Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI) Vol. 27:1-11

Exclusive breastfeeding may decrease overall diversity of the infant gut microbiome with a shift towards dominance of bacterial taxa associated with lactose metabolism

Eleanor Chen, Maitri Panchal, Daniel Song, Lauren Ung

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY Breastfeeding has repeatedly been shown to be the ideal form of feeding for infant development, including its correlation with a reduced risk of obesity, cardiovascular disease, allergies, and other health conditions. In particular, breastfeeding has been shown to influence the infant gut microbiome, which may then directly or indirectly impact overall health outcomes from infancy to adulthood. In this study, we hope to validate and further explore the effects of exclusive breastfeeding on the developing infant gut microbiota. Using the dataset generated by Rhee et al., we examined the differences in microbial diversity and composition between exclusively breastfed and non-exclusively breastfed infant gut microbiomes by comparing diversity metrics, differential and relative abundance analyses, and indicator taxa between the two feeding models. We found that the exclusively breastfed infants had lower alpha and beta diversity, and their bacterial taxa were dominated by those that were directly linked with lactose production or consumption within their gut microbiomes. Overall, these findings support the current literature surrounding the effect of feed-type on the infant gut microbiome.

INTRODUCTION

Exclusive breastfeeding (EBF) refers to when an infant receives no other source of food or liquid, including water (1). EBF is recommended for the first six months of life and accepted as the nutritional standard for healthy infant development (2). On an ideal feeding schedule, infants start with EBF for the first 6 months followed by the introduction of nutritionally appropriate and safe foods, complemented with continual breastfeeding for up to 2 years (2). Compared to other forms of feeding, EBF is preferred because breast milk supplies the infant with the correct quantity and quality of nutrients in a way that is easily and efficiently digested (3). Additionally, nutrient compositions in breast milk are self-adjust over infant maturation to support optimal growth (4).

Beyond the baseline nutrients needed for growth, breastfeeding is recognized for playing many other short- and long-term roles in infant health and development. Breastmilk consists of various bioactive factors including anti-infective immunoglobulins, white blood cells, and factors that stimulate the maturation of the small intestine that aid with digestion and absorption of nutrients (5, 6). Accordingly, breastfeeding appears to protect against gastrointestinal infections and has been identified as a potential factor in the pathophysiology of various diseases (7).

In infants, the first major event for their developing microbiota is birth and, therefore, the mode of delivery appears to play a significant role in establishing the gut microbiome (8). Beyond this, feed type, i.e., whether an infant is breastfed or formula fed, and peri-/post-natal antibiotic exposure, also both significantly affect infant microbiota development in the first year of life (9–12), and feed type alone has been shown to influence both the microbial diversity and the taxonomic composition of an infant's gut microbiota (9, 13, 14). In particular, the microbiota of breastfed infants are generally less diverse and display an increased relative abundance of Bifidobacterium and Bacteroides when compared to those of formula-fed infants (15). Our aim was thus to expand on these findings by examining the role of exclusive breastfeeding in particular, on the diversity and taxonomic composition of the infant gut microbiome. To do this, we will be using a dataset generated by Dr. Rhee et al.,

Published Online: September 2022

Citation: Eleanor Chen, Maitri Panchal, Daniel Song, Lauren Ung. 2022. Exclusive breastfeeding may decrease overall diversity of the infant gut microbiome with a shift towards dominance of bacterial taxa associated with lactose metabolism. UJEMI 27:1-11

Editor: Andy An and Gara Dexter, University of British Columbia

Address correspondence to: https://jemi.microbiology.ubc.ca/

which includes infant gut samples and associated feed information. This dataset is publically available on the European Nucleotide Archive (ENA) Browser under accession PRJEB39437.

METHODS AND MATERIALS

Dataset description. The dataset utilized in this paper was generated by Dr. Kyung Rhee et al. from the Department of Pediatrics at the University of California. It consists of stool samples from 82 infant-mother dyads collected at 2 weeks, 2 months, 4 months, 6 months, and 12 months of the infant's life, along with 171 fields of accompanying metadata, describing each subject's diet, medications and supplements, health, weight and feeding behaviours when available. Microbial sequences were obtained using 515fbc and 806r primers, following the Earth Microbiome Protocol (15) and provided as Illumina sequences for the V4 region of the 16S rRNA gene. The dataset is publicly available on the European Nucleotide Archive (ENA) Browser under accession PRJEB39437.

Preliminary data processing. The metadata and corresponding manifest file were filtered to remove samples from the mother, samples collected at time points other than 2 months, and any samples where the data for our metadata category of interest, feed, was not collected. To improve statistical power, the "Combined" feed type (n=10), which represents a mixture of formula and breastfeeding, and the "formula" feed type (n=9) were grouped together to form a larger, "non-exclusively breastfed" category (n=19) within the "feed" metadata column. The "breast" feed type (n=42), which is considered the "exclusively breastfed" category, was not altered. All metadata processing described above was performed in R (version 4.1.2) (16), and RStudio (version 1.4.1717) (17) along with the R packages: dplyr (version 1.0.7) (18) and tidyverse (version 1.3.1) (19). These steps along with all subsequent analysis in R are detailed in the supplemental R script **(script1.R**).

The rearranged metadata and corresponding sample sequences were then imported into QIIME2 (version 2021.4) (20) for downstream processing and analysis. The entire read length of 150 base pairs was retained, since all base positions maintained a median Phred quality score over 30. The Divisive Amplicon Denoising Algorithm 2 (DADA2) (21) in QIIME2 was used to correct for sequencing errors and identify unique amplicon sequence variants (ASVs). Based on the alpha-rarefaction curve (Fig. S1), a sampling depth of 15,000 was selected, which maximized the number of samples and features retained. These steps along with all subsequent analysis using QIIME2 are detailed in the supplemental QIIME2 script (**script2.txt**).

Taxonomic analysis. ASVs were assigned taxonomy using a Naive Bayes classifier (22) pre-trained on the SILVA 138 99% OTU database (23, 24) for the 515F/806R (V4) region of the 16S rRNA gene in QIIME2. Results were incorporated in downstream taxonomic abundance analyses.

Alpha and beta diversity analysis. To determine phylogenetic distances for QIIME2 core diversity analyses, a rooted phylogenetic tree was generated using MAFFT (25) for sequence alignment and FastTree (26) for phylogeny construction. Alpha diversity metrics (Shannon diversity (27), Faith's phylogenetic diversity (Faith's PD) (28) , Pielou's evenness (29) and observed features) and beta diversity metrics (Bray-Curtis (30), Jaccard (31), Weighted (32) and Unweighted (33) UniFrac distance) were calculated using QIIME2's core diversity command at a sampling depth of 15,000 as determined during data pre-processing . Significance for alpha and beta diversity metrics were determined using the Kruskal-Wallis (34) and PERMANOVA (35) tests respectively via QIIME2's group significance command. To generate a weighted UniFrac diversity Principle Coordinate Analysis (PCoA) plot, the phyloseq package (version 1.36.0) (36) in R (16) was used. QIIME2 generated outputs (features table with taxonomic classification, phylogenetic tree and metadata) were imported into R (16) and combined into a phyloseq object. The samples were then rarefied to a sampling depth of 15,000, and weighted UniFrac principal coordinate analysis was performed using the ordinate function from the phyloseq package (36) with the following parameters: method was set as "PCoA" and the distance was set as "wunifrac". The results were then visualized as a PCoA plot using the plot_ordination function in phyloseq (36).

Relative and differential abundance analysis. Relative and differential abundances between the two feed types were then calculated and compared in R using the same phyloseq object described above. Relative abundances were calculated using a user-defined function (see **script1.R**), and low-abundance ASVs (>0.005% of the total sequencing reads in the dataset) were removed. Differential abundance analysis was performed using the DESeq2 (version 1.32.0) (37), tidyverse (19), vegan (version 2.5.7) (38) and ape (version 5.5) (39) packages. DESeq2 results were filtered to retain only the taxa with an adjusted Wald test P-value of <0.05. Values were then log2 transformed and visualized using ggplot2 (version 3.3.5) (40). Between the two feed types, the relative abundances of the differentially abundant taxa were compared using the Kruskal-Wallis test statistic (34) with ggpubr (version 0.4.0) (41) and visualized using ggplot2.

Indicator Taxa Analysis. The features table with taxonomic classification and the associated metadata were imported into R from QIIME2, then grouped by species using the dplyr (18) and phyloseq (36) packages. Indicator taxa analysis was performed on the data using the indicator_multipatt function from the indicspecies package (version 1.7.9) (42) in R. This was done using default parameters except for the "duleg" argument which was set to "TRUE" in order to restrict the analysis to only individual groups without site group combinations.

RESULTS

Infant gut microbial diversity differs based on feed type. To determine the impact of feed type on the average microbial diversity of the infant gut microbiome, various alpha diversity analyses were conducted, including for observed features, Shannon's diversity, Pielou's evenness, and Faith's PD. For all alpha diversity metrics that were examined, the average microbial diversity for EBF infants was lower than for NEBF infants (**Table S1**). However, upon statistical comparison using the Kruskal-Wallis test, only the observed features and Shannon diversity metrics, which are quantitative measures of the microbial richness aspects of diversity respectively, were statistically significant (**Fig. 1A, Table S1**).

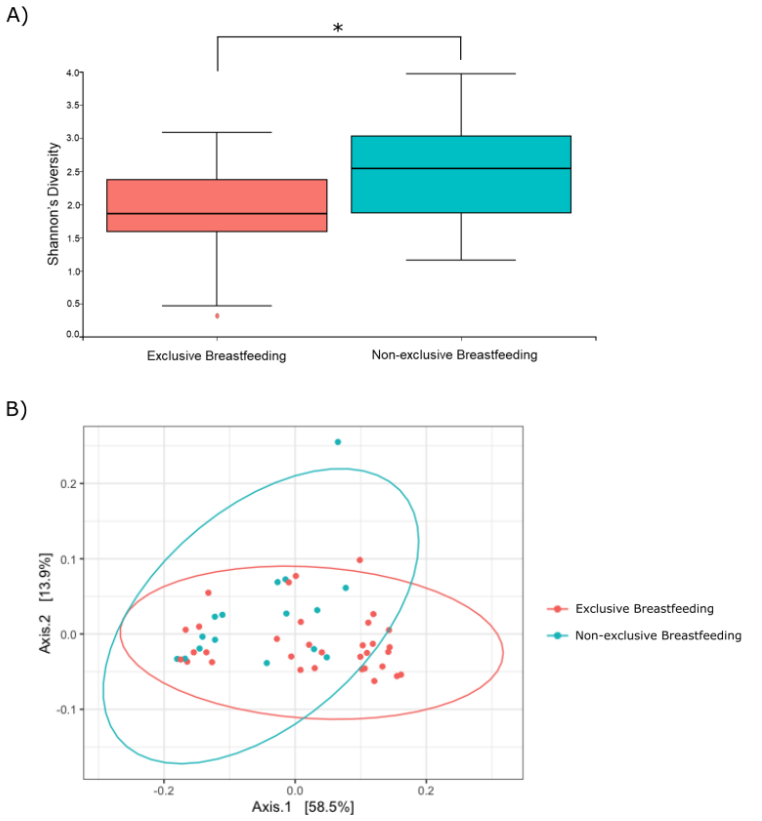

FIG. 1 Microbial diversity differs between EBF and NEBF infant guts. (A) Boxplot using Shannon's index to compare alpha diversities of EBF (red) and NEBF (blue) samples. The box represents the interquartile range, and the middle line represents the median and whiskers denote 95% confidence intervals. * indicates a significant difference, p = 0.015 (Kruskal-Wallis test). (B) Principal coordinate analysis (PCoA) plot using weighted UniFrac distances to compare beta diversity between EBF (red) and NEBF (blue) samples. EBF n=33, NEBF n=16.

To determine the impact of feed type on between-group diversity, PCoA plots were generated for Bray-Curtis, Jaccard, weighted and unweighted UniFrac distances. Although all plots show some degree of overlap, the weighted UniFrac distance metric displayed the most distinct clustering and highest variance along axis 1 (**Fig. 1B, Fig. S2**). In general, this PCoA plot shows two clusters formed by the EBF and NEBF categories, with moderate overlap seen between the two (**Fig. 1B**). To confirm differences between feed types, box-plots for the Bray-Curtis, Jaccard, weighted and unweighted UniFrac distances were generated. From this, it was found that the differences between the two groups (EBF vs NEBF) were slightly greater than within group differences (EBF vs EBF, or NEBF vs NEBF) (**Fig. S3**). Although these differences were small, pairwise PERMANOVA tests reveal that they are statistically significant (**Fig. S2**). Box plots were generated for the other beta diversity metrics as well, and similar patterns were also seen (**Table S2**). Taken together, these results suggest that feed type is correlated with differences in the microbial diversity of the infant gut, with an overall decrease in diversity for EBF infants.

***Bifidobacteria* abundance significantly differs based on feed type.** To determine whether feed type has an impact on taxonomic composition and more specifically, the abundance of specific organisms within the infant gut, we performed differential abundance analysis at the genus level, which was the most resolved taxonomic level possible. Between the EBF and NEBF categories, 10 genera were found to be differentially abundant. Three genera were found to be higher in the NEBF category, namely *Megasphaera, Proteus* and *Actinobacter*, and seven genera were higher in the EBF category, namely *Bacteroides, Bifidobacterium*, *Haemophilus, Lactobacillus, Staphylococcus, Clostridium sensu stricto*, and *Veillonella* **(Fig. 2A)**. However, relative abundance analysis for each differentially abundant genus revealed that only *Bifidobacterium* demonstrated a significantly higher abundance among EBF infants

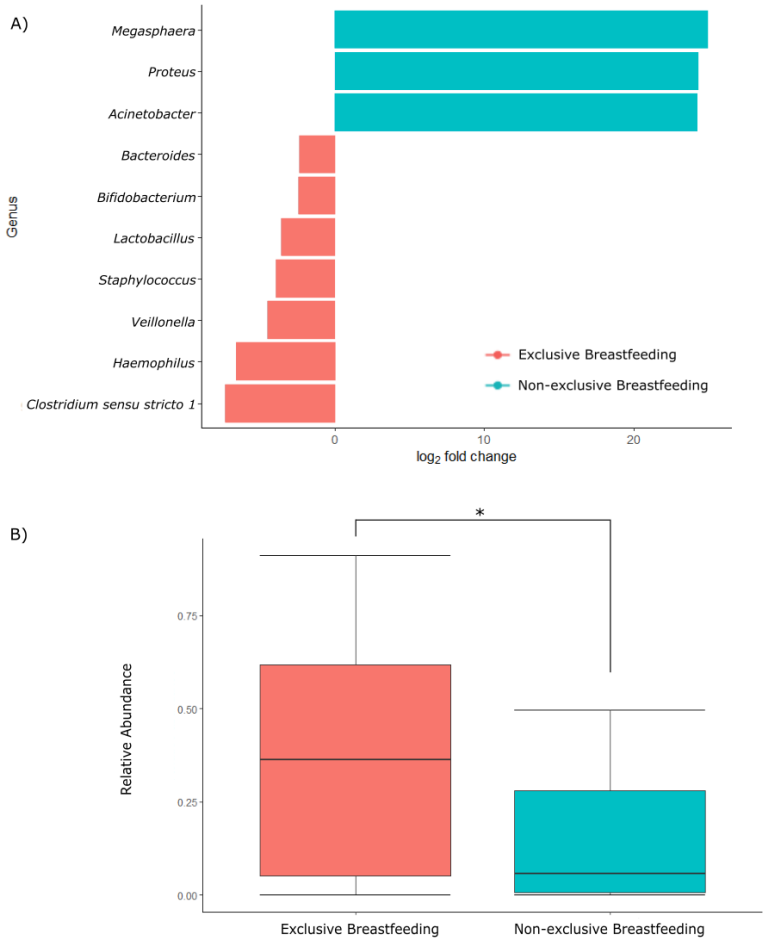

FIG. 2 Differentially abundant genera and relative abundance of *Bifidobacterium*. (A) Bar plot of differentially abundant genera between EBF (red) and NEBF (blue) samples. Bars represent log2 fold change, with negative values indicating higher abundance in the EBF group and positive values indicating higher abundance in the NEBF group. Only genera with a significant fold change (adjusted p-value < 0.05) are shown. (B) Box plot comparing the relative abundance of *Bifidobacterium* in EBF (red) and NEBF (blue) categories. The box represents the interquartile range, the middle line represents the median and whiskers denote 95% confidence intervals. * indicates a significant difference between group means, p = 0.029 (Kruskal-Wallis test). EBF n=33, NEBF n=16.

compared to NEBF infants, as confirmed by its Kruskal-Wallis p-value of 0.029 (**Fig. 2B, Table S3**).

Indicator taxa analysis shows a higher number of strongly associated species in the NEBF category. Indicator taxa analysis was performed to parse out certain species that were highly correlated with each feed category and revealed 13 species (or genera if the taxa could not be resolved to species level) that were strongly correlated with the NEBF category, and 3 species highly correlated with the EBF infants category as shown in **Table 1**. Species were determined to be good indicators if the indicator value (IV) was greater than 0.3 and the p-value was less than 0.05 (43). *Haemophillus* and *Staphylococcus* were found to be the strongest indicator species in the EBF category, with IVs greater than 0.5 and p-values less than 0.05. Within the NEBF category, *Intestinibacter* was seen to have the highest IV of 0.43. We observed that although the EBF category had a significantly lower number of indicator species compared to the NEBF category, the 3 species in the EBF category were the strongest indicator species with the highest IVs.

Table 1. Indicator taxa analysis shows more unique species in the NEBF category. 13 species were found in NEBF infants that were not found in EBF infants, and 3 species in EBF infants that were not found in the NEBF category, organized from Phylum to Species with common taxa grouped together. Species were determined to be good indicators if > 0.3 and p-value < 0.05.

Feed Type	Phylum	Class	Order	*Family*	*Genus species*	p value	Indicator Value	Good Indicator?
Exclusive Breastfed Infants	Proteobacteria	Gammaproteobacteria	Pasteurellales	*Pasteruellaceae*	*Haemophilus sp.*	0.015	0.55	Y
	Firmicutes	Bacilli	Staphylococcales	*Staphylococcaceae*	*Staphylococcus sp.*	0.035	0.53	Y
	Actinobacteriota	Actinobacteria	Micrococcales	*Micrococcaceae*	*Rothia sp.*	0.005	0.48	Y
Non-exclusive Breastfed Infants	Proteobacteria	Gammaproteobacteria	Xanthomonadales	*Xanthomonadaceae*	*Stenotrophomonas sp.*	0.005	0.33	Y
			Pseudomonadales	*Moraxellaceae*	*Acinetobacter sp.*	0.015	0.33	Y
			Enterobacterales	*Morganellaceae*	*Proteus sp.*	0.025	0.17	N
			Pseudomonadales	*Moraxellaceae*	*Acinetobacter sp.*	0.015	0.33	Y
	Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Lachnoclostridium sp.*	0.020	0.35	Y
					Tyzzerella sp.	0.005	0.32	Y
					Anaerostipes sp.	0.025	0.31	Y
			Peptostreptococcales-Tissierellales	*Peptostreptococcaceae*	*Intestinibacter sp.*	0.005	0.43	Y
					Clostridioides difficile	0.040	0.31	Y
			Clostridiales	*Clostridiaceae*	*Clostridium sensu stricto sp.*	0.015	0.22	N
			Oscillospirales	*Ruminococcaeceae*		0.010	0.22	N
		Negativicutes	Veillonellales-Selenomonadales	*Veillonellaceae*	*Negativicoccus succinicivorans*	0.015	0.27	N
					Megasphaera sp.	0.015	0.22	N
		Bacilli	Erysipelotrichales	*Erysipelotrichaceae*	*Clostridium innocuum*	0.005	0.33	Y

DISCUSSION

Prior literature has demonstrated that breastfeeding appears to be ideal for healthy infant development in terms of being protective against the development of certain diseases (e.g. asthma), overweightness and obesity in adulthood, and the development of allergies as indicated by the World Health Organization (https://www.who.int/health-topics/breastfeeding)(44). Our study explored the differences in the gut microbiome that may arise due to differential feeding types for infants at 2 months of age in order to further elucidate the benefits of breastfeeding for infants. We focused on various diversity metrics and taxonomic differences using differential abundance, relative abundance, and determining indicator taxa between the two feed groups for our analysis. The infants within our dataset,

generated by Rhee et al., were divided into those who were exclusively breastfed, termed EBF, and those who received either exclusive formula feeding or a combination of formula and breastfeeding, termed NEBF. Ultimately, we found that the EBF group had an overall reduced diversity and had taxa that were specific to lactose production or consumption. In contrast, the NEBF group had higher diversity and more closely resembled the complex adult microbiome. These results corroborate current literature.

EBF may lower overall diversity of the infant gut microbiome. The results from our diversity metrics corroborate those of previous studies that have demonstrated that diversity in EBF groups of infants tend to be lower than NEBF groups (44). Formula-fed infants typically have a gut microbiome that resembles the adult microbiome, which tends to be significantly more complex, at an earlier point in development (13). Interestingly, though higher diversity is associated with overall better health outcomes in adults, the opposite trend is proposed for infants (44, 46). Nevertheless, there is a clear difference in diversity when looking at the weighted UniFrac distance metric between the EBF and NEBF groups, and a decrease in overall diversity in the EBF group as shown by alpha diversity analysis. Together, this suggests that exclusive breastfeeding, not only breastfeeding in general, may be necessary to obtain the full benefits of breastmilk.

EBF may increase the abundance of beneficial lactose metabolism-associated Bifidobacteria. Several of the genera found to be differentially abundant in the EBF group of infants are supported in literature. The *Bifidobacterium sp.* bacteria were found to be both differentially and relatively abundant within the EBF group in our analysis (**Fig. 2B**). They are known users of the lactose in human milk oligosaccharides (HMOs), and are highly associated with exclusive breastfeeding in infants (47). Increased prevalence of *Bifidobacteria* in the infant gut is associated with many health benefits including lower rates of obesity, long-term asthma, and may prevent the development of allergies (48). Furthermore, the predominating *Bifidobacterium* bacteria population in the gut microbiome, may out-compete other bacterial species in populating the developing microbiota, and is therefore considered to be a potential reason for the decreased microbial diversity observed in breastfeeding infants (44).

Our results also demonstrate that lactate-associated bacterial genera are more prevalent within infants in the EBF group compared to the NEBF group. Specifically, *Veillonella, Lactobacillus,* and *Haemophillus,* which are more abundant in the EBF group (**Fig. 2A)**, are known, respectively, to utilize lactate, produce lactic acid, and positively correlated with duration of exclusive infant breastfeeding (47). While it is unclear in literature whether *Bacterioides* sp. is associated with infant feed or naturally present in all infants after introduction of complex foods, the presence of the *Bacteroides* genus, nevertheless, has been linked to the development of healthy and complex gut microbiomes later in life (13).

For the differential abundance in the NEBF group of infants, there is seemingly little research surrounding the correlation of *Proteus* and *Acinetobacter* genera and feed-type. The existing literature on the *Proteus* genus primarily explored the bacteria's role in pathogenicity; even this, though, has not been well-established (49). Although there is some support for *Acinetobacter* bacteria being more prevalent with infants who are non-exclusively breastfed, greater focus is placed on the opportunistic pathogenicity of *Acinetobacter* in early life in the literature (13, 50). An interesting result observed lies in the apparent differential abundance of the *Megasphaera* genus in the NEBF group, despite having been found to be strongly associated with EBF infants within the literature (47). This can likely be attributed to the limitations in the size of our dataset, which may allow outliers to greatly skew the data.

EBF indicator taxa are associated with feed-type and healthy infant development in literature. An exploration of the different indicator taxa in the EBF versus NEBF groups provided an alternative method to look at taxonomic differences in the gut microbial composition that may arise due to feed-type. Overall, the indicator taxa observed within the EBF group are well supported in the literature and associated with both healthy infant development and breastfeeding. The *Rothia* genus is one of the earliest colonizers observed in infants that tends to be found throughout the gastrointestinal tract and is a commensal group

of bacteria associated with breast milk (51, 52). *Staphylococcus* bacteria are typically associated with pathogenesis and the literature surrounding colonization in the infant gut primarily reflects this role (53). Further research is required to elucidate the role of feed type on early *Staphylococcus* colonization of the infant gut, given that the genus appeared to be both differentially abundant within the EBF group (**Fig. 2A**) and an indicator taxa for the EBF group within our analysis (**Table 1**).

Most NEBF indicator taxa are associated with feed-type and disease in literature. The prevalence of most indicator taxa in the NEBF group (namely *Lachnoclostridium*, *Clostridium innocuum*, *Tyzzerella*, and *Anaerostipes*) were also well-supported (13, 54–56). It is interesting to note that some of these taxa have been previously linked to certain diseases. *Tyzzerella* bacteria for example, have been linked to cardiovascular disease, and are typically found at lower abundances in breastfed infants (56, 57). In contrast, *Anaerostipes* bacteria have been linked to decreased rates of allergy development, and are typically observed in a healthy infant gut microbiome starting at around 4 months of age for breastfed infants, as they begin to wean-off of exclusive breastfeeding (54, 58). Further longitudinal analysis of infant feeding practices in the future may allow for a greater understanding of the positive and negative effects that feed-type may have on infant gut microbiome development.

Certain NEBF indicator taxa have not been conclusively associated with feed-type in literature. However, the role of feed-type on the distribution of some indicator taxa found in the NEBF category, *Clostridioles difficile, Intestinbacter*, *Proteus, Acinetobacter,* and *Stenotrophomonas,* does not appear to be supported conclusively by literature (49, 59, 60). *C. difficile* has been associated with negative health outcomes, and appears to be highly prevalent within the infant gut microbiome (61). The effects of this species on the development of the gut microbiome, interestingly, appears to differ based on feed-type (61). While it does not appear to greatly alter the gut microbiome of exclusively formula-fed (EFF) infants, EBF infants colonized with *C. difficile* appear to develop a more adult-like gut microbiome compared to their non-colonized EBF infant counterparts (61).

Given both the *Proteus* and *Acinetobacter* genera appeared in both our taxonomic and differential abundance analyses, further exploration into the role of these genera on the gut microbiome in the context of feed-type may be required. The *Stenotrophomonas* genus in the literature is primarily associated with breast milk, however it has been noted that it is seen in infants with a delayed onset of breastfeeding (60). Since our NEBF group of infants included those who received combined feeding, these results may be supported, but would benefit from further research and analysis. *Negativicoccus succinicivorans* is not a well-studied or established organism within the literature, and did not appear to have significant results within our study as well (62). An interesting result observed within the indicator taxa for the NEBF group was the presence of the *Clostridium sensu stricto* bacteria, which appeared to be differentially abundant within the EBF group in our analysis. Literature suggests that *Clostridium sensu stricto* may be prevalent in the composition of most infant gut microbiomes, however an imbalance is linked to pathogenesis (63). Given the contradictory results, further research for this species would likely be beneficial.

Limitations A major limitation of this study was the size of the dataset. In particular, the sample sizes between feed types were both small and unbalanced; the majority of infants were exclusively breastfed, while only a small subset were exclusively formula fed. To ensure that statistically relevant observations could be made, we were prompted to group samples by exclusive breastfeeding and non-exclusive breastfeeding, which included the combined and formula feed type. However, in doing so, we could no longer explore the differences between all three feed types (breastfed, formula fed and combined), which could have yielded more distinct results. In the weighted UniFrac PCoA plot, although clusters were observed between the NEBF and EBF groups, there was significant overlap between the NEBF and EBF groups (Figure 1B). We believe that the compilation of the combined and formula-feeding infants into one category may have ultimately resulted in the greater amount of overlap between our two groups of interest.

This limited sample size also prevented us from appropriately filtering to control for confounding variables. For instance, factors such as birth type, use of medication, use of

probiotics and maternal characteristics, like age and weight, have been previously associated with changes in the infant gut microbiome (8, 64, 65). By controlling for these variables, we could better isolate the specific effects of feed-type on the infant gut microbiome.

Finally, all infants within this dataset were born in Michigan, USA. As a result, due to their shared geographical origin and social environment, the microbiomes of these infants may be more similar to each other than to infants from other locations. Therefore, caution must be applied when extrapolating our results.

Conclusions Our study aimed to observe the effects of EBF on the developing infant gut microbiome. We found that EBF, in comparison to NEBF, is associated with a decrease in microbial diversity and a shift towards the dominance of lactose metabolism-related microbial taxa in the infant gut microbiome. More specifically, alpha and beta diversity analysis revealed a significant difference between the microbial diversity observed in the EBF and NEBF groups, with the former having a lower diversity overall. Differential and relative abundance analysis, as well as indicator taxa analysis, revealed a difference in the bacterial genera and species that occupied the EBF and NEBF groups. In particular, the taxonomic composition of the EBF group was dominated by bacterial taxa associated with lactose metabolism. Taken together, this suggests that EBF, and not just breastfeeding in general, is correlated with differences in the microbial diversity and composition of the infant gut microbiome which further supports and expands on previous literature findings.

Future Directions If possible, the generation of a larger and more robust dataset that includes or controls for potential confounding factors, as well as factors such as length of hospital stay, whether or not the infants were preterm, and a more consistent sampling schedule across the infants within the dataset would likely yield more significant and reliable results (8, 13).

In this study, we demonstrated that differential feed type has an impact on the taxonomic composition of the infant gut microbiome. To build upon the observations in our analysis, it would be interesting to explore whether differential feed type has an impact on the metabolic composition of microbial communities as well. Since differences in microbial gut metabolism can impact nutrient acquisition and may correlate with certain health outcomes later in life, this may have therapeutic implications and would be worthwhile to explore (54, 66, 67).

ACKNOWLEDGEMENTS

First, we would like to thank Dr. Rhee et al. for providing the dataset. Furthermore, we would like to thank the MICB 447 teaching team: Dr. Evelyn Sun, Dr. Stephan Koenig, Zakhar Krekhno, and Stefanie Sternagel for their support, guidance, and feedback throughout the project. Lastly, we would like to thank the Department of Microbiology, and the University of British Columbia for providing the funding and resources for this project.

CONTRIBUTIONS

Generally, experimental procedures were written and conducted as a collective effort among the authors. More specifically, the relative and differential abundance analyses were performed by E.C, while the alpha and beta diversity analyses were performed by M.P., D.S., and L.U. Indicator taxa analysis was performed by L.U. All authors contributed to the writing and the editing of the manuscript.

REFERENCES

1. **Greiner T**. 2014. Exclusive breastfeeding: measurement and indicators. *Int Breastfeed J* **9**:18.
2. **WHO** | Indicators for assessing infant and young child feeding practices. WHO. World Health Organization.
3. **Butte NF, Lopez-Alarcon MG, Garza C**. Expert Consultation on the Optimal Duration of Exclusive Breastfeeding (2001: Geneva S. 2002. Nutrient adequacy of exclusive breastfeeding for the term infant during the first six months of life. World Health Organization.
4. **Kent JC, Mitoulas LR, Cregan MD, Ramsay DT, Doherty DA, Hartmann PE**. 2006. Volume and Frequency of Breastfeedings and Fat Content of Breast Milk Throughout the Day. *Pediatrics* **117**:e387–e395.
5. **Hamosh M**. 1996. Digestion in the newborn. *Clin Perinatol* **23**:191–209.
6. **Sheard NF, Walker WA**. 1988. The role of breast milk in the development of the gastrointestinal tract. *Nutr Rev* **46**:1–8.

7. **Muscogiuri G, Cantone E, Cassarano S**, **Tuccinardi D, Barrea L, Savastano S, Colao A**. 2019. Gut microbiota: a new path to treat obesity. *Int J Obes Suppl* **9**:10–19.
8. **Lin L, Tse E, Yau E**. 2021. Mode of delivery and maternal body mass index are weakly associated with the infant gut microbiota composition. *Undergrad J Exp Microbiol Immunol* **26**.
9. **Ma J, Li Z, Zhang W, Zhang C, Zhang Y, Mei H, Zhuo N, Wang H, Wang L, Wu D**. 2020. Comparison of gut microbiota in exclusively breast-fed and formula-fed babies: a study of 91 term infants. *Sci Rep* **10**:15792.
10. **Yeung H, Leff M, Rhee KE**. 2017. Effect of Exclusive Breastfeeding Among Overweight and Obese Mothers on Infant Weight-for-Length Percentile at 1 Year. *Breastfeed Med Off J Acad Breastfeed Med* **12**:39–47.
11. **Lin L, Tse E, Yau E.** 2021. Mode of delivery and maternal body mass index are weakly associated with the infant gut microbiota composition. *Undergrad J Exp Microbiol Immunol* **26**.
12. **Levin AM, Sitarik AR, Havstad SL, Fujimura KE, Wegienka G, Cassidy-Bushrow AE, Kim H, Zoratti EM, Lukacs NW, Boushey HA, Ownby DR, Lynch SV, Johnson CC.** 2016. Joint effects of pregnancy, sociocultural, and environmental factors on early life gut microbiome structure and diversity. *Sci Rep* **6**:31775.
13. **Moore RE, Townsend SD**. 2019. Temporal development of the infant gut microbiome. *Open Biol* **9**:190128.
14. **Kim H, Sitarik AR, Woodcroft K, Johnson CC, Zoratti E.** 2019. Birth Mode, Breastfeeding, Pet Exposure, and Antibiotic Use: Associations With the Gut Microbiome and Sensitization in Children. *Curr Allergy Asthma Rep* **19**:22.
15. **Caporaso JG, Ackermann G, Apprill A, Bauer M, Berg-Lyons D, Betley J, Fierer N, Fraser L, Fuhrman JA, Gilbert JA, Gormley N, Humphrey G, Huntley J, Jansson JK, Knight R, Lauber CL, Lozupone CA, McNally S, Needham DM, Owens SM, Parada AE, Parsons R, Smith G, Thompson LR, Thompson L, Turnbaugh PJ, Walters WA, Weber L.** 2018. EMP 16S Illumina Amplicon Protocol. *protocols.io*.
16. **R Core Team**. 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
17. **RStudio Team**. 2021. RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA.
18. **Wickham H, François R, Henry L, Müller K**. 2021. dplyr: A Grammar of Data Manipulation.
19. **Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H.** 2019. Welcome to the tidyverse. *J Open Source Softw* **4**:1686.
20. **Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG.** 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* https://doi.org/10.1038/s41587-019-0209-9.
21. **Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP.** 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**:581.
22. **Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É**. 2011. Scikit-learn: Machine learning in Python. J Mach Learn Res **12**:2825–2830.
23. **Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO.** 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**:7188–7196.
24. **Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO.** 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**:D590-6.
25. **Katoh K, Standley DM.** 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol* **30**:772–780.
26. **Price MN, Dehal PS, Arkin AP.** 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS ONE* **5**:e9490.
27. **Shannon CE**. 1948. A mathematical theory of communication. *Bell Syst Tech J* **27**:379–423, 623–

656.
28. **Faith DP.** 1992. Conservation evaluation and phylogenetic diversity. *Biol Conserv* **61**:1–10.
29. **Pielou EC.** 1966. The measurement of diversity in different types of biological collections. *J Theor Biol* **13**:131–144.
30. **Sørensen T.** 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biol Skr* **5**:1–34.
31. **Jaccard P.** 1908. Nouvelles recherches sur la distribution floral. *Bull Soc Vard Sci Nat* **44**:223–270.
32. **Lozupone CA, Hamady M, Kelley ST, Knight R.** 2007. Quantitative and qualitative diversity measures lead to different insights into factors that structure microbial communities. Appl Environ Microbiol **73**:1576–1585.
33. **Lozupone C, Knight R.** 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**:8228–8235.
34. **Kruskal WH, Wallis WA.** 1952. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc* **47**:583–621.
35. **Anderson MJ.** 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecol* **26**:32–46.
36. **McMurdie PJ, Holmes S.** 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE* **8**:e61217.
37. **Love MI, Huber W, Anders S.** 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:550.
38. **Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H.** 2020. vegan: Community Ecology Package.
39. **Paradis E, Schliep K.** 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**:526–528.
40. **Wickham H.** 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
41. **Kassambara A.** 2020. ggpubr: "ggplot2" Based Publication Ready Plots.
42. **Legendre P, Legendre L.** 2012. *Numerical Ecology*, p. 499. *In* Third. Elsevier.
43. **Fortunato CS, Eiler A, Herfort L, Needoba JA, Peterson TD, Crump BC**. 2013. Determining indicator taxa across spatial and seasonal gradients in the Columbia River coastal margin. *ISME J* **7**:1899–1911.
44. **Ma J, Li Z, Zhang W, Zhang C, Zhang Y, Mei H, Zhuo N, Wang H, Wang L, Wu D**. 2020. Comparison of gut microbiota in exclusively breast-fed and formula-fed babies: a study of 91 term infants. *Sci Rep* **10**:15792.
45. **Durack J, Lynch SV.** 2019. The gut microbiome: Relationships with disease and opportunities for therapy. *J Exp Med* **216**:20–40.
46. **Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, Almeida M, Arumugam M, Batto J-M, Kennedy S, Leonard P, Li J, Burgdorf K, Grarup N, Jørgensen T, Brandslund I, Nielsen HB, Juncker AS, Bertalan M, Levenez F, Pons N, Rasmussen S, Sunagawa S, Tap J, Tims S, Zoetendal EG, Brunak S, Clément K, Doré J, Kleerebezem M, Kristiansen K, Renault P, Sicheritz-Ponten T, de Vos WM, Zucker J-D, Raes J, Hansen T, MetaHIT consortium, Bork P, Wang J, Ehrlich SD, Pedersen O.** 2013. Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**:541–546.
47. **Laursen MF, Andersen LBB, Michaelsen KF, Mølgaard C, Trolle E, Bahl MI, Licht TR.** Infant Gut Microbiota Development Is Driven by Transition to Family Foods Independent of Maternal Obesity. *mSphere* **1**:e00069-15.
48. **Arboleya S, Watkins C, Stanton C, Ross RP**. 2016. Gut Bifidobacteria Populations in Human Health and Aging. *Front Microbiol* **7**:1204.
49. **Hamilton AL, Kamm MA, Ng SC, Morrison M.** 2018. Proteus spp. as Putative Gastrointestinal Pathogens. *Clin Microbiol Rev* **31**:e00085-17.
50. **Alderete TL, Jones RB, Shaffer JP, Holzhausen EA, Patterson WB, Kazemian E, Chatzi L, Knight R, Plows JF, Berger PK, Goran MI.** 2021. Early life gut microbiota is associated with rapid infant growth in Hispanics from Southern California. *Gut Microbes* **13**:1961203.
51. **Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO.** 2007. Development of the Human Infant Intestinal Microbiota. *PLOS Biol* **5**:e177.
52. **Roswall J, Olsson LM, Kovatcheva-Datchary P, Nilsson S, Tremaroli V, Simon M-C, Kiilerich P, Akrami R, Krämer M, Uhlén M, Gummesson A, Kristiansen K, Dahlgren J, Bäckhed F.** 2021. Developmental trajectory of the healthy human gut microbiota during the first 5 years of life. *Cell Host Microbe* **29**:765-776.e3.
53. **Yang I, Corwin EJ, Brennan PA, Jordan S, Murphy JR, Dunlop A.** 2016. The Infant Microbiome: Implications for Infant Health and Neurocognitive Development. *Nurs Res* **65**:76–88.
54. **Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, Li Y, Xia Y, Xie H, Zhong H, Khan MT, Zhang J, Li J, Xiao L, Al-Aama J, Zhang D, Lee YS, Kotowska D, Colding C, Tremaroli V, Yin Y, Bergman S, Xu X, Madsen L, Kristiansen K, Dahlgren J, Wang J.** 2015. Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**:690–703.
55. **Li N, Yan F, Wang N, Song Y, Yue Y, Guan J, Li B, Huo G.** 2020. Distinct Gut Microbiota and Metabolite Profiles Induced by Different Feeding Methods in Healthy Chinese Infants. *Front*

Microbiol **11**:714.

56. **Stewart CJ, Ajami NJ, O'Brien JL, Hutchinson DS, Smith DP, Wong MC, Ross MC, Lloyd RE, Doddapaneni H, Metcalf GA, Muzny D, Gibbs RA, Vatanen T, Huttenhower C, Xavier RJ, Rewers M, Hagopian W, Toppari J, Ziegler A-G, She J-X, Akolkar B, Lernmark A, Hyoty H, Vehik K, Krischer JP, Petrosino JF**. 2018. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**:583–588.
57. **Ascher S, Reinhardt C.** 2018. The gut microbiota: An emerging risk factor for cardiovascular and cerebrovascular disease. *Eur J Immunol* **48**:564–575.
58. **Turroni F, Milani C, Duranti S, Lugli GA, Bernasconi S, Margolles A, Di Pierro F, van Sinderen D, Ventura M**. 2020. The infant gut microbiome as a microbial organ influencing host well-being. Ital J Pediatr **46**:16.
59. **Dedrick S, Sundaresh B, Huang Q, Brady C, Yoo T, Cronin C, Rudnicki C, Flood M, Momeni B, Ludvigsson J, Altindis E.** 2020. The Role of Gut Microbiota and Environmental Factors in Type 1 Diabetes Pathogenesis. *Front Endocrinol* **0**.
60. **Gonzalez E, Brereton NJB, Li C, Lopez Leyva L, Solomons NW, Agellon LB, Scott ME, Koski KG**. 2021. Distinct Changes Occur in the Human Breast Milk Microbiome Between Early and Established Lactation in Breastfeeding Guatemalan Mothers. *Front Microbiol* **12**:557180.
61. **Drall KM, Tun HM, Morales-Lizcano NP, Konya TB, Guttman DS, Field CJ, Mandal R, Wishart DS, Becker AB, Azad MB, Lefebvre DL, Mandhane PJ, Moraes TJ, Sears MR, Turvey SE, Subbarao P, Scott JA, Kozyrskyj AL.** 2019. Clostridioides difficile Colonization Is Differentially Associated With Gut Microbiome Profiles by Infant Feeding Modality at 3–4 Months of Age. *Front Immunol* **10**:2866.
62. **Church DL, Simmon KE, Sporina J, Lloyd T, Gregson DB.** 2011. Identification by 16S rRNA Gene Sequencing of Negativicoccus succinicivorans Recovered from the Blood of a Patient with Hemochromatosis and Pancreatitis. *J Clin Microbiol* **49**:3082–3084.
63. **Zhou Y, Shan G, Sodergren E, Weinstock G, Walker WA, Gregory KE.** 2015. Longitudinal Analysis of the Premature Infant Intestinal Microbiome Prior to Necrotizing Enterocolitis: A Case-Control Study. *PLOS ONE* **10**:e0118632.
64. **Levin AM, Sitarik AR, Havstad SL, Fujimura KE, Wegienka G, Cassidy-Bushrow AE, Kim H, Zoratti EM, Lukacs NW, Boushey HA, Ownby DR, Lynch SV, Johnson CC.** 2016. Joint effects of pregnancy, sociocultural, and environmental factors on early life gut microbiome structure and diversity. *Sci Rep* **6**:31775.
65. **Vandenplas Y, Carnielli VP, Ksiazyk J, Luna MS, Migacheva N, Mosselmans JM, Picaud JC, Possner M, Singhal A, Wabitsch M.** 2020. Factors affecting early-life intestinal microbiota development. *Nutrition* **78**:110812.
66. **Vojinovic D, Radjabzadeh D, Kurilshikov A, Amin N, Wijmenga C, Franke L, Ikram MA, Uitterlinden AG, Zhernakova A, Fu J, Kraaij R, van Duijn CM.** 2019. Relationship between gut microbiota and circulating metabolites in population-based cohorts. *Nat Commun* **10**:5813.
67. **Ho NT, Li F, Lee-Sarwar KA, Tun HM, Brown BP, Pannaraj PS, Bender JM, Azad MB, Thompson AL, Weiss ST, Azcarate-Peril MA, Litonjua AA, Kozyrskyj AL, Jaspan HB, Aldrovandi GM, Kuhn L.** 2018. Meta-analysis of effects of exclusive breastfeeding on infant gut microbiota across populations. *Nat Commun* **9**:4169.