Multiomic Analyses of Nascent Preterm Infant Microbiomes Differentiation Suggest Opportunities for Targeted Intervention

Stephanie B. Orchanian, Julia M. Gauglitz, Stephen Wandro, Kelly C. Weldon, Megan Doty, Kristina Stillwell, Shalisa Hansen, Lingjing Jiang, Fernando Vargas, Kyung E. Rhee, Julie C. Lumeng, Pieter C. Dorrestein, Rob Knight, Jae H. Kim, Se Jin Song,* and Austin D. Swafford

The first week after birth is a critical time for the establishment of microbial communities for infants. Preterm infants face unique environmental impacts on their newly acquired microbiomes, including increased incidence of cesarean section delivery and exposure to antibiotics as well as delayed enteral feeding and reduced human interaction during their intensive care unit stay. Using contextualized paired metabolomics and 16S sequencing data, the development of the gut, skin, and oral microbiomes of infants is profiled daily for the first week after birth, and it is found that the skin microbiome appears robust to early life perturbation, while direct exposure of infants to antibiotics, rather than presumed maternal transmission, delays microbiome development and prevents the early differentiation based on body site regardless of delivery mode. Metabolomic analyses identify the development of all gut metabolomes of preterm infants toward full-term infant profiles, but a significant increase of primary bile acid metabolism only in the non-antibiotic treated vaginally birthed late preterm infants. This study provides a framework for future multiomic, multibody site analyses on these high-risk preterm infant populations and suggests opportunities for monitoring and intervention, with infant antibiotic exposure as the primary driver of delays in microbiome development.

1. Introduction

It is now well-established that the human microbiome, the community of microbes that inhabit nearly every organ in the human body[1,2] but particularly the skin, mouth, and gut,[3,4] is linked to numerous facets of human health and disease including inflammatory bowel disease, asthma, obesity, dental caries, atopic dermatitis, type 1 and 2 diabetes, cancer, and even mental health or neurological conditions such as schizophrenia and Parkinson's disease.[3,5,6] Diet, geography, medication, hygiene, housing, and physical activity each shape the structure and diversity of the community, but the earliest effects that create our baseline microbiome begin at birth.[3,7,8]

Birthmode, antibiotic exposure, gestational age, and feeding are each known to impact the development of the infant gut microbiome through extensive

S. B. Orchanian,[+] J. M. Gauglitz, S. Wandro, K. C. Weldon, S. Hansen, K. E. Rhee, P. C. Dorrestein, R. Knight, S. J. Song, A. D. Swafford
Center for Microbiome Innovation
University of California San Diego
La Jolla, CA 92093, USA
E-mail: sjsong@eng.ucsd.edu

J. M. Gauglitz, K. C. Weldon, F. Vargas, P. C. Dorrestein
Skaggs School of Pharmacy and Pharmaceutical Sciences
University of California San Diego
La Jolla, CA 92093, USA

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adbi.202101313.

[+]Present address: Department of Molecular Biology and Biochemistry, University of California Irvine, Irvine, CA 92697, USA

[++]Present address: Neonatal Intensive Care Unit, Kapi'olani Medical Center for Women & Children, Honolulu, HI 96826, USA

[+++]Present address: Department of Pediatrics, University of California San Francisco, San Francisco, CA 94143, USA

[++++]Present address: Perinatal Institute, Department of Pediatrics, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, Cincinnati, OH 45267, USA

J. M. Gauglitz, K. C. Weldon, F. Vargas, P. C. Dorrestein
Collaborative Mass Spectrometry Innovation Center
University of California San Diego
La Jolla, CA 92093, USA

M. Doty,[++] K. Stillwell,[+++] K. E. Rhee, P. C. Dorrestein, R. Knight, J. H. Kim[++++]
Department of Pediatrics
School of Medicine
University of California San Diego
La Jolla, CA 92093, USA

L. Jiang
Division of Biostatistics
University of California San Diego
La Jolla, CA 92093, USA

F. Vargas
Division of Biological Sciences
University of California San Diego
La Jolla, CA 92093, USA

J. C. Lumeng
Department of Pediatrics
Medical School
University of Michigan
Ann Arbor, MI 48109, USA

characterization by 16S rRNA gene amplicon sequencing.[9–15] Infants born via Cesarean section are initially colonized by skin microbiota from their mothers and other adults, whereas infants born via vaginal delivery primarily receive a mix of vaginal and stool microbes.[12,16,17] This difference is reported to delay gut colonization by common adult-associated microbes,[12,18,19] and is associated with adverse health effects,[20–22] but the precise mechanisms remain unclear. Gestational age at birth also profoundly impacts development of the infant gastrointestinal tract, and is associated with additional short and long-term health outcomes,[23,24] perhaps partly attributable to altered gut microbiome development.[21,25] However, studies of gestational age are often confounded by birthmode and antibiotic exposure because lower gestational age is associated with birth by cesarean section and high antibiotic use.[26–28] Studies of the impact of antibiotics rarely examine proximal effects, and there is likely considerable under-reporting of antibiotic exposure in infants born via vaginal delivery,[29,30] and over-reporting of antibiotic exposure in infants born via Cesarean section, because studies typically assume infants born via Cesarean section are either exposed to antibiotics via their universally antibiotic-treated mothers or otherwise affected by antibiotic effects (e.g., alteration in breast milk microbiota or metabolite composition of breast milk).[31,32]

After birth, infants are exposed to their first forms of food: colostrum, breast milk (either from their mother or a donor), and/or formula.[18,33] Breast milk is a complex mixture including micronutrients, macronutrients, and bioactives, such as human milk oligosaccharides (HMOs), along with immune factors and breast milk microbiota, which change in composition with infant age and differ between mothers based on maternal characteristics.[34–36] HMOs, a group of diverse compounds unique to each mother,[37] cannot be metabolized by human cells but are converted into other metabolites including short-chain fatty acids (SCFA), a primary energy source, by gut microbes.[38] Although formulas contain HMOs, these HMOs are less complex than in breast milk; formula also lacks diverse immune factors found in breast milk.[39] Different types of infant diet and feeding practices, including differences between direct and indirect breast-feeding,[40] have been shown to correlate with the gut microbiome and later health implications.[33,41–44] However, the reported relationships between the type of milk (breast milk versus formula) an infant receives and their microbiomes vary among studies. Over larger time scales, taxonomic groups change based on diet such that gut samples in a longitudinal study were more similar to each other in the first sample than the final sample collected at age 2.5 years,[45] with developmental milestones such as the introduction of solid foods or the weaning off of breast milk resulting in a specific decrease in *Bifidobacteria,* the bacteria that reduce the conversion of HMOs to SCFA.[46]

Premature birth correlates with disparities in the metabolomic and microbial compositions in infant microbiomes,[47–49] with delayed colonization of "traditional" bacteria in very-low birth weight preterm (VLBW) infants.[50] Establishment of the microbiome in the first days to weeks after birth can have important implications later in life, due to the far-reaching health associations. Infants born late preterm (LP) and with VLBW are at higher risk for infant mortality, sepsis, hypoglycemia, feeding difficulties, and long-term neurodevelopmental impairment[47,48,51,52] compared to infants born at full-term (FT), yet the mechanisms for these risks are poorly understood. Their higher incidence of delivery by cesarean section, routine administration of antibiotics to both mothers and preterm infants and frequent use of formula places them at high risk for developing an abnormal or delayed gut microbial community.[19,33,45] Furthermore, although differences in the microbiome have been observed in the gut communities of FT compared to those of preterm or VLBW infants,[17,53,54] few studies have examined the development and diversification of distinct, body-site specific (e.g., skin, oral, nares, and stool) infant microbiomes, and only on the order of weeks and months.[12,15,54] Distinct oral and gut communities were not observed until 15 d after birth in a small cohort of six preterm infants (five without antibiotic exposure), or as late as six weeks in a larger cohort of 162 FT infants,[15] while disparities between the gut, oral, and skin microbial communities and metabolomes between preterm and VLBW infants in the neonatal intensive care unit (NICU) remains largely unexamined.

To provide additional resolution and further this understanding, we conducted 16S rRNA gene amplicon sequencing and untargeted metabolomic analyses on gut, oral, and skin samples collected daily for the first week after birth from LP infants born either via cesarean section or vaginally in San Diego and who were fed mother's milk supplemented with either donor human milk (DHM) or formula. We also examined the impact of birthweight in shaping the development of the microbial community in VLBW infants born via Cesarean section who were housed in the NICU during this same period of collection. Together these cohorts enabled us to characterize the effect of early life exposures, including delivery mode, antibiotics and type of milk on the initial colonization and differentiation of their gut, oral, and skin microbiota and metabolomes during this critical window of holobiont development. We show that although the microbiome at each body site changed over the first week of life, robust differentiation among these microbiomes was only observed in non-antibiotic exposed late preterm infants regardless of delivery mode. In addition, we find that the metabolomic profile of each body site is dominated initially by host chemistry, but microbial influence begins to appear within the first week after birth, with a stronger impact in non-antibiotic exposed infants.

R. Knight
Department of Computer Science and Engineering
University of California San Diego
La Jolla, CA 92093, USA

R. Knight
Department of Bioengineering
University of California San Diego
La Jolla, CA 92093, USA

2. Results

2.1. Study Groups and Exclusion Criteria

Eighty-five preterm infants (60 LP and 25 VLBW) were recruited from the UC San Diego NICU for approximately daily sampling of the oral, skin, and feces during the duration

of their stay, resulting in a total of 1799 samples (**Figure** 1A). Untargeted metabolomics data were generated for 1682 samples and 16S rRNA gene amplicon sequencing was performed on 1654 samples, with paired data generated for 1547 samples. For this investigation, we focused on changes in the microbial community and metabolome as a function of birth weight, delivery mode and antibiotic use in the first 7 d after birth, resulting in the exclusion of 482 16S and 495 metabolomic samples. Insufficient samples were available for examining VLBW infants born via cesarean section without antibiotic administration, or born vaginally over time resulting in the exclusion of an additional 100 16S and 107 metabolomic samples. Finally, 130 samples with fewer than 500 16S sequencing reads were excluded (Figure 1A, **Table 1**). A total of 1080 samples remained for metabolomic analyses and 942 samples remained for 16S analyses from 75 infants, consisting of three cohorts: 29 late preterm infants born via cesarean section (LP-C), 28 late preterm infants born vaginally (LP-V), and 18 very low birth weight preterm infants born via cesarean section (VLBW-C) (Figure 1B) with samples available from at least 30 infants across all three cohorts for each body site for each day of sampling (Figure S1, Supporting Information).

2.2. Determination of Significant Drivers of Microbial Community Separation

Stepwise redundancy analysis (RDA)[55] of robust Aitchison distances between all samples from the 16S sequencing data,

which account for the compositional nature of these data,[56] revealed that as expected, individual had the strongest impact on microbial community composition (stepwise ANOVA, $R^2 = 0.196$, $F = 4.8$, $p = 0.002$) (Table S1, Supporting Information). To better parse out the more subtle effects on composition, we removed the individual from the model and re-ran the stepwise redundancy analysis. Body site had the next strongest impact on the microbial community (stepwise ANOVA, $R^2 = 0.170$, $F = 97.14$, $p = 0.002$), followed by antibiotic exposure ($R^2 = 0.095$, $F = 121.59$, $p = 0.002$), age (day after birth) ($R^2 = 0.043$, $F = 59.16$, $p = 0.002$), birth weight ($R^2 = 0.016$, $F = 23.04$, $p = 0.002$), and delivery mode ($R^2 = 0.014$, $F = 20.51$, $p = 0.002$). Using stepwise RDA within body sites (again removing individual from the formula) revealed significant effects for all primary variables of interest. In stool, the strongest effect on the microbial community was delivery mode ($R^2 = 0.141$, $F = 50.05$, $p = 0.002$) followed by significant effects from antibiotic exposure, age, and delivery weight (Table S1, Supporting Information). LP infants were also randomized into two groups that received either supplemental donor breast milk or supplemental formula (Table 1, Experimental Section). This randomization group had a small, but significant effect on the stool microbial community ($R^2 = 0.013$, $F = 3.69$, $p = 0.002$). Within the oral community, antibiotic exposure had the strongest effect ($R^2 = 0.249$, $F = 104.35$, $p = 0.002$) followed by age, birth weight, and delivery mode (Table S1, Supporting Information). For skin, age had the strongest effect ($R^2 = 0.114$, $F = 43.11$, $p = 0.002$), followed by birth weight and delivery mode (Table S1, Supporting Information).
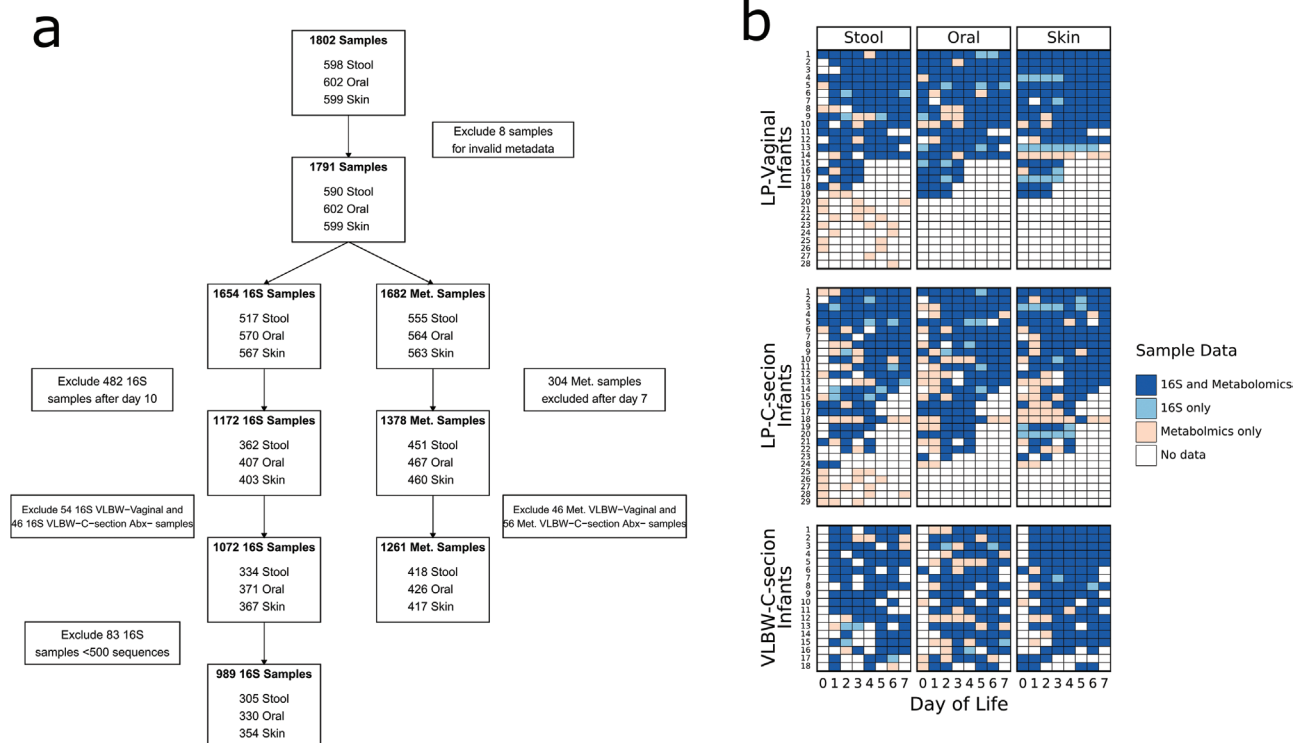

Figure 1. Sample counts and workflow diagram for 16S rRNA gene amplicon and untargeted metabolomics data. a) Sample filtering steps displaying the remaining samples after each filtering criterion. b) Sample data density for each infant over the first week after birth for stool, oral, and skin. Each column represents one infant and fill color indicates presence of 16S and/or metabolomics data. Infants divided by birth mode and birth weight.

Table 1. Demographics of preterm infant cohorts. Group significance determined by chi-square test when appropriate.

	LP-Vaginal	LP-C-section	VLBW-C-section	Group significance
Number of infants	28	29	18	
Birth gestational age (weeks)	35.0 (0.86)	35.0 (0.85)	28.4 (2.06)	1.45×10^{-29}
Birth weight (grams)	2548.8 (471.49)	2555.2 (557.49)	1079.2 (280.3)	4.22×10^{-17}
Maternal age (years)	32.9 (5.27)	32.8 (5.69)	31.5 (7.78)	n.s.
% Female	39.3	41.4	55.6	n.s.
Ethnicity				
White	28.6	44.8	44.4	n.s.
Hispanic	42.9	31	33.3	n.s.
African American	3.6	17.2	22.2	n.s.
Other	25	6.9	0	0.013124
Randomization				
Donor breast milk	15	15	n/a	
Formula	13	14	n/a	
16S Samples				
Stool n	101	119	85	
% antibiotics exposed	31.7	20.2	100	
Oral n	117	126	87	
% antibiotics exposed	29.9	20.6	100	
Skin n	116	136	102	
% antibiotics exposed	25	20.6	100	
Metabolomics samples				
Stool n	131	140	89	
% antibiotics exposed	29	20	100	
Oral n	114	142	109	
% antibiotics exposed	29.8	21.1	100	
Skin n	108	139	108	
% antibiotics exposed	32.4	20.1	100	

2.3. Alpha and Beta Diversity Changes over the First Week after Birth

The microbiome of each infant is influenced by their initial exposures outside the womb, and our multibody-site analyses enabled us to assess the differentiation of each body site over time in our cohorts. As previous reports have detailed the impact of antibiotics and delivery mode on the infant gut community,[12,54,57] we subdivided individuals into five groups for comparison: LP-V/Abx$^{(+/-)}$, LP-C/Abx$^{(+/-)}$, and VLBW-C/Abx$^{(+)}$; too few VLBW-C subjects were not administered antibiotics to enable adequate comparisons in the VLBW cohort. No differences in alpha diversity of the microbes (Shannon index) were found across any of the groups over time, with a contraction in alpha diversity of all microbiomes to a similarly-restricted community over the first week after birth (Figure S2, Supporting Information). We then compared changes within each body site by comparing robust Aitchison distances over time to the earliest available samples (day 0 for LP infants and day 1 for VLBW infants) within each infant. Skin microbiomes changed significantly in all cohorts over time (linear mixed effects models, $p = 0.001$ for LP-V; $p = 0.003$ for LP-C; and $p = 0.032$ for VLBW-C). Oral microbial communities significantly changed in all cohorts (linear mixed effects models $p = 0.001$ for LP-V; $p = 0.002$ for LP-C; $p = 0.002$ for VLBW-C). The stool microbiome was only significantly changed in the LP-C and VLBW-C groups (linear mixed effects models $p = 0.0001$
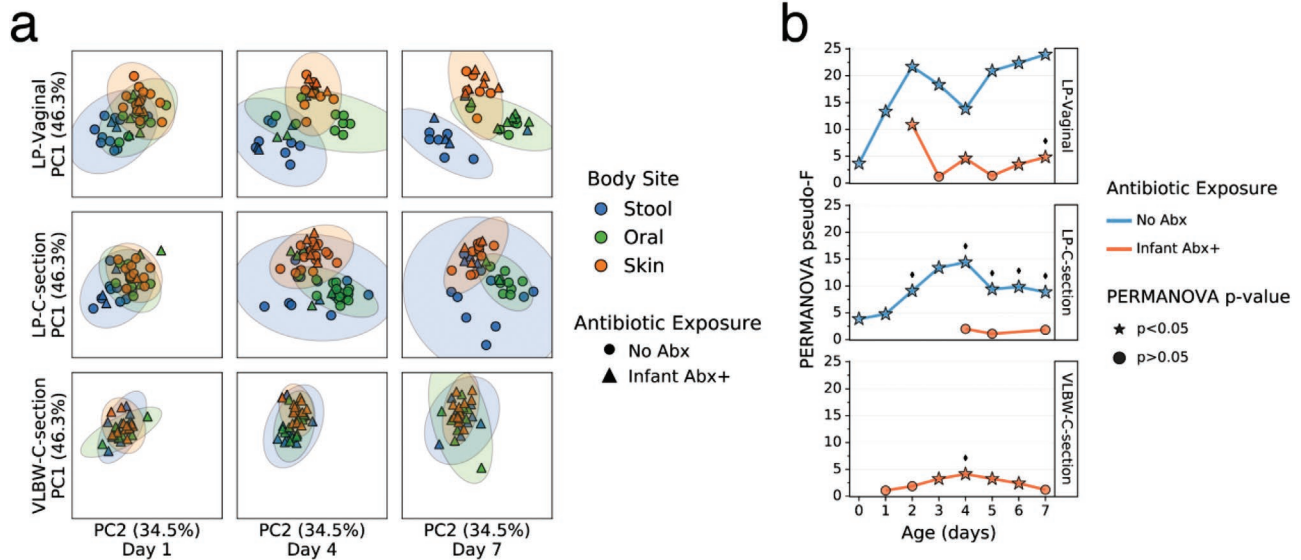
Figure 2. Impact of antibiotics exposure on the differentiation of the microbial community of preterm infants over the first week after birth. a) Robust Aitchison principal components analysis (PCA) based on 16S V4 amplicon sequence variants (ASVs) of the gut (blue), oral (green), and skin (orange) microbiomes collected at day 1, 4, and 7 after birth from late preterm infants (LPs) born vaginally (LP-Vaginal) or by cesarean section (LP-C-section) and very-low birth weight (VLBW) preterm infants (VLBW-C-section), with samples from children exposed to antibiotics as reported by clinical metadata (Infant Abx$^{(+)}$; triangles) and those without exposure to antibiotics (No Abx$^{(-)}$; circles) differentiated. b) PERMANOVA for differences in sample type (stool, oral, skin) based on Aitchison distances between all within-group samples for LP-Vaginal, LP-C-section, and VLBW-C-section preterm infants daily over the first week after birth, separated by Abx exposure. Diamonds above points indicate significant Permdisp tests ($p < 0.05$) which indicate significantly different dispersion among body sites. Refer to Table S3 (Supporting Information) for sample sizes in each comparison.

for LP-C; $p = 0.025$ for VLBW-C), however, the LP-V infant stool at day 1 already had relatively high distance from day 0 samples.

2.4. Changes in Infant Microbiome Differentiation

As clear changes occurred in most body sites and groups relative to their origins, we tested whether these changes were reflective of niche differentiation by body site. Robust Aitchison principal components analysis (PCA) revealed a pattern wherein the microbiomes were significantly differentiated at 1 d since birth in both LP/Abx$^{(-)}$ groups (V and C), but not in the VLBW-C/Abx$^{(+)}$ group (**Figure 2**a,b). Over the first week after birth, differences emerged in the degree of separation among the three body sites, as assessed by PERMANOVA pseudo-F score. Increased body site separation was seen only in the Abx$^{(-)}$ groups while the Abx$^{(+)}$ groups were not significantly differentiated or were significantly differentiated at fewer time points with a low pseudo-F score (Figure 2b). Classifying samples by maternal exposure to antibiotics showed a stark difference between infants receiving antibiotics compared to those assumed to be exposed via maternal antibiotic administration perinatally (Figure S4, Supporting Information). However, stratification of the data after applying the principle of metabolome-informed microbiome analyses[58] to provide empirical evidence of maternal antibiotic transfer demonstrated a pattern most similar to the classification of direct infant exposure only (Figure S5, Supporting Information).

2.5. Comparison of Infant Microbiomes to Adult Microbiomes

As clear differentiation occurred over the first week after birth in the Abx$^{(-)}$ groups, we evaluated the proportion of each microbiome in our analytical groups that could be attributed to their more mature equivalents and potential sources. We trained a Bayesian microbial-source tracking algorithm on 16S data from the gut ($n = 4434$), oral ($n = 2550$), skin ($n = 1975$), and vaginal ($n = 427$) communities from 11 public studies of adults age 20–80, stool samples from a cohort of FT infants age 0.5–4 months ($n = 87$) (list of studies and details in Table S2, Supporting Information[3,59–63]) and determined the percentages of each community that could be assigned to the expected body site or source of origin. Approximately half of the skin microbial community in the LP-V and VLBW-C groups was attributable to adult skin from the first day after birth, while this fraction exceeded 75% in the LP-C groups (**Figure 3**, right columns). By 6 d after birth, this proportion increased to nearly 100%, after an initial period of attribution to the vaginal microbiome in LP-V groups. The oral microbiome of all infants except LP-V/Abx$^{(-)}$ was primarily attributable to skin for all infant samples for the first 3 d since birth, however by day 4, Abx$^{(-)}$ infant samples were almost entirely attributable to the adult oral community while the Abx$^{(+)}$ infant samples remained primarily attributable to skin (Figure 3, center columns) with a trend towards increased oral attribution near the end of sampling for the LP-V/Abx$^{(+)}$ group. The brief dominance of oral attribution at day 4 in LP-C/Abx$^{(+)}$ was driven by a large relative abundance of *Streptococcus* in a minority of samples at that time point (Figure 3b, center panel). Attribution to adult stool was
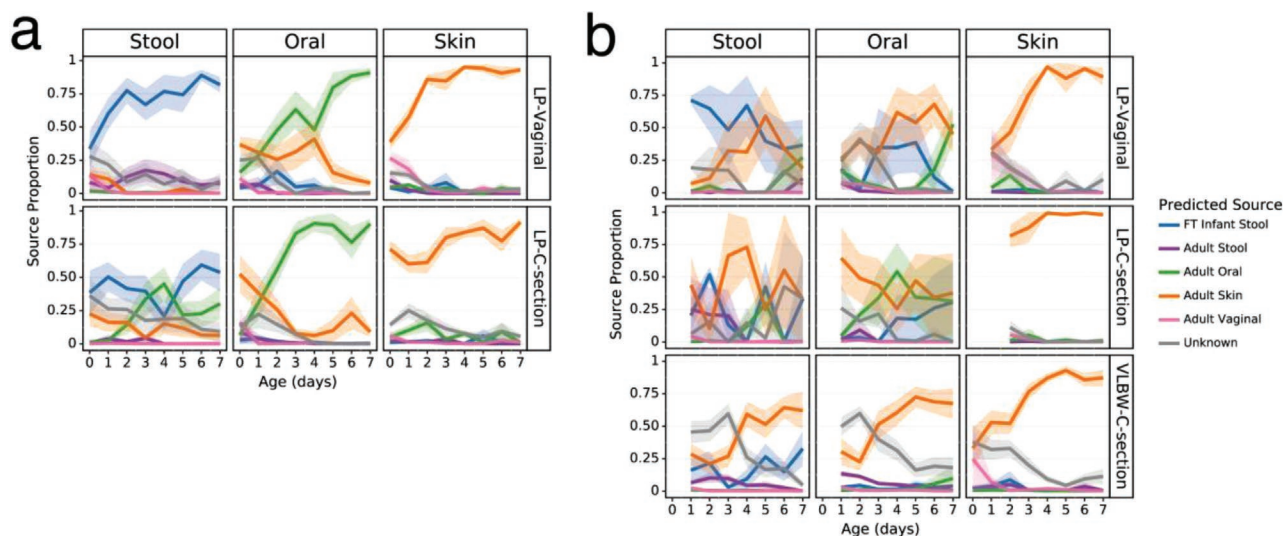
**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
BIOLOGY**

www.advanced-bio.com

Figure 3. Antibiotics and delivery mode alter maturation of microbial communities over the first week after birth. SourceTracker2 analysis showing proportion of LP and VLBW infant microbial communities attributed to adult or full term (FT) infant microbiomes based on 11 public studies in Qiita and a cohort of 87 FT infants not exposed to antibiotics aged 0.5–4 months (Experimental Section). Lines represent the mean (+/− SEM) proportion of the microbial profile attributed to a specific source for all samples from preterm infants within each cohort displayed by birth-mode (rows) and body site (column). a) Preterm infant samples not exposed to antibiotics. b) Preterm infant samples after exposure to antibiotics. Refer to Table S3 (Supporting Information) for sample sizes of the lP and VLBW infants at each time point.

highest in the gut microbiome of LP-V/Abx$^{(-)}$ samples with a slight increase over time, though never exceeding 25% attribution (Figure 3a). Instead, the gut microbiome of these infants rapidly increased to resemble the gut of FT non-antibiotic exposed infants born vaginally (FT-V/Abx$^{(-)}$) with delayed and reduced pattern of increasing attribution observed in LP-C/Abx$^{(-)}$ samples (Figure 3a, left panels). In contrast, while the gut microbiome of LP-V/Abx$^{(+)}$ samples initially resembled FT-V/Abx$^{(-)}$ stool, it quickly declined and the microbiome of all three groups of Abx$^{(+)}$ samples was <50% attributable to FT-V/Abx$^{(-)}$ stool at day 7 (Figure 3b, upper left panel). Grouping samples based on maternal, rather than infant antibiotic administration, obscured these differences and a universal pattern of increasing attribution to the expected body site for all infant groups (Figure S6, Supporting Information).

2.6. Differentially Abundant Microbes due to Age, Antibiotics, and Breast Milk Supplementation

To identify specific changes in microbial abundances associated with differentiation and development, we used Songbird,[64] a tool that overcomes the compositionality of microbiome data by ranking features based on log-fold changes with respect to variables of interest. We constructed models for each body site separately using only non-antibiotic exposed LP infants using age, delivery mode, and randomization (i.e., breast milk or formula supplementation) as explanatory variables (**Figure 4**). Each of these models resulted in a $Q^2 > 0$, indicating that the model including the variable attained higher predictive accuracy than the baseline model not including the variable. *Bifidobacteria* in stool and *Rothia* in oral samples were most highly associated with increased age, while the common urinary tract genus *Ureaplasma*, was strongly associated with oral and stool samples

from younger infants. The genera *Aggregatibacter* and *Serratia* were most associated with formula supplementation and *Staphylococcus* was most associated with donor human milk supplementation (Figure S7, Supporting Information) in stool. In contrast, no microbial taxa were found to be highly associated with these factors in skin (Songbird $Q^2 > 0$).

To identify differentially abundant microbes associated with infant antibiotic use, separate models were constructed by body site including both the antibiotic-exposed LP and VLBW infants with antibiotic exposure as the explanatory variable (Figure 4c). In stool, *Corynebacterium* and an unidentified genus in the class Bacilli were most associated with antibiotic exposure, while the genera *Serratia, Neisseria, Actinobacullus,* and *Bifidobacterium* were most associated with non-antibiotic exposed samples. In skin samples, *Ureaplasma* and the common vaginal microbe *Gardnerella* were most associated with antibiotic exposure.

2.7. Metabolomic Changes in the First Week after Birth

Given the major microbiome changes observed during the first week after birth, we sought to examine whether there was evidence of these changes in the metabolomic profiles of each body site within each infant for our cohorts. PCoA of the composition of unique metabolites, assessed by Jaccard distance, revealed that each body site within each cohort was significantly different throughout the first week after birth as expected for unique host compartments (Figure S8, Supporting Information). However, when comparing Jaccard distances over time to the earliest available samples (day 0 for LP infants and day 1 for VLBW infants) within each infant, a different pattern emerged. While changes were not significantly different in any cohort for the skin profile (linear mixed model $p > 0.05$) the profile of oral samples was significantly different with time
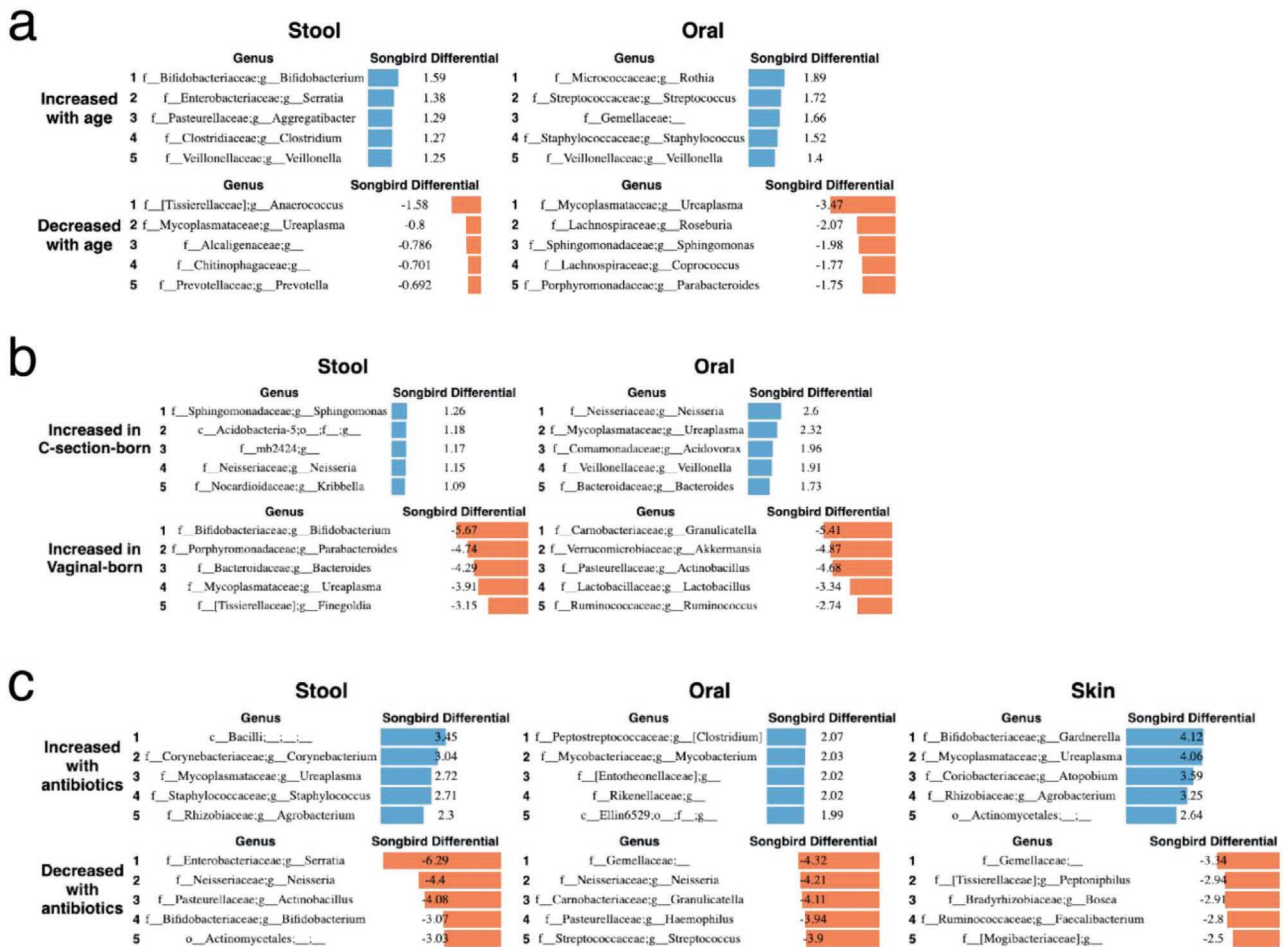
Figure 4. Top differentially abundant microbes based on Songbird differentials. Top 5 genera positively and negatively associated with a) age and b) delivery mode in non-antibiotic-exposed samples from LP infants in stool and oral microbial communities based on 16S sequencing. The songbird model fit for stool ($Q^2 = 0.072$) and oral ($Q^2 = 0.014$). Top 5 genera positively and negatively associated with c) antibiotics based on stool, oral, and skin samples from LP and VLBW infants based on 16S sequencing. The Songbird model fit for stool ($Q^2 = 0.222$), oral ($Q^2 = 0.316$), and skin ($Q^2 = 0.038$). Results for skin are not shown in (a) and (b) because the Songbird model was unable to fit to data. Songbird differential values describe the log-fold change of genera with respect to the metadata category given.

for the LP-C cohort (linear mixed model $p = 0.001$; **Figure 5**a). In contrast, the stool metabolome of each infant changed significantly over the first week after birth in all cohorts (linear mixed model $p = 0.003$ for LP-V, $p = 0.018$ for LP-C, and $p = 0.001$ for VLBW-C; Figure 4a). Differential abundance analysis via Songbird revealed that several metabolites increased with time including diet-related metabolites such as oleic acid and conjugated linoleic acid, while several primary conjugated bile acids, such as taurocholic acid and glycocholic acid, decreased over time (Figure 5b).

2.8. Bile Acid Changes in Stool Metabolome

We further investigated the changes in the log ratio of the relative abundances of non-microbially modified primary bile acids versus secondary, i.e., microbially modified, bile acids in infant stool samples over time. As primary bile acids are transformed into secondary bile acids through microbial metabolism, the

ratio of these should give a measure of changes in microbial metabolism in our groups. We observed a significant change in the ratio of secondary to primary bile acids in stool samples over the first week after birth only in the LP-V/Abx$^{(-)}$ group (linear mixed effects, $p = 0.001$, slope = 0.188; Figure 5c). The LP-C infants show a similar increasing trend, but had a high ratio at day 0 for unknown reasons. The log-ratio for LP-C/Abx$^{(+)}$ infants mostly remained constant, while the LP-V/Abx$^{(+)}$ infants showed an increasing trend, but the model was not significant due to low sampling density and higher variance. This data supports the idea that in the absence of antibiotic exposure, the gut microbiome of LP infants differentiates and gut-specific microbial metabolism occurs within the first week after birth.

2.9. Microbe-Associated Changes in Stool Metabolome

Given our observation of increasing microbial differentiation by body site over the first week after birth, we hypothesized that
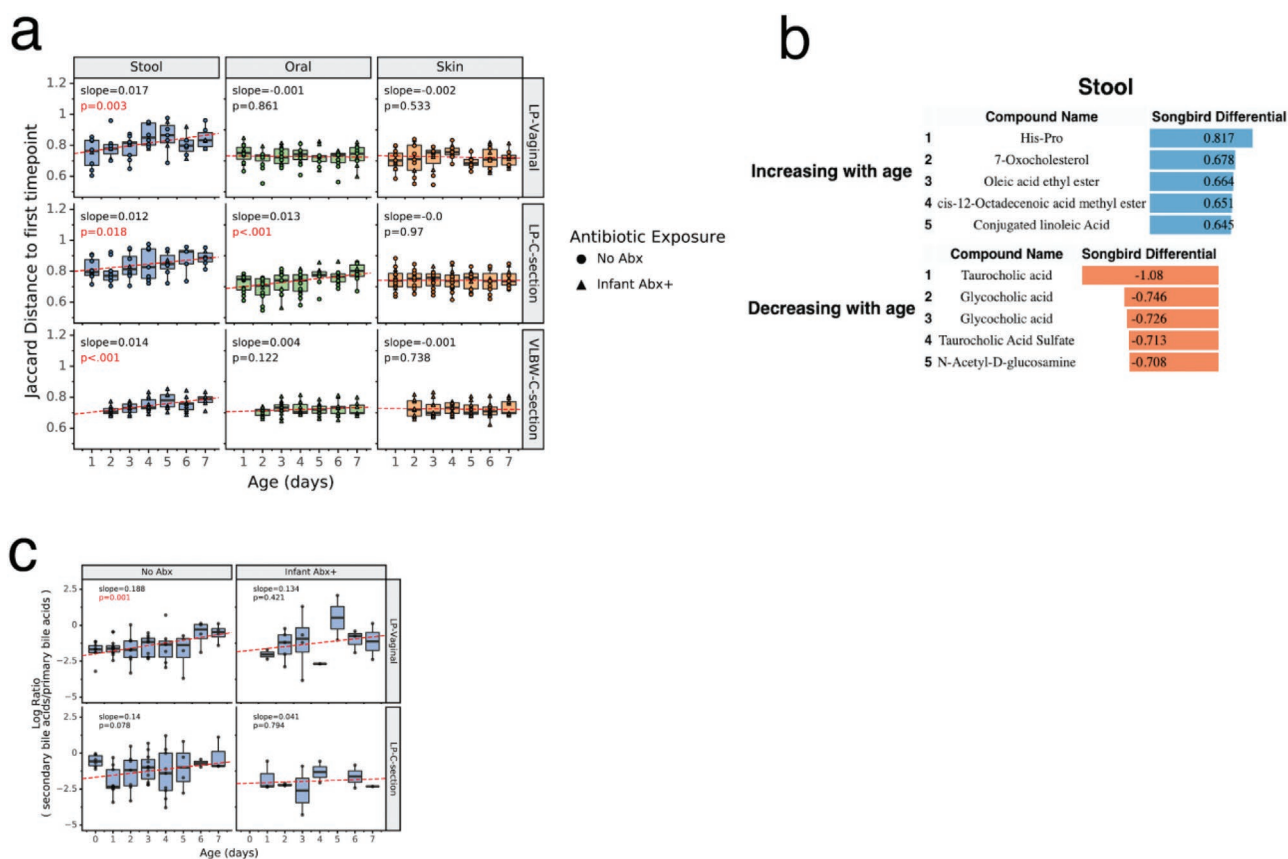
Figure 5. Metabolomics profiles change over the first week after birth. a) Within-infant metabolome distance to first sample over time. For each infant, the Jaccard distance to the first sample (day 0 for LP infants, day 1 for VLBW infants) is plotted over time for each body site. Boxplots show median and interquartile range with whiskers extending to the furthest value within 1.5 times the edge of the interquartile range. Red lines show the fitted linear mixed effect model with individual as a random effect and age as a fixed effect. *P*-value and slope of the age variable shown. Stool samples for all cohorts (LP-Vaginal, LP-C-section, VLBW C-section) and oral samples for LP-C-section infants have a significant increasing slope, corresponding to increased distance from the first time point over the first week after birth. All data points are shown with shape indicating infant antibiotic exposure. b) Top 5 identified metabolites positively and negatively associated with age based on Songbird model including LP infants. Songbird model fit ($Q^2 = 0.017$. c) Log ratios of secondary bile acids over primary bile acids, identified using GNPS annotations to structural family matches. Linear mixed effects models were run in the same way as panel (a). Refer to Table S3 (Supporting Information) for sample sizes in each comparison.

microbes would have an increasing effect on the infant metabolome over that time. We compared the composition of matched 16S and metabolomic samples from LP infants using Mantel tests with data divided by infant day after birth, antibiotic exposure, and delivery mode, each observed to be major factors influencing the microbial profile. Mantel tests for the first 4 d after birth were largely nonsignificant, however, days 5, 6, and 7 in LP-V/Abx[−] infants were significant with an increasing trend in the Mantel correlation. Furthermore, the microbial community and metabolome were significantly correlated in all LP groups except LP-C/Abx[+] infants (Figure S9, Supporting Information).

3. Conclusion

Research regarding the development of the human microbiome has increased dramatically in the last decade, and will increase as researchers and clinicians identify new methods of intervention and establish best practices during critical early windows of microbial community niche differentiation. LP and VLBW infants represent two of the most at-risk populations for adverse impacts of early life exposures, and even a brief NICU stay may increase these risks via inadequate exposure to the necessary milieu of microbes, exposure to antibiotics, and/or delays in feeding.[65] Early life provides a unique opportunity for potential microbiome-directed interventions and for observing microbiomes with few exposures compared to adults,[8,21,45] although key factors in early development such as antibiotic use and diet play a significant role for the gut microbiome regardless of age.[3]

Here, by implementing daily sampling, and a multi-omic, multi-body site analysis in the context of publicly available human data from infants and adults in Qiita[66] and publicly available information in GNPS,[67] we were able to examine the impact of these factors on the developing microbiomes of preterm infants in the first week after birth. The skin microbiome in all infants appeared to develop a resemblance to adults regardless of exposure or delivery mode, suggesting that at least one niche may be resilient to these effects. The lack of

oral versus stool microbiome differentiation and development in the VLBW infants, who are typically both born via Cesarean section and administered antibiotics, is of concern, and suggests that these infants may be able to benefit most from any interventions, e.g., vaginal-seeding,[68] or compensatory actions. Crucially, by separating the effects of presumed intrapartum antibiotic exposure from clinically reported direct antibiotic administration, we demonstrate that only the latter form of exposure appears to have a significant impact on the differentiation of the microbial communities and the establishment of the normal stool and oral microbial communities. While we saw empirical evidence that antibiotic exposure was occurring in infants from maternal use and/or other sources, reclassifying antibiotic-exposed samples per body site did not impact these results. Together these data suggest that future studies should separate these groups to better elucidate differential impacts of antibiotics on the infant microbiome via maternal versus direct exposure when examining short or long-term impacts of delivery mode.

Using compositionally aware methods enabled us to identify specific taxa enriched in infants supplemented with formula or DHM. Our observation that members of the *Staphylococcus* genus were associated with DHM supplementation is perhaps unsurprising given the source. *Staphylococcus* infections remain a leading cause of infections in the NICU, and it is poorly understood what promotes this vulnerability, though removal of *S. aureus* in particular appears to reduce the risk of infection. This suggests the need to consider treating upstream sources of milk[69,70] (the human milk donor, or screening of human milk for *S. aureus*), which may mitigate the need for this vulnerable preterm population to directly receive antibiotic treatment that would disrupt microbiome development. In addition, the increased ratio of *Aggregatibacter* in formula-supplemented infants, observed previously,[33] may indicate an early risk exposure for the later development of caries[71,72] as well as localized aggressive periodontitis[73–75] which may require treatment with antibiotics. Limiting antibiotics to a local application may ameliorate the global dysfunctional differentiation observed with systemic antibiotic exposure in our cohorts. Together, our results suggest that there is opportunity for treatment options specific to the type of feeding supplementation for preterm infants in the NICU.

An organ-specific metabolomic profile dominated in oral, stool, and skin samples in the first week after birth. However, the increasing correlation between the microbial community and metabolome of stool samples over the first week after birth, especially in the LP-V/Abx$^{(-)}$ group, suggests that as the microbial community expands and establishes itself, a greater degree of influence is exerted on the metabolomic profile, as also described in Bittinger et al.[76] The late preterm infant cohort reflects a low microbial biomass host, mirroring conditions explored in Quinn et al.,[77] where germ-free and specific pathogen free mice were used to demonstrate that microbes profoundly reshaped the metabolomic profile of each organ. In addition, an increasing trend in the ratio of secondary to primary bile acids, again primarily in the LP-V/Abx$^{(-)}$ group provides further evidence that the microbial community is actively engaged in modifying this axis early in human development. The lack of this trend is most evident in the LP-C/Abx$^{(+)}$ group,

suggesting that the combination of antibiotics and birth by Cesarean section may result in not just delayed microbiome development, but metabolomic development as well. Microbes enriched for producing secondary bile acids may therefore be a key component of successful seeding efforts to ameliorate the impact of early antibiotic exposure and birth mode.

Ultimately, the microbes living both on and in humans and their corresponding metabolites are influenced by and, in turn, influence the holobiont's collective health. Future studies into the causes of health disparities and potential interventions should ensure that this critical window of development is captured for the oral and stool microbiome, including the metabolome, rather than solely focusing on the gut microbial community, to ensure robust clinical interpretation. As the identification of metabolites continues to improve through the public sharing, curation, and aggregation of data using tools such as ReDU[78] and MASST,[79] the data in this study will remain available for providing context for future studies. Our study adds to the body of knowledge of preterm infant development and provides a model for future investigations as we continue to modify and improve our standard of care of preterm infants to optimize their long-term health.

4. Experimental Section

Cohorts: The data presented in this study originate from three independent cohorts: LP, VLBW, and FT infants (included for comparison), including birth by both vaginal delivery and cesarean section. For all analyses, the gut, oral, and skin microbiome samples were separated by sample type as well as by delivery mode, in order to prevent possible confounding effects. Due to independent collection efforts, some methods are data set specific, as described below.

Recruitment: For each study participant, a parent provided a written informed consent, and the study was approved by the Institutional Review Board at the University of California, San Diego [LP cohort: IRB approval number 151713; VLBW cohort: IRB approval number 151689] or the University of Michigan Institutional Review Board [FT: University of Michigan IRB HUM00103575].

LP: Pregnant English and Spanish speaking women at 34–36 6/7 weeks gestation who were expected to, or had delivered, prior to 37 weeks gestation were approached by a member of the research team. Infants outside of the gestational age criteria were excluded, as well as infants with major congenital anomalies including gastrointestinal anomalies, known or suspected metabolic disorder requiring a specialized diet, and infants who had received any formula prior to randomization.

Seventy-seven families representing 86 babies (seven sets of twins, one set of triplets) were approached regarding the study. Of those approached, 57 of the families consented, resulting in enrollment of 60 infants including one set of triplets, and four sets of twins. There were two infants consented but not randomized in error and one infant consented who was not randomized due to delivery after 37 weeks gestation.

VLBW: English or Spanish speaking mothers who were likely to deliver an infant prematurely were approached by a member of the research team from June 2016 until May 2017. Consent and enrollment in the study were finalized if the infant birth weight was 500–1500 g and between 23 and 34 weeks gestational age. Infants with medical problems incompatible with life or birth weight <500 g were excluded as the outcomes for these infants are poor. Most consents with the mother were initialized before the birth of their infant, however occasionally they were consented shortly after the delivery if the birth was imminent and unexpected. The first 25 patients from Hillcrest Medical Center were enrolled prior to establishment of a consistent oral colostrum care policy

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
BIOLOGY**

www.advanced-bio.com

and the latter 25 patients were enrolled from the Jacobs Medical Center after the oral colostrum care policy was put into place.

FT: Infant–mother dyads were recruited from the community when infants were between two weeks and two months of age. Mothers provided written informed consent for themselves and their infants. Inclusion criteria were: Child was born at 37.0–42.0 weeks gestation, with weight appropriate for gestational age, and no significant perinatal or neonatal complications. Exclusions were: 1) non-fluency in English in the parent; 2) foster child; 3) mother < 18 years old; 4) medical problems or known diagnosis affecting current or future eating, growth, or development; 5) child protective services involvement in the neonatal period; 6) infant does not consume at least two ounces in one feeding from an artificial nipple and bottle at least once per week.

Sample Collection: All samples were collected using sterile dual-tip swabs (BD Swube; (Becton, Dickinson and Company, Sparks, MD). One swab was used for 16S sequence analysis and the other for untargeted metabolomics analysis.

LP: Gloves and a mask were worn while collecting the samples to try to prevent contamination. Three double-tip swabs were used to gently swab the inside of the cheek, the skin of the axilla, and stool from the diaper. Once the collection was completed, the swabs were placed in the freezer, typically within 15–30 min of collection, and stored at −80 °C until analysis. Swabs were collected once daily until the infant was discharged or had reached 21 d after birth. Mother's breast milk and donor milk from the LP cohort were collected into 2 mL cryovials and stored at −80 °C until analysis.

VLBW: For sample collection a research team member or RN working with the infant used one double-tip swab to gently swab the inside of the cheek, the skin of the axilla, and stool from the diaper each day for the first 7 d after birth. Once the collection was completed, the swabs were stored at −80 °C until analysis.

rRNA Gene Sequencing: DNA was extracted using the MO BIO PowerSoil DNA extraction kit according to Earth Microbiome Project (EMP) standard protocols[80] (http://www.earthmicrobiome.org/emp-standard-protocols/). PCR targeting the V4 region of the 16S rRNA bacterial gene was performed with the 515F/806R primers, using the protocol described in ref. [81]. Amplicons were barcoded and pooled in equal concentrations for sequencing. The amplicon pool was purified with the MO BIO UltraClean PCR Clean-up kit and 2×150 bp sequencing was performed on the MiSeq sequencing platform at the Institute for Genomic Medicine at UCSD.

Amplicon Sequencing Data Processing and Analysis: Raw sequencing data were deposited in Qiita[66] and processed using the default parameters. Briefly, forward reads were demultiplexed and quality control filtered, followed by trimming to 150 bp and then processed with deblur[82] and filtered to remove amplicon sequence variants (ASVs) with a total count below 10 across all samples. The resulting ASVs were inserted into the Greengenes tree[83] via SEPP (SATé-Enabled Phylogenetic Placement using q2-fragment-insertion in QIIME2.[84,85] The feature table was collapsed to the genus level for use in differential abundance testing.

Metabolomics Data Acquisition: Swab tips were transferred to 96-well deep well plates, cutting or breaking the wooden swab handle slightly above the cotton swab. Breast milk samples were extracted in a final concentration of 80% MeOH:20% water (Optima LC-MS grade; Fisher Scientific, Fair Lawn, NJ).

Untargeted mass spectrometry data for VLBW sample sets were collected using a ultrahigh performance liquid chromatography system (Vanquish, Thermo) coupled to an orbitrap mass spectrometer (Q-Exactive, Thermo). Reverse chromatographic separation was accomplished using a Kinetic C18 column. The column compartment was held at 40 °C with a flow rate of 500 μL min⁻¹. Mobile phase composition was (A) 0.1% formic acid in water and (B) 0.1% formic acid in acetonitrile. The elution gradient used was: 0–1.0 min isocratic at 5% B, 1.0–9.0 min 100% B, 9.0–11.0 min isocratic at 100% B, 11.0–11.5 min 5% B, and 11.5–12.5 min isocratic at 5% B. Data-dependent acquisition for Full MS: Resolution set at 35000, AGC target 5e4, scan range 100–1500 m/z. For dd-MS/MS: Resolution at 17500, AGC target 2e4, loop count 5, and (N)CE/stepped (N)CE 20, 30, 40. Data-dependent

acquisition settings were set at: minimum AGC target 2.00e4, and Apex trigger 3–15 s.

LP data were collected using the data-dependent acquisition method outlined in Gauglitz et al.[58] Extracts were dried down using centrifugal evaporation (Labconco) and resuspended in 50% MeOH:50% water (Optima LC-MS grade; Fisher Scientific, Fair Lawn, NJ,), with 1×10^{-6} M sulfadimethoxine. Untargeted metabolomics was carried out using an ultrahigh-performance liquid chromatography system (UltiMate 3000, Thermo Scientific, Waltham, MA) coupled to a Maxis Q-TOF (Bruker Daltonics, Bremen, Germany) mass spectrometer with a Kinetex C18 column (Phenomenex Torrance, CA, USA). A linear gradient was applied: 0–0.5 min isocratic at 5% B, 0.5–8.5 min 100% B, 8.5–11 min isocratic at 100% B, 11–11.5 min 5% B, 11.5–12 min 5% B, where mobile phase A is water with 0.1% formic acid (v/v) and phase B is acetonitrile 0.1% formic acid (v/v) (LC-MS grade solvents, Fisher Chemical). Electrospray ionization in positive mode was used.

MS1 Feature Finding and Data Processing: qToF files (.d) were exported using DataAnalysis (Bruker) as .mzXML files after lock mass correction using hexakis (1H, 1H, 2H-difluoroethoxy)phosphazene (Synquest Laboratories, Alachua, FL), with m/z 622.029509. Data quality was assessed by qualitatively evaluating the m/z error and retention time of the LC-MS standard solution (i.e., mixture of six compounds), which was analyzed at least once in every 96-well plate.

MS1 feature finding was performed on the .mzXML files in MZmine2 (version MZmine-2.37.corr16.4).[86] The mzMINE parameters used for feature finding are as follows: mass detection (centroid; ms1: 1.5E3; MS2: 90); ADAP Chromatogram builder (min group size in # of scans: 4; group intensity threshold: 5E3; min highest intensity: 2E3; m/z tolerance: 0.001 m/z to 20 ppm); chromatogram deconvolution (LMS: chromatographic threshold of 96%, search minimum in RT range (min) of 0.03, minimum relative height of 5%, minimum absolute height of 2E3, min ratio of peak top/edge of 1 and peak duration range (min) of 0–2; m/z center calculation set to auto; m/z range for MS2 scan pairing (Da) of 0.02 and RT range for MS2 scan pairing (min) of 0.15); isotope peaks grouper (m/z tolerance set to 0.0015 m/z or 10 ppm; retention time tolerance of 0.05, maximum charge of 3 and representative isotope set to most intense); order peak lists; join aligner (m/z tolerance set at 0.0015 m/z or 15 ppm; weight for m/z of 2; retention time tolerance of 0.2 min; weight for RT of 1). A filter was used such that only features present in at least two samples were included. The output was a data matrix of variables (i.e., MS1 features that triggered MS2 scans) by samples, exported for GNPS (.mgf and .csv quant table). Tables were normalized by sample wherein metabolite relative abundances were determined by dividing the integrated intensity for each feature by the sum of the intensities across all features for that sample.

MS2 Molecular Networking: The resulting .mgf and .csv quantification table produced by MZmine2 were imported into the GNPS interface for use in the feature based molecular networking workflow.[67,87] The parameters were set at: Precursor Ion Mass Tolerance 0.02 Da, Fragment Ion Mass Tolerance 0.02 Da, Minimum Pairs Cosine 0.7, Minimum Matched Fragment Ions 6, Maximum shi between precursors 500 Da, Network TopK 10, Maximum Connected Component Size (Beta) 100, Library Search Min Matched Peaks 6, Score Threshold 0.7. The resulting annotations fall under the Metabolomics Standards Initiative metabolite identification levels 1–3.[88]

Diversity Analyses: Alpha diversity was measured by Shannon diversity after rarefying the ASV table to a depth of 1250 sequences.[89] Beta diversities for ASVs were calculated using robust Aitchison distances.[90,91] PCA of these distances were plotted and distances were used to calculate body site differentiation by PERMANOVA as implemented by skbio. To determine the overall contribution of metadata factors to microbial diversity, stepwise redundancy analysis was run using the ordistep function from the vegan package in R. The data used in the analysis were the robust Aitchison distances on samples from each body site separately. The metadata factors included in the ordistep formula were antibiotic use, age, and delivery mode.

Metabolomics Analysis: Beta diversity of normalized metabolite intensities was calculated with Jaccard distances. Metabolite profile

change over time was determined by calculating the Jaccard distance of each infant sample to the sample from the same body site at day 0 for LP infants and day 1 for VLBW infants. These distances to the earliest time point over time were split by birth-weight, delivery-mode, and antibiotic exposure and used to calculate linear mixed models using individual as a random effect and age as a fixed effect. Slopes and p-values of the age variable in linear mixed models were reported.

Differential Abundance Testing: Differential abundances of microbes and metabolites were determined by Songbird. Songbird takes in a feature table, sample metadata, and a formula to calculate differential scores for each feature as a measure of differential abundance depending on metadata variables specified in the formula. Songbird also supplies a Q2 that behaves similar to an R2. Following advice from the Songbird creators, only models with positive Q2 values were included in results. The genus collapsed table was used for differential abundance testing of microbes. Songbird was run separately for each body site. To determine differential abundances due to delivery-mode, age, and formula/donor breast milk supplementation, only non-antibiotic exposed LPI infant samples were included. The formula in this case was "age + delivery-mode + milk_supplementation". To determine the effect of antibiotics, all LP and VLBW infant samples were included and the formula "age + delivery-mode + milk_supplementation + antibiotics" was used. Songbird differentials for metabolite abundances were calculated using normalized metabolite abundances with all LP infant samples using the formula "age + delivery-mode + milk_supplementation." Log ratios of secondary to primary bile acids were calculated using Qurro.[92] Linear mixed effects models (age as fixed variable and individual as random variable) were run on these log ratios split by delivery mode and antibiotic exposure over time and the slope, and p-value of the age variable in these models were reported.

SourceTracker 2: To compare infant microbiomes to healthy adult microbiomes, 16S-V4 sequencing data from healthy adults, including 4434 stool, 2550 oral, 1975 skin samples, and 427 vaginal samples were downloaded (https://msystems.asm.org/content/5/1/e00630-19). All samples that indicated exposure to antibiotics within the last year were excluded. The LP and VLBW infants were compared to healthy FT infants, using the cohort of 100 FT vaginally born infants from a publicly available data set in Qiita. Only stool samples were available from FT infants, not skin or oral samples. All data were processed in the same manner as for preterm infants described above, with an additional filtering to remove all ASVs present in less than 1% of samples to reduce complexity. All adult samples from all three body sites and stool samples from FT infants were merged into a single database that was used as a source for Sourcetracker2[93] analysis on each preterm infant sample. The relative contribution of each source (adult stool, adult oral, adult skin, and FT infant stool) was determined for each preterm infant sample. Preterm infant samples were grouped by delivery mode, birth weight (LP vs VLBW), and antibiotic status, and the mean and standard error was calculated for each group over time and plotted.

Mantel Test: Correlation between microbe abundances and metabolite intensities were calculated using a Mantel test as implemented by skbio. The genus collapsed feature table and normalized metabolite profile were subsampled to only include samples present in both datasets. Samples from all three body sites were included in this analysis. Samples were divided by age, delivery-mode, and antibiotic exposure, and then the Mantel test was run on each data set independently. The Mantel correlation and p-value were plotted over time for each sample set.

Statistical Analysis: The methods used for each specific component of the analyses outlined above are detailed in those respective sections, but a general summary follows.

Preprocessing: Amplicon sequencing data were quality and noise-filtered within Qiita using recommended settings, removing low abundance ASVs (<10 across the data set). Sequences were further rarefied to 1250 reads per sample for calculation of alpha diversity, but no other analyses used a rarefied data set. Metabolomic feature relative abundances determined by MS1 feature finding using mzMINE were calculated using per-sample normalization.

Data presentation: Tabular data were summarized in Microsoft Excel, manipulated using the Pandas package in Python, and plotted using the Python packages *matplotlib* and *plotnine*, a Python implementation of the R package *ggplot2*. For within-individual distances, boxplots were generated showing the median and interquartile range with whiskers extending to the furthest value within 1.5 times the edge of the interquartile range. For PCA or PCoA plots, confidence ellipses were calculated in *matplotlib* showing the boundary of 3 standard deviations from the centroid. For alpha diversity and SourceTracker 2 analysis results, line graphs were generated showing the mean with shaded areas indicating the standard error of the mean.

Sample size: From 85 infants enrolled in the study, 1799 samples were obtained, 942 16S amplicon sequencing tables (from 61 infants) and 1080 metabolomic feature tables (from 75 infants) across 3 body sites and 7 days of life were analyzed. The exact sample sizes per body site, per day varied considerably and a detailed breakdown of subgroups compared for clinical characteristics can be found in Table 1, while the subgroups analyzed for metagenomic and metabolomic attributes are listed in Table S3 (Supporting Information) and shown in the flowchart in Figure 1.

Statistical methods: Permutational multivariate analysis of variance (PERMANOVA), stepwise redundancy analysis (RDA), linear mixed models, and Mantel tests were all performed with a testing value of $p = 0.05$ to determine significance (https://github.com/ucsd-cmi/ preterm_infant/tree/master/results/tables). Multiple test correction was not required for the analyses conducted.

Software: PERMANOVA, linear mixed models, and Mantel tests were performed using the *scikit-learn* and *skbio* Python libraries. Stepwise redundancy analysis was run in R using the ordistep function in the *vegan* package.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Conflict of Interest

J.H.K. is a medical advisor for Evolve BioSystems and a consultant for Nutricia. R.K., A.D.S., and S.J.S. are directors at the Center for Microbiome Innovation at UC San Diego, which receives industry research funding for various microbiome initiatives, but no industry funding was provided for this study. All other authors declare no conflict of interest.

Author Contributions

Data Availability Statement

Keywords

[1] G. D. Poore, E. Kopylova, Q. Zhu, C. Carpenter, S. Fraraccio, S. Wandro, T. Kosciolek, S. Janssen, J. Metcalf, S. J. Song, J. Kanbar, S. Miller-Montgomery, R. Heaton, R. Mckay, S. P. Patel, A. D. Swafford, R. Knight, *Nature* **2020**, *579*, 567.

[2] D. Nejman, I. Livyatan, G. Fuks, N. Gavert, Y. Zwang, L. T. Geller, A. Rotter-Maskowitz, R. Weiser, G. Mallel, E. Gigi, A. Meltser, G. M. Douglas, I. Kamer, V. Gopalakrishnan, T. Dadosh, S. Levin-Zaidman, S. Avnet, T. Atlan, Z. A. Cooper, R. Arora, A. P. Cogdill, M. A. W. Khan, G. Ologun, Y. Bussi, A. Weinberger, M. Lotan-Pompan, O. Golani, G. Perry, M. Rokah, K. Bahar-Shany, et al., *Science* **2020**, *368*, 973.

[3] D. McDonald, E. Hyde, J. W. Debelius, J. T. Morton, A. Gonzalez, G. Ackermann, A. A. Aksenov, B. Behsaz, C. Brennan, Y. Chen, L. DeRight Goldasich, P. C. Dorrestein, R. R. Dunn, A. K. Fahimipour, J. Gaffney, J. A. Gilbert, G. Gogul, J. L. Green, P. Hugenholtz, G. Humphrey, C. Huttenhower, M. A. Jackson, S. Janssen, D. V. Jeste, L. Jiang, S. T. Kelley, D. Knights, T. Kosciolek, J. Ladau, J. Leach, et al., *mSystems* **2018**, *3*, https://doi.org/10.1128/mSystems.00031-18.

[4] L. M. Proctor, H. H. Creasy, J. M. Fettweis, J. Lloyd-Price, A. Mahurkar, W. Zhou, G. A. Buck, M. P. Snyder, J. F. Strauss, G. M. Weinstock, O. White, C. Huttenhower, *Nature* **2019**, *569*, 623.

[5] A. Paun, J. S. Danska, *Pediatr. Diabetes* **2016**, *17*, 469.

[6] D. C. Baumgart, S. R. Carding, *Lancet* **2007**, *369*, 1627.

[7] Y. He, W. Wu, H.-M. Zheng, P. Li, D. McDonald, H.-F. Sheng, M.-X. Chen, Z.-H. Chen, G.-Y. Ji, Z.-D.-X. Zheng, P. Mujagond, X.-J. Chen, Z.-H. Rong, P. Chen, L.-Y. Lyu, X. Wang, C.-B. Wu, N. Yu, Y.-J. Xu, J. Yin, J. Raes, R. Knight, W.-J. Ma, H.-W. Zhou, *Nat. Med.* **2018**, *24*, 1532.

[8] T. Yatsunenko, F. E. Rey, M. J. Manary, I. Trehan, M. G. Dominguez-Bello, M. Contreras, M. Magris, G. Hidalgo, R. N. Baldassano, A. P. Anokhin, A. C. Heath, B. Warner, J. Reeder, J. Kuczynski, J. G. Caporaso, C. A. Lozupone, C. Lauber, J. C. Clemente, D. Knights, R. Knight, J. I. Gordon, *Nature* **2012**, *486*, 222.

[9] A. Salava, V. Aho, E. Lybeck, P. Pereira, L. Paulin, I. Nupponen, A. Ranki, P. Auvinen, S. Andersson, A. Lauerma, *Exp. Dermatol.* **2017**, *26*, 861.

[10] R. G. Ubeja, C. Bhat, *Int. J. Clin. Pediatr. Dent.* **2016**, *9*, 326.

[11] I. Adlerberth, E. Lindberg, N. Aberg, B. Hesselmar, R. Saalman, I.-L. Strannegård, A. E. Wold, *Pediatr. Res.* **2006**, *59*, 96.

[12] M. G. Dominguez-Bello, E. K. Costello, M. Contreras, M. Magris, G. Hidalgo, N. Fierer, R. Knight, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11971.

[13] M.-M. Grölund, O.-P. Lehtonen, E. Eerola, P. Kero, *J. Pediatr. Gastroenterol. Nutr.* **1999**, *28*, 19.

[14] M.-M. Grönlund, Ł. Grześkowiak, E. Isolauri, S. Salminen, *Gut Microbes* **2011**, *2*, 227.

[15] D. M. Chu, J. Ma, A. L. Prince, K. M. Antony, M. D. Seferovic, K. M. Aagaard, *Nat. Med.* **2017**, *23*, 314.

[16] F. Asnicar, S. Manara, M. Zolfo, D. T. Truong, M. Scholz, F. Armanini, P. Ferretti, V. Gorfer, A. Pedrotti, A. Tett, N. Segata, *mSystems* **2017**, *2*, https://doi.org/10.1128/mSystems.00164-16.

[17] P. Ferretti, E. Pasolli, A. Tett, F. Asnicar, V. Gorfer, S. Fedi, F. Armanini, D. T. Truong, S. Manara, M. Zolfo, F. Beghini, R. Bertorelli, V. De Sanctis, I. Bariletti, R. Canto, R. Clementi, M. Cologna, T. Crifò, G. Cusumano, S. Gottardi, C. Innamorati, C. Masè, D. Postai, D. Savoi, S. Duranti, G. A. Lugli, L. Mancabelli, F. Turroni, C. Ferrario, C. Milani, et al., *Cell Host Microbe* **2018**, *24*, 133.

[18] M. A. Hall, C. B. Cole, S. L. Smith, R. Fuller, C. J. Rolles, *Arch. Dis. Child.* **1990**, *65*, 185.

[19] J. Penders, C. Thijs, C. Vink, F. F. Stelma, B. Snijders, I. Kummeling, P. A. van den Brandt, E. E. Stobberingh, *Pediatrics* **2006**, *118*, 511.

[20] K. Kristensen, L. Henriksen, *J. Allergy Clin. Immunol.* **2016**, *137*, 587.

[21] N. T. Mueller, R. Whyatt, L. Hoepner, S. Oberfield, M. G. Dominguez-Bello, E. M. Widen, A. Hassoun, F. Perera, A. Rundle, *Int. J. Obes.* **2015**, *39*, 665.

[22] W. A. Walker, *Nestle Nutr. Inst. Workshop Ser.* **2017**, *88*, 23.

[23] P. Shah, N. Kaciroti, B. Richards, W. Oh, J. C. Lumeng, *Pediatrics* **2016**, *138*, e20153496.

[24] F. Fouhy, C. Watkins, C. J. Hill, C.-A. O'Shea, B. Nagle, E. M. Dempsey, P. W. O'Toole, R. P. Ross, C. A. Ryan, C. Stanton, *Nat. Commun.* **2019**, *10*, 1517.

[25] A. P. Francis, M. G. Dominguez-Bello, *Trends Microbiol.* **2019**, *27*, 567.

[26] D. A. Chernikova, J. C. Madan, M. L. Housman, M. Zain-Ul-Abideen, S. N. Lundgren, H. G. Morrison, M. L. Sogin, S. M. Williams, J. H. Moore, M. R. Karagas, A. G. Hoen, *Pediatr. Res.* **2018**, *84*, 71.

[27] K. M. Puopolo, W. E. Benitz, T. E. Zaoutis, *Pediatrics* **2018**, *142*, https://doi.org/10.1542/peds.2018-2896.

[28] S. Salvatore, M. E. Baldassarre, A. Di Mauro, N. Laforgia, S. Tafuri, F. P. Bianchi, E. Dattoli, L. Morando, L. Pensabene, F. Meneghin, D. Dilillo, V. Mancini, V. Talarico, F. Tandoi, G. Zuccotti, M. Agosti, *J. Pediatr.* **2019**, *212*, 44.

[29] B. Demoré, D. Le Govic, N. Thilly, J.-M. Boivin, C. Pulcini, *Clin. Microbiol. Infect.* **2017**, *23*, 486.

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**ADVANCED
BIOLOGY**
www.advanced-bio.com

[30] Y. Zhao, Y. Zhou, Q. Zhu, B. Xia, W. Ma, X. Xiao, H. Shi, Y. Zhang, *Environ. Int.* **2019**, *123*, 70.

[31] A. B. Dunn, S. Jordan, B. J. Baker, N. S. Carlson, *MCN Am. J. Matern. Child Nurs.* **2017**, *42*, 318.

[32] H. Hermansson, H. Kumar, M. C. Collado, S. Salminen, E. Isolauri, S. Rautava, *Front. Nutr.* **2019**, *6*, 4.

[33] A. Parra-Llorca, M. Gormaz, C. Alcántara, M. Cernada, A. Nuñez-Ramiro, M. Vento, M. C. Collado, *Front. Microbiol.* **2018**, *9*, 1376.

[34] S. Perrone, M. Longini, I. Zollino, F. Bazzini, M. Tassini, A. Vivi, C. Bracciali, M. Calderisi, G. Buonocore, *Nutrition* **2019**, *62*, 158.

[35] N. J. Andreas, B. Kampmann, K. Mehring Le-Doare, *Early Hum. Dev.* **2015**, *91*, 629.

[36] M. R. Asbury, J. Butcher, J. K. Copeland, S. Unger, N. Bando, E. M. Comelli, V. Forte, A. Kiss, L. LeMay-Nedjelski, P. M. Sherman, A. Stintzi, C. Tomlinson, P. W. Wang, D. L. O'Connor, *Cell Host Microbe* **2020**, *28*, 669.

[37] L. Bode, *Glycobiology* **2012**, *22*, 1147.

[38] D. A. Sela, J. Chapman, A. Adeuya, J. H. Kim, F. Chen, T. R. Whitehead, A. Lapidus, D. S. Rokhsar, C. B. Lebrilla, J. B. German, N. P. Price, P. M. Richardson, D. A. Mills, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 18964.

[39] C. R. Martin, P.-R. Ling, G. L. Blackburn, *Nutrients* **2016**, *8*, 279.

[40] S. Moossavi, S. Sepehri, B. Robertson, L. Bode, S. Goruk, C. J. Field, L. M. Lix, R. J. de Souza, A. B. Becker, P. J. Mandhane, S. E. Turvey, P. Subbarao, T. J. Moraes, D. L. Lefebvre, M. R. Sears, E. Khafipour, M. B. Azad, *Cell Host Microbe* **2019**, *25*, 324.

[41] C. Cai, Z. Zhang, M. Morales, Y. Wang, E. Khafipour, J. Friel, *Free Radical Biol. Med.* **2019**, *142*, 146.

[42] R. L. Shen, T. Thymann, M. V. Østergaard, A. C. Støy, P. Krych, D. S. Nielsen, C. Lauridsen, B. Hartmann, J. J. Holst, D. G. Burrin, P. T. Sangild, *Am. J. Physiol. Gastrointest. Liver Physiol.* **2015**, *309*, G310.

[43] M. D. Rossiter, C. K. Colapinto, M. K. A. Khan, J.-L. D. McIsaac, P. L. Williams, S. F. L. Kirk, P. J. Veugelers, *Matern. Child Health J.* **2015**, *19*, 2048.

[44] B. Li, R. Y. Wu, R. G. Horne, A. Ahmed, D. Lee, S. C. Robinson, H. Zhu, C. Lee, M. Cadete, K. C. Johnson-Henry, E. Landberg, M. Alganabi, T. Abrahamsson, P. Delgado-Olguin, A. Pierro, P. M. Sherman, *Mol. Nutr. Food Res.* **2020**, *64*, 2000519.

[45] J. E. Koenig, A. Spor, N. Scalfone, A. D. Fricker, J. Stombaugh, R. Knight, L. T. Angenent, R. E. Ley, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, Suppl 1, 4578.

[46] E. J. de Muinck, P. Trosvik, *Nat. Commun.* **2018**, *9*, 2233.

[47] N. E. Younge, C. B. Newgard, C. M. Cotten, R. N. Goldberg, M. J. Muehlbauer, J. R. Bain, R. D. Stevens, T. M. O'Connell, J. F. Rawls, P. C. Seed, P. L. Ashley, *Sci. Rep.* **2019**, *9*, 8167.

[48] C. J. Stewart, N. D. Embleton, E. C. L. Marrs, D. P. Smith, T. Fofanova, A. Nelson, T. Skeath, J. D. Perry, J. F. Petrosino, J. E. Berrington, S. P. Cummings, *Microbiome* **2017**, *5*, 75.

[49] C. J. Hill, D. B. Lynch, K. Murphy, M. Ulaszewska, I. B. Jeffery, C. A. O'Shea, C. Watkins, E. Dempsey, F. Mattivi, K. Tuohy, R. Paul Ross, C. Anthony Ryan, P. W. O'Toole, C. Stanton, *Microbiome* **2017**, *5*, 21.

[50] H. Yoshioka, K. Fujita, H. Sakata, K. Murono, K.-I. Iseki, *Bifidobact. Microflora* **1991**, *10*, 11.

[51] A. A. Fanaroff, in *Neonatology: A Practical Approach to Neonatal Diseases* (Eds: G. Buonocore, R. Bracci, M. Weindling), Springer, Milan **2012**, pp. 94–101.

[52] S. Wandro, S. Osborne, C. Enriquez, C. Bixby, A. Arrieta, K. Whiteson, *mSphere* **2018**, *3*.

[53] H. Sakata, H. Yoshioka, K. Fujita, *Eur. J. Pediatr.* **1985**, *144*, 186.

[54] E. K. Costello, E. M. Carlisle, E. M. Bik, M. J. Morowitz, D. A. Relman, *MBio* **2013**, *4*, e00782.

[55] G. Falony, M. Joossens, S. Vieira-Silva, J. Wang, Y. Darzi, K. Faust, A. Kurilshikov, M. J. Bonder, M. Valles-Colomer, D. Vandeputte, R. Y. Tito, S. Chaffron, L. Rymenans, C. Verspecht, L. De Sutter, G. Lima-Mendez, K. D'hoe, K. Jonckheere, D. Homola, R. Garcia, E. F. Tigchelaar, L. Eeckhaudt, J. Fu, L. Henckaerts, A. Zhernakova, C. Wijmenga, J. Raes, *Science* **2016**, *352*, 560.

[56] C. Martino, J. T. Morton, C. A. Marotz, L. R. Thompson, A. Tripathi, R. Knight, K. Zengler, *mSystems* **2019**, *4*.

[57] N. A. Bokulich, J. Chung, T. Battaglia, N. Henderson, M. Jay, H. Li, A. D. Lieber, F. Wu, G. I. Perez-Perez, Y. Chen, W. Schweizer, X. Zheng, M. Contreras, M. G. Dominguez-Bello, M. J. Blaser, *Sci. Transl. Med.* **2016**, *8*, 343ra82.

[58] J. M. Gauglitz, J. T. Morton, A. Tripathi, S. Hansen, M. Gaffney, C. Carpenter, K. C. Weldon, R. Shah, A. Parampil, A. L. Fidgett, A. D. Swafford, R. Knight, P. C. Dorrestein, *mSystems* **2020**, *5*, https://doi.org/10.1128/mSystems.00635-19.

[59] J. E. Bisanz, M. K. Enos, G. PrayGod, S. Seney, J. M. Macklaim, S. Chilton, D. Willner, R. Knight, C. Fusch, G. Fusch, G. B. Gloor, J. P. Burton, G. Reid, *Appl. Environ. Microbiol.* **2015**, *81*, 4965.

[60] L. A. David, A. C. Materna, J. Friedman, M. I. Campos-Baptista, M. C. Blackburn, A. Perrotta, S. E. Erdman, E. J. Alm, *Genome Biol.* **2014**, *15*, R89.

[61] J. C. Clemente, E. C. Pehrsson, M. J. Blaser, K. Sandhu, Z. Gao, B. Wang, M. Magris, G. Hidalgo, M. Contreras, Ó. Noya-Alarcón, O. Lander, J. McDonald, M. Cox, J. Walter, P. L. Oh, J. F. Ruiz, S. Rodriguez, N. Shen, S. J. Song, J. Metcalf, R. Knight, G. Dantas, M. G. Dominguez-Bello, *Sci. Adv.* **2015**, *1*, https://doi.org/10.1126/sciadv.1500183.

[62] J. G. Caporaso, C. L. Lauber, E. K. Costello, D. Berg-Lyons, A. Gonzalez, J. Stombaugh, P. Knights, P. Gajer, J. Ravel, N. Fierer, J. I. Gordon, R. Knight, *Genome Biol.* **2011**, *12*, R50.

[63] M. G. Dominguez-Bello, K. M. De Jesus-Laboy, N. Shen, L. M. Cox, A. Amir, A. Gonzalez, N. A. Bokulich, S. J. Song, M. Hoashi, J. I. Rivera-Vinas, K. Mendez, R. Knight, J. C. Clemente, *Nat. Med.* **2016**, *22*, 250.

[64] J. T. Morton, C. Marotz, A. Washburne, J. Silverman, L. S. Zaramela, A. Edlund, K. Zengler, R. Knight, *Nat. Commun.* **2019**, *10*, 2719.

[65] C. Rougé, O. Goldenberg, L. Ferraris, B. Berger, F. Rochat, A. Legrand, U. B. Göbel, M. Vodovar, M. Voyer, J.-C. Rozé, D. Darmaun, H. Piloquet, M.-J. Butel, M.-F. de La Cochetière, *Anaerobe* **2010**, *16*, 362.

[66] A. Gonzalez, J. A. Navas-Molina, T. Kosciolek, D. McDonald, Y. Vázquez-Baeza, G. Ackermann, J. DeReus, S. Janssen, A. D. Swafford, S. B. Orchanian, J. G. Sanders, J. Shorenstein, H. Holste, S. Petrus, A. Robbins-Pianka, C. J. Brislawn, M. Wang, J. R. Rideout, E. Bolyen, M. Dillon, J. G. Caporaso, P. C. Dorrestein, R. Knight, *Nat. Methods* **2018**, *15*, 796.

[67] M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W.-T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C.-C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrewe, T. Northen, R. J. Dutton, D. Parrot, et al., *Nat. Biotechnol.* **2016**, *34*, 828.

[68] N. T. Mueller, M. G. Dominguez-Bello, L. J. Appel, S. K. Hourigan, *BJOG* **2020**, *127*, 301.

[69] A. M. Milstone, A. Voskertchian, D. W. Koontz, D. F. Khamash, T. Ross, S. W. Aucott, M. M. Gilmore, S. E. Cosgrove, K. C. Carroll, E. Colantuoni, *JAMA, J. Am. Med. Assoc.* **2019**, *323*, 319.

[70] C. Barbosa-Cesnik, K. Schwartz, B. Foxman, *JAMA, J. Am. Med. Assoc.* **2003**, *289*, 1609.

[71] W. M. Avila, I. A. Pordeus, S. M. Paiva, C. C. Martins, *PLoS One* **2015**, *10*, e0142922.

[72] I. L. Kriswandini, I. Diyatri, Tantiana, P. Nuraini, T. Berniyanti, I Putri, T. Pnbn, *Infect. Dis. Rep.* **2020**, *12*, 8722.

[73] O. Oettinger-Barak, S. G. Dashper, D. V. Catmull, G. G. Adams, M. N. Sela, E. E. Machtei, E. C. Reynolds, *J. Oral Microbiol.* **2013**, *5*, 20320.

[74] K. G. Bhat, P. Khot, S. Patil, G. Pattar, S. Majukar, *J. Oral Maxillofac. Pathol.* **2019**, *23*, 231.

[75] C. H. Åberg, P. Kelk, A. Johansson, *Virulence* **2015**, *6*, 188.

[76] K. Bittinger, C. Zhao, Y. Li, E. Ford, E. S. Friedman, J. Ni, C. V. Kulkarni, J. Cai, Y. Tian, Q. Liu, A. D. Patterson, D. Sarkar, S. H. J. Chan, C. Maranas, A. Saha-Shah, P. Lund, B. A. Garcia, L. M. Mattei, J. S. Gerber, M. A. Elovitz, A. Kelly, P. DeRusso, D. Kim, C. E. Hofstaedter, M. Goulian, H. Li, F. D. Bushman, B. S. Zemel, G. D. Wu, *Nat. Microbiol.* **2020**, *5*, 838.

[77] R. A. Quinn, A. V. Melnik, A. Vrbanac, T. Fu, K. A. Patras, M. P. Christy, Z. Bodai, P. Belda-Ferre, A. Tripathi, L. K. Chung, M. Downes, R. D. Welch, M. Quinn, G. Humphrey, M. Panitchpakdi, K. C. Weldon, A. Aksenov, R. da Silva, J. Avila-Pacheco, C. Clish, S. Bae, H. Mallick, E. A. Franzosa, J. Lloyd-Price, R. Bussell, T. Thron, A. T. Nelson, M. Wang, E. Leszczynski, F. Vargas, et al., *Nature* **2020**, *579*, 123.

[78] A. K. Jarmusch, M. Wang, C. M. Aceves, R. S. Advani, S. Aguire, A. A. Aksenov, G. Aleti, A. T. Aron, A. Bauermeister, S. Bolleddu, *bioRxiv* **2019**.

[79] M. Wang, A. K. Jarmusch, F. Vargas, A. A. Aksenov, J. M. Gauglitz, K. Weldon, D. Petras, R. da Silva, R. Quinn, A. V. Melnik, J. J. J. van der Hooft, A. M. Caraballo-Rodriguez, L. F. Nothias, C. M. Aceves, M. Panitchpakdi, E. Brown, F. Di Ottavio, N. Sikora, E. O. Elijah, L. Labarta-Bajo, E. C. Gentry, S. Shalapour, K. E. Kyle, S. P. Puckett, J. D. Watrous, C. S. Carpenter, A. Bouslimani, M. Ernst, A. D. Swafford, E. I. Zuniga, et al., *Nat. Biotechnol.* **2020**, *38*, 23.

[80] L. R. Thompson, J. G. Sanders, D. McDonald, A. Amir, J. Ladau, K. J. Locey, R. J. Prill, A. Tripathi, S. M. Gibbons, G. Ackermann, J. A. Navas-Molina, S. Janssen, E. Kopylova, Y. Vázquez-Baeza, A. González, J. T. Morton, S. Mirarab, Z. Z. Xu, L. Jiang, M. F. Haroon, J. Kanbar, Q. Zhu, S. J. Song, T. Kosciolek, N. A. Bokulich, J. Lefler, C. J. Brislawn, G. Humphrey, S. M. Owens, J. Hampton-Marcell, et al., *Nature* **2017**, *551*, 457.

[81] W. Walters, E. R. Hyde, D. Berg-Lyons, G. Ackermann, G. Humphrey, A. Parada, J. A. Gilbert, J. K. Jansson, J. G. Caporaso, J. A. Fuhrman, A. Apprill, R. Knight, *mSystems* **2016**, *1*, https://doi.org/10.1128/msystems.00009-15.

[82] A. Amir, D. McDonald, J. A. Navas-Molina, E. Kopylova, J. T. Morton, Z. Zech Xu, E. P. Kightley, L. R. Thompson, E. R. Hyde, A. Gonzalez, R. Knight, *mSystems* **2017**, *2*, https://doi.org/10.1128/msystems.00191-16.

[83] T. Z. DeSantis, P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, G. L. Andersen, *Appl. Environ. Microbiol.* **2006**, *72*, 5069.

[84] E. Bolyen, J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. K. Cope, R. Da Silva, C. Diener, P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, et al., *Nat. Biotechnol.* **2019**, *37*, 852.

[85] S. Mirarab, N. Nguyen, T. Warnow, *Pac. Symp. Biocomput.* **2012**, 247.

[86] T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, *BMC Bioinf.* **2010**, *11*, 395.

[87] L.-F. Nothias, D. Petras, R. Schmid, K. Dührkop, J. Rainer, A. Sarvepalli, I. Protsyuk, M. Ernst, H. Tsugawa, M. Fleischauer, F. Aicheler, A. A. Aksenov, O. Alka, P.-M. Allard, A. Barsch, X. Cachet, A. M. Caraballo-Rodriguez, R. R. Da Silva, T. Dang, N. Garg, J. M. Gauglitz, A. Gurevich, G. Isaac, A. K. Jarmusch, Z. Kameník, K. B. Kang, N. Kessler, I. Koester, A. Korf, A. Le Gouellec, et al., *Nat. Methods* **2020**, *17*, 905.

[88] MSI Board Members, S.-A. Sansone, T. Fan, R. Goodacre, J. L. Griffin, N. W. Hardy, R. Kaddurah-Daouk, B. S. Kristal, J. Lindon, P. Mendes, N. Morrison, B. Nikolau, D. Robertson, L. W. Sumner, C. Taylor, M. van der Werf, B. van Ommen, O. Fiehn, *Nat. Biotechnol.* **2007**, *25*, 846.

[89] C. E. Shannon, *Bell Syst. Techn. J.* **1948**, *27*, 379.

[90] J. Aitchison, *J. R. Stat. Soc., Ser. B* **1982**, *44*, 139.

[91] J. Aitchison, *Biometrika* **1983**, *70*, 57.

[92] M. W. Fedarko, C. Martino, J. T. Morton, A. González, G. Rahman, C. A. Marotz, J. J. Minich, E. E. Allen, R. Knight, *NAR: Genomics Bioinf.* **2020**, *2*, lqaa023.

[93] D. Knights, J. Kuczynski, E. S. Charlson, J. Zaneveld, M. C. Mozer, R. G. Collman, F. D. Bushman, R. Knight, S. T. Kelley, *Nat. Methods* **2011**, *8*, 761.