

Abbreviations used

VDAART: Vitamin D Antenatal Asthma Reduction Trial
MaAsLin: Multivariate Associations with Linear Models

protect against autoimmune and inflammatory disease in animal models.¹¹⁻¹³ *Lactobacillus* strains suppress production of total IgE, and decreased relative abundance of *Lactobacillus* in the infant gut has been associated with allergic disease development.¹⁴⁻¹⁶ Although these studies suggest that specific bacterial taxa in the gut may be important for protection against immune-mediated disorders, the role of entire microbial communities in the proper functioning and maturation of the immune system must also be examined. The infant gut microbiome sets the stage for the composition of gut flora by age 3 years and onward into adulthood, with potential implications for health both in early and in later life. However, before we proceed with trying to understand the effects of the early infant microbiome on later health, elucidation of the factors that have an impact on the establishment of this early life microbiome needs to be carried out.

Colonization of the infant gut may begin before, during, or after delivery. Detection of microbes in cord blood and amniotic fluid suggests that the fetal gastrointestinal tract may harbor a limited prenatal microbiome.^{17,18} During birth and in the immediate postnatal period, the infant is exposed to microbes from maternal sources (originating from vaginal, stool, and skin microbiota) as well as the environment.^{19,20} Many of these microbes have the potential to rapidly colonize the infant gut. In fact, bacteria are culturable from the stools of newborns within hours after birth.²¹ Aspects of an infant's prenatal, birth, and postnatal experience may all contribute to the types of microbial flora that populate the infant gut.

Factors associated with microbial composition and diversity in the infant gut are understudied, and existing reports in this area rely mainly on small, restricted groups of infants (most fewer than 20 subjects) and diseased infants. As a result, these studies are likely underpowered, and findings from these reports may not be generalizable to a broader population. Our aim in this work was to determine how prenatal and early life factors (circumstances of labor and delivery, gestational age at birth, breast-feeding, and home exposures) impact the gut microbiome in a relatively large, ethnically diverse study population of infants at age 3 to 6 months. Although we studied the relationship of these predictors to overall microbial diversity, community-level microbiome data (bacterial coabundance groupings, multivariate analyses, and hierarchical taxonomy) are the main focus of this work. We assessed the relative abundance of bacterial taxa by 16S rRNA gene sequencing in stool samples from more than 300 infants born to mothers in the Vitamin D Antenatal Asthma Reduction Trial (VDAART), a clinical trial of prenatal vitamin D supplementation and outcomes in pregnant women and their offspring.²²

METHODS**VDAART and ancillary study**

For our analyses, we used characteristics and stool samples of infants born to mothers enrolled in VDAART, a 2-arm, double-blind, placebo controlled, randomized, clinical trial of vitamin D, to determine whether higher vitamin D intake and levels in the pregnant mother will prevent asthma and allergy in

childhood.²² Pregnant women (who had or whose partner had allergies/asthma) were randomized (n = 880) during the first trimester of pregnancy (10-18 weeks) to 1 of 2 treatment arms of a 4-year clinical trial: 4000 IU vitamin D + prenatal vitamins or 400 IU vitamin D + prenatal vitamins. Forty-one percent of infants in VDAART had a maternal history of asthma, whereas 23% had a paternal history of asthma (7% of subjects had both maternal and paternal history of asthma).²² Of the 810 infants, 333 infants donated stool samples.

Infants' characteristics

Characteristics of infants enrolled in VDAART (including the subset enrolled in the gut microbiome ancillary study) were ascertained through medical records, questionnaires, and data obtained at the time of delivery. Information on antibiotics administration during labor, mode of delivery, and sex was gathered at the time of delivery. Gestational age at birth was calculated on the basis of estimated conception date (using data from first ultrasound) and date of delivery. Infants' racial/ethnic group was derived from initial enrollment data on race/ethnicity of each biological parent. The "black" race/ethnicity category included infants of African ancestry only, as well as those with mixed African and Hispanic ancestry. Infants of Native American or Asian descent were included in the "Other" race category. Questionnaires administered to the mother at 6 months after birth were used to assess breast-feeding. We categorized infants into 1 of 3 groups: exclusively breast-fed, breast-fed with some formula supplementation, and exclusively formula-fed. Infants in the exclusively breast-fed category and those breast-fed with some formula supplementation were each compared with exclusively formula-fed infants (reference category). Information about the specific timing of solid food introduction was gathered at the 18-month follow-up questionnaire (58% of infants). If these data were missing, we assessed whether solid food had been introduced at the 6-month questionnaire. Additional characteristics (gestational diabetes, infant day care attendance, cigarette smoking during pregnancy, maternal body mass index at first obstetrician visit, vitamin D treatment group, cord blood vitamin D level) as well as home characteristics (dog ownership, cat ownership, mold/mildew on home walls or surfaces, and water damage in the home) were also examined as potential predictors of the infant gut microbiome.

Vitamin D assessment in cord blood

Liquid chromatography-mass spectrometry was performed to measure 25(OH)vitamin D level in 261 infants' umbilical cord blood samples. For details, see this article's [Online Repository](http://www.jacionline.org) at www.jacionline.org.

Stool sample collection and sequencing of bacterial 16S rRNA gene

Mothers were asked to collect ½ teaspoon of their infant's stool from a dirty diaper using a tongue depressor, and to store the sample in the freezer, and bring it frozen to the next clinical visit. Stool samples were not collected if the infant had taken antibiotics within the past 7 days. Participants stored their infant's stools sample in their home freezer for 1 to 2 days before bringing it to their clinical visit, whereupon the samples were immediately stored at -80°C. A total of 333 infant stool samples (age 3-6 months) were collected. DNA extractions were performed on stool samples, and the bacterial rRNA 16S gene (V3-V5 hypervariable regions) was amplified. Pyrosequencing data (Roche 454 Titanium) were produced by the Genome Center at Washington University in St Louis, Missouri. For quality filtering and bioinformatic processing of sequences, see this article's [Online Repository](http://www.jacionline.org) at www.jacionline.org.

Statistical analyses of microbiome data

Diversity analysis. Shannon diversity index and richness were computed using the VEGAN package in R statistical software (<http://www.r-project.org>). The Shannon diversity index accounts for both the number (richness) and distribution (evenness) of taxa.²³ The Shannon diversity index was

TABLE I. Characteristics of infants with age 3-6 months fecal flora sampling*

Characteristic	Infants in VDAART cohort (N = 810)	Infants with fecal flora sampling (N = 333)
Gestational age (wk), mean $\pm$ SD	39.0 $\pm$ 1.9	39.1 $\pm$ 1.9
Child's race		
Black	350 (43.1)	149 (44.7)
Hispanic	112 (13.8)	42 (12.6)
White	215 (26.5)	92 (27.6)
Other	133 (16.4)	50 (15.0)
Male	422 (52.1)	180 (54.1)
Breast-fed, exclusively first 6 mo†	126 (17.0)	71 (21.7)
Breast-fed and formula-supplemented first 6 mo†	287 (38.8)	107 (32.9)
Formula-fed, exclusively first 6 mo†	328 (44.2)	147 (45.4)
C-section birth	240 (29.6)	108 (32.4)
Maternal antibiotics (during labor)†	310 (38.5)	133 (40.1)
Infant antibiotics (first days of life)†	78 (9.7)	26 (7.8)
Age at fecal flora sampling (wk), mean $\pm$ SD	NA	20.1 (4.6)
Cord blood vitamin D (ng/mL), mean $\pm$ SD		22.7 (11.9)

Values are n (%) unless indicated otherwise.

NA, Not applicable/available.

*All comparisons were $P > .05$ for entire VDAART cohort vs subset with stool samples for microbiome analysis.

†Numbers reflect some missing data.

normally distributed, and was analyzed as a continuous outcome in univariate linear regression models. A multiple regression model, using statistically significant ($P < .05$) univariate characteristics, was then constructed.

Coabundance groupings. To capture coabundance groupings (or factor representations) of bacterial taxa within the infant gut, we computed spearman correlation coefficients for the top 25 taxa identified using the Compositionally Corrected by REnormalization and PErmutation package in R biostatistical software, and used the correlation matrix for principal-factor analysis (with varimax rotation). Factor loadings were applied to 16S sequencing counts, to compute an individual's factor score for each bacterial coabundance group. Scores for bacterial coabundance groupings were used as the outcome variables in linear regression models.

Multivariate associations. We used Multivariate Associations with Linear Models (MaAsLin) to further investigate predictors associated with the infant gut microbiome, as a way to capture additional taxa not identified as differentially abundant using the other statistical analysis techniques.^{24,25} This analysis software uses boosted, additive, linear models to detect associations between one group of data (metadata/the predictors) and another group (microbial abundance of specific taxa). In this multivariate approach, all measured taxa are considered as outcomes simultaneously. Results of an MaAsLin analysis show the association between a specific predictor and abundance of one member of the microbial community.

Hierarchical (tree) structure of microbiome

We computed a taxonomic tree structure for each infant's gut microbiome sample (down to the genus level) with the Human Microbiome Project Trees package in R statistical software.²⁶ Infant characteristics showing the strongest and most consistent associations with overall diversity (C-section, breast-feeding) were used as a basis of comparison for taxonomic tree structure (for hierarchical tree comparisons, see [Methods](#) and [Results](#) sections and figures in this article's Online Repository at www.jacionline.org).

RESULTS

Infants' characteristics

Infants with stool samples showed similar characteristics to infants in the whole VDAART cohort ([Table I](#)). Approximately half of the infants were black. More than half of all infants were breast-fed (55%) in the first 6 months of life. Approximately one-third (33%) of infants were breast-fed with formula

supplementation, 22% were exclusively breast-fed, and 45% were exclusively formula-fed at 6 months. One-third of the infants were born by C-section. Solid foods had been introduced for 66% of the infants. The mean age at stool sampling was 5 months (20.1 weeks). Use of antibiotics during labor was relatively common (40%). Average relative abundance of the top 20 most abundant microbial taxa observed in the infant gut is shown by race/ethnic group ([Fig 1](#)), mode of delivery ([Fig 2](#)), and breast-feeding ([Fig 3](#)).

Microbial diversity

The mean Shannon diversity index was 2.2 ± 0.46 (range, 0.5-3.15). Richness (total number of bacterial taxa detected in an infant's stool sample) ranged from 9 to 55, with a median level of 32. A multiple regression model for Shannon diversity (including all variables with $P < .05$ in the univariate analysis) explained 26% (adjusted $R^2 = 26\%$) of the diversity level in infant gut ([Table II](#)). Variables associated with the Shannon index had similar relationships to richness (data not shown). White and other race/ethnicity was associated with decreased diversity as compared with black infants. C-section delivery was associated with higher bacterial diversity in the infant gut. Breast-feeding (both exclusive and breast-feeding with formula supplementation) predicted lower diversity, although the magnitude of the association was greater for exclusive breast-feeding. Age at fecal flora sampling had a small, but statistically significant effect on diversity in univariate models. Solid food intake was associated with increased diversity in the infant gut. Vitamin D treatment group and antibiotic administration to the infant within days after birth were not associated with microbial diversity. Dog and cat ownership was associated with reduced diversity in univariate, but not adjusted models.

Because the infant microbiome in the first 6 months undergoes rapid changes, we examined the possibility that age at fecal flora sampling modifies the effects of mode of delivery or breast-feeding on infant gut diversity. Tests for interactions between C-section and age of fecal flora sampling and breast-feeding and age at fecal flora sampling were not significant

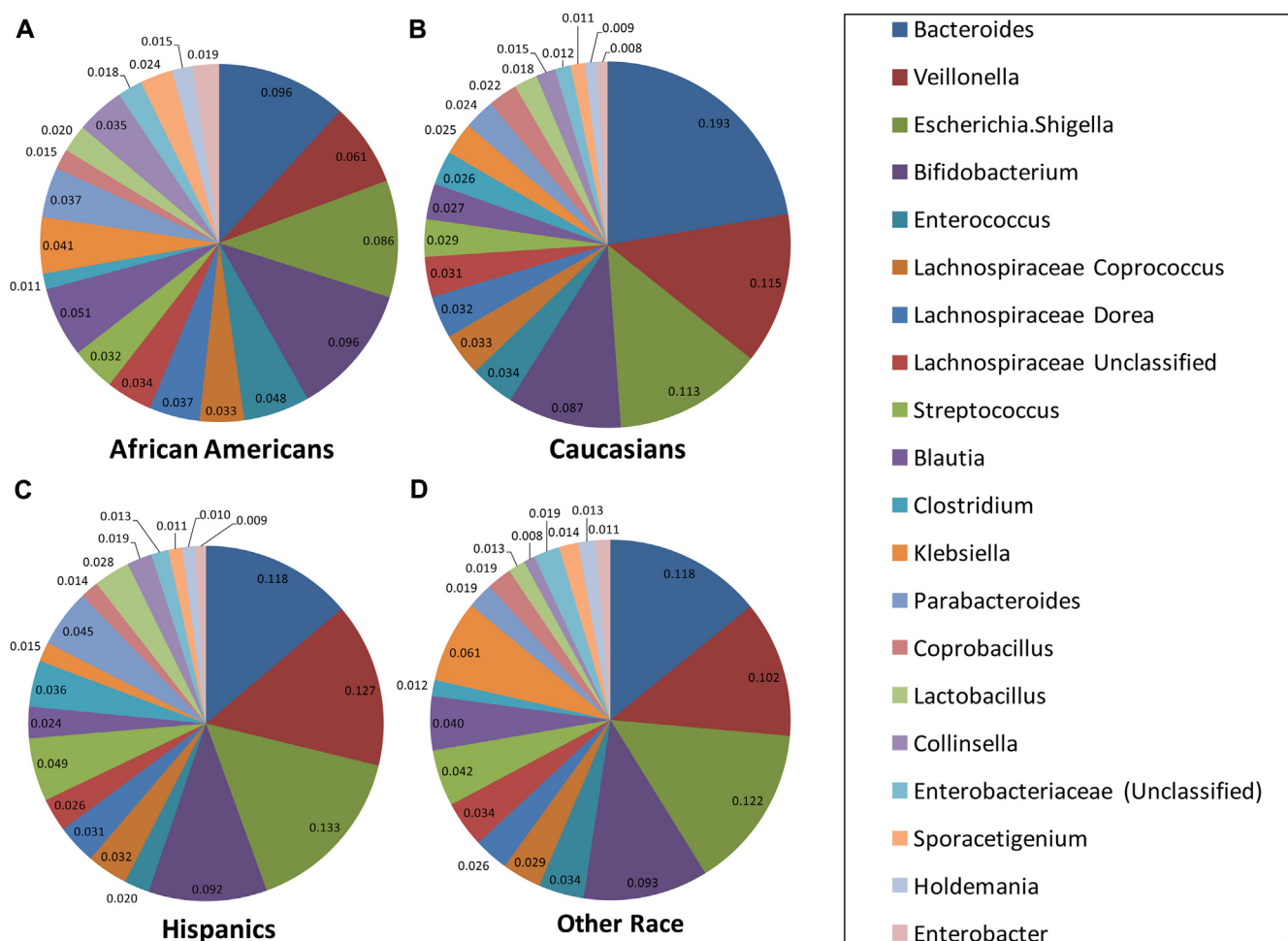


FIG 1. Race and the infant gut microbiome at age 3 to 6 months: composition of the top 20 most abundant taxa. **A**, African Americans. **B**, Caucasians. **C**, Hispanics. **D**, Other race.

($P = .72$ and $P = .53$, respectively) in models for infant gut microbial diversity. A stratified analysis (see Table E9 in this article's Online Repository at www.jacionline.org) by age at fecal flora sampling (categorized at age 12–16 weeks, >16 weeks to 20 weeks, and >20 weeks) demonstrates that mode of delivery, breast-feeding, and race all show similar magnitudes of association with infant gut diversity across this relatively wide age range. Lack of effect modification by fecal flora sampling age suggests that combining all age groups within the same analyses is acceptable for this study.

C-section delivery was associated with reduced odds of breast-feeding in the first 6 months of life (odds ratio, 0.57; 95% CI, 0.36–0.91). We tested for an interaction between these 2 variables in our multiple regression model for infant gut diversity. We did not detect a significant interaction ($P = .84$ for interaction term), suggesting that the impact of breast-feeding on infant gut diversity does not vary by mode of delivery.

Bacterial coabundance groupings

Factor analysis of the top 25 most abundant bacterial taxa in the infant gut (see Table E1 in this article's Online Repository at www.jacionline.org) revealed 4 factors associated with bacterial coabundance (a 4-factor solution was identified on the basis of

a scree plot). The first factor captured coabundance of *Lachnospiraceae* genera (positive loadings, 0.4–0.8), along with *Unclassified (U.) Clostridiales* (loading 0.65), both of which are in the *Firmicutes* phylum. *Bacteroides* was negatively associated with this *Lachnospiraceae/U. Clostridiales* coabundance grouping (loading, -0.31). For the second factor, positive loadings for *Klebsiella*, *Enterobacter*, and *Unclassified Enterobacteriaceae* were most substantial (0.46, 0.76, and 0.86, respectively). The third factor was defined mainly by *Bacteroides* abundance (0.70 for factor loading), with negative factor loadings for *Escherichia/Shigella* (-0.44) and *Bifidobacterium* (-0.47). The fourth factor showed substantial positive loadings for *Veillonella* and *Clostridium*. For ease of interpretability, these coabundance groupings (factors) are labeled in the following way: Factor 1: *Lachnospiraceae* (+), *U. Clostridiales* (+); Factor 2: *Klebsiella* (+), *Enterobacter* (+); Factor 3: *Bacteroides* (+), *Escherichia/Shigella* (–), *Bifidobacterium* (–); and Factor 4: *Veillonella* (+) and *Clostridium* (+). In Table III, factor scores for these 4 bacterial coabundance groups were used as outcomes in multiple regression models with infant characteristics as independent predictors. C-section delivery was associated with higher coabundance scores for factor 1 and factor 4, but with lower scores for factor 3. Breast-feeding (both exclusive and formula-supplemented breast-feeding) was linked to lower factor

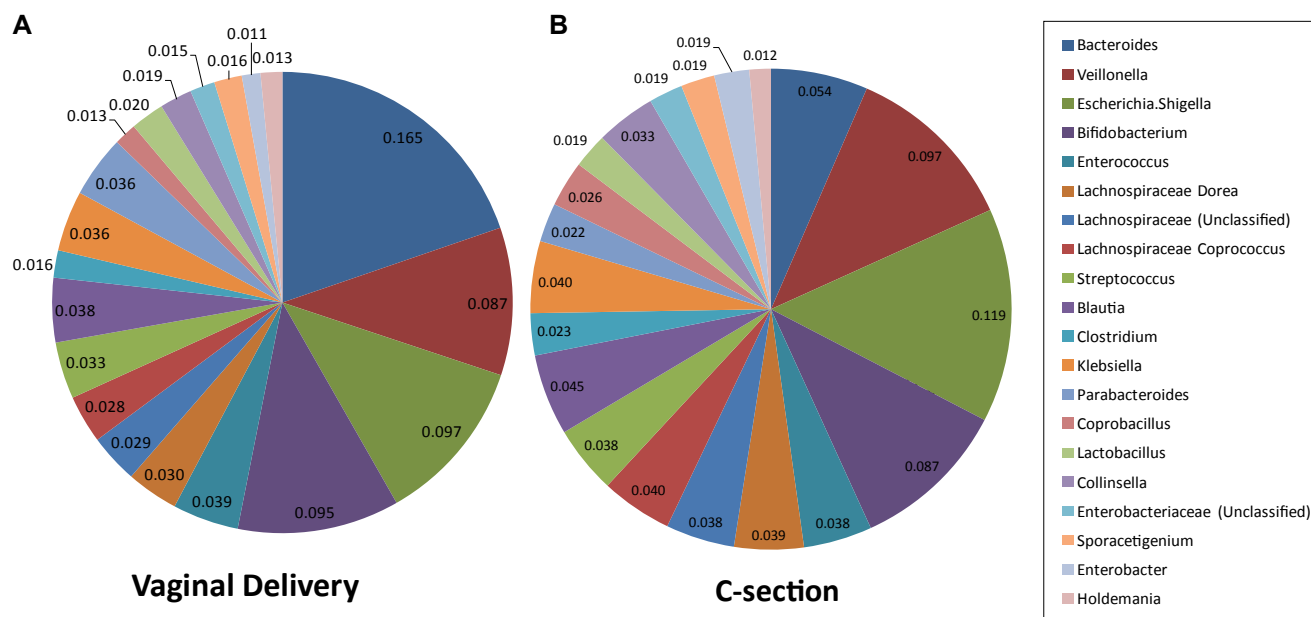


FIG 2. Mode of delivery and the infant gut microbiome at age 3 to 6 months: composition of the top 20 most abundant taxa. **A**, Vaginal delivery. **B**, C-section.

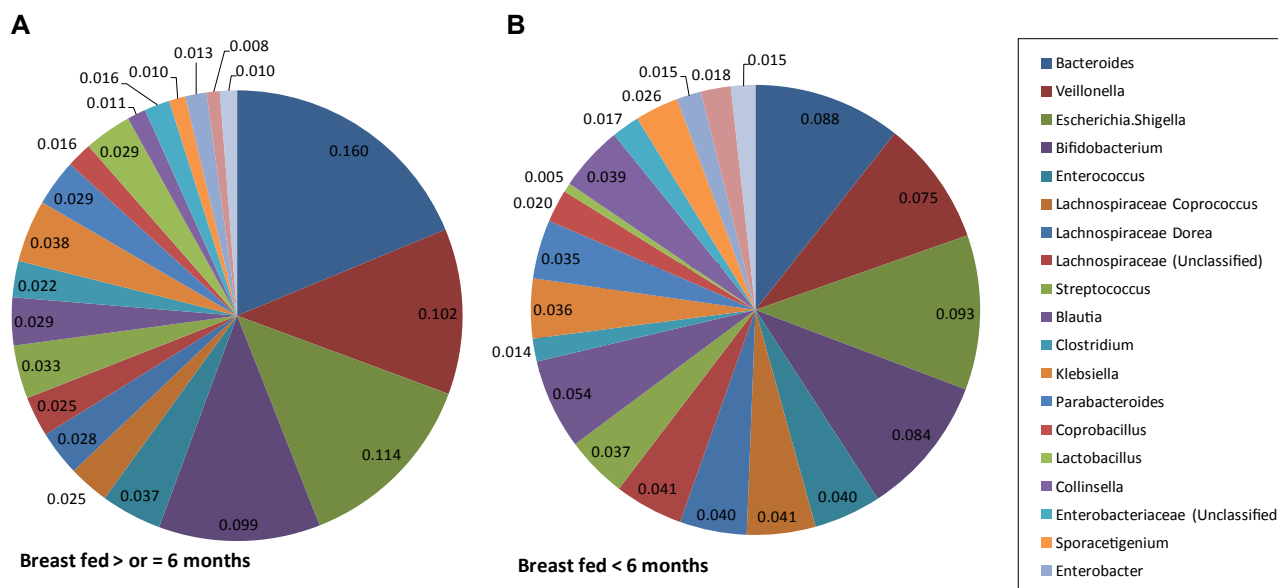


FIG 3. Breast-feeding and the infant gut microbiome at age 3 to 6 months: composition of the top 20 most abundant taxa. **A**, Breast fed ≥ 6 months. **B**, Breast fed < 6 months.

1 scores, but only exclusive breast-feeding was associated with reduced factor 2 scores. Antibiotics administration during labor was associated with higher scores for factor 1 ($P = .07$), but with lower scores for factor 3. Race/ethnicity had an impact on 2 of the 4 bacterial coabundance scores. White infants demonstrated higher factor 3 and a trend for lower factor 2 scores as compared with black infants. All racial groups (white, Hispanic, and other) had higher factor 4 scores as compared with black infants. Cat ownership was associated with higher factor 3 scores and lower factor 4 scores. Solid food intake showed a borderline

statistically significant association with increased factor 3 scores. Cord blood vitamin D levels (ng/mL) were associated with higher scores for factor 1 (Table IV).

Some prenatal characteristics (smoking in pregnancy, gestational diabetes) and environmental variables (water damage, visible mold/mildew in the home) occurred at low frequency in our population (1% to 5%); therefore, we did not have the power to examine the relationships between these variables and bacterial coabundance groupings in the infant gut microbiome.

TABLE II. Predictors of the infant gut microbiome at age 3-6 months (n = 325)

Characteristic	Overall diversity (Shannon diversity index)				
	Univariate associations		Multiple regression model		
	β	P value	β	P value	Model-adjusted R ²
Child's race*					
Hispanic	−0.232	.003	−0.09	.26	26%
White	−0.351	<.0001	−0.15	.02	
Other	−0.258	.004	−0.10	.18	
Breast-fed exclusively†	−0.531	<.0001	−0.45	<.001	
Breast-fed supplemented with formula†	−0.185	.0005	−0.13	.02	
Solid foods introduced (Y/N)	0.127	.02	0.11	.04	
C-section (Y/N)	0.146	.006	0.11	.04	
Maternal Abx labor (Y/N)	0.105	.04	0.008	.85	
Age at fecal sampling (wk)	0.013	.02	0.009	.11	
Maternal BMI at first Ob visit	0.012	.0009	0.006	.10	
Pet dog (Y/N)	−0.140	.01	−0.03	.50	
Pet cat (Y/N)	−0.153	.04	−0.05	.54	
Child Abx in first days of life (Y/N)	0.078	.40			
Treatment group	0.026	.61			
Smoking during pregnancy (Y/N)	0.162	.39			
Gestational diabetes (Y/N)	0.205	.14			
Gestational age (wk)	−0.02	.08			
Female (vs male)	−0.018	.73			
Day care attendance (Y/N)	0.021	.76			
Mold/mildew on walls (Y/N)	−0.087	.39			
Water damage in home (Y/N)	−0.086	.39			
Cord blood vitamin D level (ng/mL)	−0.004	.06			

For all models, beta coefficients and P value bolded if $P < .05$.

Abx, Antibiotics; BMI, body mass index; N, no; Ob, obstetrician; Y, yes.

*Black race/ethnicity used as reference.

†Breast-feeding data from first 6 mo of life; comparison group is exclusively formula-fed.

Multivariate analysis (MaAsLin) results

Multivariate analysis (adjusted for all covariates) showed similar findings, overall, to other statistical analysis methods used. For breast-feeding as a predictor, MaAsLin (see Table E2 in this article's Online Repository at www.jacionline.org) helped to identify decreased abundance of taxa within the family *Clostridiales* (*Coprococcus*, *Ruminococcus* and *Eubacterium*, *Lachnospiraceae* *Dorea*, *Lachnospiraceae* *Blautia*), showing considerable overlap with the lower *Lachnospiraceae*/*Clostridiales* bacterial coabundance scores observed for breast-feeding. These relationships with specific taxa in our multivariate analysis were observed only for exclusive breast-feeding (Table E2), but not for breast-feeding with formula supplementation (see Table E3 in this article's Online Repository at www.jacionline.org). Although breast-feeding was associated with increased *Lactobacillus* ($P < .05$) and a trend toward increased *Bifidobacterium* in univariate analyses, these associations were not observed in the multivariate models. For C-section (see Table E4 in this article's Online Repository at www.jacionline.org), the strongest findings by multivariate analysis were for decreased *Bacteroides* and *Parabacteroides* abundance. With regard to racial comparisons, *Megasphaera* and *Lactococcus* were significantly higher in black infants (see Table E5 in this article's Online Repository at www.jacionline.org). Cord blood vitamin D was associated with increased *Lachnobacterium* (reiterating the finding of increased *Lachnospiraceae*/*Clostridiales* coabundance scores with higher vitamin D). In addition, decreased *Lactococcus* was observed for higher cord blood vitamin D (see Table E6 in this article's Online Repository at www.jacionline.org). Although some

associations were observed for cat ownership and bacterial coabundance grouping analysis (Table III), these findings were not reflected in the multivariate analysis results. None of the taxa showed significant associations with pet ownership after adjustment for multiple comparisons (see Tables E7 and E8 in this article's Online Repository at www.jacionline.org).

DISCUSSION

In the present study, we identified predictors of the infant gut microbiome at age 3 to 6 months, in one of the largest, most ethnically diverse study populations to date. In addition, compared with other studies investigating the microbiome in early infancy, this is a relatively healthy cohort of infants. In this early life window, the infant's gut microbiota is plastic and undergoing rapid change, with potential implications for innate and adaptive immune system maturation, and for future overall health. Therefore, it is critical to understand the factors associated with microbial communities populating the infant gut during this key developmental time period. Through the use of complementary analytical methodologies, we examined the relationships between infant characteristics and gut microbiota at age 3 to 6 months. As a first step, we identified predictors of microbial diversity. Next, we determined specific compositional differences in gut microbiota by infant characteristic, using bacterial coabundance groupings (based on the top 25 most abundant taxa). Last, we used multivariate analysis to assess the relationship between multiple independent predictors and the relative abundance of all detectable taxa simultaneously.

TABLE III. Infant predictors of bacterial coabundance factor scores (linear regression results, n = 325)

Dominant phyla for grouping		Multiple regression models for bacterial coabundance groupings							
Dominant genera	Firmicutes <i>Lachnospiraceae</i> (+) <i>Clostridiales</i> U (+)	Proteobacteria <i>Klebsiella</i> (+) <i>Enterobacter</i> (+)	Bacteroidetes <i>Bacteroides</i> (+) <i>Escherichia</i> (–) <i>Bifidobacterium</i> (–)	Firmicutes <i>Veillonella</i> (+) <i>Clostridium</i> (+)					
		Factor 1 score		Factor 2 score		Factor 3 score		Factor 4 score	
		β	P value	β	P value	β	P value	β	P value
Infant characteristics									
Gestational age (wk)		–0.03	.43	–0.02	.37	0.001	.97	–0.02	.37
Child's race*									
Hispanic		0.06	.70	–0.23	.09	0.07	.69	0.54	.003
White		0.08	.56	–0.14	.20	0.29	.05	0.42	.002
Other		0.10	.47	0.05	.64	–0.05	.78	0.31	.04
Male (vs female)		–0.31	.002	0.04	.62	0.06	.55	0.02	.96
Breast-fed exclusively†		–0.85	<.001	–0.27	.02	–0.23	.17	0.22	.07
Breast-fed supplemented with formula†		–0.28	.02	0.01	.90	0.01	.83	0.08	.50
Solid foods introduced (Y/N)		0.08	.47	0.15	.13	0.23	.07	–0.003	.90
C-section (Y/N)		0.27	.02	0.09	.33	–0.29	.01	0.14	.05
Maternal Abx labor (Y/N)		0.19	.07	–0.04	.64	–0.22	.04	–0.006	.71
Infant Abx, first days of life (Y/N)		–0.02	.95	–0.05	.80	0.18	.42	–0.35	.06
Age at fecal flora sampling (wk)		0.01	.34	–0.02	.06	–0.009	.48	0.004	.71
Dog ownership (Y/N)		–0.01	.91	–0.14	.16	–0.03	.80	–0.09	.40
Cat ownership (Y/N)		0.29	.07	–0.21	.11	0.50	.003	–0.31	.03
Daycare attendance (Y/N)		–0.02	.91	–0.12	.31	0.02	.90	0.13	.31
Maternal BMI		0.003	.69	–0.006	.26	–0.002	.74	0.004	.56

Beta and P values bolded if $P < .05$.

Abx, Antibiotics; BMI, body mass index; N, No; Y, yes.

*Black race/ethnic group as the reference category.

†Breast-feeding data from first 6 mo of life; reference group is exclusively formula-fed.

TABLE IV. Cord blood vitamin D bacterial coabundance factor scores (linear regression results)* (n = 261)

Dominant phyla for grouping		Bacterial coabundance groupings							
Dominant genera	Firmicutes <i>Lachnospiraceae</i> (+) <i>Clostridiales</i> U (+)	Proteobacteria <i>Klebsiella</i> (+) <i>Enterobacter</i> (+)	Bacteroidetes <i>Bacteroides</i> (+) <i>Escherichia</i> (–) <i>Bifidobacterium</i> (–)	Firmicutes <i>Veillonella</i> (+) <i>Clostridium</i> (+)					
		Factor 1 score		Factor 2 score		Factor 3 score		Factor 4 score	
		β	P value	β	P value	β	P value	β	P value
Cord blood vitamin D (ng/mL)		0.014	.004	–0.003	.40	0.007	.19	–0.006	.19

*Models adjusted for race/ethnicity, gestational age, breast-feeding, C-section, maternal antibiotics during labor, infant antibiotics in the first days of life. Beta and P values are in boldface if $P < .05$.

Overall mean diversity within the infant gut microbiome was consistent with other studies of fecal flora within this age range.^{27,28} On average, the infant gut microbiota were dominated by gram-negative anaerobes (*Bacteroides*, *Veillonella*), with significant representations of gram-positive anaerobes (*Bifidobacterium*) and facultative microorganisms (*Escherichia/Shigella*). Extraction of 4 factors from 16S sequence data across all infants' stool samples revealed a latent phylum-specific substructure of microbiota in the infant gut that demonstrated the following relationships among taxa: (1) *Lachnospiraceae* co-occur with other *Clostridiales* (Firmicutes), (2) *Klebsiella* and *Enterobacteriaceae* (both Proteobacteria) show coabundance, (3) *Bacteroides* (Bacteroidetes) is associated with decreased abundance of *Bifidobacterium* (Actinobacteria) and *Escherichia/Shigella* (Proteobacteria), and last (4) *Veillonella* and *Clostridium* (both Firmicutes) show coabundance. This latent

substructure of the infant gut microbiome, identified using factor analysis, was associated with 4 predictors: mode of delivery, breast-feeding, race, and vitamin D. Furthermore, these groupings capture taxa with a prior association with disease status.

Mode of delivery

Mode of delivery was consistently one of the most important factors associated with the infant gut microbiome in our study. The relationship between delivery mode and the microbial composition of the infant gut likely originates from different source contributions of microbiota (infants born vaginally are exposed to both maternal fecal and vaginal flora, whereas infants born by C-section may have increased colonization from maternal skin or environmental microbes).²⁹ Surprisingly, infants born by C-section had higher Shannon indices than did infants born

vaginally. On the surface, this finding is counterintuitive, and conflicts with other reports of decreased gut diversity with C-section.^{30,31} It may be that the increased evenness of the overall taxonomic profiles in C-section-born infants (as a result of lower predominance of *Bacteroides*) may partially explain the relationship between C-section and increased Shannon diversity.

The relationship between C-section delivery and lower abundance of *Bacteroides* was consistent across all our statistical models (bacterial coabundance, multivariate analyses, and hierarchical tree comparison (see this article's [Online Repository](http://www.jacionline.org) at www.jacionline.org), and has been reported previously in other studies.^{29,30,32} However, in addition to this association, our bacterial coabundance analysis may indicate that with decreased *Bacteroides* abundance, *Proteobacteria* (*Escherichia*) thrive more in the gut microbiota of infants born by C-section, replacing *Bacteroidetes* as the dominant gram-negative phylum. This alternative gram-negative profile (dominant *Proteobacteria* with minimal *Bacteroides*) may have important consequences for immune stimulation by pathogen-associated molecular patterns (ie, heterogeneity of endotoxin TLR4 stimulation), or even potential implications for microbial function in the gut. Jakobsson et al³⁰ have demonstrated that lower *Bacteroidetes* abundance in C-section-born infants is associated with lower levels of T_H1-associated chemokines (CXCL10 and CXCL11) during infancy, as well as reduced T_H1 responses, which may ultimately increase the risk of asthma or allergic disease.

Breast-feeding

Breast-feeding was another independent predictor linked to altered gut microbiota across different types of microbiome analyses. Breast-feeding has the potential to shape the gut microbiome as a source of prebiotics, or by directly interacting with microbes via its IgA content.³³ Breast-feeding was associated with reduced diversity (potentially driven by reduced evenness of taxa), a finding that is consistent with other reports.^{32,34} We also observed a trend for increased *Bifidobacterium* with breast-feeding, which may have resulted from increased availability of prebiotics in breast milk (oligosaccharides)³⁵ that favor proliferation of this taxa. Interestingly, breast-feeding was associated with a lower score for the main *Firmicutes* coabundance grouping (representing *Lachnospiraceae* and other *Clostridiales*), and in multivariate analysis, lower abundance of specific genera within *Clostridiales* (*Ruminococcus*, *Coprococcus*, *Eubacterium*). Lower abundance of *Clostridia* in breast-fed infants has also been reported in much smaller, more homogenous study populations and is replicated here in a much larger, ethnically diverse cohort.^{31,32} In emerging studies, *Clostridia* (*C. difficile*) abundance in the infant gut at age 1 month has been linked to wheeze and sensitization in early life, and total *Clostridia* abundance is associated with atopic dermatitis risk.^{36,37} If the protective effects of breast-feeding on allergic disease risk are indeed mediated by the gut microbiome, *Clostridia* may have an important role. Exclusive, rather than formula-supplemented, breast-feeding may also be crucial, as our bacterial coabundance grouping analysis showed the greatest impact on *Clostridiales* reduction for exclusive breast-feeding, and our multivariate findings showed reductions in *Clostridiales* for only those infants who were exclusively breast-fed.

Race/ethnicity

Race/ethnicity was the third main factor associated with alterations in the infant gut microbiome in our work and, as a predictor of gut microbiota, is probably the least understood. Gut microbiota may vary by race because of differences in dietary patterns, personal care product use, or host genetics. White infants showed reduced overall diversity of the gut microbiome but also demonstrated higher abundance of *Bacteroides*. In keeping with the overall negative correlations we have seen for *Bacteroides* and *Proteobacteria*, white infants demonstrated a trend toward lower scores for the *Proteobacteria* bacterial coabundance group, represented mainly by *Enterobacteriaceae* (*Klebsiella* and *Enterobacter*). Lower abundance of taxa with potential protective benefits (ie, *Bacteroides*) in the microbiome of minority children could, at least partially, explain the increased asthma incidence and morbidity in these populations. Further investigation is needed to uncover why infant gut microbial communities vary by race. Although these relationships were observed after adjusting for potential confounding due to factors such as breast-feeding or mode of delivery, there is almost certainly residual confounding by factors tightly associated with race (ie, factors linked to socioeconomic status), which could not be accounted for in these analyses. For example, race was associated with breast-feeding type (Hispanics and whites were more likely to exclusively breast-feed than black subjects), suggesting that race could also be linked to other unmeasured infant dietary trends.

Cord blood vitamin D

Our findings suggest that vitamin D levels *in utero* (and shortly after birth) could influence the microbial community that colonizes the infant gut. We observed higher levels of *Lachnospiraceae/U. Clostridiales* in infants with higher cord blood vitamin D, and multivariate analysis also revealed an increased abundance of *Lachnospiraceae*, but a decreased abundance of *Lactococcus*. Although very few studies have examined the impact of vitamin D on the developing gut microbiome, data from animal models suggest that vitamin D may influence microbial community composition. In vitamin D receptor knockout mice, *Lactobacillus* was depleted, whereas *Clostridium* and *Bacteroides* were enriched.³⁸ In an animal model of ulcerative colitis (where mice lacked the enzyme to produce the active form of vitamin D (1,25 (OH) 2D3)), supplementation with either the active form of vitamin D (1,25 (OH) 2D3) or antibiotic treatment ameliorated the disease, suggesting that microbial dysbiosis may arise when vitamin D is insufficient, thereby promoting severe inflammation in the gut.³⁹ Future work is needed to understand whether vitamin D plays a key role in dictating which microbes colonize the gut, or whether it potentially modifies the effects of the microbiome, without directly determining community composition.

Conclusions

Until further translational studies are conducted, the precise role of altered gut communities on immune system function and risk for allergic disease will remain unknown. In our study, infant characteristics that may potentially increase the risk of asthma and allergic disease development (black race/ethnicity, C-section birth, and formula feeding) were associated with alterations in the infant gut microbiome profile. Both black race/ethnicity and

C-section were independent predictors of lower levels of *Bacteroides*. Formula feeding was linked to higher levels of *Clostridia* in the infant gut, which may be associated with increased risk of allergic disease. Future work is required to uncover how alterations in gut microbial communities influence diseases of immune origin, and will necessitate functional genetic and possibly metabolomic approaches to complete our understanding of how the gut microbiome mediates asthma and allergic disease risk. Greater taxonomic resolution (down to species and/or strain level if possible) will also be critical. Last, it is important to note that this work is a cross-sectional study of the infant gut microbiome conducted at age 3 to 6 months, when the microbiome is still highly unstable. Longitudinal analysis to assess the impact of early life factors on the developing microbiome over time is needed, and, with respect to health outcomes analyses, repeated measures in infancy/early childhood will likely give a more definitive test of the microbiome's relationship to asthma and allergy.

The VDAART investigators acknowledge the enormous contributions to the study by Dr Robert Strunk, an esteemed colleague and friend. Dr Strunk was instrumental in the design, implementation, and success of VDAART and continuation study. We are forever grateful for his guidance, input, and collegiality.

Key messages

- Race, mode of delivery, and formula-feeding, which are all asthma risk factors, are associated with altered taxonomic composition of the infant gut microbiome.
- Work is ongoing to determine whether factors associated with decreased asthma risk (white race/ethnicity, vaginal delivery, breast-feeding) are mediated through changes in the early infant gut microbiome and whether these changes in the microbiome persist as the children get older.

REFERENCES

- Maynard CL, Elson CO, Hatton RD, Weaver CT. Reciprocal interactions of the intestinal microbiota and immune system. *Nature* 2012;489:231-41.
- Young VB. The intestinal microbiota in health and disease. *Curr Opin Gastroenterol* 2012;28:63-9.
- Guarner F, Malagelada JR. Gut flora in health and disease. *Lancet* 2003;361:512-9.
- Salminen S, Bouley C, Boutron-Ruault MC, Cummings JH, Franck A, Gibson GR, et al. Functional food science and gastrointestinal physiology and function. *Br J Nutr* 1998;80:S147-71.
- Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* 2009;9:313-23.
- Rask C, Evertsson S, Telemo E, Wold AE. A full flora, but not monocolonization by *Escherichia coli* or lactobacilli, supports tolerogenic processing of a fed antigen. *Scand J Immunol* 2005;61:529-35.
- Strauch UG, Obermeier F, Grunwald N, Gunster S, Dunger N, Schultz M, et al. Influence of intestinal bacteria on induction of regulatory T cells: lessons from a transfer model of colitis. *Gut* 2005;54:1546-52.
- Sudo N, Sawamura S, Tanaka K, Aiba Y, Kubo C, Koga Y. The requirement of intestinal bacterial flora for the development of an IgE production system fully susceptible to oral tolerance induction. *J Immunol* 1997;159:1739-45.
- Abrahamsson TR, Jakobsson HE, Andersson AF, Bjorksten B, Engstrand L, Jenmalm MC. Low gut microbiota diversity in early infancy precedes asthma at school age. *Clin Exp Allergy* 2014;44:842-50.
- Kalliomaki M, Kirjavainen P, Eerola E, Kero P, Salminen S, Isolauri E. Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *J Allergy Clin Immunol* 2001;107:129-34.
- Ochoa-Reparaz J, Mielcarz DW, Ditrio LE, Burroughs AR, Begum-Haque S, Dasgupta S, et al. Central nervous system demyelinating disease protection by the human commensal bacteroides fragilis depends on polysaccharide A expression. *J Immunol* 2010;185:4101-8.
- Mazmanian SK, Round JL, Kasper DL. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* 2008;453:620-5.
- Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* 2005;122:107-18.
- Murosaki S, Yamamoto Y, Ito K, Inokuchi T, Kusaka H, Ikeda H, et al. Heat-killed lactobacillus plantarum L-137 suppresses naturally fed antigen-specific IgE production by stimulation of IL-12 production in mice. *J Allergy Clin Immunol* 1998;102:57-64.
- Shida K, Takahashi R, Iwadata E, Takamizawa K, Yasui H, Sato T, et al. Lactobacillus casei strain shirota suppresses serum immunoglobulin E and immunoglobulin G1 responses and systemic anaphylaxis in a food allergy model. *Clin Exp Allergy* 2002;32:563-70.
- Johansson MA, Sjogren YM, Persson JO, Nilsson C, Sverremark-Ekstrom E. Early colonization with a group of lactobacilli decreases the risk for allergy at five years of age despite allergic heredity. *PLoS One* 2011;6:e23031.
- DiGiulio DB, Romero R, Amogan HP, Kusanovic JP, Bik EM, Gotsch F, et al. Microbial prevalence, diversity and abundance in amniotic fluid during preterm labor: a molecular and culture-based investigation. *PLoS One* 2008;3:e3056.
- Jimenez E, Fernandez L, Marin ML, Martin R, Odriozola JM, Nueno-Palop C, et al. Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section. *Curr Microbiol* 2005;51:270-4.
- Gronlund MM, Grzeskowiak L, Isolauri E, Salminen S. Influence of mother's intestinal microbiota on gut colonization in the infant. *Gut Microbes* 2011;2:227-33.
- Matamoros S, Gras-Leguen C, Le Vacon F, Potel G, de La Cochetiere MF. Development of intestinal microbiota in infants and its impact on health. *Trends Microbiol* 2013;21:167-73.
- Bezirtzoglou E. The intestinal microflora during the first weeks of life. *Anaerobe* 1997;3:173-7.
- Litonjua AA, Lange NE, Carey VJ, Brown S, Laranjo N, Harshfield BJ, et al. The Vitamin D Antenatal Asthma Reduction Trial (VDAART): rationale, design, and methods of a randomized, controlled trial of vitamin D supplementation in pregnancy for the primary prevention of asthma and allergies in children. *Contemp Clin Trials* 2014;38:37-50.
- Morgan XC, Huttenhower C. Chapter 12: human microbiome analysis. *PLoS Comput Biol* 2012;8:e1002808.
- Morgan XC, Huttenhower C. Meta-omic analytic techniques for studying the intestinal microbiome. *Gastroenterology* 2014;146:1437-48.e1.
- Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012;13:R79.
- La Rosa PS, Shands B, Deych E, Zhou Y, Sodergren E, Weinstock G, et al. Statistical object data analysis of taxonomic trees from human microbiome data. *PLoS One* 2012;7:e48996.
- Valles Y, Artacho A, Pascual-Garcia A, Ferrus ML, Gosalbes MJ, Abellan JJ, et al. Microbial succession in the gut: directional trends of taxonomic and functional change in a birth cohort of Spanish infants. *PLoS Genet* 2014;10:e1004406.
- Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A* 2010;107:11971-5.
- Biasucci G, Rubini M, Riboni S, Morelli L, Bessi E, Retetangos C. Mode of delivery affects the bacterial community in the newborn gut. *Early Hum Dev* 2010;86:13-5.
- Jakobsson HE, Abrahamsson TR, Jenmalm MC, Harris K, Quince C, Jernberg C, et al. Decreased gut microbiota diversity, delayed bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section. *Gut* 2014;63:559-66.
- Azad MB, Konya T, Maughan H, Guttman DS, Field CJ, Chari RS, et al. Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. *CMAJ* 2013;185:385-94.
- Hesla HM, Stenius F, Jaderlund L, Nelson R, Engstrand L, Alm J, et al. Impact of lifestyle on the gut microbiota of healthy infants and their mothers—the ALADDIN birth cohort. *FEMS Microbiol Ecol* 2014;90:791-801.
- Hanson LA. Session 1: feeding and infant development breast-feeding and immune function. *Proc Nutr Soc* 2007;66:384-96.
- Azad MB, Konya T, Persaud RR, Guttman DS, Chari RS, Field CJ, et al. Impact of maternal intrapartum antibiotics, method of birth and breastfeeding on gut

- microbiota during the first year of life: a prospective cohort study. *BJOG* 2016;123: 983-93.
35. Bode L. Human milk oligosaccharides: prebiotics and beyond. *Nutr Rev* 2009;67: S183-91.
36. van Nimwegen FA, Penders J, Stobberingh EE, Postma DS, Koppelman GH, Kerkhof M, et al. Mode and place of delivery, gastrointestinal microbiota, and their influence on asthma and atopy. *J Allergy Clin Immunol* 2011;128: 948-55, e1-3.
37. Penders J, Gerhold K, Stobberingh EE, Thijs C, Zimmermann K, Lau S, et al. Establishment of the intestinal microbiota and its role for atopic dermatitis in early childhood. *J Allergy Clin Immunol* 2013;132:601-7.e8.
38. Jin D, Wu S, Zhang YG, Lu R, Xia Y, Dong H, et al. Lack of vitamin D receptor causes dysbiosis and changes the functions of the murine intestinal microbiome. *Clin Ther* 2015;37:996-1009.e7.
39. Ooi JH, Li Y, Rogers CJ, Cantorna MT. Vitamin D regulates the gut microbiome and protects mice from dextran sodium sulfate-induced colitis. *J Nutr* 2013;143:1679-86.

METHODS

Stool sample collection and sequencing of bacterial 16S rRNA gene

Mothers enrolled in VDAART were given a stool sample collection kit when their infants were between age 3 and 6 months. One to 2 days before their clinical appointment, mothers were asked to collect ½ teaspoon of their infant's stool from a dirty diaper using a tongue depressor, and to store the sample in the freezer. Stool samples were not collected if the infant had taken antibiotics within the past 7 days. A total of 333 infant stool samples were collected. DNA extractions were performed on stool samples, and the bacterial rRNA 16S gene (V3-V5 hypervariable regions) were amplified. Pyrosequencing data (Roche 454 Titanium) were produced by the Genome Center at Washington University in St Louis, Missouri. Quality filtering, trimming, and chimera checking were performed as previously described.^{E1} In brief, the protocol allows 1 mismatch in the primer and 0 mismatches in the barcode. Chimera slayer software was used to remove chimeric reads.^{E2} All samples that passed quality control screening ($n = 325$) were classified from phylum to genus level at a confidence threshold of 0.5 using the Ribosomal Database Project Naive Bayesian Classifier version 2.2, training set 6.^{E3} The reads whose taxonomic assignments were lower than a 0.5 confidence threshold were assigned to the unclassified group. All samples were scaled to 1000 reads before analysis.

Vitamin D assessment in cord blood

Liquid chromatography-mass spectrometry was performed to measure 25(OH)vitamin D level in 261 infants' umbilical cord blood samples collected. Calibration was performed using control solutions (1-200 ng/mL) diluted from standard stock solutions provided by Calbiochem (San Diego, Calif). Samples from study subjects were prepared and analyzed through a turbulent flow LC system (Cohesive Technologies, Franklin, Mass) followed by traditional laminar flow chromatography. The analysis was performed using a TSQ Quantum Ultra triple mass-spectrometer (Thermo Finnigan Corp, San Jose, Calif).

Hierarchical (tree) structure of microbiome

We computed a taxonomic tree structure for each infant's gut microbiome sample (down to the genus level) with the Human Microbiome Project Trees package in R statistical Software. Comparison of taxonomic trees using this method accounts for the hierarchical nature of the data, and allows for visualization of microbial community differences at multiple taxonomic levels (phylum, class, order, family, and genus). Known taxonomic relationships from the Ribosomal Database Project were used to construct taxonomic trees, and tree validity was assessed for each stool sample. Infant characteristics showing the strongest and most consistent associations with overall diversity (C-section, breast-feeding) were used as a basis of comparison for taxonomic tree structure. For these comparisons (based on object-oriented data analysis),^{E4} maximum likelihood estimation was used to derive a central or modal tree for one group (ie, infants delivered by C-section) versus another (ie, infants delivered vaginally). After computing and plotting the MLE trees for each group, we used the Human Microbiome Project Trees package to determine whether the central trees computed for each group were different from an overall central/modal tree (P values were computed to reject the null hypothesis that central trees from each group were the same as a grand central tree across all subjects). Plots of central trees at the family level (rather

than the genus level, for ease of visibility) were color coded, with blue representing the branches with the highest confidence, and red representing branches with the lowest confidence level. Results of this analysis are presented in this article's [Online Repository](http://www.jacionline.org) at www.jacionline.org.

RESULTS

Hierarchical (tree) structure of microbiome

A check of tree validity demonstrated a valid hierarchical tree structure for each infant's stool sample. Each infant's microbiome tree structure served as input for calculating overall central or modal hierarchical trees. On the basis of a pairwise likelihood ratio test comparing the central, modal tree for C-section (vs the central tree for vaginal delivery), we rejected the null hypothesis that the tree distribution parameters are the same regardless of mode of delivery ($P = .018$, for 1000 bootstraps). Similarly, in a likelihood ratio test comparing central, hierarchical taxonomic trees of breast-fed versus formula-fed infants (or infants breast-fed <6 months), we also detected a difference in tree structure that had borderline statistical significance ($P = .06$). Plots of the central trees revealed differences in the individual branches that are likely responsible for the overall differences in tree structure (Figs E1 and E2). For instance, in the central tree plot for infants born by C-section, the confidence value for the tree branch representing the *Bacteroidetes* family is much less than it is for the vaginal delivery tree (where this branch of the tree is featured predominantly). Conversely, tree branches for *Fusobacteriaceae*, *Clostridiaceae*, and *Peptostreptococcaceae* are featured at a greater confidence level in the central tree for C-section delivery. Although the root tree branches are essentially the same for comparison of the central tree plots by breast-feeding, bacterial taxonomic families in the leaves of the tree have different levels of confidence by breast-feeding. For instance, the central tree structure for infants breast-fed at least 6 months shows greater confidence in *Lactobacillus*, *Fusobacteriaceae*, and *Enterobacteriaceae* and lower confidence in the *Rikenellaceae* branch. Hierarchical tree comparisons for race/ethnicity (blacks vs all other races) and cord blood vitamin D level (>30 ng/mL vs ≤30 ng/mL) were not statistically significant ($P > .2$).

REFERENCES

- E1. Bezirtoglou E. The intestinal microflora during the first weeks of life. *Anaerobe* 1997;3:173-7.
- E2. Litonjua AA, Lange NE, Carey VJ, Brown S, Laranjo N, Harshfield BJ, et al. The Vitamin D Antenatal Asthma Reduction Trial (VDAART): rationale, design, and methods of a randomized, controlled trial of vitamin D supplementation in pregnancy for the primary prevention of asthma and allergies in children. *Contemp Clin Trials* 2014;38:37-50.
- E3. Morgan XC, Huttenhower C. Chapter 12: human microbiome analysis. *PLoS Comput Biol* 2012;8:e1002808.
- E4. Morgan XC, Huttenhower C. Meta'omic analytic techniques for studying the intestinal microbiome. *Gastroenterology* 2014;146:1437-48.e1.

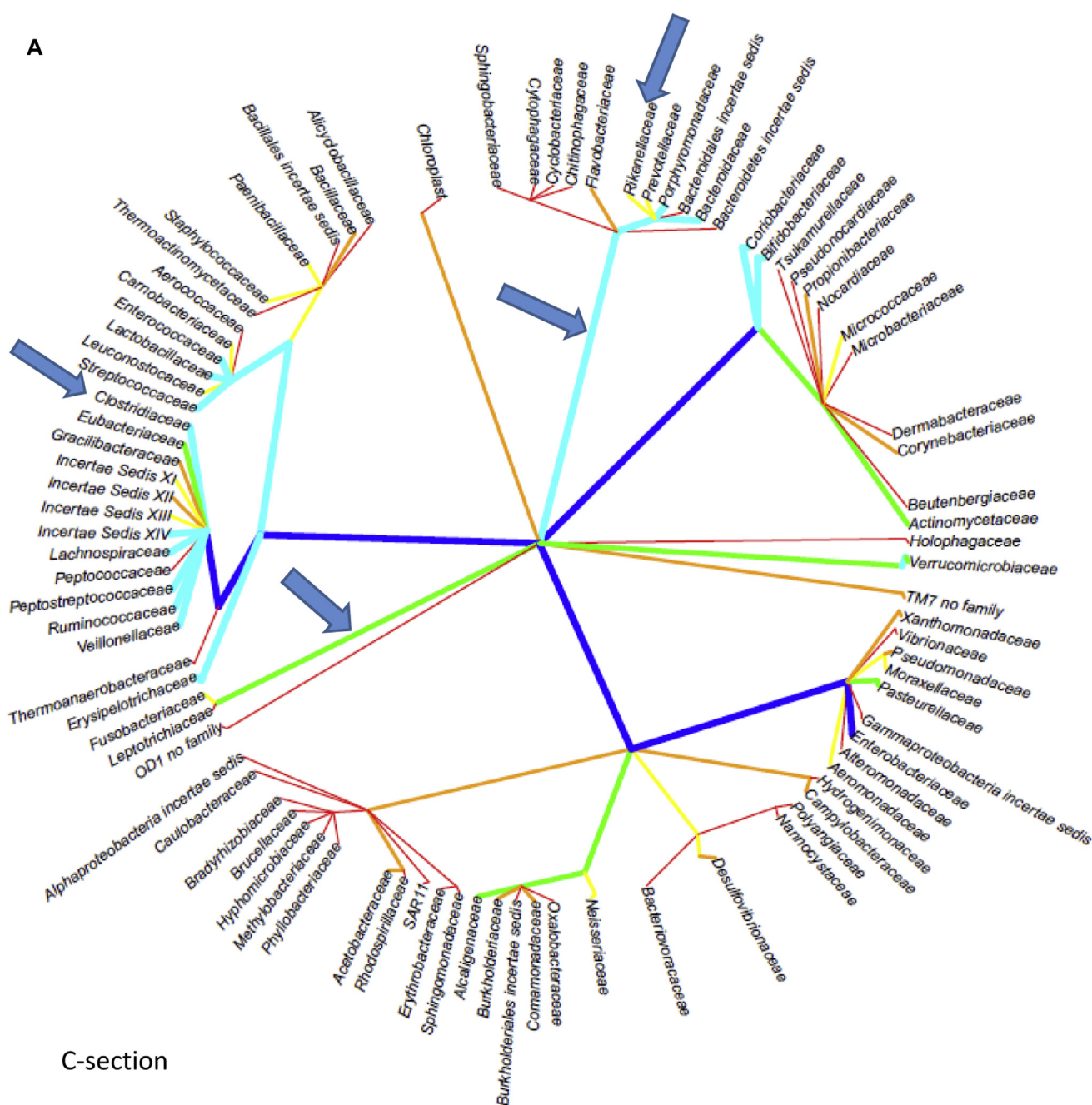


FIG E1. Central Human Microbiome Project Trees structure (family level) for C-section (**A**) and vaginal delivery (**B**); differences in hierarchical tree branches (above a confidence value of 10) with C-section are shown with blue arrows ($P = .018$ for comparison between 2 trees).

B

A phylogenetic tree illustrating the relationships between various bacterial families. The tree is rooted at the center and branches outwards. The branches are color-coded: blue for major clades, red for intermediate clades, yellow for smaller clades, green for specific families, and cyan for others. The families listed include:

- Bacillales incertae sedis
- Sporolactobacillaceae
- Staphylococcaceae
- Thermoactinomycetaceae
- Aerococcaceae
- Carnobacteriaceae
- Lactobacillaceae
- Leuconostocaceae
- Streptococcaceae
- Clostridiaceae
- Eubacteriaceae
- Gracilibacteraceae
- Incertae Sedis XI
- Incertae Sedis XII
- Incertae Sedis XIII
- Incertae Sedis XIV
- Lachnospiraceae
- Peptococcaceae
- Peptostreptococcaceae
- Ruminococcaceae
- Veillonellaceae
- Natranaerobiaceae
- Erysipelotrichaceae
- Fusobacteriaceae
- Leptotrichaceae
- Planctomycetaceae
- Alphaproteobacteria incertae sedis
- Caulobacteraceae
- Aurantimonadaceae
- Bradyrhizobiaceae
- Brucelliaceae
- Hypomicrobiaceae
- Phyllobacteriaceae
- Rhizobiaceae
- Acetivibrioaceae
- Rhodospirillaceae
- Erythrobacteraceae
- Sphingomonadaceae
- Alcaligenaceae
- Burkholderiaceae
- Burkholderiales incertae sedis
- Oxalobacteraceae
- Comamonadaceae
- Methylophilaceae
- Mitrospiraceae
- Nitrosomonadaceae
- Rhodocyclaceae
- Desulfobacteraceae
- Desulfosarcinaceae
- Desulfotomaculaceae
- Korarchaeaceae
- Geobacteraceae
- Campylobacteraceae
- Hydrogenimonadaceae
- Aeromonadaceae
- Alteromonadaceae
- Enterobacteriaceae
- Pasteurellaceae
- Moraxellaceae
- Pseudomonadaceae
- Xanthomonadaceae
- TM7 no family
- Opitutaceae
- Verrucomicrobiaceae
- Actinomycetaceae
- Beutenbergiaceae
- Bogoriellaceae
- Cellulomonadaceae
- Corynebacteriaceae
- Dermabacteriaceae
- Dermacoccaceae
- Dermatophilaceae
- Intrasporangiaceae
- Microbacteriaceae
- Nocardiaceae
- Propionibacteriaceae
- Tsukamurellaceae
- Bifidobacteriaceae
- Coniothermaceae
- Hydrognotomaceae
- Bacteroidales incertae sedis
- Bacteroidaceae
- Bacteroidaceae
- Porphyromonadaceae
- Prevotellaceae
- Rikenellaceae
- Cytomorphaceae
- Flavobacteriaceae
- Chitinophagaceae
- Cytophaga
- Cytophaga
- Flammovirga
- Sphingobacteriaceae
- Anaerolineaceae
- Sphaerobacteraceae
- Family XII
- Trueperaceae

Vaginal Delivery

FIG E1. (Continued).

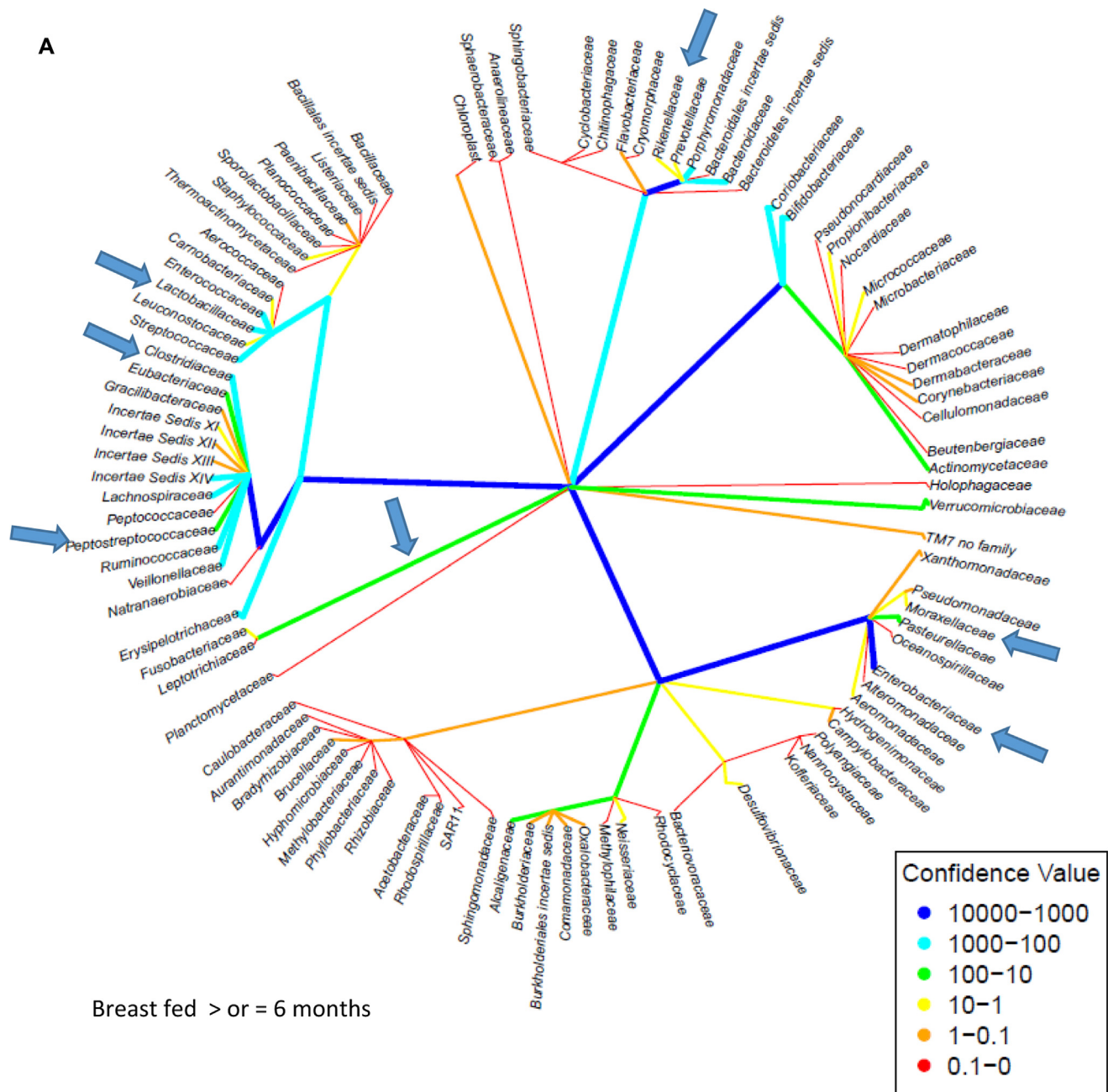
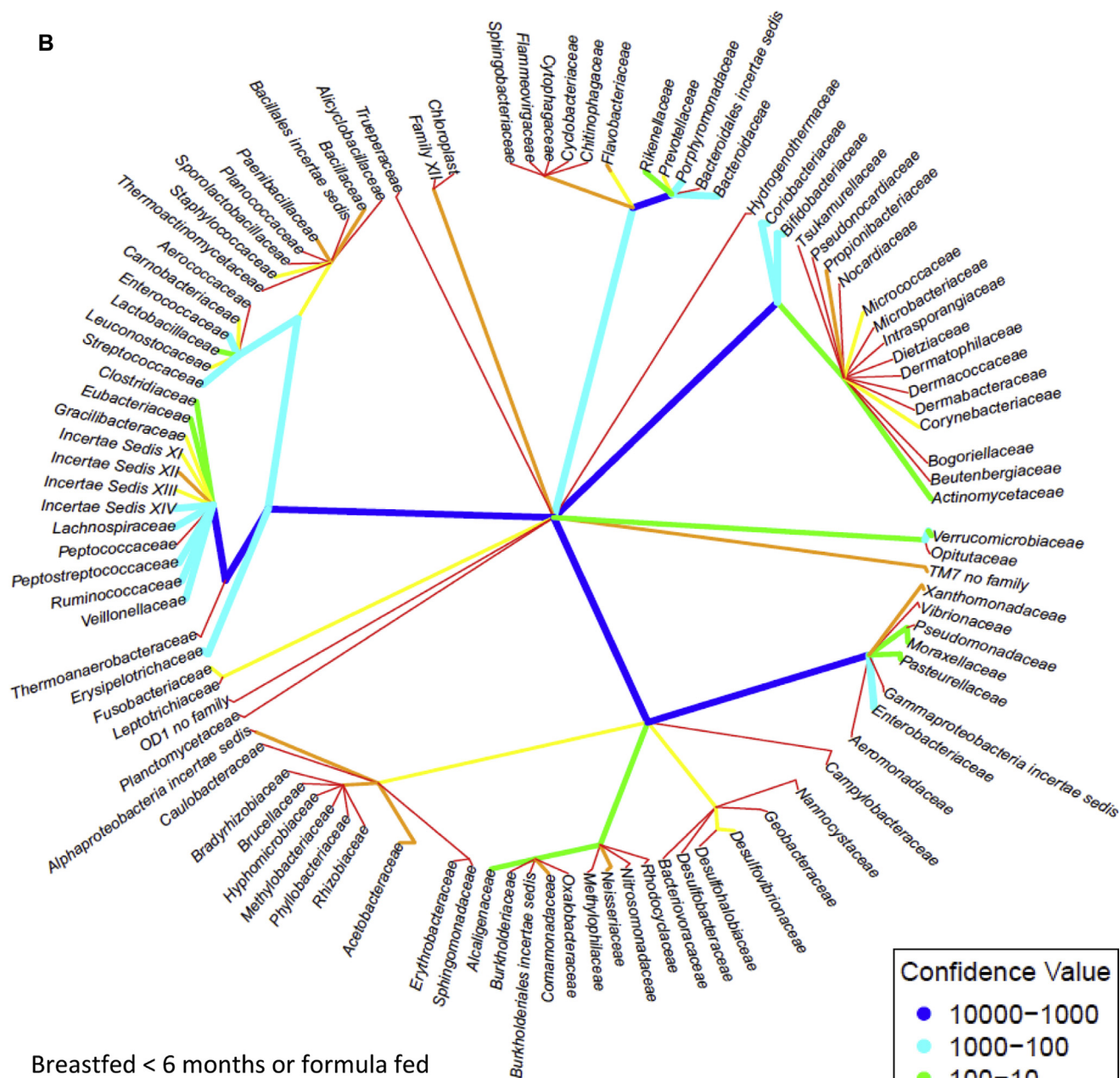


FIG E2. Central Human Microbiome Project Trees structure (family level) for infants breast-fed for 6 or more months (**A**) and breast-fed for less than 6 months or formula-fed (**B**); differences (above a confidence value of 10) in hierarchical tree with breast-feeding are shown with *blue arrows*.

B



Breastfed < 6 months or formula fed

FIG E2. (Continued).

TABLE E1. Bacterial coabundance factor analysis (of top 25 taxa)

Taxon	Bacterial coabundance factor loadings			
	Factor 1	Factor 2	Factor 3	Factor 4
<i>Bacteroides</i>	− 0.31	−0.28	0.70	−0.22
<i>Escherichia Shigella</i>	−0.23	−0.28	− 0.44	−0.07
<i>Bifidobacterium</i>	−0.28	−0.17	− 0.47	0.00
<i>Veillonella</i>	−0.08	0.08	0.15	0.66
<i>Blautia XIV</i>	0.23	0.02	0.06	−0.08
<i>Enterococcus</i>	−0.02	0.05	−0.29	−0.11
<i>Klebsiella</i>	−0.05	0.46	−0.01	0.02
<i>Streptococcus</i>	−0.09	0.00	−0.20	0.18
<i>Lachnospiraceae Dorea</i>	0.59	−0.08	−0.14	−0.23
<i>Lachnospir. Coprococcus</i>	0.41	0.00	0.06	0.24
<i>Lachnospiraceae U.</i>	0.82	−0.02	−0.06	−0.09
<i>Parabacteroides</i>	−0.05	−0.17	0.28	−0.23
<i>Collinsella</i>	−0.03	0.05	0.08	−0.23
<i>Lactobacillus</i>	−0.20	−0.05	−0.12	0.15
<i>Clostridium</i>	0.03	0.12	−0.07	0.36
<i>Coprobacillus</i>	0.16	−0.02	−0.14	0.02
<i>Sporacetigenium</i>	0.14	0.06	−0.13	0.08
<i>Enterobacteriaceae U.</i>	−0.04	0.86	−0.06	−0.05
<i>Enterobacter</i>	−0.07	0.76	0.00	−0.17
<i>Holdemania</i>	0.29	−0.01	−0.03	−0.10
<i>Megasphaera</i>	−0.09	0.06	0.06	−0.10
<i>Eubacterium</i>	0.21	−0.03	0.08	0.06
<i>Akkermansia</i>	0.06	0.07	0.00	−0.12
<i>Clostridiales U.</i>	0.65	−0.02	0.07	0.01
<i>Lactococcus</i>	0.04	0.05	−0.03	−0.15

U. denotes “unclassified” at the genus level. Factor loadings >0.3 are in boldface.

TABLE E2. Taxa associated with exclusive breast-feeding in multivariate analysis*

Taxonomic level	Feature	Value	Coefficient	N	N not 0	Q value
Order	Clostridiales	exclus_brstfed	−0.0336	325	262	2.71×10^{-5}
Genus	Eubacterium	exclus_brstfed	−0.05168	325	211	9.54×10^{-5}
Genus	Lachnospiraceae Dorea	exclus_brstfed	−0.04496	325	188	9.54×10^{-5}
Genus	Lachnospiraceae Blautia	exclus_brstfed	−0.09946	325	247	9.54×10^{-5}
Family	Lachnospiraceae	exclus_brstfed	−0.08523	325	295	9.54×10^{-5}
Genus	Escherichia	exclus_brstfed	0.101134	325	301	9.54×10^{-5}
Genus	Lactococcus	exclus_brstfed	−0.00823	325	157	0.002101
Genus	Coprococcus	exclus_brstfed	−0.03186	325	205	0.002101
Family	Peptostreptococcaceae	exclus_brstfed	−0.02355	325	206	0.004193
Genus	Shigella	exclus_brstfed	0.001246	325	9	0.004898
Genus	Collinsella	exclus_brstfed	−0.05468	325	115	0.006761
Genus	Ruminococcus	exclus_brstfed	−0.03676	325	220	0.00682
Genus	Staphylococcus	exclus_brstfed	0.004348	325	76	0.013895
Genus	Actinobaculum	exclus_brstfed	0.001051	325	2	0.037126
Family	Pasteurellaceae	exclus_brstfed	0.005379	325	62	0.037333
Family	Ruminococcaceae	exclus_brstfed	−0.04783	325	257	0.048881
Genus	Ruminococcus	exclus_brstfed	−0.07437	325	282	0.048881
Family	Erysipelotrichaceae	exclus_brstfed	−0.03576	325	229	0.048881

*Analysis adjusted for C-section, race, sex, maternal antibiotics during labor, gestational age, and treatment group.

TABLE E3. Taxa associated with breast-feeding with formula supplementation in multivariate analysis*

Taxonomic level	Feature	Value	Coefficient	N	N not 0	Q value
Genus	Megasphaera	Breast-fed + formula	0.004281	325	157	0.125102
Genus	Clostridium	Breast-fed + formula	−0.00348	325	117	0.31282
Genus	Achromobacter	Breast-fed + formula	0.000444	325	2	0.31282
Genus	Pediococcus	Breast-fed + formula	0.000985	325	4	0.406929
Family	Aerococcaceae	Breast-fed + formula	−0.00146	325	29	0.410452
Genus	Sutterella	Breast-fed + formula	0.016912	325	89	0.410452
Family	Stenotrophomonas	Breast-fed + formula	0.002109	325	36	0.423436
Family	Lactobacillaceae	Breast-fed + formula	0.000775	325	2	0.452165
Genus	Escherichia	Breast-fed + formula	−0.03608	325	301	0.452165

*Analysis adjusted for C-section, race, sex, maternal antibiotics during labor, gestational age, and treatment group.

TABLE E4. Taxa associated with C-section in multivariate analysis*

Taxonomic level	Feature	Value	Coefficient	N	N not 0	Q value
Genus	Bacteroides	C-section (yes)	−0.18421	325	267	8.34×10^{-5}
Genus	Parabacteroides	C-section (yes)	−0.05381	325	149	0.104271

*Analysis adjusted for breast-feeding, race, sex, maternal antibiotics during labor, gestational age, and treatment group.

TABLE E5. Taxa associated with race*

Taxonomic level	Feature	Value	Coefficient	N	N not 0	Q value
Genus	Megasphaera	White (vs black)	−0.00768	325	157	0.006
Genus	Megasphaera	White (vs all other races)	0.0065	325	157	0.005
Genus	Lactococcus	White (vs all other races)	0.0064	325	157	0.014

*Analysis adjusted for breast-feeding, C-section, sex, maternal antibiotics during labor, gestational age, and treatment group.

TABLE E6. Taxa associated with cord blood vitamin D level (ng/mL)*

Taxonomic level	Feature	Value	Coefficient	N	N not 0	Q value
Genus	Lactococcus	Cord blood vitamin D (ng/mL)	−0.00032	325	157	0.03
Genus	Saccharomonospora	Cord blood vitamin D (ng/mL)	1.86×10^{-5}	325	1	0.03
Family	Methylobacteriaceae	Cord blood vitamin D (ng/mL)	1.86×10^{-5}	325	1	0.03
Genus	Lachnobacterium	Cord blood vitamin D (ng/mL)	4.93×10^{-5}	325	4	0.039

*Analysis adjusted for race, breast-feeding, C-section, sex, maternal antibiotics during labor, gestational age, and treatment group.

TABLE E7. Top 10 taxa in infant gut microbiome associated with dog ownership*

Taxonomic level	Feature	Value	Coefficient	N	N not 0	Q value
Genus	Butyricoccus	Dog (yes)	0.002921	325	8	0.246027
Family	Bradyrhizobiaceae	Dog (yes)	0.000477	325	4	0.731713
Genus	Gluconacetobacter	Dog (yes)	0.001198	325	5	0.748448
Genus	Erwinia	Dog (yes)	−0.00449	325	134	0.748448
Family	Erysipelotrichaceae	Dog (yes)	0.0005	325	1	0.748448
Genus	Yersinia	Dog (yes)	0.000177	325	1	0.748448
Family	Pseudomonadaceae	Dog (yes)	0.000614	325	1	0.748448
Family	Aerococcaceae	Dog (yes)	0.001445	325	29	0.748448
Genus	Clostridium	Dog (yes)	0.010475	325	238	0.748448
Genus	Polaromonas	Dog (yes)	0.000184	325	1	0.759363

*Analysis adjusted for race, breast-feeding, C-section, sex, maternal antibiotics during labor, gestational age, and treatment group.

TABLE E8. Top 10 taxa in infant gut microbiome associated with cat ownership*

Taxonomic level	Feature	Value	Coefficient	N	N not 0	Q value
Family	Porphyromonadaceae	Cat (yes)	0.00722	325	3	0.476653
Genus	Hydrogenophaga	Cat (yes)	0.000362	325	1	0.476653
Genus	Alicyclobacillus	Cat (yes)	0.000282	325	1	0.476653
Genus	Psychrobacter	Cat (yes)	0.001777	325	1	0.476653
Genus	Leucobacter	Cat (yes)	0.0003	325	1	0.476653
Genus	Janthinobacterium	Cat (yes)	0.000327	325	1	0.476653
Genus	Yersinia	Cat (yes)	0.0003	325	1	0.476653
Family	Pseudomonadaceae	Cat (yes)	0.001039	325	1	0.476653
Family	Lachnospiraceae	Cat (yes)	0.050055	325	295	0.61248
Genus	Dysgonomonas	Cat (yes)	0.016777	325	35	0.850989

*Analysis adjusted for race, breast-feeding, C-section, sex, maternal antibiotics during labor, gestational age, and treatment group.

TABLE E9. Relationships between infant predictors (race, breast-feeding, and C-section) by age at fecal flora sampling

Predictors	Overall diversity (Shannon diversity index)					
	Fecal flora sampling at 12-16 wk (N = 74)		Fecal flora sampling at >16 wk and <20 wk (N = 77)		Fecal flora sampling at ≥20 wk (N = 154)	
	β	<i>P</i> value	β	<i>P</i> value	β	<i>P</i> value
Child's race, white*	−0.27	.07	−0.20	.25	−0.23	.005
Breast-fed (first 6 mo)	−0.27	.03	−0.20	.13	−0.21	.004
C-section (yes/no)	0.11	.34	0.11	.34	0.09	.20

*Comparison to all other racial groups.