Association Between Breast Milk Bacterial Communities and Establishment and Development of the Infant Gut Microbiome

Pia S. Pannaraj, MD, MPH; Fan Li, PhD; Chiara Cerini, MD; Jeffrey M. Bender, MD; Shangxin Yang, PhD; Adrienne Rollie, MS; Helty Adisetiyo, PhD; Sara Zabih, MS; Pamela J. Lincez, PhD; Kyle Bittinger, PhD; Aubrey Bailey, MS; Frederic D. Bushman, PhD; John W. Sleasman, MD; Grace M. Aldrovandi, MD

IMPORTANCE Establishment of the infant microbiome has lifelong implications on health and immunity. Gut microbiota of breastfed compared with nonbreastfed individuals differ during infancy as well as into adulthood. Breast milk contains a diverse population of bacteria, but little is known about the vertical transfer of bacteria from mother to infant by breastfeeding.

OBJECTIVE To determine the association between the maternal breast milk and areolar skin and infant gut bacterial communities.

DESIGN, SETTING, AND PARTICIPANTS In a prospective, longitudinal study, bacterial composition was identified with sequencing of the 16S ribosomal RNA gene in breast milk, areolar skin, and infant stool samples of 107 healthy mother-infant pairs. The study was conducted in Los Angeles, California, and St Petersburg, Florida, between January 1, 2010, and February 28, 2015.

EXPOSURES Amount and duration of daily breastfeeding and timing of solid food introduction.

MAIN OUTCOMES AND MEASURES Bacterial composition in maternal breast milk, areolar skin, and infant stool by sequencing of the 16S ribosomal RNA gene.

RESULTS In the 107 healthy mother and infant pairs (median age at the time of specimen collection, 40 days; range, 1-331 days), 52 (43.0%) of the infants were male. Bacterial communities were distinct in milk, areolar skin, and stool, differing in both composition and diversity. The infant gut microbial communities were more closely related to an infant’s mother’s milk and skin compared with a random mother (mean difference in Bray-Curtis distances, 0.012 and 0.014, respectively; P < .001 for both). Source tracking analysis was used to estimate the contribution of the breast milk and areolar skin microbiomes to the infant gut microbiome. During the first 30 days of life, infants who breastfed to obtain 75% or more of their daily milk intake received a mean (SD) of 27.7% (15.2%) of the bacteria from breast milk and 10.3% (6.0%) from areolar skin. Bacterial diversity (Faith phylogenetic diversity, P = .003) and composition changes were associated with the proportion of daily breast milk intake in a dose-dependent manner, even after the introduction of solid foods.

CONCLUSIONS AND RELEVANCE The results of this study indicate that bacteria in mother’s breast milk seed the infant gut, underscoring the importance of breastfeeding in the development of the infant gut microbiome.

Published online May 8, 2017.

© 2017 American Medical Association. All rights reserved.
Microbial colonization of the infant gut plays an important role in lifelong health.1,2 In healthy newborns, the gut microbiome composition shows large-scale longitudinal changes until age 3 years, when it settles into an adult-like anaerobic pattern.3-5 Perturbations in the microbiome are associated with susceptibility to autoimmune diseases, such as diabetes, inflammatory bowel disease, atopy, and other conditions.6-8 The mechanisms of acquisition and progression toward a “normal” infant microbiome are poorly understood. Colonization of the infant gut is a complex process dependent on multiple overlapping factors, including age,9 mode of delivery,10-12 type of feeding,10,12,13 and environmental exposures.14

Multiple studies have documented differences in the stool microbiota of breastfed compared with nonbreastfed individuals during infancy and adulthood.13,15,16 Studies show that breastfeeding confers protection against respiratory and gastrointestinal tract infections and allergic diseases in addition to reducing the risk of chronic diseases, such as diabetes, obesity, and inflammatory bowel disease.17 Most microbiota research on the breastfeeding effects to date has focused on infant stool. Little is known about the vertical transfer of breast milk microbes from mother to infant.18

Breast milk contains a diverse population of bacteria19,20 that are hypothesized to seed the infant’s gut via breastfeeding. We set out to evaluate the association between the maternal bacterial community and the infant’s stool in healthy mother-infant pairs. In this report, we show the extent that maternal milk and areolar skin contribute to the developing infant microbiome. We also describe the predicted functional profile of milk and infant stool bacterial communities.

Methods

Study Participants

We recruited healthy mothers and their full-term infants in the community around Children’s Hospital Los Angeles, Los Angeles, California, and All Children’s Hospital, St Petersburg, Florida, between January 1, 2010, and February 28, 2015. We enrolled mother-infant pairs into the following infant age groups: 0 to 7 days (initial colonization period), 8 to 30 days (first period of full maternal milk supply), 31 to 90 days (before introduction of solids), 91 to 180 days (period of solid food introduction), and 181 to 365 days (after solid food introduction). Mother-infant pairs were excluded if either had a health condition or if the infant had received antibiotics. Institutional review boards at Children’s Hospital Los Angeles and All Children’s Hospital approved the study; written informed consent was obtained from all mothers, and financial compensation was provided.

At the initial study visit, study personnel interviewed mothers to collect clinical data, including basic demographics, pregnancy and delivery history, current or recent medications (including antibiotics), and infant feeding characteristics (age at formula and solid food introduction and frequency of breastfeeding, formula, and solid foods). On all follow-up visits, we asked about new medical conditions, antibiotic use, and feeding characteristics, using the Research Electronic Data Capture (REDCap Consortium; https://www.project-redcap.org/) database.

Sample Collection, DNA Extraction, and Sequencing

We collected breast milk and areolar skin swabs from mothers and stool from infants’ soiled diapers during clinic or home study visits. Mothers expressed milk with a manual or electric breast pump using sterilized connectors into sterile bottles. Longitudinal samples were obtained when possible. The DNA extraction, library preparation, and sequencing of the V4 region of the 16S ribosomal RNA (rRNA) gene on a MiSeq sequencer (Illumina Inc) were performed using standard procedures with quality control (eMethods and eFigure 1 in the Supplement).

Statistical Analysis

We compared differences in a diversity using Faith phylogenetic diversity21; β diversity, using unweighted UniFrac and Bray-Curtis; and community composition between groups based on feeding characteristics. To evaluate sources of infant stool microbiome variation, we modeled infant age, demographics, delivery type, geographic collection site, and feeding characteristics with the permutational multivariate analysis of variance using distance matrices (adonis, R vegan package; R Foundation). Interaction terms were tested and not statistically significant. To predict the source of bacterial communities in infant stool, we used SourceTracker, version 0.9.5, software in Quantitative Insights Into Microbial Ecology22 to compare sequences in breast milk and areolar skin as input samples (source samples) with the sequences in infant stool (sink sample). To compare paired breast milk and infant stool samples, we performed a Wilcoxon rank sum test on UniFrac, Jaccard index, and Bray-Curtis distances between true mother-infant pairs and randomly paired mothers and infants matched by infant age within each run. Operational taxonomic unit (OTU) sharing was defined as the percentage of mother-infant dyads in which a given OTU was found in both members; permutation testing with randomly shuffled mother-infant pairings was used to assess significance. Oligotyping analysis was performed using sequences mapped to OTUs shared within mother-infant dyads as previously described.23
Statistical analyses were performed using R statistical software, version 3.0.3 (R Foundation).

Random forest models were used to predict taxonomy that classified the infant stool bacterial community into classes based on feeding method. We used a random forest machine-learning algorithm to determine a ranked list of all bacterial taxa in the order of age-discriminatory importance. Stool from a randomly chosen subset of vaginally born, exclusively breastfed infants (n = 42) at multiple time points throughout the first year of life was used to train the model. A sparse model with 44 predictors (57.5% variance explained) was selected on the basis of 10-fold cross validation. Relative microbiota maturity was calculated by microbiota age of a child minus microbiota age of healthy children of a similar chronologic age from the model.

We used Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) to predict metagenome function using 16S rRNA marker gene sequences and referenced to published complete genome sequences. Both R, version 3.0.3 and STAMP, version 2.0.9 were used for statistical analyses of the functional profiles. A false discovery rate P value ≤ .10 was considered significant for any analyses with multiple comparisons. A standard P value ≤ .05 was considered significant for all other analyses.

Results

Study Population

We enrolled 228 healthy individuals from 121 families, including 107 mother-infant pairs, 2 mothers only, and 12 infants only. Median infant age at the time of specimen collection was 40 days (range, 1-331 days). Family and feeding characteristics are reported in the Table and eTable 1 in the Supplement, respectively. Eighty (66.1%) of our families were Hispanic/Latino. Seventy-eight (64.5%) infants were born by vaginal delivery. We analyzed at least 2 longitudinal samples from 73 (60.3%) families. At the time of sample collection (n = 569), 484 (85.1%) samples showed that the infants were primarily breastfed (defined as receiving breast milk for >75% of their daily milk intake), including 298 (52.4%) who were exclusively breastfed (defined as never having received formula).

Breast Milk, Areolar, and Infant Stool Microbiomes

Distinct bacterial communities were present in milk, areolar skin, and stool, which differed in both composition and diversity (Figure 1). Proteobacteria (Moraxellaceae, Enterobacteriaceae, and Pseudomonadaceae) constituted the dominant phylum in milk, Firmicutes (Staphylococcaceae and Streptococcaceae) were dominant in areolar skin, and Proteobacteria (Enterobacteriaceae) and Actinobacteria (Bifidobacteriaceae) accounted for more than 50% of the community in infant stool. Milk and areolar samples collected from the right and left breast were similar (eFigure 2 in the Supplement); therefore, samples from the right breast were used for analyses if samples from both sides were collected.

The breast milk bacterial community as the infant aged is shown in Figure 2. The breast milk bacterial communities remained unchanged in a diversity (within-sample diversity) throughout the first year of life. β Diversity (between-sample diversity) increased between mothers in the first 6 months after delivery and then decreased. There were no significant differences in breast milk or areolar bacteria by race/ethnicity, infant sex, mode of delivery, or geographic location of enrollment site.

Taxonomic Changes in Infant Stool as a Function of Age, Feeding Type, and Other Factors

Bacterial diversity in infant stool increased with age, with convergence between individual infants near 12 months (Figure 2). Actinobacteria, specifically Bifidobacteriaceae (r = 0.31, P < .001), typically increased and Proteobacteria, specifically Enterobacteriaceae (r = 0.27, P < .001), decreased as the infant aged. Bacterial abundance changes occurred in direct or inverse relationships with each other in a complex network (eFigure 3 in the Supplement). Members of the Bifidobacteriaceae and Enterobacteriaceae families were dominant in the 3 observed clusters.

To evaluate factors contributing to the infant stool microbiome, we applied a permutational multivariate analysis of variance using the unweighted UniFrac and Bray-Curtis distance matrices (eTable 2 in the Supplement). Age accounted for 23% of the variation between infant stool samples.
Changes in the infant gut bacterial community were associated with the proportion of breastfeeding in a dose-dependent manner. The percentage of daily milk intake was associated with altered community composition, measured by unweighted UniFrac distance ($R^2 = 0.008, P = .002$). Differences in gut phylogenetic diversity did not reach statistical significance (nonparametric $t$ test, $P = .06$), but different compositions were observed between primarily and nonprimarily

### Infant Gut Microbiota Composition Association With Amount of Daily Breastfeeding

We used SourceTracker to estimate the proportion of bacteria in infant stool that originated from breast milk and areolar skin (Figure 3). Overall, primarily breastfed infants received more bacteria from breast milk and areolar skin compared with those not primarily breastfed (breast milk, 18.5% vs 5.7%, $P = .001$, Wilcoxon rank sum; areolar, 5.2% vs 0.001%, $P = .01$, Wilcoxon rank sum). During the first 30 days of life, infants who were primarily breastfed received mean (SD) 27.7% (15.2%) of the bacteria from breast milk and 10.4% (6.0%) from areolar skin. The remaining 61.9% (16.2%) came from other sources that we did not characterize. The contribution of bacteria from mother’s milk and areolar skin was highest during the first month of life and decreased as the infant aged.

Because bacterial communities in milk and areolar skin remained distinct for each mother, we hypothesized that an infant’s stool microbiota would be more similar to the infant’s mother’s microbiota compared with random mothers. Comparison using multiple distance metrics showed closer distances between infant stool and breast milk or areolar skin among true mother-infant pairs compared with random pairs (eTable 3 in the Supplement). In total, 26 of 478 OTUs were more significantly shared within mother-infant pairs vs random pairs (eTable 4 in the Supplement). Oligotyping analysis showed significant sharing of Streptococcus, Veillonella, and Rothia (eTable 5 in the Supplement).

## References

2. Brzezinski A, et al. Fetal colonization following complete or near-sterility at birth. C. Breast milk had significantly higher amounts of Moraxellaceae and Pseudomonadaceae (false discovery rate [FDR] $P < .001$), whereas areolar skin swabs were enriched in Streptococcaceae and Staphylococcaceae (FDR $P < .001$) and infant stool comprised greater amounts of Bifidobacteriaceae (FDR $P < .001$). Boxes indicate the 25th to 75th percentile; line within the box, median; error bars, 1.5 × interquartile range.

---

### Figure 1. Microbial Communities in Breast Milk, Areolar Skin, and Infant Stool

#### A. Principal coordinate (PC) analysis

PC1, 18.4%

#### B. Phylogenetic diversity

#### C. Bacterial composition

- **Actinobacteria**
  - Bifidobacteriaceae
  - Corynebacteriaceae
  - Micrococccaceae
- **Bacteroidetes**
  - Bacteroidaceae
- **Firmicutes**
  - Clostridiaceae
  - Firmicutes
  - Gemellaceae
  - Micrococcaceae
  - Staphylococcaceae
  - Streptococcaceae
- **Proteobacteria**
  - Enterobacteriaceae
  - Pasteurellaceae
  - Pseudomonadaceae
  - Rothia
  - Bradyrhizobiaceae
  - Other
- **Other**
Figure 2. Evaluation of Breast Milk and Infant Stool During the First Year of Life

A. Relative abundance of most bacterial taxa in 238 breast milk samples from 94 mothers remained stable with age. B. Infant stool composition changed with increasing infant age in a nonrandom pattern. A total of 259 stool samples from 112 infants are shown here. C. α and β diversity (α diversity, Faith phylogenetic diversity [21]) increased in infant stool (P = .01) but not breast milk (P = .06) with infant age. Differences between individuals (β diversity, unweighted UniFrac) in breast milk increased during the first 6 months of life (P < .001) and then decreased. β diversity of infant stool steadily decreased (P < .001) during the first 6 months and then began to level off after 6 months of life. Nonparametric t tests were used for comparisons of α and β diversity between age groups.
breastfed infants (eFigure 4 in the Supplement). We used random forest modeling to determine the bacterial taxa that best discriminate between the infant’s lifetime feeding pattern (i.e., exclusively vs nonexclusively breastfed). Erysipelotrichaceae, Bacteroidaceae, and Ruminococcaceae were the most prominent families found in nonexclusively breastfed infants that discriminated between the 2 groups (eFigure 5A in the Supplement). Ruminococcaceae and Bacteroidaceae were more abundant among formula feeders at the time of the sample collection (eFigure 5B in the Supplement).

**Age at Solid Food Introduction**

Following milk introduction, the next major change in the infant’s diet is the introduction of solid foods. Solid food introduction in infants aged 4 to 6 months was associated with a change in infant gut microbiome composition (eFigure 6 in the Supplement), but not phylogenetic diversity (nonparametric t test, \( P > .99 \)). We interrogated our data to determine whether early solid food introduction changed the pattern of the microbiota trajectory, using a random forest regression model.4 Early solid food introduction (<4 months) in our infant cohort led to a rapid maturation of the infant stool microbiota (Kruskal-Wallis, \( P = .02 \)) (Figure 4). The amount of daily breastfeeding as a proportion of total milk intake continued to influence the infant stool microbiome diversity and membership even after solid foods were introduced (eFigure 7 in the Supplement).

**Predicted Metagenome Function of Breast Milk and Infant Gut Bacterial Community**

Given the effect of breast milk microbiota on the infant gut community, we sought to gain insight into the genes present in breast milk bacterial communities using PICRUSt.24 Breast milk contains bacteria with predicted high abundance in gene families associated with membrane transport and carbohydrate, amino acid, and energy metabolism (eFigure 8 in the Supplement). We compared metagenome predictions of infant stool microbiota based on the amount of breast milk intake. Infants who were primarily breastfed had a lower representation of genes involved in energy metabolism, sphingolipid metabolism, and glycogen synthesis and metabolism (eFigure 9 in the Supplement). Infants with earlier introduction of solid food had microbiota with increased function related to xenobiotics biodegradation and metabolism.

**Discussion**

We characterized bacterial communities in mother-infant pairs and provide data suggesting that bacteria from mothers’ breast milk and areolar skin are transferred to their infants’ guts. Key support for this idea comes from comparisons of bacterial lineages in the infants’ guts with their mothers vs random mothers, which showed more shared lineages in the correct pairs. Our data do not exclude the possibility that bacteria from infants’ stool seed the mother’s microbiome. It is likely that bacterial cycle between the mother and infant in a constant exchange. However, we favor the idea that transfer of bacteria primarily occurs from mother to infant. Breast milk is an early source of bacteria and nutrition introduced to the infant gut within a few hours of birth. Bacteria from mother’s milk and skin are most prominent in their infants’ guts in the first month of life, accounting for nearly 40% of the gut bacteria in primarily breastfed infants. Mother to infant microbe transmission was compromised in infants who were not primarily...
breastfed. Other sources of the infant microbiome include the mother’s gut and vaginal bacteria as well as the environment. In mice, cross-fostering experiments initiated within 48 hours of birth have shown that the nursing mother, rather than the birth mother, dictates the infant microbiome composition, which persists after weaning and for life.26 In the Human Microbiome Project, Ding and Schloss26 found breastfeeding during infancy to be a major life-history characteristic that affects bacterial composition in adults. Early life may represent a critical window for bacterial imprinting of breast milk bacteria leading to nonrandom community assembly.1,27 Our model showed that the complex ecologic network that forms during the first 6 months of life is built on interdependent bacterial relationships in the infant gut. Breast milk bacteria that seed the gut first influence and select for bacteria that follow, leaving a footprint that can be detected even in adulthood.16

Our oligotype analysis indicated that breast milk provides Veillonella and Rothia, bacteria genera that have been associated with a lower incidence of asthma.6 These early bacterial seeding events may be a mechanism by which breastfeeding protects children.

The breast milk microbiota become more divergent between mothers during the first 6 months of her infant’s life. The divergence lessens after 6 months when breast milk is typically no longer an infant’s sole food source. The infant gut microbiota display increasing α diversity and reduced β diversity as the infant ages—a trajectory also observed by others.5,16 Mother-infant sharing decreases as the infants age. Milk bacteria are detectable in infant stool at least through 6 months of life, but differences in the gut microbiota and their function are evident throughout the entire first year. Some studies have suggested that even small amounts of formula supplementation would shift the microbiota from a breastfed pattern to a formula pattern.28 Our multivariate and diversity analyses indicate that the shift occurs in a dose-dependent manner. The proportion of breast milk intake also decreases with introduction of solid foods. A recent study suggests that cessation of breastfeeding, rather than introduction of solid foods, is the major driver in the development of an adult microbiota.15 In our cohort, introduction of solid foods prior to 6 months led to early maturation of the microbiota. Furthermore, continued breastfeeding after solid food introduction suppressed the diversification and enrichment of bacteria typically associated with solid foods. These findings further support the current World Health Organization and American Academy of Pediatrics recommendation for exclusive breastfeeding during the first 6 months with continued breastfeeding until at least 12 months.17

Our predicted metagenome functional analysis shows that breast milk harbors bacteria with prominent carbohydrate, amino acid, and energy metabolism functions. Breastfeeding has been shown to strongly protect against obesity.29 Microbiota harvest energy from the diet, and energy storage in the host is divergent between obese and lean humans and mice.7,30 Infants in our cohort who were not primarily breastfed had a higher abundance of the Bacteroidaceae, which has been associated with subsequent obesity in most pediatric studies of early gut colonization.31-33 although not all studies agreed.34 Erysipelotrichaceae and Ruminococcaceae were also increased in infants not exclusively breastfed. Both families are enriched in older children and adults with higher body mass indexes.35,36 Furthermore, continued breastfeeding as the primary source of milk intake after introduction of solids appeared to suppress increases in multiple families within the phylum Firmicutes that are butyrate producers. Butyrate can serve as an energy substrate or signaling molecule to contribute to the de novo production of lipids. Firmicutes and butyrate have also been associated with an obese phenotype in mice and humans.27-33 These pathways need to be explored further.

Limitations

This study has limitations. We did not sequence communities from maternal mouth, skin other than areolar, vagina, or stool, which may have contributed additional bacteria to infants. The origin of breast milk bacteria is unclear. One hypothesis is that milk bacteria come from the infant’s oral cavity. We did not collect infant oral specimens. However, we did not see differences in the microbiota of breast milk by delivery type in our cohort, as has been similarly reported by other investigators.40 Our recruitment focused on breastfeeding participants, and we did not enroll large numbers of strictly formula-fed infants. Some participants were unable to adhere to our longitudinal collection schedule. Sequencing of the 16S rRNA gene is limited in analysis at the strain level. SourceTracker is an estimation/prediction tool; the results may differ if the model included all potential sources of microbiota or additional longitudinal samples. We used PICRUSt to estimate bacterial gene content; because the input is 16S rRNA gene data, we did not capture any eukaryotic or viral contributions to the metagenome.24 PICRUSt also cannot distinguish variation at the strain level and different strains of 1 bacteria can vary in gene counts. Metagenomic shotgun sequencing would permit strain and more accurate functional analysis, but is costly. Finally, we focused on bacterial communities, yet milk is a complex substance with many bioactive components, including human milk oligosaccharides, which may promote the persistence of specific bacterial lineages.

Conclusions

Our study confirms a bacterial community in breast milk and tracks that community from mothers into the infant gut. Breast milk bacteria influence the establishment and development of the infant microbiome with continued impact after solid food introduction. Furthermore, breast milk contributes bacteria associated with a decreased risk for developing allergic diseases. Our results emphasize the importance of breastfeeding in the assembly of the infant gut microbiome.
California, Los Angeles (Pannaraj, Bender); Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Bittinger, Bailey, Bushman); Division of Pediatric Allergy and Immunology, Department of Pediatrics, Duke University School of Medicine, Durham, North Carolina (Sleasman); Department of Pediatrics, University of California, Los Angeles (Aldrovandi).

**Author Contributions:** Drs Pannaraj and Aldrovandi had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Pannaraj, Yang, Bushman, Sleasman, Aldrovandi. Acquisition, analysis, or interpretation of data: Pannaraj, Li, Cerini, Bittinger, Yang, Rollie, Adsetiyo, Zabih, Lincez, Bailey, Sleasman, Aldrovandi. Drafting of the manuscript: Pannaraj, Li, Cerini, Zabih, Sleasman, Aldrovandi. Administrative, technical, or material support: Bittinger, Bailey. Statistical analysis: Bushman, Sleasman, Aldrovandi. Obtained funding: Pannaraj, Sleasman, Aldrovandi. Supervision: Cerini, Bushman, Sleasman, Aldrovandi.

**Conflict of Interest Disclosures:** Dr Pannaraj receives funding from AstraZeneca for clinical research and Pfizer for vaccine education. No other disclosures were reported.

**Funding/Support:** This work was supported by National Institutes of Health grants K23 HD072774-02 (Dr Pannaraj), K12 HD052954-09 (Dr Bender), ROI AI052845 (Dr Bushman), ROI AI014071 (Dr Aldrovandi), and UMAI106716 (Dr Aldrovandi), and University of Pennsylvania Center for AIDS Research grant P30 AI045008.

**Role of the Funder/Sponsor:** The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

**REFERENCES**