectation-maximization microbial source tracking (FEAST).[29] The first 2 days of life showed a prominent maternal vaginal source in the oral and skin sites of infants exposed to vaginal fluids; however, within the first few days, a large proportion of the microbiota colonizing the infants' sites was shared with the corresponding maternal site regardless of birth mode or seeding status (Figure 4). Selection by the specific body site was evidenced by lack of overrepresentation of *Lactobacillus*, a dominant member of the mother's vagina, among infants born vaginally or exposed to vaginal gauze compared with CS-born babies. Not surprisingly, we found that the infant oral microbiota most resembled that of the mother's mouth and areolae (Figures 4H and 4K) and that the infant skin microbiota resembled that of the mother's skin (Figure 4O), consistent with exposure patterns and differential selection exerted by different body sites in the baby.

Interestingly, we observed a notable taxonomic overlap between the maternal vagina and other maternal body sites, especially feces, on the day of giving birth: nearly 30% of the bacterial ASVs in vaginal samples were shared with feces (5.5% with feces alone and 24.5% with feces and some other body sites) and 22.3% with more distant body sites such as arm skin, mouth, and nose (Figure 5A; Methods S14). These trends showing the pluripotent nature of the perinatal vaginal microbiome held true when examining the mothers in different countries from this study, despite variations in specific proportions (Figures S2B and S2C; Methods S15). In contrast, women who were not pregnant, from the Human Microbiome Project (HMP) study, shared less than 20% of vaginal ASVs with other body sites, predominantly with skin, and none with fecal samples (Figure 5B). These results point to the importance of maternal sources of microbes on the developing infant microbial consortium.

## DISCUSSION

This intervention study expands the findings of previous smaller studies, demonstrating that microbial differences associated with delivery mode can be reduced by exposure to a vaginal microbial source at birth. The study only included scheduled CSs on healthy mothers (mostly because of multiple previous CSs and malposition presentations) because infants born by emergency CS after rupture of the chorioamniotic membrane are likely exposed to maternal microbes, given enough time before the CS procedure.[30]

Using advanced and longitudinally aware methods, we found that birth mode significantly differentiated infant gut and skin microbiome development and that seeding worked to adjust the trajectory of CS-delivered infants through partial restoration of microbiome features associated with vaginal delivery. For example, differential abundance analyses confirmed previous findings showing that, in the gut, *Bacteroides* and *Parabacteroides*, common gut-associated genera, are highly associated with vaginally born infants. Our study further shows that seeding works to effectively

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



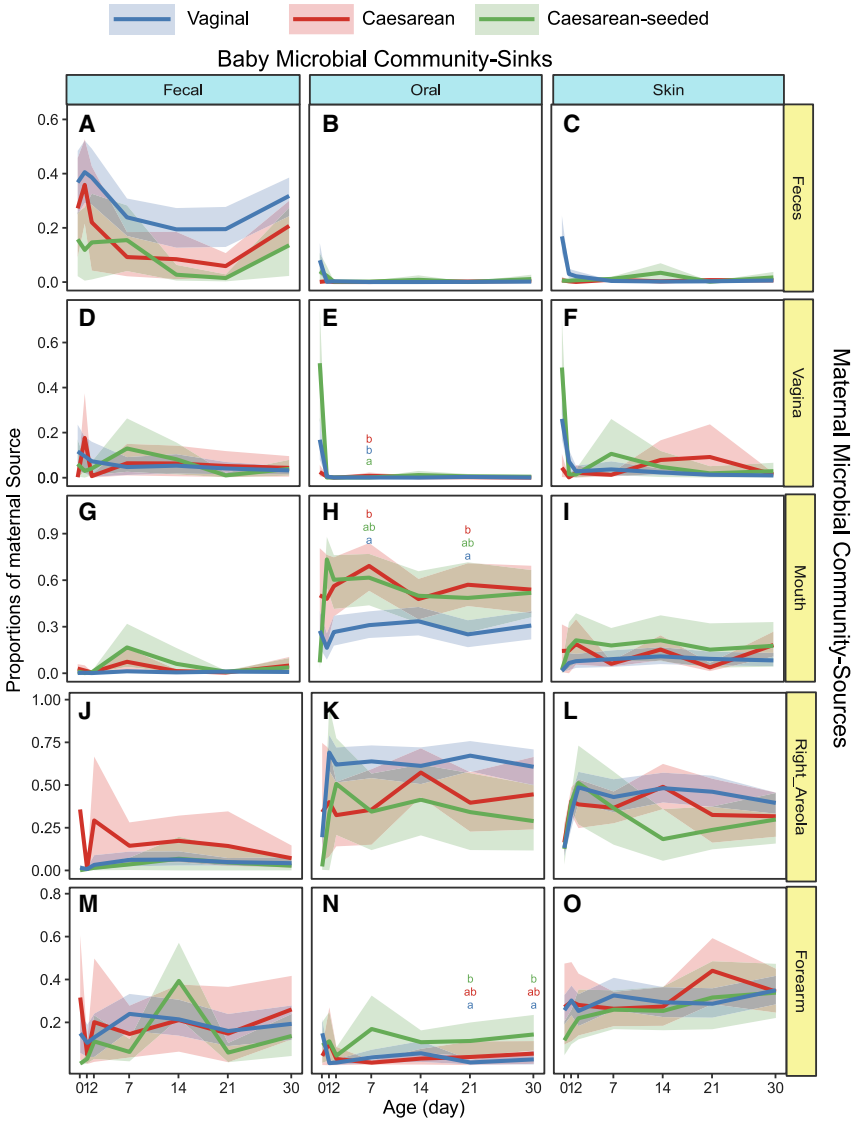



**Figure 4. Microbial source tracking of the neonate microbiome (first month) through fast expectation-maximization microbial source tracking (FEAST)**

(A–O) Contributions (y axes) of various maternal sources (rows) to the infant microbial community (columns) are estimated across age in days (x axes) for the first month of life in 15 mother-baby pairs. Error bars show 95% confidence interval of the mean calculated by bootstrapping; Dunn's tests based on Kruskal-Wallis were performed on each time point by each maternal source for each baby sink, and significant differences are marked by different letters in each panel. The vaginal source, prominent on day 0 for mouth and skin in babies exposed to vaginal fluid (vaginal and CS seeded; E and F), was not prominent later in any baby site. Baby site-specific communities resemble the corresponding maternal site (A, H, and O), consistent with specific site selection of bacteria. The maternal right areola appears as a source for baby oral bacteria (K), which likely means that baby oral bacteria is transmitted to the mother's areola during lactation.

restore these and other genera associated with vaginal birth. However, there are several other genera that do not appear to establish well in seeded infants (e.g., *Bilophila*). Although we observed a significant association of *Enterococcus* with CS-born infants (which has been noted in previous studies as a potential opportunistic pathogen), we did not see a weakened association of this genus or most other CS-associated genera with seeded babies. Further research is needed to determine

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



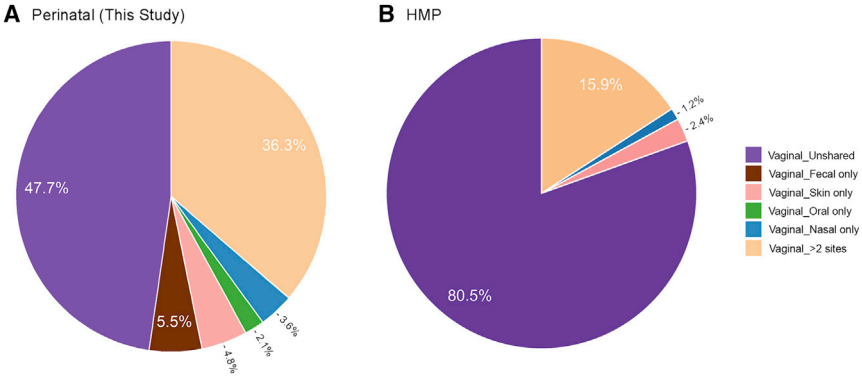



**Figure 5. Proportions of bacterial vaginal ASVs shared with other body sites in the mothers of the current study on the day of delivery and in non-pregnant women**

(A) V4 sequences from vaginal swabs and gauze obtained from 97 parturient mothers in this study on the day of birth. Current study data were sequenced by Illumina HiSeq and processed by QIIME2 using the same pipeline as for the HMP data.

(B) HMP V4 data from vaginal swabs obtained from 105 non-pregnant women; ASVs included in the analyses were present in at least 10% of the samples in the respective body site. Roche 454 V3V5 sequences were trimmed to obtain the V4 region.

See also Figures S2B and S2C and Methods S13 and S14.

why certain gut genera may show higher effectiveness for seeding but other taxa may exhibit more resilience after a seeding procedure and the roles of these microbes in the developing infant microbiome.

An interesting facet of our study is the finding that vaginal seeding led to converging microbial compositions in the infant gut despite the exposure coming from a vaginal source. The same pattern was observed in the skin environment. Our results clearly indicate that, from very early time points, the microbiota of an infant largely resembles the same maternal site, supporting the idea of strong site selection occurring from very early ages (i.e., that different body sites will select for specific microbes out of a diverse population). This is further supported by the finding that Lactobacillaceae, the most dominant family of the mother's vagina, was not identified as one of the most differentially abundant among infants at any of the three body sites observed. Site selection is also consistent with the recent evidence of successful engraftment after fecal microbiota transplant from the mother to CS neonates[31] and with previous evidence of fecal bacteria in the infant gut.[32,33] Indeed, bacterial transfer from homologous sites from the mother and other family members surely occurs after birth. However, this may only be part of the story. Our results show that, unlike in non-pregnant women, more ASVs from the vaginal microbiome from parturient women overlap with those in other body sites, mostly the proximal rectum (which, in mammals, is next to the reproductive canal), but also more distant sites. This strongly suggests a pluripotent capacity of vaginal fluids to seed different sites of the baby's body. Transmission and colonization by these pioneer species may then modulate the succession that proceeds, influencing engraftment of later colonizers to each body site.[34] Major changes in the vaginal microbiota during pregnancy have been described[35], although the changes in the last semester have not been deeply characterized. This begs the question of whether the vaginal microbiome becomes specifically primed during pregnancy to deliver key pioneer colonizers tailored toward multiple body sites of the infant. This hypothesis is supported by previous work demonstrating the bi-phasic dynamics in gestational changes in which after decreasing diversity in the first two thirds of gestation, in the last gestational trimester diversity increases at the expense of *Lactobacillus* from week 24 of

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



pregnancy until birth;[36] increase in vaginal diversity continues in the postpartum vaginal tract for up to 1 year following birth.[37]

This study provides solid evidence that deliberate, early microbial seeding can help naturalize the microbiome developmental trajectory of CS-born infants. Although overall trajectories do appear to head toward convergence over time, studies show that early perturbations during the crucial developmental window of very early life seem to have irreversible consequences.[38–41] Restoring natural exposure at birth may be one way to reduce the risk of CS-associated diseases such as obesity, asthma, allergies, and immune disfunction. However, randomized clinical trials with large cohorts are needed to gain conclusive evidence for microbial restoration at birth improving health outcomes.[42] Moreover, in light of recent research showing that oral administration of maternal fecal microbes is also effective in restoring the microbiome in CS-delivered infants,[20] future research investigating the effects of exposure to both sources explicitly compared with either single source will help determine the best routes for restoring the neonate microbiome. In this study, we exposed infants to freshly collected maternal vaginal/perineal microbes, but it is unknown how storage would alter the microbiota composition. More research is needed to determine whether it is best that they receive their own mother's microbiome or achieve defined universal cocktails that can be used to restore neonates.

### Limitations of study

This study is limited by the cohort size, particularly in countries outside of the United States; the follow-up time of the first year of life because any longer-term consequences of seeding were not assessed; and by 16S rDNA amplicon sequencing, which excludes functional characterization as well as fungi and viruses. Future studies capturing longer time frames, larger and broader cultural and geographic representation, and additional data types are needed to gain a better understanding of how seeding affects the microbiome and, ultimately, the health of CS-delivered infants.

## 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
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects
- METHOD DETAILS
  - Microbial restoration procedure
  - Sample collection and processing
  - Sequencing and data processing
  - Bacterial source tracking
  - Alpha diversity
  - Beta diversity and convex hull
  - Dimensionality reduction
  - Effect size analysis
  - Differential abundance and effectiveness score
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.medj.2021.05.003.

## ACKNOWLEDGMENTS

This work was partially supported by the C&D Research Fund, the Emch Fund for Human Microbiome Studies, and CIFAR FS20-078 #125869 (to M.G.D.-B.); Chilean CONICYT PIA/ANILLO grant ACT172097 and Chilean SOCHIPE Project 022019 (to P.R.H.); and the Norwegian Institute of Public Health (2019-0350), the Emerald Foundation, an NIH Pioneer award (1DP1AT010885), the National Institute of Justice (2016-DN-BX-4194), the San Diego Digestive Diseases Research Center (NIDDK 1P30DK120515), Janssen Pharmaceuticals (20175015), and by in-kind donations from Illumina, MoBio/QIAGEN, and the Center for Microbiome Innovation at University of California, San Diego (to R.K.). We acknowledge the contribution of students, technicians, and MDs who participated in obtaining the samples and preparing the metadata: Sukhleen Bedi, Allison Horan, and Yi Cai from NYUSM; Noraliz Garcia, Hebe Rosado, Selena Marie Rodriguez, and Keimari Mendez from UPR; Maricruz Mojica, Magaly Magariños, and Myriam Corrales from Universidad Real y Pontificia San Francisco Javier de Chuquisaca and the University Hospital in Sucre, Bolivia; and Mauricio Sandoval, Marlene Ortiz, and Carolina Serrano from the PUC. We thank Gail Ackermann for help with curating the metadata and uploading sequences to Qiita and James T. Morton for providing advice regarding use of Songbird differentials.

## AUTHOR CONTRIBUTIONS

M.G.D.-B. designed the study. M.G.D.-B., S.J.S., J.W., P.R.H., C.D.H., N.H., E.A., D.N., C.G., V.M., W.S., M.J., J.C., H.S., M.C.C., I.G.-M., F.G.R., J.I.R.-V., M.C.-R., J.F.R.-C., and R.K. collected and processed specimens. M.G.D.-B. and R.K. sequenced and generated data. M.G.D.-B., S.J.S., J.W., C.M., L.J., W.K.T., L.S., D.M., and R.K. analyzed data and performed and oversaw the bioinformatics and statistical analyses. M.G.D.-B. and R.K. have unrestricted access to all data. M.G.D.-B., S.J.S., and R.K. drafted the manuscript. All authors reviewed, agreed to submit the final manuscript, read and approved the final draft, and take full responsibility for its content, including the accuracy of the data and the fidelity of the trial to the registered protocol and its statistical analyses.

## DECLARATION OF INTERESTS

New York University has a U.S. patent (10357521) on behalf of M.G.D.-B., related to methods for restoring the microbiota of newborns.

## INCLUSION AND DIVERSITY

We worked to ensure that the study questionnaires were prepared in an inclusive way. One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science. The author list of this paper includes contributors from the location where the research was conducted who participated in the data collection, design, analysis, and/or interpretation of the work.

Received: January 29, 2021
Revised: April 1, 2021
Accepted: May 13, 2021
Published: June 17, 2021

## REFERENCES


1. Gensollen, T., Iyer, S.S., Kasper, D.L., and Blumberg, R.S. (2016). How colonization by microbiota in early life shapes the immune system. Science *352*, 539–544.
2. Al Nabhani, Z., and Eberl, G. (2020). Imprinting of the immune system by the microbiota early in life. Mucosal Immunol. *13*, 183–189.
3. Dominguez-Bello, M.G., Costello, E.K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., and Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc. Natl. Acad. Sci. USA *107*, 11971–11975.
4. Bokulich, N.A., Chung, J., Battaglia, T., Henderson, N., Jay, M., Li, H., D Lieber, A., Wu, F., Perez-Perez, G.I., Chen, Y., et al. (2016). Antibiotics, birth mode, and diet shape microbiome maturation during early life. Sci. Transl. Med. *8*, 343ra82.
5. Yassour, M., Vatanen, T., Siljander, H., Hämäläinen, A.M., Härkönen, T., Ryhänen, S.J., Franzosa, E.A., Vlamakis, H., Huttenhower, C., Gevers, D., et al.; DIABIMMUNE Study Group (2016). Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. Sci. Transl. Med. *8*, 343ra81.
6. 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.
7. 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.
8. Malamitsi-Puchner, A., Protonotariou, E., Boutsikou, T., Makrakis, E., Sarandakou, A., and Creatsas, G. (2005). The influence of the mode of delivery on circulating cytokine concentrations in the perinatal period. Early Hum. Dev. *81*, 387–392.
9. Stokholm, J., Blaser, M.J., Thorsen, J., Rasmussen, M.A., Waage, J., Vinding, R.K., Schoos, A.M., Kunøe, A., Fink, N.R., Chawes, B.L., et al. (2018). Maturation of the gut microbiome and risk of asthma in childhood. Nat. Commun. *9*, 141.
10. Andersen, V., Möller, S., Jensen, P.B., Møller, F.T., and Green, A. (2020). Caesarean Delivery and Risk of Chronic Inflammatory Diseases (Inflammatory Bowel Disease, Rheumatoid Arthritis, Coeliac Disease, and Diabetes Mellitus): A Population Based Registry Study of 2,699,479 Births in Denmark During 1973-2016. Clin. Epidemiol. *12*, 287–293.
11. Blustein, J., Attina, T., Liu, M., Ryan, A.M., Cox, L.M., Blaser, M.J., and Trasande, L. (2013). Association of caesarean delivery with child adiposity from age 6 weeks to 15 years. Int. J. Obes. *37*, 900–906.
12. Ardic, C., Usta, O., Omar, E., Yıldız, C., and Memis, E. (2021). Caesarean delivery increases the risk of overweight or obesity in 2-year-old children. J. Obstet. Gynaecol. *41*, 374–379.
13. Cox, L.M., Yamanishi, S., Sohn, J., Alekseyenko, A.V., Leung, J.M., Cho, I., Kim, S.G., Li, H., Gao, Z., Mahana, D., et al. (2014). Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. Cell *158*, 705–721.
14. Martinez, K.A., 2nd, Devlin, J.C., Lacher, C.R., Yin, Y., Cai, Y., Wang, J., and Dominguez-Bello, M.G. (2017). Increased weight gain by C-section: Functional significance of the primordial microbiome. Sci. Adv. *3*, eaao1874.
15. Olszak, T., An, D., Zeissig, S., Vera, M.P., Richter, J., Franke, A., Glickman, J.N., Siebert, R., Baron, R.M., Kasper, D.L., and Blumberg, R.S. (2012). Microbial exposure during early life has persistent effects on natural killer T cell function. Science *336*, 489–493.
16. Livanos, A.E., Greiner, T.U., Vangay, P., Pathmasiri, W., Stewart, D., McRitchie, S., Li, H., Chung, J., Sohn, J., Kim, S., et al. (2016). Antibiotic-mediated gut microbiome perturbation accelerates development of type 1 diabetes in mice. Nat. Microbiol. *1*, 16140.
17. Moya-Pérez, A., Luczynski, P., Renes, I.B., Wang, S., Borre, Y., Anthony Ryan, C., Knol, J., Stanton, C., Dinan, T.G., and Cryan, J.F. (2017). Intervention strategies for cesarean section-induced alterations in the microbiota-gut-brain axis. Nutr. Rev. *75*, 225–240.
18. Braniste, V., Al-Asmakh, M., Kowal, C., Anuar, F., Abbaspour, A., Tóth, M., Korecka, A., Bakocevic, N., Ng, L.G., Kundu, P., et al. (2014). The gut microbiota influences blood-brain barrier permeability in mice. Sci. Transl. Med. *6*, 263ra158.
19. Dominguez-Bello, M.G., De Jesus-Laboy, K.M., Shen, N., Cox, L.M., Amir, A., Gonzalez, A., Bokulich, N.A., Song, S.J., Hoashi, M., Rivera-Vinas, J.I., et al. (2016). Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. Nat. Med. *22*, 250–253.
20. 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.
21. 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.
22. Jiang, L., Zhong, Y., Elrod, C., Natarajan, L., Knight, R., and Thompson, W.K. (2020). BayesTime: Bayesian Functional Principal Components for Sparse Longitudinal Data. arXiv, arXiv:2012.00579. https://arxiv.org/abs/2012.00579.
23. Martino, C., Shenhav, L., Marotz, C.A., Armstrong, G., McDonald, D., Vázquez-Baeza, Y., Morton, J.T., Jiang, L., Dominguez-Bello, M.G., Swafford, A.D., et al. (2020). Context-aware dimensionality reduction deconvolutes gut microbial community dynamics. Nat. Biotechnol. *39*, 165–168.
24. Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., and Egozcue, J.J. (2017). Microbiome Datasets Are Compositional: And This Is Not Optional. Front. Microbiol. *8*, 2224.
25. Morton, J.T., Marotz, C., Washburne, A., Silverman, J., Zaramela, L.S., Edlund, A., Zengler, K., and Knight, R. (2019). Establishing microbial composition measurement standards with reference frames. Nat. Commun. *10*, 2719.
26. Falony, G., Joossens, M., Vieira-Silva, S., Wang, J., Darzi, Y., Faust, K., Kurilshikov, A., Bonder, M.J., Valles-Colomer, M., Vandeputte, D., et al. (2016). Population-level analysis of gut microbiome variation. Science *352*, 560–564.
27. Halfvarson, J., Brislawn, C.J., Lamendella, R., Vázquez-Baeza, Y., Walters, W.A., Bramer, L.M., D'Amato, M., Bonfiglio, F., McDonald, D., Gonzalez, A., et al. (2017). Dynamics of the human gut microbiome in inflammatory bowel disease. Nat. Microbiol. *2*, 17004.
28. Zaneveld, J.R., McMinds, R., and Vega Thurber, R. (2017). Stress and stability: applying the Anna Karenina principle to animal microbiomes. Nat. Microbiol. *2*, 17121.
29. Shenhav, L., Thompson, M., Joseph, T.A., Briscoe, L., Furman, O., Bogumil, D., Mizrahi, I., Pe'er, I., and Halperin, E. (2019). FEAST: fast expectation-maximization for microbial source tracking. Nat. Methods *16*, 627–632.
30. Azad, M.B., Konya, T., Maughan, H., Guttman, D.S., Field, C.J., Chari, R.S., Sears, M.R., Becker, A.B., Scott, J.A., and Kozyrskyj, A.L.; CHILD Study Investigators (2013). Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. CMAJ *185*, 385–394.
31. 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.
32. 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.
33. Helve, O., Korpela, K., Kolho, K.-L., Saisto, T., Skogberg, K., Dikareva, E., Stefanovic, V., Salonen, A., de Vos, W.M., and Andersson, S. (2019). Maternal Fecal Transplantation to Infants Born by Cesarean Section: Safety and Feasibility. Open Forum Infect. Dis. *6*, S68.
34. Martínez, I., Maldonado-Gomez, M.X., Gomes-Neto, J.C., Kittana, H., Ding, H., Schmaltz, R., Joglekar, P., Cardona, R.J., Marsteller, N.L., Kembel, S.W., et al. (2018). Experimental evaluation of the importance of colonization history in early-life gut microbiota assembly. eLife *7*, e36521.

35. Stout, M.J., Zhou, Y., Wylie, K.M., Tarr, P.I., Macones, G.A., and Tuuli, M.G. (2017). Early pregnancy vaginal microbiome trends and preterm birth. Am. J. Obstet. Gynecol. *217*, 356.e1–356.e18.

36. Rasmussen, M.A., Thorsen, J., Dominguez-Bello, M.G., Blaser, M.J., Mortensen, M.S., Brejnrod, A.D., Shah, S.A., Hjelmsø, M.H., Lehtimäki, J., Trivedi, U., et al. (2020). Ecological succession in the vaginal microbiota during pregnancy and birth. ISME J. *14*, 2325–2335.

37. DiGiulio, D.B., Callahan, B.J., McMurdie, P.J., Costello, E.K., Lyell, D.J., Robaczewska, A., Sun, C.L., Goltsman, D.S.A., Wong, R.J., Shaw, G., et al. (2015). Temporal and spatial variation of the human microbiota during pregnancy. Proc. Natl. Acad. Sci. USA *112*, 11060–11065.

38. Huh, S.Y., Rifas-Shiman, S.L., Zera, C.A., Edwards, J.W., Oken, E., Weiss, S.T., and Gillman, M.W. (2012). Delivery by caesarean section and risk of obesity in preschool age children: a prospective cohort study. Arch. Dis. Child. *97*, 610–616.

39. Pistiner, M., Gold, D.R., Abdulkerim, H., Hoffman, E., and Celedón, J.C. (2008). Birth by cesarean section, allergic rhinitis, and allergic sensitization among children with a parental history of atopy. J. Allergy Clin. Immunol. *122*, 274–279.

40. Sevelsted, A., Stokholm, J., Bønnelykke, K., and Bisgaard, H. (2015). Cesarean section and chronic immune disorders. Pediatrics *135*, e92–e98.

41. Thavagnanam, S., Fleming, J., Bromley, A., Shields, M.D., and Cardwell, C.R. (2008). A meta-analysis of the association between Caesarean section and childhood asthma. Clin. Exp. Allergy *38*, 629–633.

42. Mueller, N.T., Dominguez-Bello, M.G., Appel, L.J., and Hourigan, S.K. (2020). 'Vaginal seeding' after a caesarean section provides benefits to newborn children: FOR: Does exposing caesarean-delivered newborns to the vaginal microbiome affect their chronic disease risk? The critical need for trials of 'vaginal seeding' during caesarean section. BJOG *127*, 301.

43. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. *37*, 852–857.

44. Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. *6*, 1621–1624.

45. Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., Prill, R.J., Tripathi, A., Gibbons, S.M., Ackermann, G., et al.; Earth Microbiome Project Consortium (2017). A communal catalogue reveals Earth's multiscale microbial diversity. Nature *551*, 457–463.

46. Amir, A., McDonald, D., Navas-Molina, J.A., Kopylova, E., Morton, J.T., Zech Xu, Z., Kightley, E.P., Thompson, L.R., Hyde, E.R., Gonzalez, A., and Knight, R. (2017). Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems *2*, e00191-16.

47. Janssen, S., McDonald, D., Gonzalez, A., Navas-Molina, J.A., Jiang, L., Xu, Z.Z., Winker, K., Kado, D.M., Orwoll, E., Manary, M., et al. (2018). Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. mSystems *3*, e00021-18.

48. McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. *6*, 610–618.

49. Schiffer, L., Azhar, R., Shepherd, L., Ramos, M., Geistlinger, L., Huttenhower, C., Dowd, J.B., Segata, N., and Waldron, L. (2019). HMP16SData: Efficient Access to the Human Microbiome Project Through Bioconductor. Am. J. Epidemiol. *188*, 1023–1026.

50. James, G.M., Hastie, T.J., and Sugar, C.A. (2000). Principal component models for sparse functional data. Biometrika *87*, 587–602.

51. Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. Appl. Environ. Microbiol. *71*, 8228–8235.

52. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al.; SciPy 1.0 Contributors (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods *17*, 261–272.

53. Anderson, M.J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA). https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat07841.

54. Aitchison, J. (1983). Principal Component Analysis of Compositional Data. Biometrika *70*, 57–65.

55. Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., et al. (2019). vegan: Community Ecology Package. https://cran.r-project.org/web/packages/vegan/index.html.

56. Aitchison, J. (1982). The Statistical-Analysis of Compositional Data. J. Royal Stat. Soc. B Stat. Met. *44*, 139–177.

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



CellPress

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Biological samples | | |
| Stool, mouth, and skin specimen from infants | This Study | N/A |
| Stool, mouth, skin, vagina, nasal specimen from mothers | This Study | N/A |
| Critical commercial assays | | |
| DNeasy PowerSoil HTP 96 Kit | QIAGEN | 12955-4 |
| Illumina MiSeq/HiSeq sequencing PE150 | Genewiz, New York University Genome Technology Center, and UC San Diego Institute for Genomic Medicine Genomics Facility | N/A |
| Deposited data | | |
| Raw sequencing data | European Bioinformatics Institute | ENA: ERP120105, ENA: ERP016173 and ENA: ERP120109 |
| Raw sequencing data and processed data | Qiita | Qiita: Study 10894, Qiita: 10249, and Qiita: 1718 |
| Source codes, tests, and notebooks to generate the figures | This study | https://github.com/knightlab-analyses/seeding-study |
| Oligonucleotides | | |
| 515f forward primer: GTGYCAGCMGCCGCGGTAA | EMP protocol | N/A |
| 806r reverse primer: GGACTACNVGGGTWTCTAAT | EMP protocol | N/A |
| Software and algorithms | | |
| Qiime2-2019.8 | Bolyen et al.[43] | N/A |
| FEAST | Shenhav et al.[29] | https://github.com/ETaSky/FEAST branch: removewritingfiles |
| BayestTime | Jiang et al.[22] | https://github.com/biocore/bayestime |
| Gemelli | Martino et al.[23] | https://github.com/biocore/gemelli |
| stepwise-rda.R | This study | https://github.com/knightlab-analyses/seeding-study |
| Songbird | Morton et al.[25] | https://github.com/biocore/songbird |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by Dr. Maria Gloria Dominguez-Bello, mg.dominguez-bello@rutgers.edu

### Materials availability

DNA or samples that are available upon request, as collaboration, subject to availability.

### Data and code availability

Sequence data have been deposited at European Bioinformatics Institute (EBI) under study accession number ENA: ERP120105, ENA: ERP016173 and ENA: ERP120109. Any Supplementary Information and Source Data files are available in the online version of the paper. Source codes, tests, and notebooks to generate the figures are available at https://github.com/knightlab-analyses/seeding-study.

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Human subjects

The study recruitment was conducted in the period 2017 to 2019, at participating centers in the mainland US, Puerto Rico, Chile, Spain, and Bolivia. Physician-assessed healthy mothers who were set to deliver vaginally or by scheduled CS were offered to participate in this study during a physician's visit. Inclusion criteria included healthy, non-obese mothers with no pregnancy complications. Mothers delivering vaginally included GBS positive subjects who received antibiotics during labor. Mothers scheduled to have a CS were divided into two groups based on their willingness to have their newborns swabbed with the gauze: seeded CS (CS-seeded) and unseeded CS (CS). The mothers in the CS groups had intact amniotic membranes at the time of delivery. For the CS-seeded group, mothers had to have negative results for standard-of-care tests for Group B Streptococcus (GBS, standard test at 36 weeks by culturing) and STDs (including HIV and Chlamydia), no signs of vaginosis or viral infections as determined by their obstetrician, and a vaginal pH < 4.5 at 1-2 h preceding the procedure. The infants from this group received swabbing of a gauze soaked with maternal vaginal fluids for microbial restoration at birth. No mock gauze was applied to the unseeded CS babies. All mothers received standard-of-care treatment, including preventive perinatal antibiotics (Beta lactams: mostly Cephalosporins or Penicillin) for mothers who underwent CS section or for vaginally delivering GBS positive mothers (Table 1). The study was approved by the Institutional Review Boards from New York University School of Medicine (S14-00377), University of Puerto Rico Rio Piedras (1011-107) campus, Pontificia Universidad Católica de Chile (180814027), and Hospital Universitario y Politécnico La Fe, Spain (2015/0024), Universidad Mayor, Real y Pontificia de San Francisco Xavier de Chuquisaca, Bolivia (02/2014). Written informed consent was obtained from all participants. This research did not require authorization from the FDA or an equivalent regulatory organization.

## METHOD DETAILS

### Microbial restoration procedure

Restoration procedures were the same in all participating centers. Within the hour prior to the procedure, maternal vaginal pH was measured using a sterile swab and a paper pH strip (Fisher). Once the pH was confirmed to be < 4.5, an 8x8 cm four-layered gauze (Fisherbrand Cat # 22028558) was folded like a fan, and then in half, inserted in a tampon applicator and wet with sterile saline solution using a sterile Pasteur pipette. The gauze was inserted in the maternal vagina for one hour. Immediately before the CS surgery started, the gauze was extracted and placed in a sterile collector and kept at room temperature. As soon as the baby was brought to the neonate lamp and within 1 minute after delivery, the infant was swabbed with the gauze, starting on the lips, followed by the face, thorax, arms, legs, genitals and anal region, and finally the back. The swabbing took approximately 15 s. The neonatologist then proceeded to perform the standard detailed examination of the newborn.

### Sample collection and processing

Sampling with sterile swabs in different body sites took place within the first hours after birth in all babies (including the vaginal gauze exposed CS group, who were sampled after the gauze swabbing procedure), then at day 1-3, weekly for the first month and monthly up to the first year. Sampled body sites included oral mucosa, right arm region and feces of the baby, and the same sites of the mother plus nasal, right areola, vaginal swabs. All samples were transported to the laboratory with ice packs within two hours of collection and stored at −80°C until further processing.

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003





Gauze samples for microbiota determination were obtained from one $cm^2$ from the center of the gauze. Together with swabs and gauzes, 399 control blanks and 249 reagent blanks were included and processed.

### Sequencing and data processing

DNA extraction, amplicon generation, and sequencing were performed as described in the protocols for the Earth Microbiome Project (https://earthmicrobiome.org).[44,45] Briefly, DNA was from samples using the DNeasy PowerSoil HTP 96 Kit, and the V4 region of the 16S rRNA gene was amplified using the 515f/806r primers and prepared for sequencing http://press.igsb.anl.gov/earthmicrobiome/protocols-and-standards/16s/. Sequencing was performed at the University of Colorado Biofrontiers Sequencing Facility, New York University Genome Technology Center, or UC San Diego Institute for Genomic Medicine Genomics Facility using the Illumina MiSeq or HiSeq sequencing instrument. Raw reads were de-multiplexed and quality filtered using QIIME2 v2019.10 with default parameters.[43] Quality-filtered reads were clustered into amplicon sequence variants (ASVs) using deblur v1.1.0.[46] A phylogenetic tree was constructed through insertion[47] with the Greengenes v13_8 as a reference backbone.[48]

Data used from the human microbiome project (HMP) included 16S rRNA gene sequence data from 105 women (Methods S14) identified based on metadata provided in the HMP16SData R package.[49] Sequences, which spanned the of V3-V5 region of 16S rRNA gene, were downloaded, trimmed to the corresponding 100nt of the V4 region generated in the current study, and processed as described above.

### Bacterial source tracking

To estimate the sources of the microbial communities observed in each of the three infant groups at different body sites and time points, we used FEAST,[29] based on expectation-maximization (EM) estimation, for bacterial source tracking. Samples from each body site in the infants were designated as sinks, and samples collected within the first month after birth from the vagina, stool, skin, mouth and nose of the corresponding mother were tagged as sources. The analysis was performed on rarified count tables with 5000 reads per sample with 1000 EM iterations.

### Alpha diversity

Shannon Diversity was calculated on rarified count tables with 5000 reads/sample using QIIME2 v2019.10.[43] The birth mode effect over time in Shannon diversity was analyzed initially using a linear mixed-effect model and then using a novel longitudinal method, Bayesian Sparse Functional Principal Components Analysis (SFPCA).[22] Generally, a functional principal components analysis is used to investigate longitudinal data with highly non-linear temporal trends,[50] and Bayesian SFPCA performs dimensionality reduction on longitudinal alpha diversity measurements to reveal changes in microbial diversity over time, estimating both mean trajectories, as well as subject-level variation around this mean. Bayesian SFPCA expresses repeated-measurements from each baby in the form of a smooth function that represents the entire time course as a single observation, and then uses a reduced rank mixed-effects framework to handle scenarios where datapoints are collected at irregular and sparse time points. The final inference on birth mode effect is done based on the estimated weights that each baby receives on the first principal component function, which captures the different growth rates in alpha diversity among babies with different birth modes.

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003



### Beta diversity and convex hull

The infant microbiome variability over time of infants for each birth mode were compared to the mothers through a calculation of the per subject convex hull volume from Principal Coordinates derived from unweighted UniFrac distances.[51] The convex hull is the volume produced by considering a set of points to define the outer surface of a shape (like stretching wrapping paper around those points), so a larger convex hull volume indicates greater variability across the set of samples around which the convex hull is built. The unweighted UniFrac distance was calculated on rarefied tables with 1000 reads/sample and dimensionality reduction on the distances was performed through Principal Coordinates Analysis (PCoA) for each body site. Per subject convex hull volume from the individuals' data points in each ordination using scipy v1.3.1.[52] The resulting volumes were assessed by body-site pairwise for statistical significance between mothers and infants of different birth modes with a Bonferroni corrected Mann-Whitney rank test using scipy v1.3.1.[52]

### Dimensionality reduction

We used compositional tensor factorization (CTF) to produce a dimensionality reduction aware of the repeated-measure structure of the longitudinal experimental design and the compositional nature of microbiome data.[23] To reduce sparsity data was split into two time periods being days 0-2 and 7-360 of life and subjects with more than two missing time points were removed. CTF was run on the filtered data through gemelli v0.0.6 (https://github.com/biocore/gemelli) on default parameters. The resulting ordinations first principal component (i.e., Axis 1) was plotted for each sample type across time. To evaluate the statistical significance, a permutational multivariate analysis of variance (PERMANOVA)[53] was performed on the CTF based Aitchison distances[54] separately for each time point and sample type using scikit-bio v0.5.5. Bonferroni p value correction was applied to each comparison with PERMANOVA to correct for multiple comparisons.

### Effect size analysis

In order to calculate the relative effect size of all recorded metadata within the CTF ordinations, a stepwise redundancy analysis was performed (stepwise RDA). The stepwise analysis was performed on the first three principal components of the CTF sample ordination through the ordistep function in vegan v2.4-2.55 The ordiR2-step function was run with 5000 permutation steps, with a permutation *p-value*s limit of 0.1 and otherwise following the procedure of Falony et al.[26]

### Differential abundance and effectiveness score

Differential abundance analysis was performed through Songbird at days 2, 30, and 180 on each infant body site (i.e., feces, mouth, and skin).[25] Optimized model parameters were determined for each model with respect to the main factor of birth mode and the covariate of country by the cross-validation (CV) minimization. The models were then compared by a Q2-value defined as 1 – model CV / baseline CV. A positive Q2-value was observed for all models indicating good predictive accuracy (Methods S16). Differentials were obtained with the reference class as vaginal and contrast variables as CS and CS-seeded, producing the columns log(CS/Vaginal) and log(CS-seeded/Vaginal). The differentials from Songbird in this contrast setup can give a rank of how much each ASV is associated with a contrast variable relative to the reference variable. For example, for log(CS/Vaginal), positive valued ASVs are more associated with a CS birth and negative values with vaginal birth. However, in order to describe the probability of association of a given ASV relative to all three groups (vaginal, CS, CS-seeded), the inverse additive log-ratio (alr) transformation[56]

Please cite this article in press as: Song et al., Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding, Med (2021), https://doi.org/10.1016/j.medj.2021.05.003





was applied to the songbird differentials. We then defined a score for each microbe's "seeding effectiveness" as a measure of efficacy of seeding. We defined the score as

$$seeding\ effectiveness = \beta^{*}_{vaginal} * \beta^{*}_{CS-seeded} + \frac{\beta^{*}_{CS}}{1 + N_{classes}}$$

where the score can range from zero to one, with zero indicating the least effective seeding and one indicating the most effective seeding. Effectively seeded microbes would be those shared by vaginal and CS-seeded infants. Effectively seeded microbes can also be associated with only Caesarean infants, indicating that seeded infants no longer contained that microbe after seeding. Poorly seeded microbes would be those shared between CS and CS-seeded infants. Poorly seeded microbes can also be associated only with Vaginal birth indicating that seeding did not graft that microbe effectively.

The class probabilities were plotted as a ternary plot and colored by their 'seeding effectiveness' score (Figures 1C–1E, 2C–2E, and 3C–3E). The differentials for Vaginal versus Caesarean born infants were plotted at the genus taxonomic level for the top and bottom 20% of ASVs based on the centered songbird differentials and colored by the mean 'seeding effectiveness' score (Figures 1C–1E, 2C–2E, and 3C–3E).

## QUANTIFICATION AND STATISTICAL ANALYSIS

Details on statistical tests, n numbers can be found in the figure legends and further details can be found in the method details for specific measurement.

**Med, Volume *2***

## Supplemental information

# Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding

**Se Jin Song, Jincheng Wang, Cameron Martino, Lingjing Jiang, Wesley K. Thompson, Liat Shenhav, Daniel McDonald, Clarisse Marotz, Paul R. Harris, Caroll D. Hernandez, Nora Henderson, Elizabeth Ackley, Deanna Nardella, Charles Gillihan, Valentina Montacuti, William Schweizer, Melanie Jay, Joan Combellick, Haipeng Sun, Izaskun Garcia-Mantrana, Fernando Gil Raga, Maria Carmen Collado, Juana I. Rivera-Viñas, Maribel Campos-Rivera, Jean F. Ruiz-Calderon, Rob Knight, and Maria Gloria Dominguez-Bello**

# Supplementary Figures

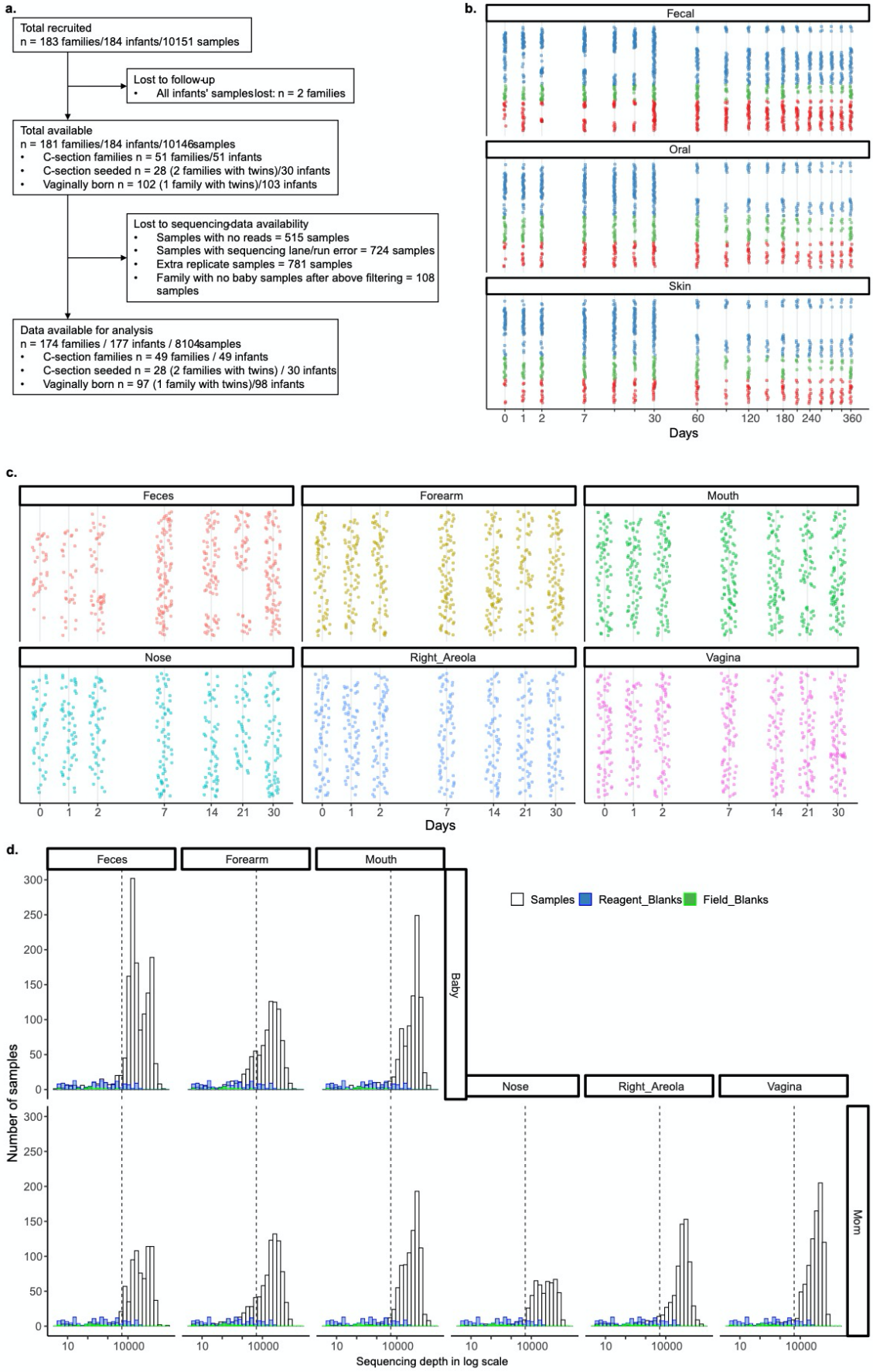



**Figure S1. Longitudinal sampling of mother-infant pairs, Related to STAR Methods.** (a) Number of families, infants and samples from the current study. (b) Longitudinal sampling of infant samples by birth modes and body

sites. Sampling with sterile swabs in different body sites took place within the first hour after birth in all babies, (including the vaginal gauze used to expose CS-born neonates) who were sampled within the hour after birth (after gauze swabbing of the CS-seeded babies), then at day 1-3, weekly for the first month, and monthly for up to the first year. Each row along y-axis is an individual baby. Each point represent a sample for one baby. The points are colored by birth modes, vaginal (blue), cesarean-seeded (green), and cesarean (red). On average, each baby contributed 18, 17, and 21 samples (across three body sites and multiple time points for the first year) for vaginal, cesarean, and cesarean-seed groups. (c) Longitudinal sampling of maternal samples by body sites within the first month after delivery. Each row along y-axis is an individual mom. On average each mom contributed 17 samples (across six body sites and multiple time points for the first month). (d) Distribution of number of reads per sample by different body sites in moms or babies Reagent blanks (blue), and field blanks (green) presentation were overplayed on each panel, and show much lower depth than the samples, indicating good overall quality of the sequences and lack of contamination. Dashed line marked the 5000 reads per sample position.

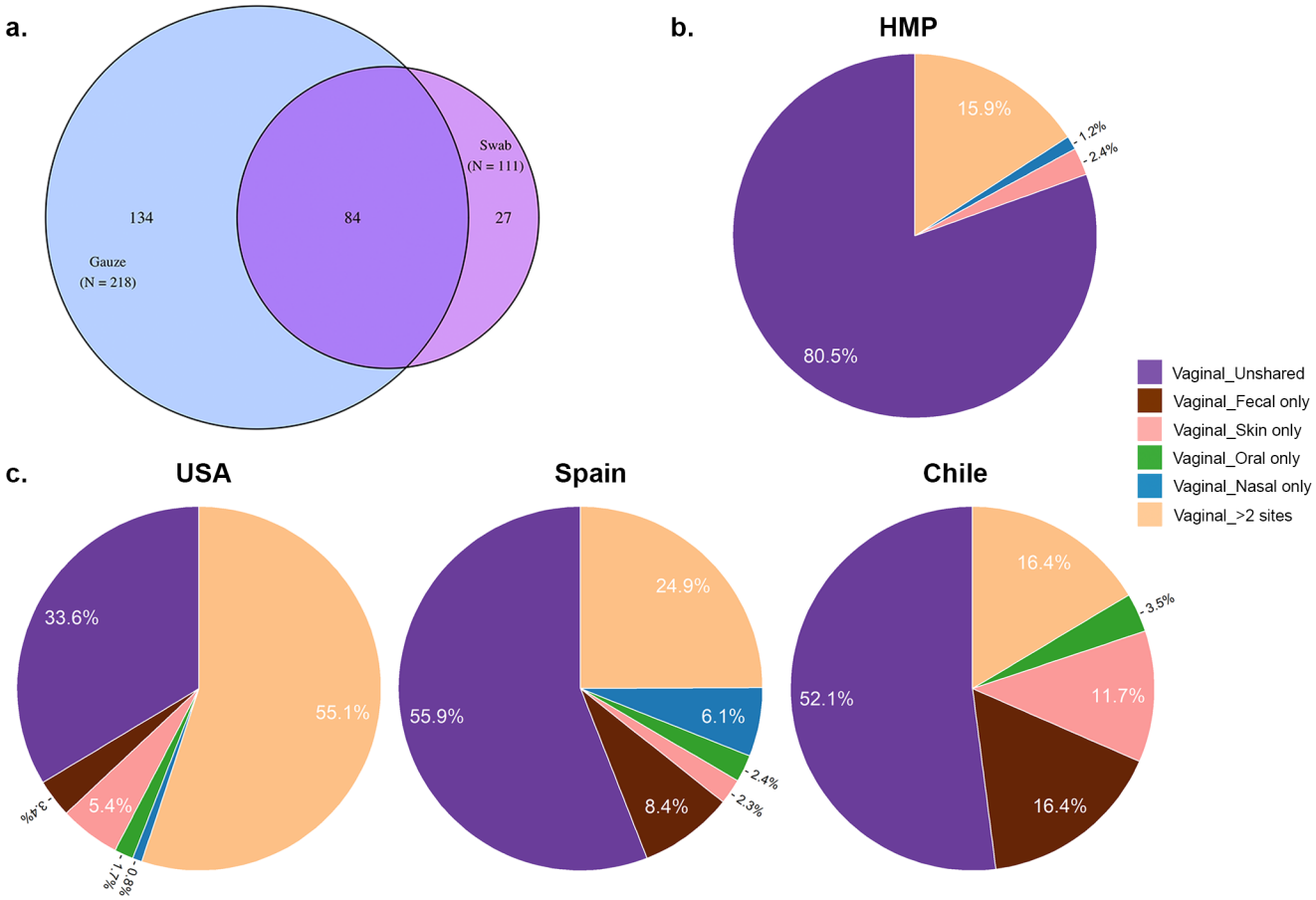



**Figure S2. Pluripotential nature of perinatal vaginal microbiome (related to Figure 5, Supplementary Methods S13-S14).** (a) Number of ASVs shared between perinatal vaginal swabs and vaginal gauzes. Gauzes show higher ASVs richness than vaginal swabs. Proportions of bacterial vaginal ASVs shared with other body sites. (b) in HMP data of non-pregnant women (105 women), (c) in parturient mothers at the day of delivery from USA (53 mothers), Spain (24 mothers) and Chile (20 mothers). HMP data was reprocessed by extracting V4 sequences and analyzed using the same pipeline.

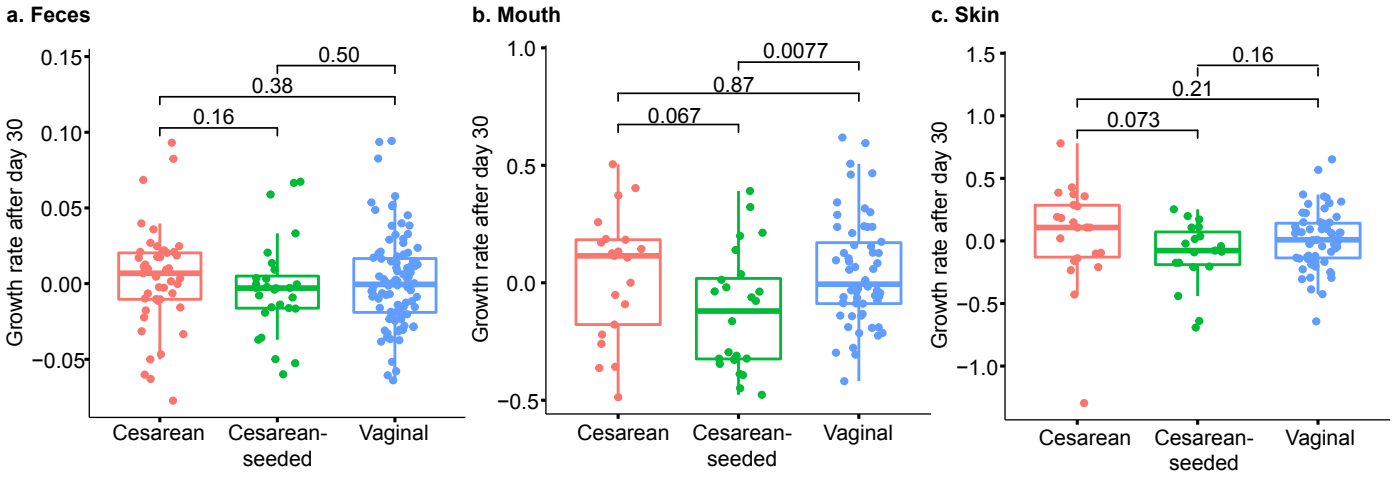



**Figure S3. Bayesian Sparce Functional PCA (SFPCA) analyses on Shannon alpha diversity from 1 to 12 months of age, Related to STAR Methods.** Bayesian Sparse Functional Principal Components Analysis (SFPCA) performed on Shannon alpha diversity across time did not differ by birth mode using Wilcoxon rank-sum test. The rate of growth of the Shannon diversity after day 30 (y-axes) is shown across birth modes (x-axis) for fecal (left), oral (middle), and skin (right) samples.

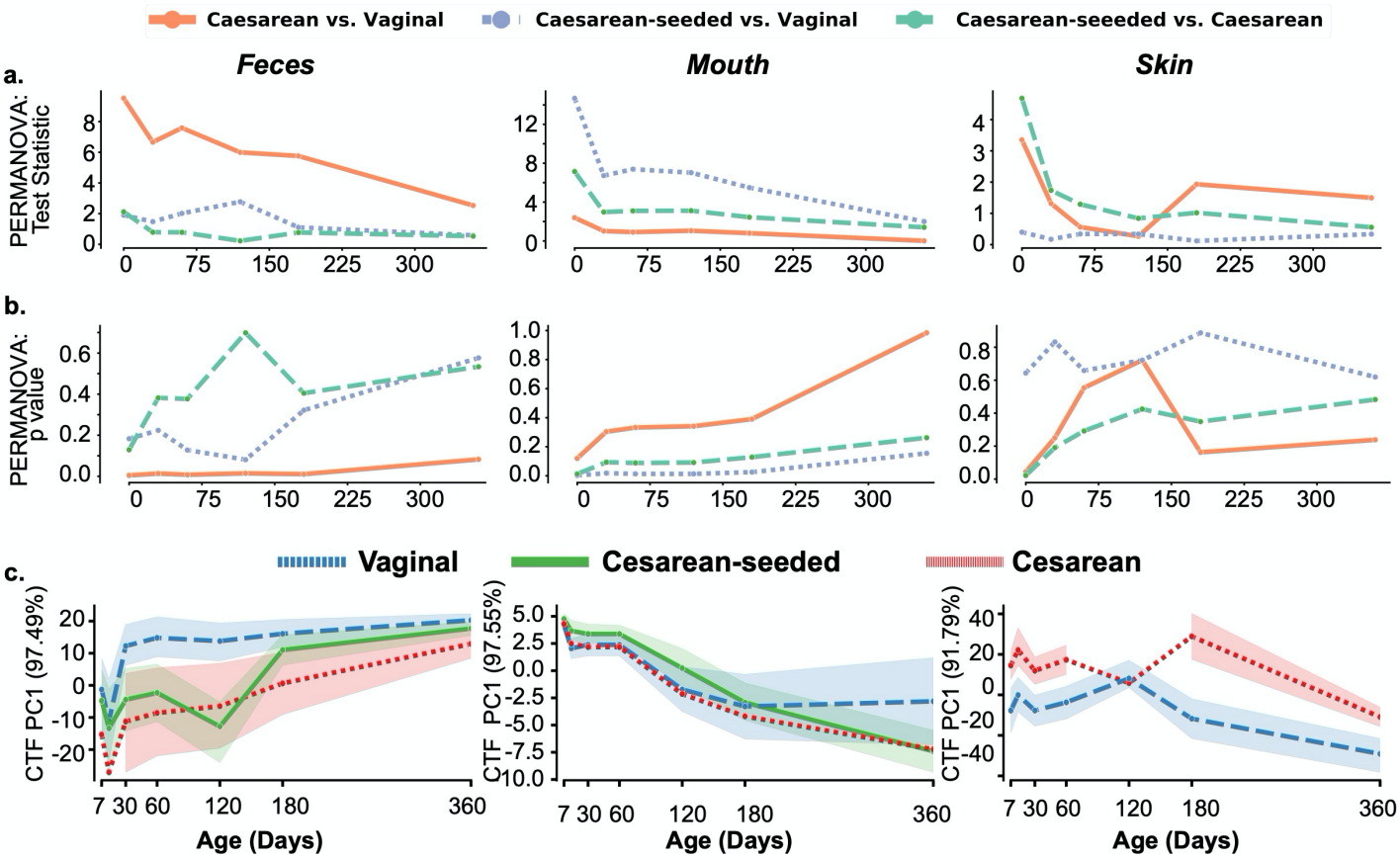



**Figure S4. Compositional Tensor Factorization identifies the partial restoration of microbiome among cesarean-seeded babies, Related to Figure 1,2, 3.** PERMANOVA of Aitchison distances from Compositional Tensor Factorization (CTF) during the first year of life. (a) PERMANOVA test statistics and (b) Bonferroni corrected p-values are plotted across age in days (x-axes). PERMANOVA plots are colored by compared pairs, Caesarean vs. Vaginal, Caesarean-seeded vs. Vaginal, and Caesarean-seeded vs. Caesarean. (c) Compositional Tensor Factorization (CTF) in the USA cohort. CTF ordination plot as in Figure 1a but only with the 101 US infants shows the same trends as the whole dataset, with vaginally born and seeded babies clustering together and separately from Cesarean-born infants. Comparison of Vaginal (blue; n=62), Cesarean (red; n=23), Cesarean-seeded (green; n=16) with CTF first principal component (y-axes) of infant samples over age in days (x-axes); error bars show the standard error of the mean. There were not enough sequences after filtering the samples in the skin of Cesarean-seeded babies.