Determinants of Early Child Development and Growth in Guatemalan Infants: An Analysis of Disruptions in Co-Development of the Gut Microbiome and Host

by Lauren E. Louck, M.S.

A dissertation presented to the faculty of the Gerald J. and Dorothy R. Friedman School of Nutrition Science and Policy at Tufts University, Boston, MA, USA

In partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Biochemical and Molecular Nutrition Program Defense date: April 25, 2025

Thesis committee chair and advisor:
Mei Chung, PhD, MPH
Friedman School of Nutrition Science and Policy, Tufts University

Committee members:
Taylor C. Wallace, PhD, CFS, FACN
Think Health Group, Inc., Washington, D.C., USA; Friedman School of Nutrition Science and Policy, Tufts University

Naisi Zhao, Dr.PH
Tufts University School of Medicine, Public Health and Community Medicine

Lynne L. Ausman, DSc, RD
Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA, USA

ABSTRACT

Background: Malnutrition in the first 1000 days can lead to lifelong developmental and growth deficits. Indigenous Mayans in Guatemala have some of the highest rates of malnutrition and stunting in the world. While research suggests an association of gut microbiota composition with early child development (ECD) and growth outcomes in healthy human infants, higher risk populations and the role of early localized gastrointestinal (GI) inflammation have been less studied. The **main objective** of this research was to determine if diet diversity, fecal microbiome composition, and GI inflammation were associated with infant development scores in a sub-population of Guatemalan infants enrolled in The Saqmolo' Project. The **central hypothesis** was more optimal ECD and growth scores would be associated with higher diet diversity, a more mature GI microbiome composition, and lower levels of GI inflammation.

Methods: Endline data were used from the Saqmolo' Project Microbiome Sub-Study of 11-16-month-old indigenous Mayan Guatemalan infants. The Saqmolo' Project was a 6-month randomized controlled trial with infants receiving one egg a day compared to standard care controls. Fecal microbiome composition was assessed with 16S rRNA sequencing and QIIME2. Caregiver Reported Early Child Development Instruments (CREDI) long form measured ECD. Anthropometric Z-scores and WHO criteria for stunting were used for growth outcomes. The World Health Organization (WHO) Infant and Young Child Feeding indicators questionnaire measured minimum diet diversity and most frequently consumed milk type with additional questionnaires. Fecal microbiome co-abundance groups were created using guild-based clustering. **Aim 1 explored the association of egg intake versus multiple micronutrients powders, diet diversity, and most frequently consumed milk type with fecal microbiome composition and fecal calprotectin (ng/ml).** MaAsLin2, DESeq2, and LEfSe were used to assess the relationships between dietary factors with microbiome composition. Fecal calprotectin and dietary associations were evaluated with linear regression. **Aim 2 examined the association of fecal microbiome composition with child development metrics** using MaAsLin2**. Aim 3 explored the association of fecal calprotectin with fecal microbiome composition** using MaAsLin2, **and ECD and growth metrics** using linear regression.

Results: In **Aim 1**, among 200 infants (mean age 12.7 months, 49% female), microbiome composition differences between egg and control groups were more pronounced in non-breastfed than breastfed infants. LEfSe analysis identified 23 ASVs with higher and four ASVs with lower abundance in non-breastfed infants. Neither egg intake nor minimum diet diversity (MDD) was associated with specific ASVs or co-abundance groups. Fecal calprotectin levels were not significantly associated with treatment (β = -57.85, 95% CI: -119.08 to 98.52, P = 0.281) or MDD (β = -10.28, 95% CI: -187.01 to 71.31, P = 0.337), but were lower in non-breastfed infants (β = -174.70, 95% CI: -314.16 to -35.23, P = 0.014). **In Aim 2**, CREDI overall (co-abundance group 18: MaAsLin2 coefficient = 1.22, Q = 0.216) and language domain (co-abundance group 18: MaAsLin2 coefficient = 1.95, Q = 0.0226) were positively associated with co-abundance groups in non-breastmilk-fed infants. In **Aim 3**, fecal calprotectin was associated with higher relative abundance of co-abundance groups 8 (MaAsLin2 coefficient = 0.578, q = 0.105) and 28 (MaAsLin2 coefficient = 0.588, q = 0.245), and lower relative abundance of co-abundance group 1 (MaAsLin2 coefficient = -0.392, q = 0.136). Associations between fecal calprotectin and CREDI scores, Z-scores, or stunting in linear regression models were not significant (all P > 0.05).

Conclusion: The findings underscore the influence of breastfeeding on gut microbiome stability, with non-breastfed infants exhibiting greater microbial shifts in response to egg intervention. The positive association of the co-abundance group with CREDI language scores indicates these microbes could contribute to neurodevelopmental benefits. Localized GI inflammation does not appear to have a significant impact on ECD or growth outcomes, though it was linked to gut microbiome composition.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ACRONYMS AND ABBREVIATIONS

ANOVA	Analysis of Variance
ASVs	Amplicon sequencing variants
BDNF	Brain-derived neurotrophic factor
BMI	Body Mass Index
CAG	Co-abundance Group
CREDI	Caregiver Reported Early Child Development Instruments
ECD	Early Childhood Development
ELISA	Enzyme-Linked Immunosorbent Assay
FDR	False Discovery Rate
GABA	Gamma-aminobutyric acid
GI	Gastrointestinal
GLP-1	Glucagon-like peptide
HAZ	Head circumference-for-age
HMO	Human milk oligosaccharides
IGF-1	Insulin-like growth factor 1
LAZ	Length-for-age
LMIC	Low- and Middle-Income Countries
MaAsLin2	Multivariate Associations with Linear Models
PCoA	Principal Coordinate Analysis
QIME2	Quantitative Insights into Microbial Ecology version 2
QQ	Quantile-quantile
SCFA	Short chain fatty acids
SD	Standard Deviation

WAL	Weight-for-length
WAZ	Weight-for-age
WHO	World Health Organization

CHAPTER 1: Introduction

While research suggests an association between gut microbiota composition and early child development outcomes in healthy human infants, causation and mechanistic links have not been proven. Research is mixed on whether microbial diversity impacts cognitive development outcomes but does suggest microbial co-abundance patterns may be a driving factor. Limited research has been done to elucidate interacting factors between diet, milk type, complimentary foods, gut microbiome composition, and development. The current research on ECD and gut microbiome associations is limited to healthy infants and higher risk populations have been less studied. Lastly, gut dysbiosis is a known risk factor for systemic inflammation, yet the impact of localized GI inflammation on gut microbiome composition and co-development has not been investigated in a dietary context. This study will contribute to elucidating the role of the developing gut microbiome on the co-developing infant host and contribute to prevention of developmental deficits that impact long-term health outcomes, economic productivity, and quality of life.

Main Objective

The main objectives of the proposed research are to determine whether 1) diet diversity; 2) fecal microbiome composition; and 3) fecal calprotectin are associated with infant development scores, in a sub-population of indigenous Mayan infants enrolled in The Saqmolo' Project. The *central hypothesis* is that more optimal development scores will be associated with higher diet diversity, a more mature gastrointestinal (GI) microbiome composition, and lower levels of localized GI inflammation.

Study Aims

A diagram of Specific Aims can be seen in **Figure 1.1**. Each aim is described below.

- **Aim 1:** Determine the association of egg intake versus standard care controls, and diet diversity with 1) endline fecal microbiome composition; and 2) localized GI inflammation (measured with fecal calprotectin).
    - *Hypothesis 1:* Egg intake and higher diet diversity scores will be associated with a more mature GI microbiome composition and lower levels of GI inflammation.
- **Aim 2:** Determine whether endline fecal microbiome composition is associated with differences in 1) early child development scores; 2) growth anthropometrics (z-scores for weight-for-age, length-for-age, weight-for-length, and head circumference-for-age); and 3) incidence of stunting and wasting.
    - *Hypothesis 2:* A more mature endline fecal microbiome composition and microbial community structure will be associated with 1) higher early child development scores; 2) more optimal growth anthropometrics; and 3) lower incidence of stunting and wasting.
- **Aim 3:** Determine whether endline localized GI inflammation (measured with fecal calprotectin) is associated with differences in 1) endline fecal microbiome composition; 2) early child development scores; 3) growth anthropometrics (z-scores for weight-for-age, length-for-age, weight-for-length, and head-circumference-for-age); and 4) stunting and wasting incidence.
    - *Hypothesis 3:* Lower endline fecal calprotectin concentrations will be associated with 1) a more mature fecal microbiome composition; 2) higher early child development scores; 3) more optimal growth anthropometric scores; and 4) lower incidence of stunting and wasting.

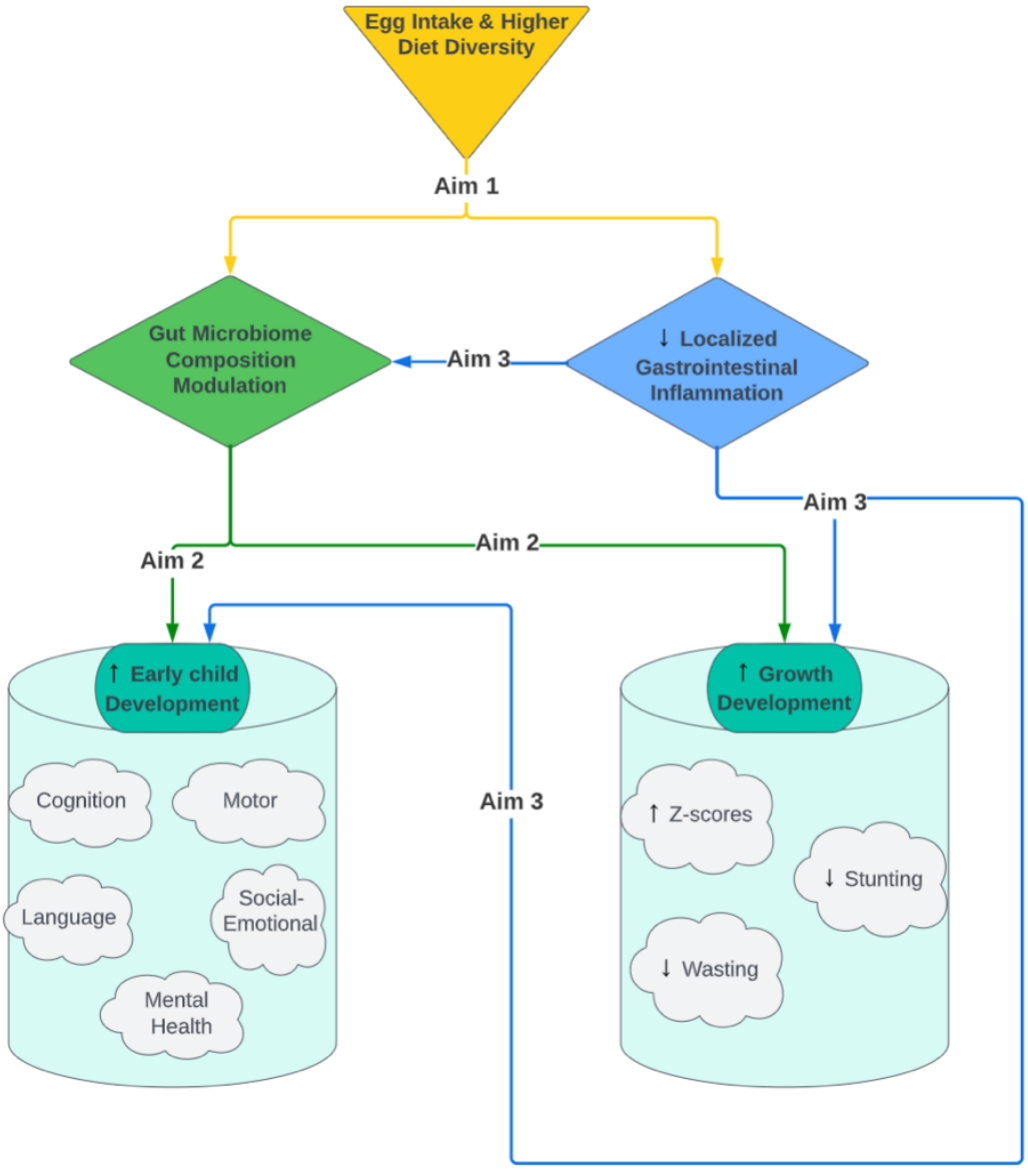

Figure 1.1. Specific aims of the research.

This figure depicts the proposed aims of the research. The specific aims are to determine whether 1) diet diversity; 2) gut microbiome composition; and 3) localized GI inflammation are associated with infant development scores (ECD and growth).

CHAPTER 2: Review of the Literature

Development during the first 1000 days after birth sets the stage for a healthy adult life. There are multiple stages, known as sensitive development periods, which require specific inputs for healthy development. Numerous environmental inputs affect and alter infant development, including nutritional (diet diversity, breastfeeding practices, and macro- and micronutrient intakes), environmental (living conditions, healthcare access, and caregiver education), medical (early life stress, birth delivery mode, antibiotic use, diarrheal episodes), and genetic factors. Early child development (ECD) is sensitive to these nutritional and environmental stimuli, and adequate nutrition is critical to its evolution during the first 1000 days of life.[1-6]

Infant and Gut Microbiome Co-Development

The gut microbiome and infant host are hypothesized to co-develop in tandem. Research in healthy infants points to patterns of dynamic age-related changes in gut microbiome composition and development, with the microbiome stabilizing at age 2-3 and resembling composition of an adult microbiome.[7,8] Large changes in gut microbiota development and composition parallel other developmental bursts such as brain development and growth.[7]Evidence also suggests inflammation may play a role in a less developed gut microbiome, with potential impacts on infant health,[9] cognitive function,[7,10] and physical growth.[11,12]

Infant Microbiome Characteristics

The infant gut microbiome has a distinct composition compared to the adult microbiome, with infant composition being less diverse and less stable.[7] While composition among infants varies greatly based on ethnicity and location, research consistently reports temporal progression of microbiome development in healthy infants.[13] Early intestinal colonization is highly

influenced by mode of delivery and infant diet, with the main dietary factor being breastfeeding versus no breastfeeding.[13] Early antibiotic exposure is an additional factor which heavily impacts infant gut microbiome development.[13] Early infant exposure to antibiotics is associated with an under-developed gut microbiome composition compared to healthy infants, and in some cases an increased risk for allergies, asthma, and obesity later in life.[13]

Microbiota-Gut-Brain Axis

The concept of sensitive development periods also applies to the gut microbiome and the gut-brain axis. The gut-brain axis is hypothesized to underly infant and gut microbiome co-development. It is further hypothesized that absent nutrition inputs to the gut microbiome during the first 1000 days of life may disrupt or impair optimal development of the gut-brain axis, subsequently disrupting critical phases in early child development (ECD) and growth.[7,14-18]

Mechanisms of the Gut-Brain Axis

The gut-brain axis is the biological interface through which the gut and brain communicate in a bi-directional relationship.[19] The gut-brain axis relays information and this communication occurs via neuro-immuno-endocrine mediators. Key elements involved in gut-brain axis co-development during infancy are maturation of intestinal barrier function, blood-brain barrier function, the immune system, and nutrient extraction (**Figure 2.1**).[7] Viewing the effect of gut microbiota modulation on the brain, research in mice has shown changes in neurotransmitters such as gamma-aminobutyric acid ([GABA]: an inhibitory neurotransmitter) and brain-derived neurotrophic factor (BDNF), intestinal barrier protection and integrity, bacterial metabolites absorbed by enterocytes, modulation of enteric sensory afferents, and mucosal immune regulation.[19] In the reverse, the brain effects the gut microbiota by alterations in mucus production, motility, and immune function coalescing in a bi-directional highway of communication.[19]

The immune system is a core pathway in the gut-brain axis and its development occurs in tandem with microbiota maturation, particularly at weaning.[20] During weaning there are significant fluctuations in microbiota composition and an associated inflammatory response, known as the weaning reaction. Researchers report a greater risk for inflammatory and immune disorders later in life, when the weaning reaction or gut microbiome development is disrupted.[7,8,21] Gut dysbiosis is known to upregulate inflammatory responses in the gut, instigating bi-directional inflammatory responses in other areas of the body such as the liver,[22] lungs,[23,24] and brain.[10] While systemic inflammatory responses due to gut dysbiosis have been well-studied in adults, research in infants and children is sparse and localized GI inflammation has not been studied in this context.

Studies in mice have demonstrated the functional effects of the microbiota-gut-brain axis. Age-related changes in mouse gut microbiota via young and old fecal donor mice to healthy adult mice, triggered modulation of lipids in the aging brain.[25] Specifically, expression of monounsaturated and polyunsaturated fat synthesis genes were downregulated in mice who received fecal transplants from old mice.[25] A study comparing wild-type vs germ free mice showed gut microbiota can affect both host energy and lipid metabolism, which have implications for growth and weight gain.[26] Overall, these studies demonstrate how changes to the gut microbiome can provoke functional changes in the brain and host metabolism.

Diet and Microbiota

Gut microbiota consume food eaten by the human host, metabolize the nutrients, and secrete metabolites. Certain metabolites can be absorbed by enterocytes and reach the blood stream, which may modulate immune, endocrine, incretin, and genetic processes in the host. The most widely researched gut microbiota metabolites are short chain fatty acids (SCFAs). SCFAs

are known to modulate immune responses in colonocytes, strengthen tight junctions, reduce gut permeability, provide fuel for colonocytes, and combat infiltration of pathogens.[27] Multiple studies have reported SCFAs triggering an increase in GLP-1 resulting in improved insulin secretion, thus having positive effects on blood glucose.[27] Other microbial metabolites and downstream effects are coming to light, which will be described in the following section.

While a large body of research has investigated the effect of fibers on gut microbiome composition and function,[27] there is limited research investigating the effects of proteins, lipids, and micronutrients on the gut microbiome, particularly in the context of healthy development. To date, one randomized controlled trial investigated the effect of protein (bovine colostrum supplemented with egg powder) compared to corn/soy controls on Malawian children gut microbiome composition, reporting an increase in 5 bacterial species (*Clostridium perfringens, Megamonas rupellensis, Streptococcus thermophilus, Eubacterium sp.* cL1013, and *Megasphaera paucivorans*) and decrease in 1 species (*Lactobacillus sp.* DJFRP24).[28] The effect of protein on the infant gut microbiome is highly relevant to LMIC children, considering the prevalence of protein malnutrition in this population.

Eggs Nutrients, Microbiota & Development Outcomes

Eggs are currently being considered and tested as an alternative intervention for LMIC children over traditional multiple micronutrient powders. There is evidence that micronutrient powders may elicit further gut dysbiosis in infants and rat pups, likely due to the high non-heme iron content.[29-31] Additionally, eggs have an ideal nutrient profile for optimal cognitive and growth development, boasting a complete essential amino acid profile, omega-3 fatty acids (ALA and DHA), and key micronutrients.

Dietary and supplementary amino acids have been shown to alter gut microbiota composition and function in relation to various disorders. For example, immunomodulatory effects of amino acids (L-arginine and L-tryptophan) on colonocytes have been reported, due to changes in microbiota metabolites.[32] Branched-chain amino acid supplementation has shown microbiota effects in multiple studies, eliciting changes in microbiota abundance in adults,[33] with subsequent modulation of sugar and lipid metabolites, slowed age-related changes in microbiota composition, and decreased lipopolysaccharide-binding protein.[34] Furthermore, changes have been observed in the gut microbiota associated with amino acid transporter, autism spectrum disorder-like behavioral, and electrophysiological phenotype changes in a young mouse model.[35] These studies suggest dietary amino acids can impact gut microbiota composition, with subsequent changes in metabolic, inflammatory, and behavioral outcomes.

Interesting microbiota driven metabolite and metabolic changes have also been observed with dietary lipid and micronutrient intake. For example, changes in gut microbiota composition in mice fed high saturated fat, high omega-6, or high omega-3, versus low fat diets have been reported.[36] Another study reported host resistance to diet-induced obesity to a high omega-6 diet in *Lactobacillus* colonized mice, via elevated 10-hydroxy-*cis*-12-octadecenoic acid levels.[37] Microbial and microbial metabolite alterations have also been seen with dietary intake of micronutrients such as choline,[38,39] non-heme iron,[30] and vitamin K.[40]

Cognitive Development

The successive stages of cognitive development include sensory function, language, learning and memory, and higher order cognition.[7] Neurological and cognitive development is thought to have nutrition-dependent stages. If nutrients are deficient, complete neurological development and cognitive function cannot be retrieved at later stages.[5,41]

Co-development of Gut Microbiota and Cognitive Function

In vivo research in mice has revealed numerous mechanisms by which the microbiota-gut-brain axis impacts infant and child cognitive development. One study reported significant differences in brain growth (neuronal differentiation, oligodendrocyte development, and myelination) in the cerebral cortex of germ-free mice given fecal transplants, from either low growth or high growth preterm infants.[10] Gross morphology brain changes (larger amygdala and hippocampus, immature microglia, and hypermyelination of the pre-frontal cortex) have been reported in germ-free and antibiotic treated mice, suggesting that microbiota disruption or depletion can impact both brain structure and immune function.[42] Early life stress in rat pups has shown alterations in the gut microbiome, leading to enhanced fear response which was reversed by probiotic supplementation.[43] Alterations in BDNF and serotonin neurotransmitters have been observed in young mice with altered or depleted gut microbiomes, suggesting neurochemistry can be affected by perturbations in gut microbiota.[44-49] One study points to programming of the hypothalamic-pituitary-adrenal system via microbial colonization in germ-free 9-week old mice.[50] These studies report a wide range of cognitive and neurological effects including structural, neurochemical, and behavioral alterations, demonstrating the far reaching potential of the microbiota-gut-brain axis on cognitive development and function.

A handful of studies have been conducted primarily in healthy children investigating the link between the microbiota-gut-brain axis and cognitive development. One cross-sectional study investigating neurodevelopment in 3-year-old rural Chinese children reported no differences in microbiota alpha-diversity between higher and lower neurodevelopment scores.[15] However, higher neurodevelopment scores were observed with specific microbial co-abundance groupings (higher abundance of genera *Faecalibacterium, Sutterella*, and *Clostridium cluster XIVa*).[15] In preterm neonates, depletion in the abundance of *Bacteroidota* and *Lachnospiraceae* correlated

with lower head circumference growth, a validated marker of neurodevelopment.[14] A study in 1-year-old infants investigated whether microbial composition was associated with cognitive scores and regional brain volume at 1 and 2 years of age. The investigators reported no difference in regional brain volumes. However, higher alpha-diversity was associated with overall lower cognitive scores, and particularly in the visual and expressive language scales.[17] An ancillary study of the Vitamin D Antenatal Asthma Reduction Trial cohort found a significant association between less favorable gut microbiome composition (dominated by *Clostridiales* or *Bacteroides*) and lower communication, personal and social, and fine motor skills at 3 years of age.[16] There are two studies in infants which report differences in temperament and behavioral outcomes associated with gut microbiome composition.[51,52] The majority of infant studies to date point to the importance of specific composition of gut microbiota, rather than microbiota diversity, for healthy neurodevelopment. Mediators of the relationship such as microbial transcriptomics, proteomics, and metabolomics, have not been explored in human infants.

Co-development of Gut Microbiota and Growth

Physical growth heavily relies on adequate nutrient status, and deficiencies often lead to stunting, wasting and other suboptimal growth anthropometrics.[1,2,6] There is an additional link between gut microbiome composition and child growth. While the mechanisms are not fully understood, with gut dysbiosis there is evidence of impaired nutrient absorption, intestinal integrity, intestinal immune responses, and greater pathogen carriage.

Evidence from germ free mice suggests the microbiome may modulate insulin-like growth factor 1 (IGF-1) secretion. Germ free mice have shown depleted IGF-1 levels, impacting growth hormone release and stunting linear growth.[11] Supplementation with certain microbiota

strains have rescued IGF-1 secretion and growth in chronically undernourished infant mice.[53] Supplementation with short-chain fatty acids after antibiotic treatment in mice restores IGF-1, bone mass and growth deficits induced by antibiotic use.[54] Gut dysbiosis is further associated with impaired growth development, as gut dysbiosis is closely linked with infections and pathogens, thus contributing to and exacerbating severe acute malnutrition.[11,28,55]

Studies in human infants and children also suggest a link between malnutrition, gut microbiome composition, and anthropometric growth. A longitudinal study investigating Bangladeshi children from birth to 2 years of age, reported immature gut microbiota composition in the presence of severe acute malnutrition and suboptimal z-scores.[56] Both immature microbiota composition and growth improved in Bangladeshi children, after a nutrition intervention (intervention n = 31; standard diet controls n = 31).[56] A diet-based randomized controlled trial in infants reported infants receiving a protein hydrolysate formula vs cow's milk showed faster gut microbiota maturation and increased alpha-diversity, while abundance of *Clostridia* at 3-4 months was negatively correlated with weight-gain velocity.[55] Specific human milk oligosaccharides were shown to be correlated with Gambian infant growth, morbidity, and microbiota composition, with *Bacteroides* being increased in infants with abnormal calprotectin levels.[57] While correlations have been reported, mechanistic and causal associations between diet, gut microbiome composition, and growth outcomes are lacking in human infants.

Gastrointestinal Inflammation and Early Child Development

As previously mentioned, the immune system may be an integral mechanism in gut microbiome and infant co-development. It is hypothesized that unfavorable nutritional and environmental factors can result in greater pathogen infiltration due to gut dysbiosis and compromised tight junctions.[12] This unfavorable gut environment triggers, at first, a localized

inflammatory response in the gastrointestinal (GI) system.[12] With no intervention, particularly in infants whose gut microbiomes and immune systems are still developing, the inflammatory response can become chronic and promote systemic inflammation.[11,12] It is hypothesized that this systemic inflammatory response could impair ECD and growth.[11,12] Moreover, a cyclic feedback response can occur between unfavorable gut microbiome composition and GI inflammation, with each exacerbating the other.[1111] GI inflammation is also a common co-morbidity of malnutrition in LMICs, often presenting as environmental enteric dysfunction.[11,58]

Recent studies have explored the association between GI inflammation, microbiome composition, and human infant development. A 2021 longitudinal study in healthy Polish infants concluded markers of intestinal paracellular permeability and fecal calprotectin correlated with specific gut microbiota strains, and fluctuated significantly throughout the first 2 years of life.[59] A 2021 study in the U.S. found specific bacteria associated with enteric inflammation, antibiotic resistance genes, and signatures of dysbiosis, with bacterial strains thought to be essential to infant development being significantly depleted.[9] One study showed that a low-growth phenotype human infant microbiome transplant was associated with neuroinflammation and decreased neurodevelopment and growth factors in gnotobiotic mice.[10] While these investigations into human infants and localized inflammation have explored associations, impacts of long-term causative disruption of the gut microbiome on ECD and growth has not been proven.

Malnutrition in Infants from Low- and Middle-Income Countries

Low- and middle-income countries (LMIC) are particularly affected by poor development outcomes, given their lack of nutrition resources and awareness.[60,61] According to the United Nations, many Guatemalan infants do not meet adequate dietary diversity or complementary feeding practices.[62-65] Guatemalan neonates have high prevalence of low birth

weight (15%) and children ages 18-23 months have high prevalence of chronic malnutrition (58%), iron-deficiency anemia (33%), stunting (43%) and severe stunting (23%).[60,66,67] Indigenous Mayan children in Guatemala are most severely affected by malnutrition, with twice the rate of stunting in indigenous children compared to non-indigenous children.[68,69]Additionally, environmental factors, breast-feeding and complementary feeding practices, maternal malnutrition, and limited nutrition resources lead to higher risk of malnutrition, gut dysbiosis, and suboptimal cognitive and growth development in LMIC infants. These populations also have a high risk for early antibiotic exposure, given their suboptimal living conditions and food/healthcare access.

References

1. Onyango AW, Borghi E, de Onis M, Casanovas Mdel C, Garza C. Complementary feeding and attained linear growth among 6-23-month-old children. *Public Health Nutr*. 2014;17(9):1975-1983. doi:10.1017/s1368980013002401
2. Daelmans B, Mangasaryan N, Martines J, Saadeh R, Casanovas C, Arabi M. Strengthening actions to improve feeding of infants and young children 6 to 23 months of age: summary of a recent World Health Organization/UNICEF technical meeting, Geneva, 6-9 October 2008. *Food Nutr Bull*. 2009;30(2 suppl):S236-S238. doi:10.1177/15648265090302s208
3. Bhutta ZA, Das JK, Rizvi A, et al. Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost? *Lancet*. 2013;382(9890):452-477. doi:10.1016/s0140-6736(13)60996-4
4. Schwarzenberg SJ, Georgieff MK. Advocacy for improving nutrition in the first 1000 days to support childhood development and adult health. *Pediatrics*. 2018;141(2):e20173716. doi:10.1542/peds.2017-3716
5. Cusick SE, Georgieff MK. The role of nutrition in brain development: the golden opportunity of the "first 1000 days". *J Pediatr*. 2016;175:16-21. doi:10.1016/j.jpeds.2016.05.013
6. Hoffman D, Arts M, Bégin F. The "first 1,000 days+" as key contributor to the double burden of malnutrition. *Ann Nutr Metab*. 2019;75(2):99-102. doi:10.1159/000503665
7. Cowan CSM, Dinan TG, Cryan JF. Annual Research Review: Critical windows - the microbiota-gut-brain axis in neurocognitive development. *J Child Psychol Psychiatry*. 2020;61(3):353-371. doi:10.1111/jcpp.13156

8. Stewart CJ, Ajami NJ, O'Brien JL, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature*. 2018;562(7728):583-588. doi:10.1038/s41586-018-0617-x

9. Casaburi G, Duar RM, Brown H, et al. Metagenomic insights of the infant microbiome community structure and function across multiple sites in the United States. *Sci Rep*. 2021;11(1):1472. doi:10.1038/s41598-020-80583-9

10. Lu J, Lu L, Yu Y, Cluette-Brown J, Martin CR, Claud EC. Effects of intestinal microbiota on brain development in humanized gnotobiotic mice. *Sci Rep*. 2018;8(1):5443. doi:10.1038/s41598-018-23692-w

11. Robertson RC, Manges AR, Finlay BB, Prendergast AJ. The human microbiome and child growth - first 1000 days and beyond. *Trends Microbiol*. 2019;27(2):131-147. doi:10.1016/j.tim.2018.09.008

12. Oriá RB, Murray-Kolb LE, Scharf RJ, et al. Early-life enteric infections: relation between chronic systemic inflammation and poor cognition in children. *Nutr Rev*. 2016;74(6):374-386. doi:10.1093/nutrit/nuw008

13. Moore RE, Townsend SD. Temporal development of the infant gut microbiome. *Open Biol*. 2019;9(9):190128. doi:10.1098/rsob.190128

14. Oliphant K, Ali M, D'Souza M, et al. Bacteroidota and Lachnospiraceae integration into the gut microbiome at key time points in early life are linked to infant neurodevelopment. *Gut Microbes*. 2021;13(1):1997560. doi:10.1080/19490976.2021.1997560

15. Rothenberg SE, Chen Q, Shen J, et al. Neurodevelopment correlates with gut microbiota in a cross-sectional analysis of children at 3 years of age in rural China. *Sci Rep*. 2021;11(1):7384. doi:10.1038/s41598-021-86761-7

16. Sordillo JE, Korrick S, Laranjo N, et al. Association of the infant gut microbiome with early childhood neurodevelopmental outcomes: an ancillary study to the VDAART randomized clinical trial. *JAMA Netw Open*. 2019;2(3):e190905. doi:10.1001/jamanetworkopen.2019.0905

17. Carlson AL, Xia K, Azcarate-Peril MA, et al. Infant gut microbiome associated with cognitive development. *Biol Psychiatry*. 2018;83(2):148-159. doi:10.1016/j.biopsych.2017.06.021

18. Bokulich NA, Chung J, Battaglia T, et al. Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci Transl Med*. 2016;8(343):343ra382. doi:10.1126/scitranslmed.aad7121

19. Carabotti M, Scirocco A, Maselli MA, Severi C. The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Ann Gastroenterol*. 2015;28(2):203-209.

20. Al Nabhani Z, Dulauroy S, Marques R, et al. A weaning reaction to microbiota is required for resistance to immunopathologies in the adult. *Immunity*. 2019;50(5):1276-1288.e1275. doi:10.1016/j.immuni.2019.02.014

21. Johnson CC, Ownby DR. The infant gut bacterial microbiota and risk of pediatric asthma and allergic diseases. *Transl Res*. 2017;179:60-70. doi:10.1016/j.trsl.2016.06.010

22. Tripathi A, Debelius J, Brenner DA, et al. The gut-liver axis and the intersection with the microbiome. *Nat Rev Gastroenterol Hepatol*. 2018;15(7):397-411. doi:10.1038/s41575-018-0011-z

23. Anand S, Mande SS. Diet, microbiota and gut-lung connection. *Front Microbiol*. 2018;9:2147. doi:10.3389/fmicb.2018.02147

24. Bingula R, Filaire M, Radosevic-Robin N, et al. Desired turbulence? Gut-lung axis, immunity, and lung cancer. *J Oncol*. 2017;2017:5035371. doi:10.1155/2017/5035371

25. Albouery M, Buteau B, Grégoire S, et al. Age-related changes in the gut microbiota modify brain lipid composition. *Front Cell Infect Microbiol*. 2019;9:444. doi:10.3389/fcimb.2019.00444

26. Velagapudi VR, Hezaveh R, Reigstad CS, et al. The gut microbiota modulates host energy and lipid metabolism in mice. *J Lipid Res*. 2010;51(5):1101-1112. doi:10.1194/jlr.M002774

27. Silva YP, Bernardi A, Frozza RL. The role of short-chain fatty acids from gut microbiota in gut-brain communication. *Front Endocrinol (Lausanne)*. 2020;11:25. doi:10.3389/fendo.2020.00025

28. Bierut T, Duckworth L, Grabowsky M, et al. The effect of bovine colostrum/egg supplementation compared with corn/soy flour in young Malawian children: a randomized, controlled clinical trial. *Am J Clin Nutr*. 2021;113(2):420-427. doi:10.1093/ajcn/nqaa325

29. Tang M, Frank DN, Hendricks AE, et al. Iron in micronutrient powder promotes an unfavorable gut microbiota in Kenyan infants. *Nutrients*. 2017;9(7):776. doi:10.3390/nu9070776

30. Paganini D, Zimmermann MB. The effects of iron fortification and supplementation on the gut microbiome and diarrhea in infants and children: a review. *Am J Clin Nutr*. 2017;106(Suppl 6):1688s-1693s. doi:10.3945/ajcn.117.156067

31. McMillen S, Thomas S, Liang E, Nonnecke EB, Slupsky C, Lönnerdal B. Gut microbiome alterations following postnatal iron supplementation depend on iron form and persist into adulthood. *Nutrients*. 2022;14(3):412. doi:10.3390/nu14030412

32. Wu L, Tang Z, Chen H, et al. Mutual interaction between gut microbiota and protein/amino acid metabolism for host mucosal immunity and health. *Anim Nutr*. 2021;7(1):11-16. doi:10.1016/j.aninu.2020.11.003

33. Genton L, Pruijm M, Teta D, et al. Gut barrier and microbiota changes with glycine and branched-chain amino acid supplementation in chronic haemodialysis patients. *J Cachexia Sarcopenia Muscle*. 2021;12(6):1527-1539. doi:10.1002/jcsm.12781

34. Yang Z, Huang S, Zou D, et al. Metabolic shifts and structural changes in the gut microbiota upon branched-chain amino acid supplementation in middle-aged mice. *Amino Acids*. 2016;48(12):2731-2745. doi:10.1007/s00726-016-2308-y

35. Yu Y, Zhang B, Ji P, et al. Changes to gut amino acid transporters and microbiome associated with increased E/I ratio in Chd8(+/-) mouse model of ASD-like behavior. *Nat Commun*. 2022;13(1):1151. doi:10.1038/s41467-022-28746-2

36. Schoeler M, Caesar R. Dietary lipids, gut microbiota and lipid metabolism. *Rev Endocr Metab Disord*. 2019;20(4):461-472. doi:10.1007/s11154-019-09512-0

37. Miyamoto J, Igarashi M, Watanabe K, et al. Gut microbiota confers host resistance to obesity by metabolizing dietary polyunsaturated fatty acids. *Nat Commun*. 2019;10(1):4007. doi:10.1038/s41467-019-11978-0

38. Cho CE, Aardema NDJ, Bunnell ML, et al. Effect of choline forms and gut microbiota composition on trimethylamine-N-oxide response in healthy men. *Nutrients*. 2020;12(8):2220. doi:10.3390/nu12082220

39. Hartiala J, Bennett BJ, Tang WH, et al. Comparative genome-wide association studies in mice and humans for trimethylamine N-oxide, a proatherogenic metabolite of choline and L-carnitine. *Arterioscler Thromb Vasc Biol*. 2014;34(6):1307-1313. doi:10.1161/atvbaha.114.303252

40. Ellis JL, Karl JP, Oliverio AM, et al. Dietary vitamin K is remodeled by gut microbiota and influences community composition. *Gut Microbes*. 2021;13(1):1-16. doi:10.1080/19490976.2021.1887721

41. Fox SE, Levitt P, Nelson CA, 3rd. How the timing and quality of early experiences influence the development of brain architecture. *Child Dev*. 2010;81(1):28-40. doi:10.1111/j.1467-8624.2009.01380.x

42. Luczynski P, McVey Neufeld KA, Oriach CS, Clarke G, Dinan TG, Cryan JF. Growing up in a bubble: using germ-free animals to assess the influence of the gut microbiota on brain and behavior. *Int J Neuropsychopharmacol*. 2016;19(8):pyw020. doi:10.1093/ijnp/pyw020

43. Callaghan BL, Cowan CS, Richardson R. Treating generational stress: effect of paternal stress on development of memory and extinction in offspring is reversed by probiotic treatment. *Psychol Sci*. 2016;27(9):1171-1180. doi:10.1177/0956797616653103

44. Bercik P, Denou E, Collins J, et al. The intestinal microbiota affect central levels of brain-derived neurotropic factor and behavior in mice. *Gastroenterology*. 2011;141(2):599-609, 609.e591-593. doi:10.1053/j.gastro.2011.04.052

45. Clarke G, Grenham S, Scully P, et al. The microbiome-gut-brain axis during early life regulates the hippocampal serotonergic system in a sex-dependent manner. *Mol Psychiatry*. 2013;18(6):666-673. doi:10.1038/mp.2012.77

46. Desbonnet L, Clarke G, Traplin A, et al. Gut microbiota depletion from early adolescence in mice: Implications for brain and behaviour. *Brain Behav Immun*. 2015;48:165-173. doi:10.1016/j.bbi.2015.04.004

47. Diaz Heijtz R, Wang S, Anuar F, et al. Normal gut microbiota modulates brain development and behavior. *Proc Natl Acad Sci U S A*. 2011;108(7):3047-3052. doi:10.1073/pnas.1010529108

48. Gareau MG, Wine E, Rodrigues DM, et al. Bacterial infection causes stress-induced memory dysfunction in mice. *Gut*. 2011;60(3):307-317. doi:10.1136/gut.2009.202515

49. Neufeld KM, Kang N, Bienenstock J, Foster JA. Reduced anxiety-like behavior and central neurochemical change in germ-free mice. *Neurogastroenterol Motil*. 2011;23(3):255-264. doi:10.1111/j.1365-2982.2010.01620.x

50. Sudo N, Chida Y, Aiba Y, et al. Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice. *J Physiol*. 2004;558(Pt 1):263-275. doi:10.1113/jphysiol.2004.063388

51. Aatsinki AK, Lahti L, Uusitupa HM, et al. Gut microbiota composition is associated with temperament traits in infants. *Brain Behav Immun*. 2019;80:849-858. doi:10.1016/j.bbi.2019.05.035

52. Loughman A, Ponsonby AL, O'Hely M, et al. Gut microbiota composition during infancy and subsequent behavioural outcomes. *EBioMedicine*. 2020;52:102640. doi:10.1016/j.ebiom.2020.102640

53. Schwarzer M, Makki K, Storelli G, et al. Lactobacillus plantarum strain maintains growth of infant mice during chronic undernutrition. *Science*. 2016;351(6275):854-857. doi:10.1126/science.aad8588

54. Yan J, Herzog JW, Tsang K, et al. Gut microbiota induce IGF-1 and promote bone formation and growth. *Proc Natl Acad Sci U S A*. 2016;113(47):e7554-e7563. doi:10.1073/pnas.1607235113

55. Mennella JA, Li Y, Bittinger K, et al. The macronutrient composition of infant formula produces differences in gut microbiota maturation that associate with weight gain velocity and weight status. *Nutrients*. 2022;14(6):1241. doi:10.3390/nu14061241

56. Subramanian S, Huq S, Yatsunenko T, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature*. 2014;510(7505):417-421. doi:10.1038/nature13421

57. Davis JC, Lewis ZT, Krishnan S, et al. Growth and morbidity of Gambian infants are influenced by maternal milk oligosaccharides and infant gut microbiota. *Sci Rep*. 2017;7:40466. doi:10.1038/srep40466

58. Crane RJ, Jones KD, Berkley JA. Environmental enteric dysfunction: an overview. *Food Nutr Bull*. 2015;36(1 suppl):S76-S87. doi:10.1177/15648265150361s113

59. Kaczmarczyk M, Löber U, Adamek K, et al. The gut microbiota is associated with the small intestinal paracellular permeability and the development of the immune system in healthy children during the first two years of life. *J Transl Med*. 2021;19(1):177. doi:10.1186/s12967-021-02839-w

60. Black RE, Victora CG, Walker SP, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet*. 2013;382(9890):427-451. doi:10.1016/s0140-6736(13)60937-x

61. Lu C, Black MM, Richter LM. Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional,

and country level. *Lancet Glob Health*. 2016;4(12):e916-e922. doi:10.1016/s2214-109x(16)30266-2

62. Food and Nutritional Assistance III Project. *Development of Evidence-Based Dietary Recommendations for Children, Pregnant Women, and Lactating Women Living in the Western Highlands in Guatemala*. Food and Nutritional Assistance III Project; 2013.

63. Brown K, Henretty N, Chary A, et al. Mixed-methods study identifies key strategies for improving infant and young child feeding practices in a highly stunted rural indigenous population in Guatemala. *Matern Child Nutr*. 2016;12(2):262-277. doi:10.1111/mcn.12141

64. Webb MF, Chary AN, De Vries TT, et al. Exploring mechanisms of food insecurity in indigenous agricultural communities in Guatemala: a mixed methods study. *BMC Nutr*. 2016;2(1):55. doi:10.1186/s40795-016-0091-5

65. Food and Nutrition Technical Assistance III Project. Validation of food-based recommendations developed using Optifood for groups at nutritional risk in the Western Highlands of Guatemala. Accessed August 15, 2022, https://www.fantaproject.org/countries/guatemala/optifood-report-2014

66. Ministerio de Salud Pública y Asistencia Social MSPAS, Guatemala Instituto Nacional de Estadística INE/Guatemala, Secretaría de Planificación y Programación del la Presidencia Segeplán/Guatemala, ICF International. Encuesta nacional de salud materno infantil 2014-2015: informe final. Accessed August 12, 2022, https://www.dhsprogram.com/publications/publication-fr318-dhs-final-reports.cfm

67. World Health Organization, United Nations Children's Fund. *WHO Child Growth Standards and the Identification of Severe Acute Malnutrition in Infants and Children: A*

Joint Statement by the World Health Organization and the United Nations Children's Fund. World Health Organization; 2009.

68. World Bank. International Bank for Reconstruction and Development and International Finance Corporation Country Partnership Strategy for the Republic of Guatemala. Report No. 44772-GT. Accessed December 6, 2022, http://web.worldbank.org/archive/website01270/WEB/IMAGES/GT_CPS_0.PDF

69. World Bank. Improving Health and Nutrition of Mothers and Young Children in Guatemala Accessed December 6, 2022, https://documents.worldbank.org/en/publication/documents-reports/documentdetail/699671468032739491/Improving-health-and-nutrition-of-mothers-and-young-children-in-Guatemala

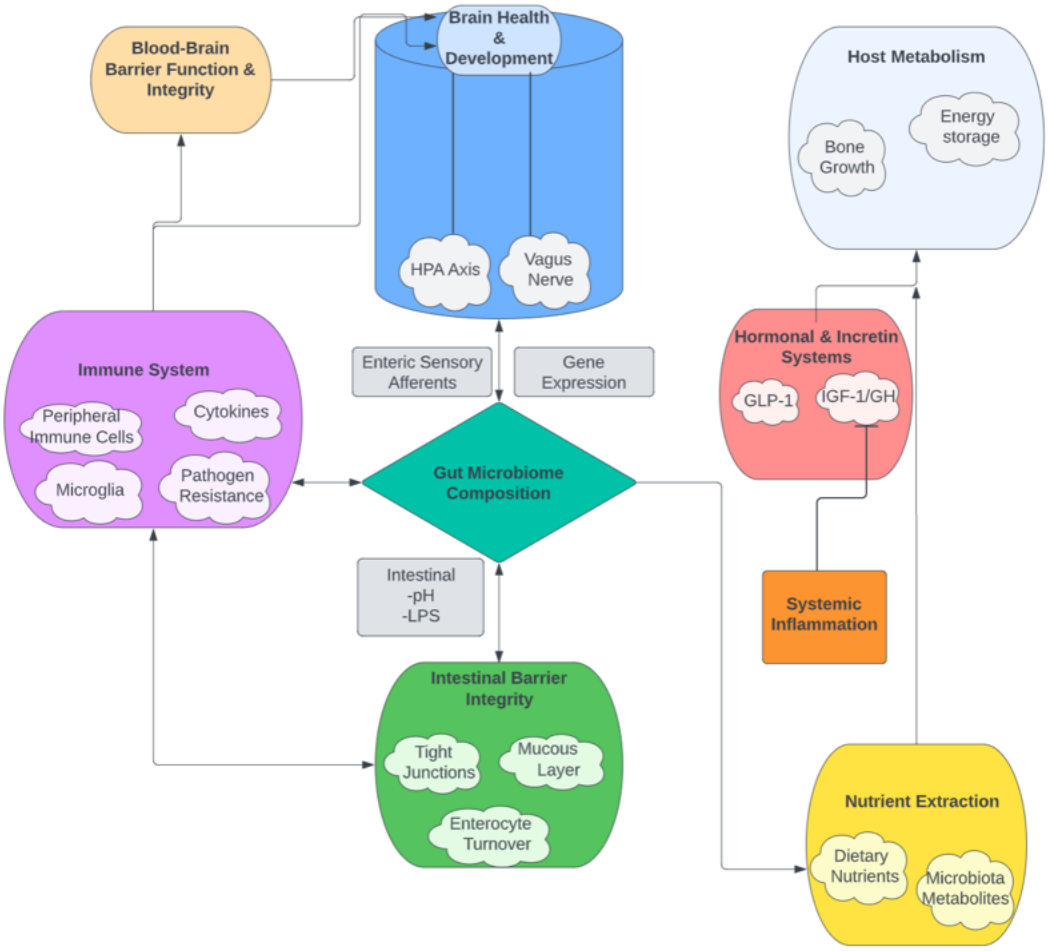

Figure 2.1. Key components of infant and gut microbiome co-development in relation to other systems in the body.

This original diagram depicts key components of infant and gut microbiome co-development in relation to other systems in the body. The key systems which develop during the first 1000 days of life include nutrient extraction, immune signaling, gut and blood-brain barrier function, and neuronal and hormonal signaling. Hormonal, incretin, and host metabolism systems are key to growth, but can be blunted by systemic inflammation, which may occur with localized GI inflammation.

CHAPTER 3:
The Saqmolo' Project: How Early Diet Shapes the Gut Microbiome in 11-16-Month-Old Guatemalan Infants

Abstract

Objective: This study examined the association of egg intake versus standard care controls, diet diversity, and milk type with gut microbiome composition and localized gastrointestinal inflammation in 11–16-month-old infants from low- and middle-income country (LMIC) population experiencing a high degree of malnutrition.

Methods: The Saqmolo' project enrolled 1,200 rural Guatemalan infants aged 6–9 months, with 200 in a microbiome sub-study. The infants were randomized to receive either an egg daily or micronutrient powder for six months. Endline fecal samples were assessed using 16S rRNA sequencing, and dietary data were collected via caregiver interviews. Beta diversity differences were assessed using principal coordinates analysis (PCoA), stratified by treatment and milk type interactions. Microbiome analyses included QIIME2, MaAsLin2, LEfSe, and co-abundance group clustering. Associations between fecal calprotectin and dietary factors were evaluated using linear regression.

Results: Among 200 infants (mean age 12.7 months, 49% female), microbiome composition differences were larger between egg and control groups in non-breastfed infants versus breastfed infants on principal component plots. LEfSe analysis identified four amplicon sequence variants (ASVs) with lower abundance and 23 ASVs with higher abundance in non-breastfed compared to breastfed infants. Neither egg intake nor minimum diet diversity (MDD) was associated with specific ASVs or co-abundance groups, though co-abundance groups 3 and 16 had the highest relative abundance. Fecal calprotectin was not significantly associated with treatment (β = -57.85, 95% CI: -119.08 to 98.52, P = 0.281) or MDD (β = -10.28, 95% CI: -187.01 to 71.31, P = 0.337). However, milk type was negatively associated with fecal calprotectin levels (β = -174.70,

95% CI: -314.16 to -35.23, P = 0.014), although the overall model was not statistically significant (P = 0.228).

Conclusions: Differences in gut microbiome composition were more pronounced in non-breastfed infants receiving eggs compared to controls. Additionally, non-breastfed infants exhibited a microbiome composition indicative of a less favorable bacterial profile relative to breastfed infants.

Introduction

Early-life gut microbiome development plays a crucial role in shaping long-term health outcomes, including immune function, metabolism, and nutritional status.[1] The gut microbiome is established through maternal transmission, mode of delivery (vaginal birth vs. C-section), early dietary exposures, and environmental factors. A well-balanced gut microbiota is essential for efficient digestion, nutrient absorption, and protection against pathogenic infections. Infants in low- and middle-income countries (LMICs) face distinct challenges in gut microbiome development due to factors such as food insecurity, poor water sanitation, and high rates of malnutrition. In rural Guatemala, where stunting and undernutrition are prevalent, diet quality and microbial exposure may significantly impact microbial colonization patterns.[2,3]

Breastfeeding is widely recognized as a primary determinant of a healthy gut microbiome, enriching beneficial bacteria such as Bifidobacterium and supporting immune function through human milk oligosaccharides (HMOs). Breastfeeding may confer greater gut microbiome resilience to environmental challenges compared to infant formula and animal milk consumption. Additionally, the transition from exclusive milk feeding to complementary feeding represents a critical period for microbiome maturation, where inadequate or poor-quality complementary foods (nutrient poor, contaminated, etc.) may disrupt microbial balance and impair nutrient absorption.[1,4,5] Nutritional interventions during complementary feeding, such as protein-rich foods, may help support gut microbiome health, yet few studies have systematically assessed the interaction between diet quality, milk type, and microbial development in malnourished populations.

Research on the infant gut microbiome in malnourished LMIC populations remains limited. Many existing studies are underpowered, often including fewer than 20 subjects, and

primarily focus on either well-nourished infants or those with severe infections.[6] Additionally, most studies have not accounted for the crucial transition from exclusive milk feeding to complementary foods.

It has been hypothesized that poor nutrition and malnutrition can alter gut microbiome composition and contribute to undernutrition by impairing nutrient metabolism. Prior research suggests that a disrupted gut microbiome in malnourished infants may lead to deficiencies in short-chain fatty acids (SCFAs),[7] which play a key role in gut barrier integrity and immune modulation. However, little is known about how specific complementary foods interact with the primary type of milk consumed (breastmilk, infant formula, or animal milk) and how this interaction influences microbiome composition and gut inflammation. Localized gastrointestinal (GI) often occurs in LMIC infant populations as well as in the presence of gut dysbiosis.

One potential dietary intervention to address malnutrition and microbiome development is egg consumption. Eggs provide high-quality protein, choline, and essential micronutrients that are critical for infant growth and development. Prior studies have investigated the impact of egg consumption on growth in LMIC settings, but its influence on the gut microbiome has not been thoroughly examined. Understanding how egg intake, diet diversity, and milk type collectively shape the gut microbiome and GI inflammation could provide valuable insights into improving infant nutrition strategies in malnourished populations.

Objectives and Hypotheses

The aim of this study was to determine the association of egg intake versus standard care controls (multiple micronutrients powder), diet diversity, and most frequently consumed milk type with 1) endline fecal microbiome composition; and 2) localized gastrointestinal (GI) inflammation (measured with fecal calprotectin) in 11–16-month-old infants from a LMIC population with a high degree of malnutrition. We hypothesized that egg intake, higher diet

diversity scores, and higher breastmilk consumption would be associated with a more mature fecal microbiome composition and lower levels of GI inflammation. To test these hypotheses, we analyzed the relative abundance of bacterial amplicon sequence variants (ASVs) via 16S rRNA gene sequencing using fecal samples from 200 indigenous Mayan infants (aged 11–16 months) who participated in The Saqmolo' Project. This study provides a unique opportunity to assess how dietary interventions shape gut microbiome development and inflammation in a high-risk malnourished population.

Materials and Methods

Study Population

The Saqmolo' project enrolled 1,200 rural Guatemalan infants between 6 to 9-months-old, 200 of which were enrolled in the microbiome sub-study. Infants from the primary study were randomized into one of two groups: a one-egg-a-day intervention compared to multiple micronutrient powder controls for 6 months. Both intervention and control groups received growth monitoring, medical care, deworming, medication, multiple micronutrients powder, and individualized complementary and responsive feeding education for caregivers from the Maya Health Alliance. Multiple micronutrient powder composition included ferrous fumarate (12.5 mg), zinc gluconate (5 mg), retinol acetate (300 mg), folic acid (160 mg), and ascorbic acid (30 mg) provided daily beginning at age 6 months. Fecal samples were collected from 200 infants for gut microbiome quantification at the end of the 6-month trial. The Saqmolo' study protocol has been published elsewhere.[8]

Dietary and Demographics Questionnaires

All data were collected by trained research study staff. Egg consumption was monitored within interviews and questionnaires during monthly in-home visits and via biweekly telephone calls to caregivers for the first 2 months. Diet diversity was measured using the World Health

Organization (WHO) Infant and Young Child Feeding indicators questionnaire,[9] which captures minimum diet diversity (MDD), minimum meal frequency (MMF), minimum acceptable diet (MAD), and breastfeeding frequency and duration. The most frequently consumed milk type was reported based on dietary questionnaires designed for the Saqmolo' Sub-study and collapsed into two categories (breastmilk versus non-breastmilk). All infants categorized as primarily consuming breastmilk were included in the breastmilk category, while all infants reported to primarily consume either fluid animal milk, powdered animal milk, or infant formula were included in the non-breastmilk category. A family care indicators questionnaire assessed home play activities, while drinking water access and source were evaluated using the WHO water access questionnaire. Poverty levels were determined using the Quick Poverty Score.

Fecal Procedures and Gut Microbiome Quantification

Two stool samples per infant (intended for fecal microbiome quantification) from each infant's diaper were collected in OmniGENE fecal collection kits (GUT-OM-200). The samples were transported same day in lunch boxes at room temperature, to the Maya Health Alliance (MHA) Wuqu'Kawok offices in Tecpan and stored at 4 ℃. Samples were transported in batches each week ~60 kilometers to the Santiago Biomedical Laboratory in Santiago Sacatepéquez on ice and stored at -20 ℃. Samples were transported on dry ice, with biomedical grade packing, using an international courier to support the cold chain. Gut microbiome bacterial composition data were generated by the Rutgers Center for Microbiome Analysis & Microbiome Core at New Jersey Institute for Food, Nutrition and Health. Fecal microbiome composition was assessed using 16S rRNA sequencing of the V4 region. Fecal calprotectin levels were analyzed at Colorado State University, College of Health, and Human Sciences Intestinal Health Laboratory. Fecal calprotectin levels were assessed with ELISA assay kits (Eagle Biosciences, Amherst, NH). Stool samples were thawed and 15mg of homogenized sample was added to the extraction

buffer (IDK Extract cat #K6999.US, Immunodiagnostiks, Bernsheim, Germany), vortexed and allowed to rest at room temperature for 10 minutes until all solid matter had sedimented. Supernatants were evaluated for calprotectin using the IDK Calprotectin ELISA kit following manufacturer's instructions (Cat # KR6927, Immunodiagnostiks, Bernsheim, Germany). Stool sample extract was diluted to a final dilution of 1:2500 and 100 µl was used in the assay. All samples were tested in duplicate. Wells were read at an absorbance at 450nm and plotted against a standard curve to determine final concentrations using a four-parameter algorithm with a linear ordinate for the optical density and a logarithmic abscissa for the concentration.

Quantitative Insights into Microbial Ecology Version 2 (QIIME2)

All statistical analyses were performed in QIIME2,[10] RStudio (*v*4.2.0)[11] and Tufts University High Performance Computing Cluster. Amplicon sequence variants (ASV) were analyzed using QIIME2 to determine fecal microbiome composition. Sequence quality control and denoising were performed using DADA2. Data were cut at length 215 during denoising. Sampling depth was set to 12,000 after inspecting rarefaction curves (see **Supplementary Figure S3.1** in **Appendix A**). Gut microbiota profiles were represented by a relative abundance table of amplicon sequence variants (ASVs), which represent unique pieces of 16S rRNA V4 sequences. Gut microbiota profiles were evaluated at the species and genus levels by binning ASVs into these taxonomic units. Sequence variants were aligned with the mafft method using the q2-aligment plugin. A phylogenetic tree was created with the FastTree method using the q2-phylogeny plugin. Alpha-diversity was measured using observed ASVs, Faith's Phylogenetic Diversity, Shannon Index, and Pielou's Evenness. Beta-diversity measures for fecal samples were calculated using Bray-Curtis distance and weighted and unweighted UniFrac distances. Principal coordinates analysis (PCoA) vectors were constructed from beta-diversity measures to

visually assess differences between high and low groups. A sensitivity analysis was performed to investigate the impact of rarefaction level and inclusion or exclusion of unidentified ASVs (see **Supplementary Figure S3.2** and **Figure S3.3**). All statistical tests were performed for the treatment group (egg versus standard care controls), MDD (high and low groups were created based on the median for statistical comparisons), and most frequently consumed milk type. QIIME2 statistical tests were reported with q-values (FDR adjusted p-values) ≤ 0.05 as significant.

Individual Point Mean + SEM PCoA Plots

Traditional individual point PCoA plots were created in R Studio using ggplot colored by treatment, most frequently consumed milk type, and subset groups of treatment and milk type (egg + breastmilk, egg + non-breastmilk, no egg + breastmilk, no egg + non-breastmilk). Mean plus SEM PCoA plots were also created for these variables. We plotted the group means with standard error bars, representing the central tendency and variability of each group on the PCoA plots. This gives a clearer summary of how each treatment or milk type group tends to cluster on average, while the error bars show the range within which individual samples vary.

ASV Associations: Multivariate Associations with Linear Models (MaASLin2)

MaASLin2[12] was used to individually test the association between relative abundance of the top 10% most prevalent ASV present in fecal samples, with treatment group, MDD, and milk type. MaASLin2 is designed specifically for microbiome data, offering advantages over traditional statistical methods by handling the high-dimensional, sparse, and compositional nature of microbiome features. It applies transformations (e.g., log) to improve normality, filters low abundance features to reduce noise, and allows for multiple covariates and random effects to control for confounding factors. Additionally, it applies multiple testing correction (e.g.,

Benjamini-Hochberg FDR) to reduce false positives, ensuring reliable results in large-scale microbiome analyses.

Abundance of the ASVs were normalized by dividing the raw count of each ASV in a sample by the total number of reads in that sample. Treatment and MDD models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, and diarrhea in the past 14 days. Treatment models were additionally adjusted for animal contact. MDD models were additionally adjusted for treatment group. Milk type models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, and diarrhea in the past 14 days, poverty score, medication, and treatment. Covariates were chosen based on a priori considerations. ASV counts were normalized and MaASLin2 coefficients were reported with q-values (FDR adjusted p-values) ≤ 0.25 as significant. After initial analysis, an ASV relative abundance table of the top 10 most abundant taxa was created. The table was stratified both by treatment group (egg versus standard care controls) and most frequently consumed milk type (breastmilk versus non-breastmilk) and relative abundance percentages were reported.

Differentially Abundant ASVs by Milk Type (DESeq2)

We conducted an analysis to identify differentially abundant ASVs between samples categorized by milk type (breastmilk vs. non-breastmilk) using DESeq2[13] in RStudio. Non-breastmilk was designated as the reference level. The DESeq2 framework was chosen for its robustness in analyzing count data, as it accounts for library size differences and variance heterogeneity. Data preprocessing included filtering low-abundance ASVs to minimize noise and normalizing the dataset through the DESeq2 pipeline. Statistical comparisons were made using the Wald test, and p-values were adjusted for multiple testing using the Benjamini-Hochberg procedure to control the false discovery rate. Results were further assessed for biological relevance by setting thresholds for log2 fold change and adjusted p-value significance. This

analysis allowed us to identify ASVs significantly enriched in either breastmilk or non-breastmilk samples, providing insights into the microbiota's role in diet-specific microbiome compositions. Only ASVs present in at least 5% of the samples were included.

Differentially Abundant ASVs by Milk Type (LEfSe)

We conducted an analysis to identify differentially abundant ASVs between samples categorized by milk type (breastmilk vs. non-breastmilk) using the LEfSe[14] algorithm. LEfSe was selected for its ability to detect bacterial features that not only differ significantly between groups but also have consistent biological relevance, ranking features based on their effect size. This analysis involved non-parametric Kruskal-Wallis tests and pairwise Wilcoxon tests and ended with linear discriminant analysis (LDA). A higher LDA score indicates greater discriminatory power between breastmilk and non-breastmilk groups. Results were visualized using LDA score plots and cladograms to illustrate ASV differences between groups.

Guild Based Co-Abundance Groupings

A binary, tree-based group identification method was used to identify and partition potential guilds without predefining the number of groups to create co-abundance groupings using the method by Wu et al.[15] In summary, correlation coefficients between ASVs were transformed into distance metrics (1 - correlation coefficient) and clustered using the Ward clustering algorithm.[16] Starting from the top of the clustering tree, a permutational multivariate analysis of variance (PERMANOVA) was conducted to determine if any two clades were significantly different.[17] For instance, if the tree splits into two main clades (A and B), each containing two sub-clades (C and D under A, and E and F under B), PERMANOVA will first test the larger clades. If a significant difference is found between them, the analysis moves to the sub-clades within each main clade. If no significant difference is found between the sub-clades

(C vs. D, E vs. F), the process concludes with two potential guilds (for example, A and B). In this analysis, individual co-abundance groups are referred to as CAGs.

After co-abundance group creation, three separate per subject relative abundance tables were created for the co-abundance groups. One table was stratified by treatment group, the next by milk type, and the third by sub-groups of egg and breastmilk, egg and non-breastmilk, control and breastmilk, and control and non-breastmilk. The co-abundance groups were used in downstream analysis to determine the association between each group with treatment group, MDD, and most frequently consumed milk type using MaAsLin2. Co-abundance models were adjusted for the same variables as the ASV MaAsLin2 models described in the previous section. The treatment and MDD cluster models were additionally analyzed stratified by milk type (breastmilk and non-breastmilk).

Statistical Analysis: Fecal Calprotectin and Diet Measures

Linear regression models were used to assess the association between fecal calprotectin, treatment group, MDD, and milk type most frequently consumed in separate models. The milk type model was adjusted for age, sex, antibiotic use, antibiotic use in the past 14 days, diarrhea in the past 14 days, water source, and treatment group. Treatment and MDD models were additionally adjusted for animal contact and milk type. All model assumptions were assessed.

Results

Participant Characteristics

The analysis included 200 infant participants. Infants were between age 11-16 months at endline, mean age was 12.7 months, and 49.0% were female. For MDD, 68.3% of the infants met the WHO MDD standard and 82.0% of the infants most frequently consumed breastmilk. Of the 34 primarily non-breastfed infants, 65% consumed infant formula, 29% consumed powdered animal milk, and 6% consumed fluid animal milk. Two fecal samples were successfully

collected for each of the 200 infants. Population characteristics by treatment group are presented in **Table 3.1**.

Diet Diversity Metrics

MDD and MAD were collinear in the overall sample. MDD and MMF are included in the determination of MAD. Since MDD and MAD were collinear in the overall sample, and most of the variability was in MDD, only MDD was used in this study for diet diversity analysis.

QIIME2: Alpha- and Beta-Diversity Comparisons

A total of 1,829 ASVs were identified. Alpha diversity was not significantly different between treatment groups or MDD using Shannon diversity indices (see **Supplementary Table S3.1** and **Supplementary Figures S3.5** and **S3.7**). Alpha diversity was significantly lower in the breastmilk group compared to the non-breastmilk group (**Figure 3.1** and **Table S3.1**) (median breastmilk group Shannon index = 2.96, p-value = 1.23E-09). Similarly, beta diversity was not significantly different between treatment groups or MDD for either unweighted or weighted Unifrac metrics (see **Supplementary Table S3.2** and **Supplementary Figures S3.9 to S3.12**). Beta diversity measured by weighted Unifrac was significantly different in the breastmilk group compared to the non-breastmilk group (q-value = 0.001).

Treatment and milk type PCoA Plots: individual point and mean + SEM

Minimal to no separation was seen in the individual point PCoA plots. Principle component 1 explained 37.26% of the variance. The treatment group individual point PCoA plot is displayed in **Supplementary Figure S3.14**, the milk type group individual point PCoA plot is displayed in **Supplementary Figure S3.15**, and the treatment milk type interaction individual point PCoA plot is displayed in **Supplementary Figure S3.16**. The mean + SEM PCoA plots more clearly demonstrated the separation in microbial composition between treatment and milk type groups (**Figure 3.2**). In the treatment milk type interaction mean + SEM PCoA plot (**Figure**

3.3), there is distinct separation between each group. Mainly, there was a large separation on the plot between egg and no egg in the non-breastmilk compared to the breastmilk group.

ASV MaAsLin2 Associations

No significant ASV associations were identified for MDD or treatment group. All significant ASVs identified were more relatively abundant in the non-breastmilk group compared to breastmilk. Significant ASV association model results for milk type groups are presented in **Figure 3.4**. The ASV highest in relative abundance stratified by treatment (35%) and control (32.6%) groups was ASV000J.Bifidobacterium. For milk type, the ASV highest in relative abundance was also ASV000J.Bifidobacterium (breastmilk: 40.7%, non-breastmilk: 11.4%). The 10 most relatively abundant ASVs stratified by treatment group and milk type are displayed in **Supplementary Tables S3.3** and **S3.4**, respectively.

LEfSe for Individual Milk Type ASVs

After FASTPAR and P-value filtering, 295 ASVs were used in the LEfSe analysis. ASVs greater than 5% prevalence in the overall sample were chosen, then ASVs with p-values < 0.001 were selected. Across all 295 ASVs, 27 were significantly differentially abundant among all participants comparing the breastmilk to the non-breastmilk group. For treatment versus control groups, 2 ASVs were significantly differentially abundant. In the breastmilk group ASV0279.Bifidobacterium was higher in the egg group. Among the non-breastmilk group, ASV001D.Bifidobacterium was lower in the egg group. When comparing milk types, 4 ASVs were lower in abundance in the breastmilk vs the non-breastmilk group, while 23 ASVs were higher in abundance. The LEfSe results and differentially abundant ASVs for breastmilk versus non-breastmilk are displayed in **Figure 3.5**.

Co-abundance Groupings

Thirty-three co-abundance groups were identified. The table of ASVs belonging to each co-abundance grouping is presented in **Supplemental Table S3.5**. The top 10 highest co-abundance group per subject relative abundance tables stratified by treatment group and milk type are presented in **Supplementary Tables S3.6** and **S3.7**. Table 3.2 displays the top 10 highest per subject relative abundance co-abundance groupings results stratified by treatment group and milk type interaction sub-groups, with the full co-abundance group results presented in **Supplementary Table S3.8**. CAG 3 had the largest per subject relative abundance percentage drop (11.67%) due to the egg intervention. CAG 6 and 22 had the largest percentage increase in per subject relative abundance (5.89% and 5.00%, respectively) due to the egg intervention in the non-breastmilk group. CAGs 3 and 6 had the largest relative abundance discrepancy between milk type groups (breastmilk: CAG 3 = 48.68%, CAG 6 = 6.88%; non-breastmilk: CAG 3 = 15.06%, CAG 6 = 20.06%). No co-abundance group relative abundance was significantly associated with either treatment group or MDD in either the non-stratified or stratified by milk type models. The model results for treatment and MDD are presented in **Supplementary Tables S3.10 and S3.11**. MaAsLin2 could not generate results for the MDD non-breastmilk stratified co-abundance group model due to the small sample size (n = 11). Consistent with individual ASV associations, the relative abundance of 13 co-abundance groups were associated with milk type (**Table 3.3**), using breastmilk as the reference. Co-abundance groups 3 and 16, which were highest in relative abundance amongst the breastmilk and non-breastmilk (respectively) stratified table were also significant in the MaAsLin2 co-abundance group results. CAGs 3 and 16 had opposite directions of association matching the results of the stratified table.

Linear Regression: Fecal Calprotectin Association with Milk type, Treatment and MDD

Linear regression analyses were conducted to evaluate the association between fecal calprotectin and treatment, MDD, and milk type, in separate models adjusting for the aforementioned covariates. Neither treatment nor MDD were significantly associated with fecal calprotectin (treatment group: β = -57.85 [95% CI = -119.08 - 98.524] p-value = 0.2811, MDD: β = -10.28 [95% CI = -187.01 - 71.313] p-value = 0.3365). There was a significant negative association between fecal calprotectin and milk type (β = -174.70 [95% CI = -314.16- -35.234], milk type p-value = 0.01439). However, the overall model p-value was not significant (p-value = 0.2282) and the model explained a small proportion of the variance (R^2 = 0.07674, Adjusted R^2 = 0.01769). The full model results are displayed in **Table 3.4**.

Discussion

This study examined the effects of an egg intervention compared to a multiple micronutrient powder control, along with diet diversity and milk type, on the gut microbiome and localized gastrointestinal (GI) inflammation in rural Guatemalan infants. Our approach leveraged high-resolution analysis by using amplicon sequence variants (ASVs) as the fundamental unit of microbial composition, allowing for sub-species level resolution. Furthermore, rather than performing taxonomic profiling, we explored bacterial ecological relationships by organizing ASVs into co-abundance groups (CAGs), or guilds, based on their co-occurrence and shared ecological functions.

Traditional alpha and beta diversity metrics revealed significantly lower diversity in the breastmilk group, with beta diversity showing distinct clustering between breastfed and non-breastfed infants. The mean + SEM Principal Coordinates Analysis (PCoA) plots confirmed these findings, providing further clarity on the interaction between treatment and milk type. The

treatment-milk type interaction PCoA plots revealed a marked separation between egg and control groups in non-breastfed infants, suggesting that microbiome composition in non-breastfed infants may be more susceptible to dietary interventions. While PCoA plots indicate compositional shifts, they do not inherently differentiate between beneficial and harmful changes. To address this, we applied guild-based co-abundance analysis to identify functionally relevant microbial assemblages.

Our results highlighted key bacterial differences. CAG3, which includes beneficial bacterial ASVs, decreased in relative abundance in the non-breastmilk group following egg intervention, whereas CAG6 and CAG16, ASV members often associated with dysbiosis, or pathogenicity, increased. This suggests that in non-breastfed infants, complementary foods should be introduced with caution, as the absence of breastmilk may lead to a more unstable microbiome composition. The relatively stable microbiota observed in breastfed infants suggests that breastmilk plays a dominant role in shaping microbial composition, potentially minimizing the impact of complementary food introduction. Prior studies have documented the resilience of breastmilk-fed infant microbiota, with human milk oligosaccharides (HMOs) selectively enriching beneficial bacteria such as Bifidobacterium and Lactobacillus, which were also more prevalent in the breastmilk group in our study.

Further, comparisons of individual ASVs revealed consistent patterns across multiple analytical approaches (MaAsLin2, LEfSe, and co-abundance group analyses). Specifically, Blautia and Lachnospiraceae ASVs often associated with dysbiotic states in early life,[18,19] were more abundant in the non-breastmilk group. Conversely, beneficial commensals such as ASVs from Veillonella, Enterococcus, Clostridioides, Bifidobacterium, Megasphaera, Lactobacillus, and Campylobacter were more prevalent in the breastmilk group. These results reinforce

previous findings that breastmilk supports a healthier microbial profile,[20,21] whereas non-breastmilk feeding may predispose infants to less favorable microbiota compositions.

The fecal calprotectin and milk type regression results suggest that non-breastmilk feeding may be associated with lower fecal calprotectin levels, though the biological significance of this finding remains uncertain. Research has shown that fecal calprotectin levels are typically higher in younger infants and decrease with age. Lower levels in non-breastfed infants in this study could reflect reduced inflammatory activity but may also indicate insufficient immune stimulation necessary for developing appropriate immune responses. Prior work has suggested that transient inflammatory responses during complementary food introduction may be beneficial for immune system maturation. Given the malnourished population studied here, more research is needed to determine whether lower fecal calprotectin levels in this context indicate improved or impaired gut health.

Study Limitations

Several limitations should be considered when interpreting our findings. First, the cross-sectional design limits our ability to infer causality or track temporal changes in gut microbiome composition. Although differences were observed at the endline, it is unclear how microbiota composition evolved throughout the intervention period. Future longitudinal studies with repeated fecal sampling would provide a more comprehensive understanding of microbiome dynamics in response to dietary changes. Second, 16S rRNA sequencing of the V4 region provides limited taxonomic resolution compared to whole-metagenome sequencing. It does not capture fungi or protists and is restricted in its ability to resolve microbial diversity at the strain level. Additionally, 16S sequencing lacks functional annotation capabilities, limiting our ability to infer metabolic potential and ecological interactions. Given the poor sanitation conditions in the study population, potential exposures to mycotoxins such as aflatoxins, which could

influence gut health, remain unexamined. Third, dietary intake was assessed via caregiver-reported questionnaires, which introduces potential recall bias and misclassification. While the World Health Organization (WHO) Infant and Young Child Feeding indicators questionnaire is a validated tool,[9] objective measures such as food diaries or biochemical markers would enhance dietary data accuracy.

Future Directions and Conclusions

Future research should prioritize longitudinal designs to assess microbiome stability over time and the persistence of diet-induced changes. Additionally, studies should include exclusively breastfed and non-breastfed infants to clarify how microbiome maturation differs across feeding patterns. Investigating the relationship between gut microbiome composition, functional microbiome pathways, and host metabolic outcomes would provide deeper insights into the mechanisms by which diet influences infant health. Moreover, further research is needed to establish reference values for fecal calprotectin levels in malnourished LMIC infant populations to aid interpretation of inflammatory biomarkers.

Overall, our findings underscore the strong influence of breastfeeding on gut microbiome stability, with non-breastfed infants exhibiting greater microbial shifts in response to dietary interventions. By leveraging high-resolution ASV-based analysis combined with guild-based co-abundance grouping, we captured ecologically relevant microbial interactions rather than solely relying on conventional taxonomic profiling. The specific microbial differences observed suggest that non-breastmilk diets may lead to less favorable microbiota compositions, with potential implications for long-term health. These results highlight the importance of developing complementary feeding strategies that consider milk type to optimize gut microbiome health in malnourished infants.

Author Contributions

LEL designed and conducted the analysis, analyzed the data, and wrote the paper. MC, TCW, NZ and LMA advised LEL throughout the entire research process. LEL had primary responsibility for final content. All authors have read and approved the final manuscript.

Funding

This work was supported by a grant from the Egg Nutrition Center to the Academy of Nutrition and Dietetics.

References

1. Cowan CSM, Dinan TG, Cryan JF. Annual Research Review: Critical windows - the microbiota-gut-brain axis in neurocognitive development. *J Child Psychol Psychiatry*. 2020;61(3):353-371. doi:10.1111/jcpp.13156
2. Black RE, Victora CG, Walker SP, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet*. 2013;382(9890):427-451. doi:10.1016/s0140-6736(13)60937-x
3. Lu C, Black MM, Richter LM. Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional, and country level. *Lancet Glob Health*. 2016;4(12):e916-e922. doi:10.1016/s2214-109x(16)30266-2
4. Moore RE, Townsend SD. Temporal development of the infant gut microbiome. *Open Biol*. 2019;9(9):190128. doi:10.1098/rsob.190128
5. Johnson CC, Ownby DR. The infant gut bacterial microbiota and risk of pediatric asthma and allergic diseases. *Transl Res*. 2017;179:60-70. doi:10.1016/j.trsl.2016.06.010
6. Sordillo JE, Zhou Y, McGeachie MJ, et al. Factors influencing the infant gut microbiome at age 3-6 months: findings from the ethnically diverse Vitamin D Antenatal Asthma Reduction Trial (VDAART). *J Allergy Clin Immunol*. 2017;139(2):482-491.e414. doi:10.1016/j.jaci.2016.08.045
7. Fontaine F, Turjeman S, Callens K, Koren O. The intersection of undernutrition, microbiome, and child development in the first years of life. *Nat Commun*. 2023;14(1):3554. doi:10.1038/s41467-023-39285-9

8. Wallace TC, Rohloff P, Jimenez EY, et al. Academy of Nutrition and Dietetics Nutrition Research Network: The Saqmolo' Project rationale and study protocol for a randomized controlled trial examining the influence of daily complementary feeding of eggs on infant development and growth in Guatemala. *J Acad Nutr Diet*. 2022;122(2):432-444. doi:10.1016/j.jand.2021.03.015

9. World Health Organization. Indicators for assessing infant and young child feeding practices part 3: country profiles. August 31, 2022. https://apps.who.int/iris/handle/10665/44368

10. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7(5):335-336. doi:10.1038/nmeth.f.303

11. RStudio Team. *RStudio: Integrated Development for R*. PBC; 2020.

12. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol*. 2021;17(11):e1009442. doi:10.1371/journal.pcbi.1009442

13. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:10.1186/s13059-014-0550-8

14. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6):R60. doi:10.1186/gb-2011-12-6-r60

15. Wu G, Xu T, Zhao N, et al. A core microbiome signature as an indicator of health. *Cell*. 2024;187(23):6550-6565.e6511. doi:10.1016/j.cell.2024.09.019

16. Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *J Classif*. 2014;31(3):274-295. doi:10.1007/s00357-014-9161-z

17. Claesson MJ, Jeffery IB, Conde S, et al. Gut microbiota composition correlates with diet and health in the elderly. *Nature*. 2012;488(7410):178-184. doi:10.1038/nature11319

18. Fallani M, Amarri S, Uusijarvi A, et al. Determinants of the human infant intestinal microbiota after the introduction of first complementary foods in infant samples from five European centres. *Microbiology (Reading)*. 2011;157(Pt 5):1385-1392. doi:10.1099/mic.0.042143-0

19. Arrieta MC, Stiemsma LT, Amenyogbe N, Brown EM, Finlay B. The intestinal microbiome in early life: health and disease. *Front Immunol*. 2014;5:427. doi:10.3389/fimmu.2014.00427

20. Yao Y, Cai X, Ye Y, Wang F, Chen F, Zheng C. The role of microbiota in infant health: from early life to adulthood. *Front Immunol*. 2021;12:708472. doi:10.3389/fimmu.2021.708472

21. Robertson RC, Manges AR, Finlay BB, Prendergast AJ. The human microbiome and child growth - first 1000 days and beyond. *Trends Microbiol*. 2019;27(2):131-147. doi:10.1016/j.tim.2018.09.008

Tables

Table 3.1. Infant characteristics by treatment group.

	Egg (N=106)	Control (N=94)	Total (N=200)	P value
Minimum Diet Diversity				0.506[1]
Yes	65 (70.7%)	49 (65.3%)	114 (68.3%)	
No	27 (29.3%)	26 (34.7%)	53 (31.7%)	
Missing	14	19	33	
Milk type most frequently consumed				0.131[1]
Non-breastmilk*	14 (13.9%)	20 (22.7%)	34 (18.0%)	
Breastmilk	87 (86.1%)	68 (77.3%)	155 (82.0%)	
Missing	5	6	11	
Age (Months)				0.872[2]
Mean (SD)	12.642 (1.044)	12.667 (1.155)	12.653 (1.094)	
Missing	0	1	1	
Sex				0.779[1]
Boy	53 (50.0%)	49 (52.1%)	102 (51.0%)	
Girl	53 (50.0%)	45 (47.9%)	98 (49.0%)	
National Poverty Likelihood Score				0.349[2]
Mean (SD)	23.501 (29.351)	27.487 (30.611)	25.374 (29.940)	
Food Insecurity Score				0.253[1]
0	56 (52.8%)	42 (44.7%)	98 (49.0%)	
1	14 (13.2%)	8 (8.5%)	22 (11.0%)	
2	7 (6.6%)	12 (12.8%)	19 (9.5%)	
3	12 (11.3%)	6 (6.4%)	18 (9.0%)	
4	3 (2.8%)	9 (9.6%)	12 (6.0%)	
5	4 (3.8%)	6 (6.4%)	10 (5.0%)	
6	4 (3.8%)	3 (3.2%)	7 (3.5%)	
7	4 (3.8%)	5 (5.3%)	9 (4.5%)	
8	2 (1.9%)	3 (3.2%)	5 (2.5%)	
Water Source				0.009[1]
Bagged/sachet water	20 (18.9%)	23 (24.5%)	43 (21.5%)	
Borehole/tubewell	15 (14.2%)	9 (9.6%)	24 (12.0%)	
Other**	12 (11.3%)	17 (18.1%)	29 (14.5%)	
Piped water (inside home or land)	53 (50.0%)	29 (30.9%)	82 (41.0%)	
Protected dug well	6 (5.7%)	16 (17.0%)	22 (11.0%)	
Premature				1.000[1]
No	82 (77.4%)	73 (77.7%)	155 (77.5%)	

	Egg (N=106)	Control (N=94)	Total (N=200)	P value
Yes	24 (22.6%)	21 (22.3%)	45 (22.5%)	
Prenatal Care				0.737[1]
No	4 (3.8%)	5 (5.3%)	9 (4.5%)	
Yes	102 (96.2%)	89 (94.7%)	191 (95.5%)	
Born Post-term				0.283[1]
No	76 (92.7%)	71 (97.3%)	147 (94.8%)	
Yes	6 (7.3%)	2 (2.7%)	8 (5.2%)	
Missing	24	21	45	
Low Birthweight (<2500g)				0.580[1]
No	77 (78.6%)	69 (82.1%)	146 (80.2%)	
Yes	21 (21.4%)	15 (17.9%)	36 (19.8%)	
Missing	8	10	18	
Antibiotics Ever				0.518[1]
No	25 (23.6%)	26 (28.0%)	51 (25.6%)	
Yes	81 (76.4%)	67 (72.0%)	148 (74.4%)	
Missing	0	1	1	
Antibiotics last 14 days				0.410[1]
No	83 (78.3%)	68 (72.3%)	151 (75.5%)	
Yes	23 (21.7%)	26 (27.7%)	49 (24.5%)	
Diarrhea last 14 days				0.570[1]
No	56 (52.8%)	54 (57.4%)	110 (55.0%)	
Yes	50 (47.2%)	40 (42.6%)	90 (45.0%)	
Animal Contact				0.026[1]
No	45 (42.5%)	25 (26.6%)	70 (35.0%)	
Yes***	61 (57.5%)	69 (73.4%)	130 (65.0%)	
Herbal Supplements				1.000[1]
No	102 (96.2%)	89 (96.7%)	191 (96.5%)	
Yes	4 (3.8%)	3 (3.3%)	7 (3.5%)	
Missing	0	2	2	
Medication				0.277[1]
No	96 (90.6%)	79 (84.9%)	175 (87.9%)	
Yes	10 (9.4%)	14 (15.1%)	24 (12.1%)	
Missing	0	1	1	

1. Fisher's Exact Test for Count Data.
2. Linear Model ANOVA
*Includes fluid animal milk, powdered animal milk, and infant formula.
**Includes unprotected dug well, protected spring, unprotected spring, tanker truck, surface water [pond, river, lake], other person, other.
***Includes regular direct contact with animals, including dog, cat, chicken, turkey, pig, cow, sheep, goat, donkeys, and other.

Table 3.2. Per subject relative abundance of the top 10 most abundant co-abundance groups by treatment and milk type interaction subgroups.

Egg + Breastmilk group	%	Control + Breastmilk group	%	Egg + Non-breastmilk group	%	Control + Non-breastmilk group	%
CAG3	48.5	CAG3	48.9	CAG16	22.4	CAG3	19.9
CAG22	13.6	CAG22	13.5	CAG22	22.2	CAG16	18.4
CAG16	5.97	CAG16	8.08	CAG17	10.2	CAG22	17.2
CAG5	5.06	CAG5	4.49	CAG3	8.22	CAG17	10.5
CAG8	4.12	CAG1	3.47	CAG6	6.88	CAG8	8.95
CAG1	3.99	CAG4	3.11	CAG18	5.37	CAG18	4.95
CAG17	3.55	CAG17	2.95	CAG8	4.05	CAG24	3.07
CAG9	3.34	CAG8	2.86	CAG1	3.10	CAG1	2.12
CAG4	3.02	CAG6	2.69	CAG24	2.88	CAG10	1.75
CAG6	1.32	CAG9	2.04	CAG9	2.19	CAG21	1.40

*Egg + Breastmilk: n = 87, Control + Breastmilk: n = 67, Control + Nonbreastmilk: n = 20, Egg + Nonbreastmilk: n = 14

Table 3.3. MaAsLin2 association of infant gut microbiome co-abundance groupings with most frequently consumed milk type (breastmilk versus non-breastmilk).

Feature	Coefficient	N.not.0	P Value	Q Value
CAG3	-2.05	186	9.74E-10	4.18E-07
CAG5	-3.44	156	2.46E-08	5.28E-06
CAG4	-4.32	130	3.19E-07	4.56E-05
CAG19	3.23	105	1.03E-06	1.10E-04
CAG25	3.28	93	2.10E-06	1.80E-04
CAG24	3.52	117	1.22E-05	8.70E-04
CAG16	2.67	178	3.08E-05	1.89E-03
CAG18	3.31	130	3.76E-05	2.01E-03
CAG17	3.16	145	8.35E-05	3.72E-03
CAG31	1.82	52	8.68E-05	3.72E-03
CAG2	2.13	164	1.08E-04	4.22E-03
CAG21	2.29	79	5.86E-04	2.10E-02
CAG26	1.78	57	7.07E-04	2.33E-02

Table 3.4. Linear regression results displaying the association between milk type most frequently consumed (breastmilk or non-breastmilk) and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	**SE**	**P Value**	**β (95% CI)**	**SE**	**P Value**
Intercept	400 (345, 455)	27.9	8.77E-32	524 (-88.2, 1135)	310.070	0.0930
Milk Type (Non-breastmilk)	-193 (-321, 63.7)	65.3	0.00360	-175 (-314, -35.2)	70.7	0.0144
Age				-20.0 (-67.4, 27.3)	24.0	0.405
Sex (Girl)				8.85 (-93.6, 111)	51.9	0.865
Antibiotics ever (Yes)				85.9 (-39.2, 211)	63.3	0.177
Water source (Borehole/tubewell)				96.3 (-89.1, 282)	93.9	0.307
Water source (Other)				145 (-31.0, 321)	89.2	0.106
Water source (Piped water (inside home or land))				93.6 (-44.5, 232)	69.9	0.183
Water source (Protected dug well)				45.2 (-145, 235)	96.2	0.639
Antibiotics 14 days (Yes)				-74.5 (-208, 59.0)	67.7	0.272
Diarrhea 14 days (Yes)				16.2 (-95.1, 128)	56.4	0.774
Treatment (Egg)				-16.0 (-122, 90.2)	53.8	0.766
Model Summary						
R^2	0.0451			0.0763		
Adjusted R^2	0.0399			0.0229		
Model P value	0.00360			0.171		
Sample size (n)	186			184		

Figures

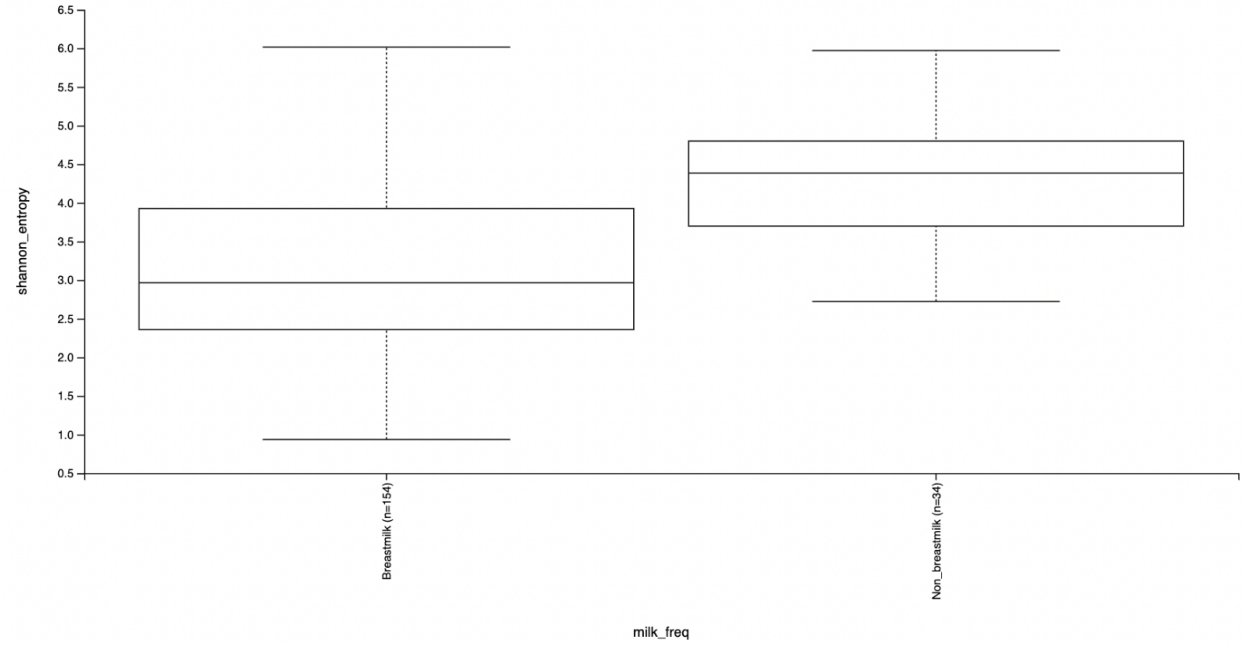

Figure 3.1. Shannon Diversity comparison between milk type most frequently consumed groups.

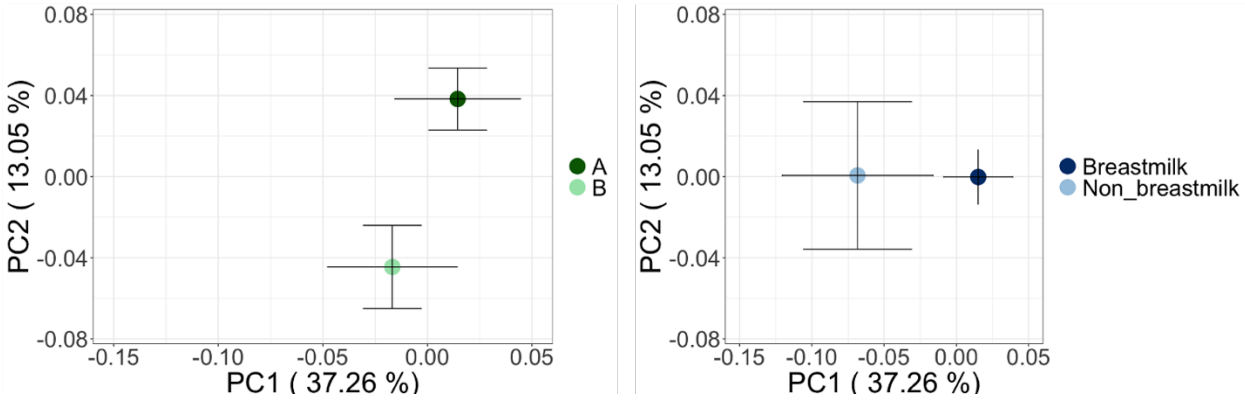

Figure 3.2. Mean + SEM PCoA plots using Bray-Curtis distance (left: colored by treatment groups [A = egg group, B = control group, n = 199], right: colored by milk type [breastmilk, non-breastmilk, n = 188]).

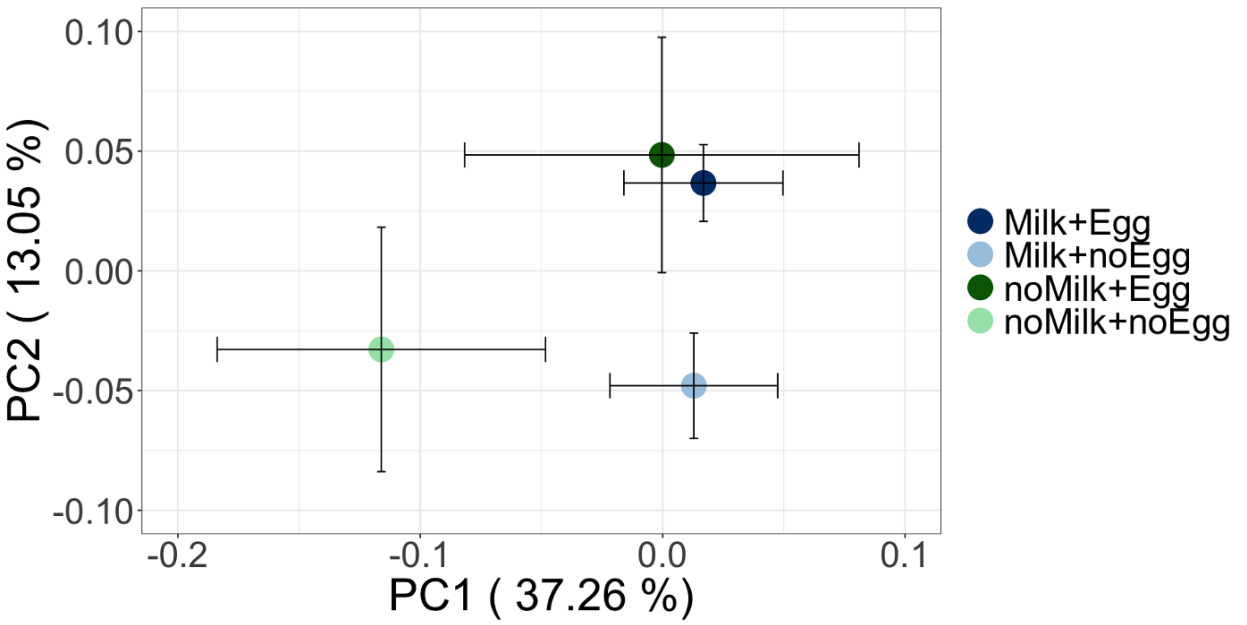

Figure 3.3. Mean + SEM PCoA plots using Bray-Curtis distance colored by treatment and milk type interactions (breastmilk + egg [n = 87], breastmilk + control [n = 67], non-breastmilk + egg [n = 14], non-breastmilk + control [n = 20]).

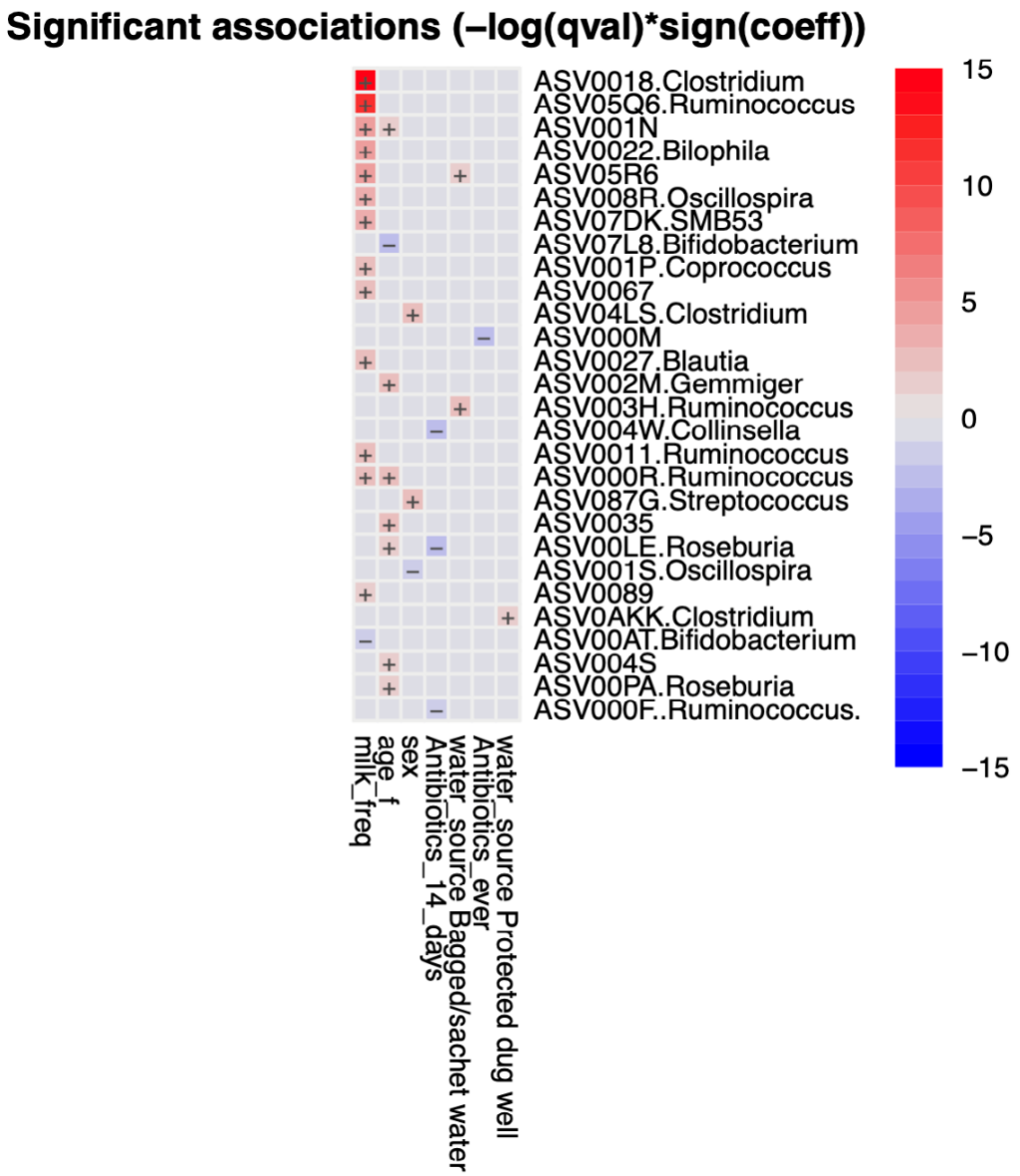

Figure 3.4. Significant ASVs associated with milk type most frequently consumed using MaAsLin2.

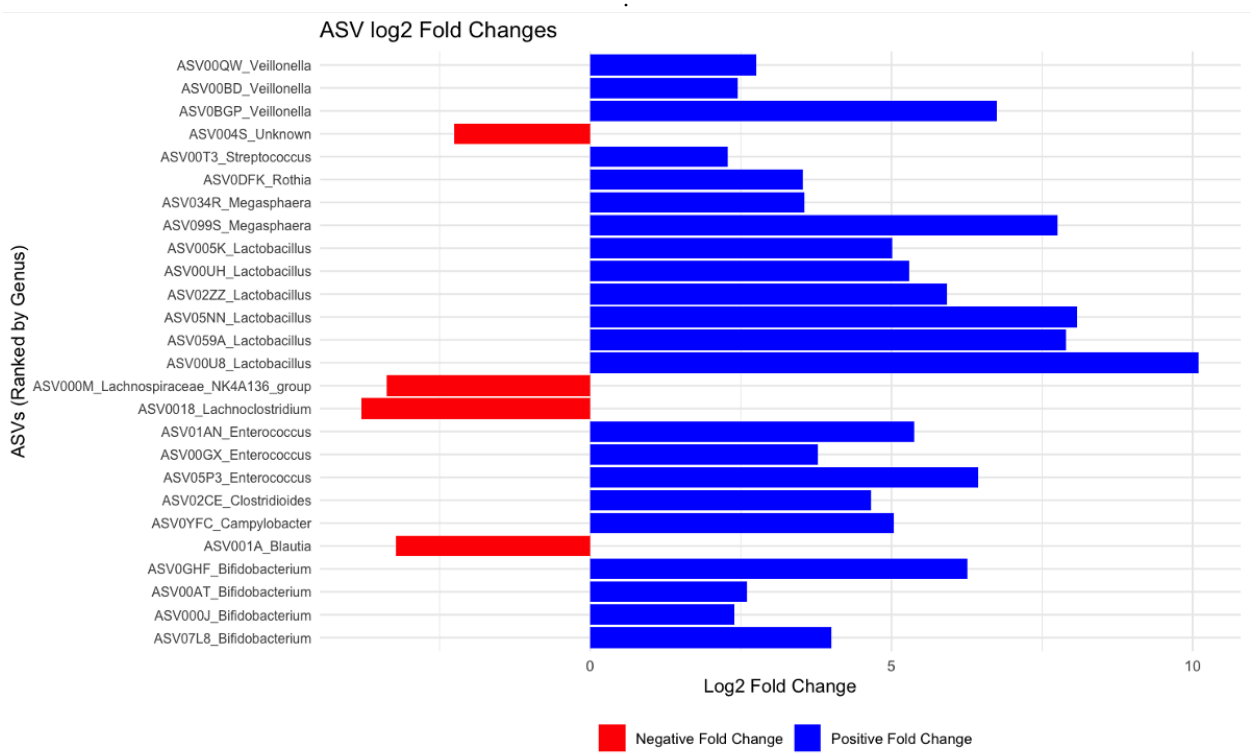

Figure 3.5. LEfSe (DESEQ2) Analysis displaying the 27 significantly differentially abundant ASVs comparing the breastmilk group to the non-breastmilk group out of 295 total ASVs.

CHAPTER 4:
The Association of Early Child Development and Growth Outcomes with the Gut Microbiome in 11-16-Month-Old Guatemalan Infants

Abstract

Objective: The main objective of this study was to determine whether fecal microbiome composition was associated with differences in early child development (ECD) scores, growth anthropometrics, and incidence of stunting.

Methods: The parent study enrolled 1,200 rural Guatemalan infants aged 6 to 9 months, with 200 participating in the microbiome sub-study. ECD was assessed using the Caregiver Reported Early Child Development Instruments (CREDI) and the Guide for Monitoring Child Development, while anthropometric measures included World Health Organization defined Z-scores, stunting, and wasting. Dietary intake, including egg consumption and primary milk type, was assessed through caregiver questionnaires and interviews. Gut microbiome composition was sequenced using 16S rRNA and analyzed using QIIME2. MaAsLin2 was used to evaluate associations between amplicon sequencing variants (ASVs), guild-based co-abundance groupings, ECD scores, and growth outcomes, adjusting for relevant covariates. MaAsLin2 co-abundance grouping models were additionally stratified by most frequently consumed milk type.

Results: Among 200 infant participants, the mean overall CREDI score was 47.4, with no significant differences between treatment groups. Stunting was observed in 48.7% of infants. Neither alpha or beta diversity comparisons showed significant differences between high and low CREDI or growth score groups. After stratification by milk type, CREDI overall (co-abundance group 18: MaAsLin2 coefficient = 1.22, Q value = 0.216) and language (co-abundance group 18: MaAsLin2 coefficient = 1.95, Q value = 0.0226) maintained a significant positive association in non-breastmilk fed infants with co-abundance groups.

Conclusion: Milk type influenced the relationship between gut microbiome composition and both ECD and growth outcomes, with breastmilk potentially modifying these associations. The

positive correlation between CAG 18 and CREDI language scores suggest beneficial gut microbes may support neurodevelopment. These findings highlight the need for further research on the microbiome's role in infant development, especially in populations with high rates of malnutrition.

Introduction

The well-established relationship between infant health, breastfeeding, and gut microbiome development underscores the critical role of early nutrition. Breastfeeding has been shown to promote healthier gut microbiome profiles, characterized by higher abundances of beneficial bacteria such as *Bifidobacterium* and *Lactobacillus*. Emerging research is beginning to unravel the intricate connections between infant growth, development, and gut microbiome composition—an area of particular importance in low- and middle-income countries (LMICs), where malnutrition, stunting, environmental contamination (e.g., unsafe water, food supply, and domestic animals), and intestinal pathogens are prevalent.

Proper nutrition is essential for physical growth, as deficiencies can lead to stunting, wasting, and other developmental concerns. Gut dysbiosis has been implicated in impaired nutrient absorption, compromised intestinal integrity, and increased pathogen load, all of which can negatively impact growth. Studies in infants and young children have demonstrated links between malnutrition, gut microbiome composition, and anthropometric outcomes.[1-3] For example, a longitudinal study of Bangladeshi children from birth to two years found that severe acute malnutrition and poor growth trajectories were associated with an immature gut microbiota.[2] While these associations are well-documented, the precise mechanisms underlying the relationships between complementary feeding, milk type, gut microbiome composition, and growth outcomes in human infants remain unclear.

Beyond physical growth, early childhood development (ECD) encompasses cognitive, linguistic, and neurological maturation, all of which may be influenced by both breast milk components and gut microbiota.[4-7] Specific bacterial taxa, such as Bifidobacterium and Bacteroides, have been linked to cognitive development,[8] while Streptococcus abundance has

been associated with worse executive function.[7] Additionally, disruptions to early microbial colonization—such as antibiotic use in the first 9 months of life—have been correlated with lower receptive vocabulary,[9] suggesting that gut microbiome alterations may impact neurodevelopmental trajectories.

Recent findings highlight the role of human milk oligosaccharides (HMOs) in shaping gut microbial communities, particularly through the enrichment of *Bifidobacterium* and *Bacteroidetes*, which contribute to both physical growth and cognitive outcomes. Notably, bacteria involved in digestion and nutrient absorption (*Bifidobacterium, Lactobacillus, Ruminococcaceae*) are also implicated in neurodevelopment via the gut-brain axis,[6,10,11] suggesting potential overlap in the microbial profiles that support both growth and cognitive function. These insights point to the potential of breastfeeding and targeted complementary feeding strategies as interventions to optimize both growth and neurodevelopment in early childhood.

Despite these advances, the relationships between ECD, growth, milk type, and gut microbiome composition remain poorly understood in malnourished populations. Most existing research on these associations has been conducted in healthy infants, leaving critical gaps in our understanding of how these dynamics unfold in at-risk populations. Our previous findings from the Saqmolo' Project suggest that milk type is a key determinant of gut microbiome composition, aligning with recent research on HMOs and child development. Further investigation is needed to clarify how early feeding practices shape the gut microbiome and, in turn, influence both growth and neurodevelopment in vulnerable populations.

The main objective of this study was to determine whether endline fecal microbiome composition was associated with differences in 1) early child development scores; 2) growth

anthropometrics (z-scores for weight-for-age [WAZ], length-for-age [LAZ], weight-for-length [WAL], and head circumference-for-age [HAZ]); and 3) incidence of stunting and wasting. We hypothesized a more mature endline fecal microbiome composition and microbial community structure would be associated with 1) higher early child development scores; 2) more optimal growth anthropometrics; and 3) lower incidence of stunting and wasting. To achieve our objective, we analyzed the relative abundance of bacterial amplicon sequence variants (ASVs) through 16S rRNA gene sequencing of fecal samples collected from 200 indigenous Mayan infants participating in The Saqmolo' Project, and their association with ECD and growth outcomes.

Material and Methods

Study Population

The study population details are described in Chapter 3: subsection *Study Population*. Briefly, the Saqmolo' Project enrolled 1,200 rural Guatemalan infants age 6 to 9 months, with 200 infants participating in a microbiome sub-study age 11-16 months at endline.

ECD, Demographic, and Anthropometric Assessments

ECD was measured using the global development scoring tools: the Caregiver Reported Early Child Development Instruments [CREDI] long form and the Guide for Monitoring Child Development.[12] Anthropometric measures included Z-scores for weight-for-age, length-for-age, weight-for-length, and head circumference-for-age. Weight was measured using a hanging Salter scale. Length and head circumference were measured using a portable stadiometer. Stunting and wasting were measured using the World Health Organization (WHO) standards with stunting considered length-for-age < -2SD and wasting weight-for-length < -2SD.[13] A family care indicators questionnaire assessed home play activities, while drinking water access and source

were evaluated using the WHO water access questionnaire.[14] Poverty levels were determined using the Quick Poverty Score.[15]

Dietary Questionnaires

Egg consumption was tracked using interviews and questionnaires during monthly home visits, along with biweekly phone check-ins with caregivers over the first two months. Diet diversity was measured using the World Health Organization (WHO) Infant and Young Child Feeding indicators questionnaire.[16] The primary type of milk consumed was determined from the dietary questionnaires and categorized into two groups: breastmilk and non-breastmilk. The non-breastmilk group included fluid animal milk, powdered animal milk, and infant formula.

Calculation of CREDI and Growth Statistics

Mean and SD were calculated for CREDI and anthropometric Z-scores stratified by treatment group and significance was tested using Analysis of Variance (ANOVA). Fisher's exact test was used to test differences between stunting in the treatment compared to the control group. After initial analyses, mean and SD were calculated for CREDI language scores stratified by most frequently consumed milk type and significance was tested using Analysis of Variance (ANOVA). Mean and SD were additionally calculated for growth scores stratified by milk type and treatment sub-groups.

Fecal Procedures, Sequencing, and Gut Microbiome Analysis

Fecal collection, transportation procedures, and gut microbiome sequencing have been previously described in Chapter 3 subheading: *Fecal Procedures and Gut Microbiome Quantification*. QIIME2 analyses are described in Chapter 3 subheading: *Quantitative Insights into Microbial Ecology version 2 (QIIME2)*. Creation of co-abundance groupings is described in Chapter 3 subheading: *Guild Based Co-Abundance Groupings.* High and low groups were

created based on the median for each ECD score and anthropometric score to use in QIIME2 for all statistical comparisons.

Multivariate Associations with Linear Models (MaASLin2): Individual ASV Associations

MaASLin2 was used to individually test the association between relative abundance of the top 10% most prevalent ASV present in fecal samples, with raw ECD scores, anthropometric Z-scores, and stunting. Abundance of the ASVs were normalized by dividing the raw count of each ASV in a sample by the total number of reads in that sample. Model equations including adjusted covariates are listed below in **Equations 4.1 to 4.3**. MaASLin2 coefficients were reported with q-values (FDR adjusted p-values) ≤ 0.25 as significant.

> ***Equation 4.1:*** *ASVs ~ CREDI scores + age + sex + milk type + antibiotics ever + water source + stunting + antibiotics_14_days + diarrhea 14 days + treatment + play activities score + poverty score*
>
> ***Equation 4.2:*** *ASVs ~ Z-scores + age + sex + milk type + antibiotics ever + water source + antibiotics_14_days + diarrhea 14 days + treatment + low birthweight + poverty score*
>
> ***Equation 4.3:*** *ASVs ~ Stunting + age + sex + milk type + antibiotics ever + water source + antibiotics_14_days + diarrhea 14 days + treatment + low birthweight + poverty score*

Co-abundance Groupings: MaAsLin2

The co-abundance groupings were used in downstream analysis using MaAsLin2 to determine the association between each co-abundance groups with ECD and growth scores. The co-abundance group models were adjusted for the same covariates as the ASV MaAsLin2 models described in the previous section. Co-abundance grouping models were additionally performed stratified by most frequently consumed milk type. Co-abundance groups are referred to as CAGs in the results.

Results

Population, CREDI and Growth Statistics

The analysis included 200 infant participants. Mean score for CREDI overall = 47.4. The mean (SD) CREDI Language Score was 49.0 (0.657) among non-breastmilk-fed infants (N = 34) and 48.7 (0.759) among breastmilk-fed children (N = 155), with a statistically significant difference between groups (p = 0.041). For stunting, 48.7% met the WHO criteria. Only 1 infant met the WHO criteria for wasting, therefore wasting was not included in further analyses. All descriptive statistics for individual CREDI module scores, stunting, anthropometric z-scores, and model covariates are listed in **Table 4.1**. **Supplementary Table S4.1** and **Table S4.2** (in **Appendix B**) display the mean for anthropometric Z-scores stratified by milk type and treatment subgroups.

QIIME2: Alpha- and Beta-Diversity Comparisons

Alpha diversity was compared between high and low CREDI, anthropometric Z-scores, and stunting. For Kruskal-Wallis test comparisons of Shannon diversity, no CREDI or growth scores were statistically different between high and low score groups (full results are displayed in **Supplementary Table S4.3**). Shannon Diversity, Faith's Phylogenetic diversity and Peilou's Evenness for all CREDI and growth scores are displayed in **Supplementary Table S4.3** and boxplots in **Supplementary Figures S4.1** to **S4.33**. PCoA plots showed no clear separation between high and low CREDI or growth scores. For beta-diversity, no CREDI or growth scores were significantly different between groups when comparing weighted Unifrac metrics using PERMANOVA tests (see **Supplementary Table S4.4**).

ASV MaAsLin2 Associations

The most significant ASV associations with CREDI score domains are as follows (*CREDI Overall:* lower relative abundance of *ASV02CE.Clostridioides* (MaAslin2 coefficient = -

0.470, q = 0.074).; *CREDI Cognitive:* lower relative abundance of *ASV02CE.Clostridioides* (MaAslin2 coefficient = -0.478, q = 0.0204).; *CREDI Language*: higher relative abundance of *ASV0067.Butyricicoccus* (MaAslin2 coefficient = 0.423, q = 0.186); *CREDI Motor*: higher relative abundance of *ASV0013.Bacteroides* (MaAslin2 coefficient = 0.728, q = 0.147); *CREDI Socioemotional*: lower relative abundance of *ASV02CE.Clostridioides* (MaAslin2 coefficient = -0.583, q = 0.00524). The full ASV results for the CREDI domain models are presented in **Supplementary Table S4.5**. Lower relative abundance of ASV02CE.*Clostridioides* was associated with each of the CREDI score modules; however, this feature was not present in co-abundance groupings significantly associated with CREDI scores when stratified by milk type.

For growth metrics, no ASVs were significantly associated with stunting. The most significant ASV associations with Z-scores are as follows (*Z-BMI*: lower relative abundance of ASV0065.*Veillonella* (MaAsLin2 coefficient = -0.685, Q = 0.228); *Z-Head-Circumference*: lower relative abundance of ASV099S.*Megasphaera* (MaAsLin2 coefficient = -0.977, Q = 0.205); *Z-Length-for-Age*: higher relative abundance of ASV004I.*Lachnoclostridium* (MaAsLin2 coefficient = 0.365, Q = 0.161; *Z-Weight-for-Length*: lower relative abundance of ASV01S6.*Lachnospira* (MaAsLin2 coefficient = -0.619, Q = 0.107). The full ASV results for growth score models are presented in **Supplementary Table S4.6**. No ASVs were the same across growth scores and none were present in the co-abundance groupings significantly associated with growth scores in the non-stratified or milk type stratified models.

Co-abundance Groupings Associations

The full list of ASVs and their module membership for guild co-abundance groups are listed in **Supplementary Table S4.5**. For CREDI scores associated with co-abundance groupings, CREDI overall, cognitive, motor, language and socioemotional scores were significantly associated with various co-abundance groups (**Table 4.2**). After stratification by

milk type, only CREDI overall (CAG 18: MaAsLin2 coefficient = 1.22, Q value = 0.216) and language (CAG 18: MaAsLin2 coefficient = 1.95, Q value = 0.0226) maintained a significant positive association with relative abundance of co-abundance group 18 in the non-breastmilk group. All significant CREDI score results from the stratified model are displayed in **Table 4.3**. For growth scores, head-circumference, Z-weight-for-age, and stunting were significantly associated with various co-abundance groupings relative abundance in the non-stratified models (**Table 4.4**). However, the associations did not remain significant after milk type stratification (**Table 4.5**).

Discussion

This study aimed to determine the association of gut microbiome composition with ECD and growth outcomes, in the context of a LMIC infant population. In summary, CREDI overall and language scores appeared to have the strongest association with gut microbiome co-abundance groupings within the non-breastmilk group.

When stratifying by milk type, previously significant associations between gut microbiome composition and both ECD and growth outcomes lost significance, likely due to differences in microbiome profiles between breastfed and non-breastfed infants. While sample size constraints may have contributed to this loss in non-breastfed models, the persistence of associations between CAG 18 and CREDI scores in the non-breastmilk group suggests a potential microbial role in language development. This especially holds true given language scores were significantly higher in non-breastfed infants, although the actual language score difference was minimal. Meanwhile, CREDI overall score and CREDI language score gained significance in the non-breastmilk group. Of note, none of the individual ASVs from the MaAsLin2 models associated with CREDI overall or CREDI language were present in CAG 18.

Given that the language score is a module in the CREDI overall score, and the significantly associated co-abundance group was the same for both the overall and language score (CAG 18), it is reasonable to assume that the language module is driving the association. The ASVs present in CAG 18 are generally considered commensal and beneficial. *Faecalibacterium* has been associated with higher neurodevelopment scores and was lower in abundance in stunted infants.[17] However, a systematic review reported variable effects of higher *Faecalibacterium* abundance on behavior and neurodevelopment.[17] There is limited research on the abundance of *Erysipelotrichaceae* in infant guts; however, it has been linked to the breakdown of human milk oligosaccharides (HMO) and has been shown to increase with increased age.[18] *Lachnospira, Ruminococcus, and Roseburia* are generally considered commensal short chain fatty acid producers which strengthen the gut barrier.[19]

It is interesting that relative abundance of CAG 18 was positively associated with language scores in the non-breastmilk group given that CAG 18 was also positively associated with language score in the breastmilk group. However, the result in the breastmilk group was not significant. It's possible this could indicate there are other nutritional factors at play such as nutrient differences in infant formula and animal milk versus breastmilk. In malnourished settings, variability in maternal nutrition may affect breastmilk composition, potentially altering its impact on infant neurodevelopment compared to formula or other milk sources. Additionally, the composition of infant formula is known to alter infant gut microbiome profiles due to the absence of HMO[3] and the interaction of micronutrients with gut bacteria.[20-22] As breastmilk is a strong influencer of a healthy gut microbiome composition in infants, breastmilk consumption may buffer or minimize the relationship between CAG 18 and language scores. While the positive association between CAG 18 and language scores was observed in both groups, its

significance only in the non-breastmilk group raises questions about potential nutritional or microbial buffering effects of breastmilk.

Previous research on the gut-brain-axis offers biological plausibility for the relationship between gut microbiota and neurodevelopment. The gut-brain axis is a biological highway by which the gut and brain communicate in a bi-directional relationship.[23] The gut and brain communicate by neuro-immuno-endocrine mediators. Neurotransmitter modulation (gamma-aminobutyric acid [GABA] and brain-derived neurotrophic factor [BDNF]) has been observed in mice given changes in microbiota composition occurring alongside changes in intestinal barrier protection and integrity, bacterial metabolites absorbed by enterocytes, modulation of enteric sensory afferents, and mucosal immune regulation in the gut.[24] Changes in the brain can have a reciprocal effect on the gut by stimulating changes in mucus production, motility, and immune function.[24] The exact mechanism between specific bacterial co-abundance groupings and language development remains unclear in human infants.

Study Limitations

As discussed in Chapter 3, the cross-sectional nature of this study precludes any determination of causality. For instance, while this analysis examines the potential influence of the gut microbiome on growth, previous studies have also explored whether stunting itself may drive changes in the microbiome rather than the other way around. Longitudinal studies would be necessary to establish temporal relationships and causality. Additionally, the relatively small sample size in the non-breastmilk group limited the statistical power to detect significant associations in the co-abundance grouping MaAsLin2 analyses when stratifying by predominant milk type. With a larger sample size, it is possible that more robust and meaningful associations between microbiome composition, ECD, and growth outcomes could be identified in the non-breastmilk group. While key covariates were adjusted for, unmeasured factors such as maternal

nutrition or mixed feeding patterns may have influenced results. Again, while 16S rRNA sequencing provides valuable insights into microbiome composition, it does not offer strain-level resolution or allow for precise functional profiling of microbial communities. Lastly, milk type was categorized as breastmilk or non-breastmilk based on caregiver-reported dietary questionnaires. This classification also does not account for mixed feeding patterns or variations in the composition of non-breastmilk sources (e.g., cow's milk vs. infant formula).

Future Directions and Conclusions

Future longitudinal studies with larger and more balanced sample sizes across milk groups would improve the robustness of these findings. Future studies should aim to include equal sample size of exclusively breastfed and non-breastfed infants to determine whether gut microbiome composition in non-breastfed infants has a greater effect on ECD and growth outcomes. Breastmilk composition should also be analyzed in future studies and compared to the nutrient profiles of other milk types. Determining the nutrient composition of milk types would further clarify the impact of milk type on the relationship between the gut microbiome and infant development outcomes. Lastly, shotgun metagenomic sequencing should be used in the future as it is needed to determine functional pathways related to growth and development.

In conclusion, milk type modified the association between the gut microbiome composition and both ECD and growth outcomes in this infant cohort. The co-abundance group of CAG 18 (including predominantly beneficial and commensal ASVs) was positively associated with CREDI overall and CREDI language outcomes. The positive association with CREDI scores indicates that the presence of these microbes could contribute to neurodevelopmental benefits, potentially through pathways such as improved gut health, reduced inflammation, or enhanced nutrient metabolism. These findings highlight the importance of the gut microbiome in

early life development and suggest that fostering a beneficial microbial environment may have positive implications for neurodevelopmental outcomes.

Author Contributions

LEL designed and conducted the analysis, analyzed the data, and wrote the paper. MC, TCW, NZ and LMA advised LEL throughout the entire research process. LEL had primary responsibility for final content. All authors have read and approved the final manuscript.

Funding

This work was supported by a grant from the Egg Nutrition Center to the Academy of Nutrition and Dietetics.

References

1. Mennella JA, Li Y, Bittinger K, et al. The macronutrient composition of infant formula produces differences in gut microbiota maturation that associate with weight gain velocity and weight status. *Nutrients*. 2022;14(6):1241. doi:10.3390/nu14061241
2. Subramanian S, Huq S, Yatsunenko T, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature*. 2014;510(7505):417-421. doi:10.1038/nature13421
3. Davis JC, Lewis ZT, Krishnan S, et al. Growth and morbidity of Gambian infants are influenced by maternal milk oligosaccharides and infant gut microbiota. *Sci Rep*. 2017;7:40466. doi:10.1038/srep40466
4. Aatsinki AK, Kataja EL, Munukka E, et al. Infant fecal microbiota composition and attention to emotional faces. *Emotion*. 2022;22(6):1159-1170. doi:10.1037/emo0000924
5. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol*. 2021;17(11):e1009442. doi:10.1371/journal.pcbi.1009442
6. Guo S, Huang K, Liu R, Sun J, Yin C. Regulation of gut microbiota through breast milk feeding benefits language and cognitive development of preterm toddlers. *Microorganisms*. 2023;11(4):866. doi:10.3390/microorganisms11040866
7. Willemsen Y, Ou Y, Belzer C, et al. A longitudinal study of the gut microbiota during the first three years of life: Links with problem behavior and executive functions at preschool age. *Dev Psychopathol*. 2024;36(4):2032-2048. doi:10.1017/s0954579423001402

8. Cho S, Samuel TM, Li T, et al. Interactions between Bifidobacterium and Bacteroides and human milk oligosaccharides and their associations with infant cognition. *Front Nutr*. 2023;10:1216327. doi:10.3389/fnut.2023.1216327

9. Slykerman RF, Neumann D, Underwood L, Hobbs M, Waldie KE. Age at first exposure to antibiotics and neurodevelopmental outcomes in childhood. *Psychopharmacology (Berl)*. 2023;240(5):1143-1150. doi:10.1007/s00213-023-06351-5

10. Falsaperla R, Sortino V, Gambilonghi F, Vitaliti G, Striano P. Human milk oligosaccharides and their pivotal role in gut-brain axis modulation and neurologic development: a narrative review to decipher the multifaceted interplay. *Nutrients*. 2024;16(17):3009. doi:10.3390/nu16173009

11. Odiase E, Frank DN, Young BE, et al. The gut microbiota differ in exclusively breastfed and formula-fed United States infants and are associated with growth status. *J Nutr*. 2023;153(9):2612-2621. doi:10.1016/j.tjnut.2023.07.009

12. Ertem IO, Dogan DG, Gok CG, et al. A guide for monitoring child development in low- and middle-income countries. *Pediatrics*. 2008;121(3):e581-e589. doi:10.1542/peds.2007-1771

13. World Health Organization, United Nations Children's Fund. *WHO Child Growth Standards and the Identification of Severe Acute Malnutrition in Infants and Children: A Joint Statement by the World Health Organization and the United Nations Children's Fund*. World Health Organization; 2009.

14. World Health Organization, United Nations Children's Fund. Water quality monitoring. Accessed November 20, 2020, https://washdata.org/monitoring/drinkingwater/water-quality-monitoring

15. Innovations for Poverty Action. Guatemala: poverty probability index. Accessed November 20, 2020, https://www.povertyindex.org/country/guatemala

16. Daelmans B, Mangasaryan N, Martines J, Saadeh R, Casanovas C, Arabi M. Strengthening actions to improve feeding of infants and young children 6 to 23 months of age: summary of a recent World Health Organization/UNICEF technical meeting, Geneva, 6-9 October 2008. *Food Nutr Bull*. 2009;30(2 suppl):S236-S238. doi:10.1177/15648265090302s208

17. Rothenberg SE, Chen Q, Shen J, et al. Neurodevelopment correlates with gut microbiota in a cross-sectional analysis of children at 3 years of age in rural China. *Sci Rep*. 2021;11(1):7384. doi:10.1038/s41598-021-86761-7

18. Borewicz K, Gu F, Saccenti E, et al. The association between breastmilk oligosaccharides and faecal microbiota in healthy breastfed infants at two, six, and twelve weeks of age. *Sci Rep*. 2020;10(1):4270. doi:10.1038/s41598-020-61024-z

19. Alcazar CG, Paes VM, Shao Y, et al. The association between early-life gut microbiota and childhood respiratory diseases: a systematic review. *Lancet Microbe*. 2022;3(11):e867-e880. doi:10.1016/s2666-5247(22)00184-7

20. Tang M, Frank DN, Hendricks AE, et al. Iron in micronutrient powder promotes an unfavorable gut microbiota in Kenyan infants. *Nutrients*. 2017;9(7):776. doi:10.3390/nu9070776

21. Paganini D, Zimmermann MB. The effects of iron fortification and supplementation on the gut microbiome and diarrhea in infants and children: a review. *Am J Clin Nutr*. 2017;106(Suppl 6):1688s-1693s. doi:10.3945/ajcn.117.156067

22. McMillen S, Thomas S, Liang E, Nonnecke EB, Slupsky C, Lönnerdal B. Gut microbiome alterations following postnatal iron supplementation depend on iron form and persist into adulthood. *Nutrients*. 2022;14(3):412. doi:10.3390/nu14030412

23. Cowan CSM, Dinan TG, Cryan JF. Annual Research Review: Critical windows - the microbiota-gut-brain axis in neurocognitive development. *J Child Psychol Psychiatry*. 2020;61(3):353-371. doi:10.1111/jcpp.13156

24. Carabotti M, Scirocco A, Maselli MA, Severi C. The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Ann Gastroenterol*. 2015;28(2):203-209.

Tables

Table 4.1. Infant CREDI and growth statistics by treatment group.

	Egg (N=106)	Control (N=94)	Total (N=200)	P Value
CREDI Score Overall				0.8161
Mean (SD)	47.424 (1.173)	47.465 (1.313)	47.443 (1.237)	
Missing	0	1	1	
CREDI Socio-Emotional Score				0.8861
Mean (SD)	48.814 (0.792)	48.798 (0.859)	48.806 (0.822)	
Missing	0	1	1	
CREDI Motor Score				0.7581
Mean (SD)	48.008 (0.785)	48.043 (0.811)	48.024 (0.795)	
Missing	0	1	1	
CREDI Language Score				0.6751
Mean (SD)	48.751 (0.694)	48.796 (0.806)	48.772 (0.747)	
Missing	0	1	1	
CREDI Cognitive Score				0.9111
Mean (SD)	48.935 (0.783)	48.921 (0.898)	48.928 (0.836)	
Missing	0	1	1	
Weight-for-Age Z-Score				0.4681
Mean (SD)	-0.963 (0.733)	-0.876 (0.931)	-0.922 (0.830)	
Missing	2	4	6	
Length-for-Age Z-Score				0.4021
Mean (SD)	-1.983 (1.024)	-1.855 (1.129)	-1.923 (1.073)	
Missing	0	1	1	
Weight-for-Length Z-Score				0.6891
Mean (SD)	0.035 (0.640)	0.077 (0.829)	0.055 (0.732)	
Missing	2	4	6	
Head-Circumference Z-Score				0.2761
Mean (SD)	-0.779 (0.829)	-0.913 (0.897)	-0.842 (0.862)	
Missing	0	1	1	
BMI Z-Score				0.8041
Mean (SD)	0.292 (0.657)	0.319 (0.838)	0.304 (0.745)	

	Egg (N=106)	Control (N=94)	Total (N=200)	P Value
Missing	2	4	6	
Stunting				0.5702
No	52 (49.1%)	50 (53.8%)	102 (51.3%)	
Yes	54 (50.9%)	43 (46.2%)	97 (48.7%)	
Missing	0	1	1	
Wasting				0.4642
0	104 (100.0%)	89 (98.9%)	193 (99.5%)	
1	0 (0.0%)	1 (1.1%)	1 (0.5%)	
Missing	2	4	6	
Milk type most frequently consumed*				0.1312
Non-breastmilk	14 (13.9%)	20 (22.7%)	34 (18.0%)	
Breastmilk	87 (86.1%)	68 (77.3%)	155 (82.0%)	
Missing	5	6	11	
Age (Months)				0.8721
Mean (SD)	12.642 (1.044)	12.667 (1.155)	12.653 (1.094)	
Missing	0	1	1	
Sex				0.7792
Boy	53 (50.0%)	49 (52.1%)	102 (51.0%)	
Girl	53 (50.0%)	45 (47.9%)	98 (49.0%)	
National Poverty Likelihood Score				0.3491
Mean (SD)	23.501 (29.351)	27.487 (30.611)	25.374 (29.940)	
Water Source**				0.0092
Bagged/sachet water	20 (18.9%)	23 (24.5%)	43 (21.5%)	
Borehole/tubewell	15 (14.2%)	9 (9.6%)	24 (12.0%)	
Other	12 (11.3%)	17 (18.1%)	29 (14.5%)	
Piped water (inside home or land)	53 (50.0%)	29 (30.9%)	82 (41.0%)	
Protected dug well	6 (5.7%)	16 (17.0%)	22 (11.0%)	
Low Birthweight (less than 2500g)				0.5802
No	77 (78.6%)	69 (82.1%)	146 (80.2%)	
Yes	21 (21.4%)	15 (17.9%)	36 (19.8%)	
Missing	8	10	18	
Antibiotics Ever				0.5182
No	25 (23.6%)	26 (28.0%)	51 (25.6%)	
Yes	81 (76.4%)	67 (72.0%)	148 (74.4%)	

	Egg (N=106)	Control (N=94)	Total (N=200)	P Value
Missing	0	1	1	
Antibiotics last 14 days				0.4102
No	83 (78.3%)	68 (72.3%)	151 (75.5%)	
Yes	23 (21.7%)	26 (27.7%)	49 (24.5%)	
Diarrhea last 14 days				0.5702
No	56 (52.8%)	54 (57.4%)	110 (55.0%)	
Yes	50 (47.2%)	40 (42.6%)	90 (45.0%)	
Animal Contact***				0.0262
No	45 (42.5%)	25 (26.6%)	70 (35.0%)	
Yes	61 (57.5%)	69 (73.4%)	130 (65.0%)	

Linear Model ANOVA
Fisher's Exact Test for Count Data
*Includes fluid animal milk, powdered animal milk, and infant formula.
**Includes unprotected dug well, protected spring, unprotected spring, tanker truck, surface water [pond, river, lake], other person, other.
***Includes regular direct contact with animals, including dog, cat, chicken, turkey, pig, cow, sheep, goat, donkeys, and other.

Table 4.2. Significant Maaslin2 associations between co-abundance groupings and CREDI scores.

Feature	Model*	Coefficient	N.not.0**	P Value	Q Value
CAG6	CREDI Overall	1.22	96	0.00280	0.0925
CAG28	CREDI Overall	0.887	91	0.00366	0.111
CAG6	CREDI Cognitive	1.37	96	0.000636	0.0258
CAG28	CREDI Cognitive	0.786	91	0.00897	0.215
CAG6	CREDI Motor	1.10	96	0.00990	0.233
CAG6	CREDI Language	1.17	96	0.00264	0.0997
CAG6	CREDI Socioemotional	1.26	96	0.00236	0.0745

*Variables are from separate models.
**Number of participants with non-zero feature.

Table 4.3. Significant MaAslin2 associations between co-abundance groupings and CREDI scores stratified by milk type.

Feature	Model*	Coefficient	N.not.0**	P Value	Q Value	Milk Type
CAG6	CREDI Overall	1.30	77	0.00321	0.769	Breastmilk
CAG28	CREDI Overall	0.856	66	0.00941	0.801	Breastmilk
CAG6	CREDI Overall	1.20	19	0.288	0.955	Non-breastmilk
CAG18	CREDI Overall	2.22	30	0.000837	0.216	Non-breastmilk
CAG28	CREDI Overall	0.989	25	0.235	0.955	Non-breastmilk
CAG6	CREDI Cognitive	1.42	77	0.00190	0.763	Breastmilk
CAG28	CREDI Cognitive	0.915	66	0.00711	0.763	Breastmilk
CAG6	CREDI Cognitive	0.881	19	0.346	0.976	Non-breastmilk
CAG28	CREDI Cognitive	0.255	25	0.717	0.990	Non-breastmilk
CAG6	CREDI Motor	1.33	77	0.00348	0.808	Breastmilk
CAG6	CREDI Motor	0.597	19	0.662	0.994	Non-breastmilk
CAG6	CREDI Language	1.14	77	0.00979	0.840	Breastmilk
CAG6	CREDI Language	1.81	19	0.0276	0.911	Non-breastmilk
CAG18	CREDI Language	0.426	100	0.280	0.999	Breastmilk
CAG18	CREDI Language	1.95	30	4.57E-05	0.0226	Non-breastmilk
CAG6	CREDI Socio-emotional	1.57	77	0.000653	0.323	Breastmilk
CAG6	CREDI Socio-emotional	-0.182	19	0.855	0.989	Non-breastmilk

*Variables are from separate models.
**Number of participants with non-zero feature.

Table 4.4. Significant MaAslin2 associations between co-abundance groupings and growth scores.

Feature	Model*	Coefficient	N.not.0**	P Value	Q Value
CAG20	Z-Head-Circumference	0.597	50	0.0130	0.231
CAG7	Z-Head-Circumference	0.574	100	0.0141	0.238
CAG10	Z-Weight-for-Age	0.715	71	0.0117	0.226
CAG15	Stunting (Yes)	1.39	79	0.00312	0.0802

*Variables are from separate models.
**Number of participants with non-zero feature.

Table 4.5. Significant MaAslin2 associations between co-abundance groupings and growth scores from milk type stratified models.

Feature	Model*	Coefficient	N.not.0**	P Value	Q Value	Milk type
CAG10	Z-Weight-for-Age	0.447	54	0.132	0.889	Breastmilk
CAG10	Z-Weight-for-Age	0.613	17	0.557	0.933	Non-breastmilk
CAG7	Z-Head-Circumference	0.603	76	0.0222	0.622	Breastmilk
CAG20	Z-Head-Circumference	0.475	34	0.0601	0.737	Breastmilk
CAG7	Z-Head-Circumference	0.991	24	0.320	0.989	Non-breastmilk
CAG20	Z-Head-Circumference	0.658	16	0.414	0.989	Non-breastmilk
CAG15	Stunting (Yes)	1.46	65	0.00657	0.579	Breastmilk
CAG15	Stunting (Yes)	1.17	14	0.299	0.897	Non-breastmilk

*Variables are from separate models.
**Number of participants with non-zero feature.

CHAPTER 5:
The Association of Early Child Development and Growth Outcomes with Localized Gastrointestinal Inflammation in 11-16-Month-Old Guatemalan Infants

Abstract

Objective: The main objective of this study was to determine whether endline localized GI inflammation (measured with fecal calprotectin) was associated with differences in endline fecal microbiome composition, early child development (ECD) scores, growth anthropometrics, and stunting.

Methods: The Saqmolo' microbiome sub-study enrolled 200 rural Guatemalan infant participants aged 11-16 months at endline. Fecal calprotectin concentration (ng/ml) was quantified using ELISA assay kits. ECD was assessed using the Caregiver Reported Early Child Development Instruments (CREDI) and the Guide for Monitoring Child Development. Anthropometric measures included World Health Organization defined Z-scores and stunting. Gut microbiome composition was quantified using 16S rRNA sequencing and analyzed using QIIME2. MaAsLin2 evaluated associations between fecal calprotectin with amplicon sequencing variants (ASVs) and guild-based co-abundance groupings. Individual linear regression models assessed the relationship of fecal calprotectin with ECD, Z-scores, and stunting incidence.

Results: Mean fecal calprotectin was 352 (346) ng/ml in 199 infant participants. Two ASVs associated with fecal calprotectin in MaAsLin2 models were also present in co-abundance group associations, including higher relative abundance of *Dialister* (ASV02ZY, MaAsLin2 coefficient = 0.323, q = 0.149) and lower relative abundance of *Clostridium* (ASV081D, MaAsLin2 coefficient = -0.254, q = 0.246). Fecal calprotectin was associated with higher abundance of co-abundance group 8 (MaAsLin2 coefficient = 0.578, q value = 0.105) and co-abundance group 28 (MaAsLin2 coefficient = 0.588, q value = 0.245), but lower abundance of co-abundance group 1 (MaAsLin2 coefficient = -0.392, q value = 0.136). No statistically significant associations were

observed between fecal calprotectin and CREDI scores, Z-scores, or stunting (all p-values > 0.05).

Conclusion: Localized GI inflammation does not have a significant impact on early childhood development or growth outcomes in this cohort. However, it is linked with differences in relative abundance of gut bacteria and microbial communities.

Introduction

Infants and children in low- and middle-income countries (LMIC) are profoundly affected by suboptimal development outcomes.[1,2] According to the United Nations, the majority of Guatemalan infants do not meet adequate dietary diversity or complementary feeding practices.[3-6] Guatemalan neonates have high prevalence of low birthweight (15%) and children ages 18-23 months have high prevalence of chronic malnutrition (58%), iron-deficiency anemia (33%), stunting (43%) and severe stunting (23%).[1,7,8] Indigenous Mayan children in Guatemala are most severely affected by malnutrition, with twice the rate of stunting in indigenous children compared to non-indigenous children.[9,10] Additionally, environmental factors, breast-feeding and complementary feeding practices, maternal malnutrition, and limited nutrition resources lead to higher risk of malnutrition, gut dysbiosis, and impaired cognitive and growth development in LMIC infants. These populations also have a high risk for early antibiotic exposure, given their less-than-optimal living conditions and food and healthcare access.

The immune system plays an important role in early child development, and it is thought to impact gut microbiome maturation, particularly at weaning.[11] During weaning there are significant fluctuations in microbiota composition and an associated inflammatory response, known as the weaning reaction. Researchers report a greater risk for inflammatory and immune disorders later in life, when the weaning reaction or gut microbiome development is disrupted,[12-14] demonstrating the importance of healthy gut microbiome maturation. Gut dysbiosis is also known to upregulate inflammatory responses in the gut, instigating bi-directional inflammatory responses in other areas of the body such as the liver,[15] lungs,[16,17] and brain.[18] Prolonged inflammation is known to impact intestinal barrier function and tight junctions, reduce epithelial surface area, and contribute to villous atrophy.[19] Infections, gastrointestinal infections and

dysbiosis contributes to intestinal inflammation. Furthermore, intestinal inflammation reduces nutrient absorption, further impairing the immune system and absorption of nutrients essential for growth and development.[19]

Infections, intestinal inflammation, and immune system activation is common in LMIC settings. Under these conditions, the body prioritizes survival (protecting against pathogenic invaders) over growth.[19] Current research considers immune system activation to be a biomarker of undernutrition, rather than an inherent aspect of undernutrition.[19,20]However, the impact and timing of localized GI inflammation, and the transition from localized GI inflammation to systemic inflammation, is less well characterized. It is currently unknown how early localized infant GI inflammation occurs in a rural LMIC population, how this impacts gut microbiome composition, and subsequent impacts on early child development outcomes. In this study, we investigated the association of localized GI inflammation with both gut microbiome composition and early child development outcomes in 11-16-month-old infants enrolled in the Saqmolo' Project randomized controlled trial.

The main objective of this study was to determine whether endline localized GI inflammation (measured with fecal calprotectin) was associated with differences in 1) endline fecal microbiome composition; 2) early child development scores; 3) growth anthropometrics (z-scores for weight-for-age, length-for-age, weight-for-length, and head-circumference-for-age); and 4) stunting and wasting incidence. We hypothesized lower endline fecal calprotectin would be associated with 1) a more mature fecal microbiome composition; 2) higher early child development scores; 3) more optimal growth anthropometric scores; and 4) lower incidence of stunting and wasting. Gut dysbiosis is a known risk factor for systemic inflammation, but the impact and timing of localized GI inflammation on gut microbiome composition and co-

development has not been investigated in a dietary context in LMIC infants. This study will elucidate the impact of localized GI inflammation on infant gut microbiome composition and host development.

Methods and Materials

Study Population & Assessments

The study population details are described in Chapter 3: subsection *Study Population*. Briefly, the Saqmolo' Project enrolled 1,200 rural Guatemalan infants aged 6 to 9 months, with 200 infants participating in the microbiome sub-study (age 11-16 months at endline). All ECD, demographic, anthropometric, and dietary questionnaires and assessments are previously described in Chapter 4 subsection: *ECD, Demographic, & Anthropometric Assessments* and Chapter 4 subsection: *Dietary Questionnaires*.

Fecal Calprotectin

Fecal calprotectin is a validated marker of localized gastrointestinal inflammation (GI).[21] It is a calcium and zinc binding protein that is secreted by neutrophils, eosinophils, and monocytes in response to inflammatory processes and can act as a defense against bacterial and fungal pathogens due to its zinc-binding property.[22-25] Fecal calprotectin is a non-invasive marker as it is eliminated in the feces and elevated levels indicate on-going intestinal inflammation. Of note, in healthy infants, calprotectin level is age-associated where levels are higher in infants less than 6 months and levels decrease with increasing age, with infants age 9-12 months and 12-18 months having median fecal calprotectin concentrations of 96.1 μg/g and 104.2 μg/g.[26]

Fecal Procedures, Sequencing, and Gut Microbiome Analysis

Fecal calprotectin levels were analyzed at Colorado State University, College of Health, and Human Sciences Intestinal Health Laboratory. Fecal calprotectin levels were assessed with ELISA assay kits (Eagle Biosciences, Amherst, NH). Fecal collection, transportation procedures, and gut microbiome sequencing have been previously described in Chapter 3 subheading: *Fecal Procedures and Gut Microbiome Quantification*. QIIME2 analyses are described in Chapter 3 subheading: *Quantitative Insights into Microbial Ecology version 2 (QIIME2)*. Creation of co-abundance groupings is described in Chapter 3 subheading: *Guild Based Co-Abundance Groupings.*

Multivariate Associations with Linear Models (MaASLin2): Individual ASV Associations

MaAsLin2[27] is specifically designed for microbiome data, providing key advantages over traditional statistical methods by accounting for its high-dimensional, sparse, and compositional nature. It enhances data normality through transformations (e.g., log transformation), filters out low abundance features to minimize noise, and accommodates multiple covariates and random effects to control for confounding variables. Additionally, it applies multiple testing corrections, such as the Benjamini-Hochberg FDR, to reduce false positives, ensuring more robust and reliable results in large-scale microbiome analyses. MaASLin2 was used to test the association between relative abundance of the top 10% most prevalent ASV present in fecal samples with fecal calprotectin. Abundance of the ASVs were normalized by dividing the raw count of each ASV in a sample by the total number of reads in that sample. Model equations including adjusted covariates are listed below in **Equation 5.1**. MaASLin2 coefficients were reported with q-values (FDR adjusted p-values) ≤ 0.25 as significant (the MaAsLin2 standard threshold).

> ***Equation 5.1:*** *ASVs ~ fecal calprotectin + age + sex + milk type + antibiotics ever + water source + antibiotics 14 days + diarrhea 14 days + treatment*

Co-Abundance Groupings: MaAsLin2

The co-abundance groupings were used in downstream analysis using MaAsLin2 to determine the association between each co-abundance group with fecal calprotectin concentration. The abundance of each group was normalized by dividing the raw count of each group by the total number of reads. In MaAsLin2 settings, the minimum prevalence was set to zero since each group inherently has an abundance greater than zero. The co-abundance group models were adjusted for the same covariates as Equation 5.1 described in the previous section. Co-abundance groups are referred to as CAGs in the results.

Statistical Analysis

All analyses were conducted using QIIME2,[28] R statistical software (*v*4.2.0)[29] and Tufts University High Performance Computing Cluster. High and low groups were created based on values above and below the median for fecal calprotectin concentration. These high and low groups were used for QIIME2 statistical comparisons. Spearman rank correlation tests were used to analyze the association between fecal calprotectin level and fecal microbiome alpha diversity measures and PERMONVA tests for beta-diversity measures using QIIME2. QIIME2 statistical tests were reported with q-values (FDR adjusted p-values) ≤ 0.05 as significant (the QIIME2 standard threshold).

Individual linear regression models tested the association between fecal calprotectin level and ECD scores (CREDI overall, CREDI cognitive, CREDI language, CREDI motor, and CREDI socioemotional), growth anthropometrics (z-scores for weight-for-age, length-for-age, weight-for-length, head circumference-for-age) and stunting incidence. In all models, all infants categorized as primarily having consumed breastmilk were included in the breastmilk category, while all infants reported primarily having consumed either fluid animal milk, powdered animal milk, or infant formula based on the WHO feeding indicators questionnaire were collapsed into

the non-breastmilk category. Play activities score was collapsed into 3 categories (0-1, 2-3, 4-6) based on the distribution, combining scores 0-6 to preserve statistical power. Water source was collapsed into 5 categories (Bagged/sachet, Borehole/tubewell, Piped water (inside home or land), Protected dug well, and Other), with "Other" containing unprotected dug well, protected spring, unprotected spring, tanker truck, surface water (pond, river, lake), other person, and other categories. Models were adjusted for the covariates listed in **Equations 5.2 to 5.4** below.

> ***Equation 5.2****: CREDI scores ~ fecal calprotectin + age + sex + milk type + antibiotics ever + water source + stunting + antibiotics 14 days + diarrhea 14 days + treatment + poverty score + play activities*
>
> ***Equation 5.3****: Z-scores ~ fecal calprotectin + age + sex + milk type + antibiotics ever + water source + low birthweight + antibiotics 14 days + diarrhea 14 days + treatment + poverty score*
>
> ***Equation 5.4****: Stunting ~ fecal calprotectin + age + sex + milk type + antibiotics ever + water source + low birthweight + antibiotics 14 days + diarrhea 14 days + treatment + poverty score*

For linear regression models, regression coefficients were reported with p-values ($p \leq 0.05$ as significant), 95% confidence intervals, standard error (SE) and R^2 values. Model fit was assessed with R^2 and adjusted R^2. All regression assumptions were evaluated. Linearity between fecal calprotectin and the outcome were assessed with scatterplots using ggplot2. Normality of residuals was assessed with QQ plots and homoscedasticity was evaluated using residuals plots.

Results

Study Population and Fecal Calprotectin Statistics

The main fecal calprotectin analyses included 199 infant participants. Mean fecal calprotectin was 352 (346) ng/ml. The percentage of female infants in the population was 49.0%. Mean age at endline was 12.7 months. The percentage of breastfed infants was 82.0%. Mean fecal calprotectin stratified by treatment and milk type subgroups were egg + breastmilk = 404

(386) ng/ml, control + breastmilk group = 395 (354) ng/ml, egg + non-breastmilk group = 156 (86.8) ng/ml, control + non-breastmilk group = 244 (208) ng/ml).

QIIME2: Alpha and Beta Diversity

A total of 1,829 ASVs were identified. Kruskal-Wallis pairwise testing showed no significant differences between high and low fecal calprotectin groups using Shannon diversity indices (high fecal calprotectin group median Shannon index = 3.27, p-value = 0.626). PERMANOVA testing showed beta diversity was significantly different between high and low fecal calprotectin groups using weighted Unifrac metrics (q-value = 0.006). However, there was no apparent clustering of the groups in the principal coordinates analysis (PCoA) plot (**Figure 5.1**).

MaAsLin2: ASV Associations with Fecal Calprotectin

Fecal calprotectin was associated with a higher relative abundance of several ASVs, including *Butyricicoccus* (ASV0YGD, MaAsLin2 coefficient = 0.297, q = 0.0579), *Faecalibacterium* (ASV006D, MaAsLin2 coefficient = 0.388, q = 0.0800), *Dialister* (ASV02ZY, MaAsLin2 coefficient = 0.323, q = 0.149), *Collinsella* (ASV004W, MaAsLin2 coefficient = 0.465, q = 0.217), and *Ruminococcus* (ASV026G, MaAsLin2 coefficient = 0.262, q = 0.246). Conversely, fecal calprotectin was associated with a lower relative abundance of two *Clostridium* ASVs (ASV04LS, MaAsLin2 coefficient = -0.623, q = 0.141; ASV081D, MaAsLin2 coefficient = -0.254, q = 0.246) and *Streptococcus* (ASV004F, MaAsLin2 coefficient = -0.330, q = 0.173). The full results are displayed in **Table 5.1**.

MaAsLin2: Co-Abundance Grouping Associations with Fecal Calprotectin

Fecal calprotectin was associated with higher relative abundance of CAG 8 (MaAsLin2 coefficient = 0.578, q value = 0.105) and CAG 28 (MaAsLin2 coefficient = 0.588, q value = 0.245). Fecal calprotectin was associated with lower relative abundance of CAG 1 (MaAsLin2

coefficient = -0.392, q value = 0.136) (see full results in **Table 5.2**). When the fecal calprotectin co-abundance grouping models were rerun stratified by milk type, there were no significant associations with the co-abundance groups for either milk type (breastmilk or non-breastmilk). For the non-breastmilk stratified models, the direction of association was the same for CAG 8 and CAG 28. However, the direction of association for CAG 1 was the opposite (MaAsLin2 coefficient = 0.258, q value = 0.979). For the breastmilk stratified models, the direction of association was the same for all CAGs and effect sizes were comparable (**Table 5.3**).

Linear Regression: Fecal calprotectin Associations with CREDI and Growth Scores

Across all unadjusted and adjusted models, no statistically significant associations were observed between fecal calprotectin and CREDI or growth outcomes (all p-values > 0.05). Results for all models are presented in **Supplementary Tables S5.1 to S5.7** (in **Appendix C**). Model scatterplots are presented in **Supplementary Figures S5.1 to S5.11**, Model QQ plots are presented in **Supplementary Figures S5.12 to S5.21**, and residuals plots are presented in **Supplementary Figures S5.22 to S5.31**. No transformations or robust standard errors were applied. A borderline significant effect was observed between fecal calprotectin concentration and both Z-BMI (**Table 5.4**) and Z-weight-for-length (

Table 5.5), where each ng/ml change in fecal calprotectin was associated with 3.20 x -04 higher Z-BMI score (95% CI: -1.37 x -05, 6.54 x -04, p = 0.0600), and each ng/ml change in fecal calprotectin was associated with 3.03 x -04 higher Z-weight-for-length score (95% CI: -1.94 x -05, 6.26 x -04, p = 0.0652). The model p-values for CREDI overall score, language, motor, cognitive, and socioemotional were statistically significant (all models = $p < 0.05$). Play activities, stunting, and age covariates were significantly associated with all CREDI domains with direction and magnitude of the effect comparable across all models. For example, higher play activity score (scores 4-6) was associated with higher CREDI overall scores ($\beta = 0.860$, 95% CI: 0.423, 1.30, p = 0.0001), stunting was associated with lower CREDI overall scores ($\beta = -0.514$, 95% CI: -0.820, -0.209, p = 0.0011) compared to non-stunted infants, and each unit change in age (months) was associated with higher CREDI overall scores ($\beta = 0.569$, 95% CI: 0.432, 0.706, p = 6.35 x -14).

Discussion

This study aimed to determine the association between localized GI inflammation (measured with fecal calprotectin) with differences in fecal microbiome composition, ECD scores and growth outcomes. In summary, fecal calprotectin has clear associations with gut microbiome composition and may be linked to weight and growth metrics. However, CREDI scores appear to be more strongly influenced by other factors such as play activities and stunting. While there appear to be statistical differences in beta diversity between high and low fecal calprotectin groups, the lack of apparent clustering in the PCoA plot as well as the lack of species richness and evenness (alpha diversity) suggests localized GI inflammation is not a major driving factor in overall diversity. However, ASV and co-abundance group associations suggest

localized GI inflammation has a stronger link with specific bacteria and bacterial groupings than diversity.

Fecal calprotectin is a marker of localized GI inflammation. It stands to reason gut bacteria higher in abundance associated with fecal calprotectin are likely to be pro-inflammatory or thrive in comparatively more aerobic environments, while species lower in abundance are likely to be anti-inflammatory in relation to the GI system. The two ASVs associated with fecal calprotectin, which were also present in significantly associated co-abundance groupings, belonged to genera Dialister and Clostridium. Fecal calprotectin concentration was associated with a greater abundance of the Dialister ASV and CAG 28 (the co-abundance group containing Dialister). Limited information is available regarding Dialister in the infant gut; however, there is some evidence to suggest Dialister bacteria are linked to inflammatory bowel disease,[30,31] supporting the results that it may act as a pro-inflammatory bacteria in the gut.

Fecal calprotectin concentration was also associated with lower relative abundance of the Clostridium ASV and CAG 1 (the co-abundance group containing Clostridium). In mice, genus Clostridium contributes to the induction of colonic regulatory T cells (Tregs), which are essential for maintaining gastrointestinal homeostasis.[32] Higher abundance of Clostridium species has been linked to the generation, homing, and functional maturation of Tregs in the colon, thereby supporting immune tolerance and preventing excessive inflammation.[33] However, in one study, Clostridium species, particularly from Clostridium cluster XIV, were found to be more abundant in infants experiencing health complications at birth compared to healthy infants.[34] While some Clostridium species may have a beneficial presence, other species appear to be more harmful when they are elevated in certain conditions, particularly in the context of health complications, formula feeding, and altered infant growth outcomes.[34]

In this study, fecal calprotectin was linked to increased abundance of CAG 8 and CAG 28 and lower abundance of CAG 1. When MaAsLin2 co-abundance group models were stratified by milk type, no significant associations were found. Though trends largely remained consistent in both the breastmilk and non-breastmilk models, CAG 1 reversed direction in the non-breastmilk group. This suggests that certain bacterial communities may be linked to localized GI inflammation (CAG 8 and CAG 28 bacterial groupings) but may also be influenced by dietary factors such as breastfeeding. Breastmilk may help stabilize microbiome differences via immunomodulatory effects or by promoting a more resilient gut microbiome. Conversely, non-breastmilk consumption may introduce external factors that influence microbiome composition and contribute to GI inflammation. The ASVs present in CAGs 1, 8, and 28 were mixed between genera typically considered beneficial, commensal, and potentially harmful or pathogenic. This may suggest that the collection of bacteria in the gut is more influential than individual bacteria.

GI inflammation was not associated with CREDI scores, Z-scores, or stunting. CREDI scores appear to be more strongly associated with other variables such as play activities in the home and stunted growth. Borderline significance was seen for Z-BMI and Z-weight-for-length models suggesting a weak relationship between GI inflammation and both weight and length metrics. However, the effect sizes were too small to be meaningful or clinically relevant. As fecal calprotectin is a marker of localized GI inflammation, it is biologically plausible that it could contribute to altered growth patterns, but the relationship remains unclear. Some studies suggest that chronic low-grade inflammation impairs nutrient absorption and stunts growth, while others propose that a mild inflammatory response in the GI system may reflect healthy immune system activation in response to complimentary food introduction.

Strengths, Limitations, and Future Directions

Although these results suggest a possible link between fecal calprotectin and growth outcomes, the small effect sizes and borderline p-values indicate that these findings should be interpreted with caution. Although our approach leveraged high-resolution analysis by using ASVs as the unit of microbial composition, 16S rRNA microbiome data are limited and cannot provide species level resolution for functional analysis. The models included and stratified by milk type in separate models – milk type appears to influence the relationship between GI inflammation and microbiome composition, though the non-breastmilk group sample size was small. The non-breastmilk group encompassed various milk types, including fluid animal milk, powdered animal milk, and infant formula, reflecting real-world dietary patterns in an LMIC setting. However, each of these non-breastmilk sources may have distinct effects on the gut microbiome and result in unique microbial guilds.

Future research should consider longitudinal analyses to determine whether fecal calprotectin concentration predicts changes in growth over time rather than relying on cross-sectional data. Moreover, studies with larger sample sizes would help clarify whether the associations can reach statistical significance and clinically relevant effect sizes. Recruiting larger sample sizes within non-breastmilk subgroups is also needed to improve statistical power and clarify the modifying role of milk type in the relationship between GI inflammation and microbiome composition. Biomarkers of systemic inflammation should be examined alongside localized GI inflammation to determine whether microbial species and their functional profiles contribute to systemic inflammation in ways that impact growth and early child development.

Conclusions

Localized GI inflammation does not appear to have a significant impact on early childhood development or growth outcomes in this cohort. However, it is linked to gut

microbiome composition, with potential differences influenced by breastmilk versus non-breastmilk consumption. Rather than individual bacterial species, microbial communities associated with fecal calprotectin may play a more pivotal role in shaping this relationship.

Author Contributions

LEL designed and conducted the analysis, analyzed the data, and wrote the paper. MC, TCW, NZ and LMA advised LEL throughout the entire research process. LEL had primary responsibility for final content. All authors have read and approved the final manuscript.

Funding

This work was supported by a grant from the Egg Nutrition Center to the Academy of Nutrition and Dietetics.

References

1. Black RE, Victora CG, Walker SP, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet*. 2013;382(9890):427-451. doi:10.1016/s0140-6736(13)60937-x

2. Lu C, Black MM, Richter LM. Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional, and country level. *Lancet Glob Health*. 2016;4(12):e916-e922. doi:10.1016/s2214-109x(16)30266-2

3. Food and Nutritional Assistance III Project. *Development of Evidence-Based Dietary Recommendations for Children, Pregnant Women, and Lactating Women Living in the Western Highlands in Guatemala*. Food and Nutritional Assistance III Project; 2013.

4. Brown K, Henretty N, Chary A, et al. Mixed-methods study identifies key strategies for improving infant and young child feeding practices in a highly stunted rural indigenous population in Guatemala. *Matern Child Nutr*. 2016;12(2):262-277. doi:10.1111/mcn.12141

5. Webb MF, Chary AN, De Vries TT, et al. Exploring mechanisms of food insecurity in indigenous agricultural communities in Guatemala: a mixed methods study. *BMC Nutr*. 2016;2(1):55. doi:10.1186/s40795-016-0091-5

6. Food and Nutrition Technical Assistance III Project. Validation of food-based recommendations developed using Optifood for groups at nutritional risk in the Western Highlands of Guatemala. Accessed August 15, 2022, https://www.fantaproject.org/countries/guatemala/optifood-report-2014

7. Ministerio de Salud Pública y Asistencia Social MSPAS, Guatemala Instituto Nacional de Estadística INE/Guatemala, Secretaría de Planificación y Programación del la Presidencia Segeplán/Guatemala, ICF International. Encuesta nacional de salud materno infantil 2014-2015: informe final. Accessed August 12, 2022, https://www.dhsprogram.com/publications/publication-fr318-dhs-final-reports.cfm

8. World Health Organization, United Nations Children's Fund. *WHO Child Growth Standards and the Identification of Severe Acute Malnutrition in Infants and Children: A Joint Statement by the World Health Organization and the United Nations Children's Fund*. World Health Organization; 2009.

9. World Bank. International Bank for Reconstruction and Development and International Finance Corporation Country Partnership Strategy for the Republic of Guatemala. Report No. 44772-GT. Accessed December 6, 2022, http://web.worldbank.org/archive/website01270/WEB/IMAGES/GT_CPS_0.PDF

10. World Bank. Improving Health and Nutrition of Mothers and Young Children in Guatemala Accessed December 6, 2022, https://documents.worldbank.org/en/publication/documents-reports/documentdetail/699671468032739491/Improving-health-and-nutrition-of-mothers-and-young-children-in-Guatemala

11. Al Nabhani Z, Dulauroy S, Marques R, et al. A weaning reaction to microbiota is required for resistance to immunopathologies in the adult. *Immunity*. 2019;50(5):1276-1288.e1275. doi:10.1016/j.immuni.2019.02.014

12. Cowan CSM, Dinan TG, Cryan JF. Annual Research Review: Critical windows - the microbiota-gut-brain axis in neurocognitive development. *J Child Psychol Psychiatry*. 2020;61(3):353-371. doi:10.1111/jcpp.13156

13. Moore RE, Townsend SD. Temporal development of the infant gut microbiome. *Open Biol*. 2019;9(9):190128. doi:10.1098/rsob.190128

14. Johnson CC, Ownby DR. The infant gut bacterial microbiota and risk of pediatric asthma and allergic diseases. *Transl Res*. 2017;179:60-70. doi:10.1016/j.trsl.2016.06.010

15. Tripathi A, Debelius J, Brenner DA, et al. The gut-liver axis and the intersection with the microbiome. *Nat Rev Gastroenterol Hepatol*. 2018;15(7):397-411. doi:10.1038/s41575-018-0011-z

16. Anand S, Mande SS. Diet, microbiota and gut-lung connection. *Front Microbiol*. 2018;9:2147. doi:10.3389/fmicb.2018.02147

17. Bingula R, Filaire M, Radosevic-Robin N, et al. Desired turbulence? Gut-lung axis, immunity, and lung cancer. *J Oncol*. 2017;2017:5035371. doi:10.1155/2017/5035371

18. Lu J, Lu L, Yu Y, Cluette-Brown J, Martin CR, Claud EC. Effects of intestinal microbiota on brain development in humanized gnotobiotic mice. *Sci Rep*. 2018;8(1):5443. doi:10.1038/s41598-018-23692-w

19. Bourke CD, Jones KDJ, Prendergast AJ. Current understanding of innate immune cell dysfunction in childhood undernutrition. *Front Immunol*. 2019;10:1728. doi:10.3389/fimmu.2019.01728

20. Bourke CD, Berkley JA, Prendergast AJ. Immune dysfunction as a cause and consequence of malnutrition. *Trends Immunol*. 2016;37(6):386-398. doi:10.1016/j.it.2016.04.003

21. Bonnín Tomàs A, Vila Vidal M, Rosell Camps A. [Fecal calprotectin as a biomarker to distinguish between organic and functional gastrointestinal disease]. *Rev Esp Enferm Dig*. 2007;99(12):689-693. Calprotectina fecal como marcador diferencial entre patología gastrointestinal orgánica y funcional. doi:10.4321/s1130-01082007001200002

22. Steinbakk M, Naess-Andresen CF, Lingaas E, Dale I, Brandtzaeg P, Fagerhol MK. Antimicrobial actions of calcium binding leucocyte L1 protein, calprotectin. *Lancet*. 1990;336(8718):763-765. doi:10.1016/0140-6736(90)93237-j

23. Boussac M, Garin J. Calcium-dependent secretion in human neutrophils: a proteomic approach. *Electrophoresis*. 2000;21(3):665-672. doi:10.1002/(sici)1522-2683(20000201)21:3<665::Aid-elps665>3.0.Co;2-u

24. Bjerke K, Halstensen TS, Jahnsen F, Pulford K, Brandtzaeg P. Distribution of macrophages and granulocytes expressing L1 protein (calprotectin) in human Peyer's patches compared with normal ileal lamina propria and mesenteric lymph nodes. *Gut*. 1993;34(10):1357-1363. doi:10.1136/gut.34.10.1357

25. Rammes A, Roth J, Goebeler M, Klempt M, Hartmann M, Sorg C. Myeloid-related protein (MRP) 8 and MRP14, calcium-binding proteins of the S100 family, are secreted by activated monocytes via a novel, tubulin-dependent pathway. *J Biol Chem*. 1997;272(14):9496-9502. doi:10.1074/jbc.272.14.9496

26. Li F, Ma J, Geng S, et al. Fecal calprotectin concentrations in healthy children aged 1-18 months. *PLoS One*. 2015;10(3):e0119574. doi:10.1371/journal.pone.0119574

27. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol*. 2021;17(11):e1009442. doi:10.1371/journal.pcbi.1009442

28. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7(5):335-336. doi:10.1038/nmeth.f.303

29. RStudio Team. *RStudio: Integrated Development for R*. PBC; 2020.

30. Anthamatten L, von Bieberstein PR, Menzi C, et al. Stratification of human gut microbiomes by succinotype is associated with inflammatory bowel disease status. *Microbiome*. 2024;12(1):186. doi:10.1186/s40168-024-01897-8

31. Khan I, Ullah N, Zha L, et al. Alteration of gut microbiota in inflammatory bowel disease (IBD): cause or consequence? IBD treatment targeting the gut microbiome. *Pathogens*. 2019;8(3):126. doi:10.3390/pathogens8030126

32. Atarashi K, Tanoue T, Shima T, et al. Induction of colonic regulatory T cells by indigenous Clostridium species. *Science*. 2011;331(6015):337-341. doi:10.1126/science.1198469

33. Kamada N, Núñez G. Regulation of the immune system by the resident intestinal bacteria. *Gastroenterology*. 2014;146(6):1477-1488. doi:10.1053/j.gastro.2014.01.060

34. Obermajer T, Grabnar I, Benedik E, et al. Microbes in infant gut development: placing abundance within environmental, clinical and growth parameters. *Sci Rep*. 2017;7(1):11230. doi:10.1038/s41598-017-10244-x

Tables

Table 5.1. Significant ASV associations with fecal calprotectin concentrations (ng/ml) using MaAsLin2.

Feature	Coefficient	N.not.0*	P value	Q value
ASV0YGD.Butyricicoccus	0.297	27	0.00164	0.0579
ASV006D.Faecalibacterium	0.388	27	0.00267	0.0800
ASV04LS.Clostridium	-0.623	97	0.00530	0.141
ASV02ZY.Dialister	0.323	20	0.00590	0.149
ASV004F.Streptococcus	-0.330	70	0.00756	0.173
ASV004W.Collinsella	0.465	94	0.0120	0.217
ASV026G.Ruminococcus	0.262	21	0.0157	0.246
ASV081D.Clostridium	-0.254	26	0.0154	0.246

*Number of participants with non-zero feature.

Table 5.2. Significant co-abundance group associations with fecal calprotectin concentrations (ng/ml) using MaAsLin2.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG8	0.578	179	0.00506	0.105
CAG1	-0.392	184	0.00841	0.136
CAG28	0.588	89	0.0192	0.245

*Number of participants with non-zero feature.

Table 5.3. Co-abundance group associations from milk type stratified models with fecal calprotectin concentrations (ng/ml) using MaAsLin2.

Feature	Breastmilk Coefficient	N.not.0*	P Value	Q Value	Non-breastmilk Coefficient	N.not.0*	P Value	Q Value
CAG8	0.613	147	0.00489	0.407	0.426	32	0.497	0.974
CAG1	-0.450	151	0.00617	0.407	0.258	33	0.509	0.979
CAG28	0.578	64	0.0344	0.528	0.777	25	0.219	0.926

*Number of participants with non-zero feature.

Table 5.4. Linear regression results displaying the association between Z-BMI and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted Model			Adjusted Model		
	β (95% CI)	**SE**	**P Value**	**β (95% CI)**	**SE**	**P Value**
Intercept	0.227 (0.0699, 0.385)	0.0798	0.00489	-0.445 (-1.87, 0.983)	0.723	0.539
Fecal Calprotectin	0.000196 (-0.000115, 0.000508)	0.000158	0.215	0.000320 (-0.0000137, 0.000654)	0.000169	0.0600
Age				0.0698 (-0.0392, 0.179)	0.0552	0.208
Sex (Girl)				0.0313 (-0.205, 0.268)	0.120	0.794
Milk Type (Non-breastmilk)				0.251 (-0.0779, 0.580)	0.166	0.134
Antibiotics ever (Yes)				-0.191 (-0.474, 0.0922)	0.143	0.185
Water source (Borehole/tubewell)				0.0461 (-0.377, 0.469)	0.214	0.830
Water source (Other)				-0.0962 (-0.499, 0.306)	0.204	0.638
Water source (Piped water)				-0.171 (-0.488, 0.145)	0.160	0.286
Water source (Protected dug well)				-0.316 (-0.755, 0.123)	0.222	0.157
Low birthweight (Yes)				-0.193 (-0.502, 0.116)	0.156	0.220
Antibiotics 14 days (Yes)				-0.0589 (-0.372, 0.255)	0.159	0.711
Diarrhea 14 days (Yes)				-0.102 (-0.355, 0.151)	0.128	0.427
Treatment (Egg)				0.0403 (-0.200, 0.281)	0.122	0.741
Poverty score				0.000312 (-0.00373, 0.00436)	0.00205	0.879
Model Summary						
R^2	0.00858			0.0841		
Adjusted R^2	0.00304			-0.00143		
Model P value	0.215			0.473		
Sample size (n)	181			165		

Table 5.5. Linear regression results displaying the association between Z-weight-for-length and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted Model			Adjusted Model		
	β (95% CI)	SE	P Value	β (95% CI)	SE	P Value
Intercept	-0.0168 (-0.172, 0.138)	0.0784	0.831	0.0897 (-1.29, 1.47)	0.699	0.898
Fecal calprotectin	0.000189 (-0.000117, 0.000495)	0.000155	0.226	0.000303 (-0.000019, 0.000626)	0.000163	0.065
Age				0.0107 (-0.0948, 0.116)	0.0534	0.841
Sex (Girl)				0.0718 (-0.157, 0.301)	0.116	0.536
Milk Type (Non-breastmilk)				0.2271 (-0.0911, 0.5453)	0.161	0.161
Antibiotics ever (Yes)				-0.1974 (-0.4712, 0.0765)	0.1386	0.156
Water source (Borehole/tube well)				0.0127 (-0.3962, 0.4217)	0.2070	0.951
Water source (Other)				-0.1267 (-0.5160, 0.2627)	0.1970	0.521
Water source (Piped water)				-0.1940 (-0.5002, 0.1123)	0.1550	0.213
Water source (Protected dug well)				-0.3225 (-0.7473, 0.1022)	0.2150	0.136
Low birthweight (Yes)				-0.2292 (-0.5281, 0.0698)	0.1513	0.132
Antibiotics 14 days (Yes)				-0.0429 (-0.3461, 0.2604)	0.1535	0.780
Diarrhea 14 days (Yes)				-0.1137 (-0.3586, 0.1312)	0.1240	0.361
Treatment (Egg)				0.0175 (-0.2154, 0.2504)	0.1179	0.882
Poverty score				0.000195 (-0.0037, 0.0041)	0.00198	0.922

Model Summary		
R^2	0.0082	0.086
Adjusted R^2	0.0027	0.0006
Model P value	0.226	0.449
Sample size (n)	181	165

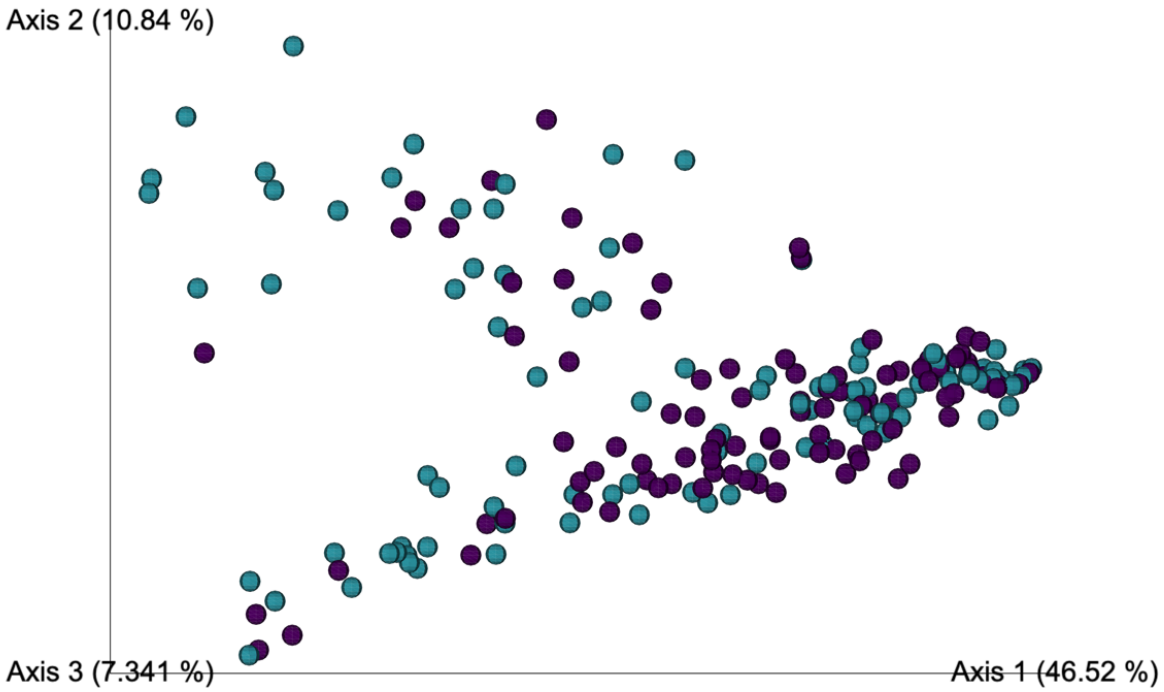

Figure 5.1. PCoA plot of high and low fecal calprotectin groups using weighted Unifrac metrics (low = green, high = purple, n = 197).

CHAPTER 6: Summary and Discussion

Complimentary feeding and the type of milk consumed are precursors which shape the infant gut microbiome and infant immune responses. This study sought to investigate whether diet diversity, fecal microbiome composition, and localized gastrointestinal (GI) inflammation were associated with infant development and growth scores. The 3 aims of this study were investigated using data from a sub-population of indigenous Mayan infants enrolled in The Saqmolo' Project Microbiome Sub-Study. The Saqmolo' Project Microbiome Sub-study recruited 200 infants from the parent study collecting endline fecal samples at ages 11-16 months. The parent study recruited 1,200 infants ages 6-9 months in a 6-month randomized controlled trial comparing the effect of an egg intervention to a multiple micronutrients powder control on early child development (ECD) and growth. The data from this trial were used to assess the impact of egg intake and diet diversity on fecal microbiome composition and localized GI inflammation (Aim 1), examined associations between fecal microbiome composition with early child development and growth metrics (Aim 2), and investigated links between localized GI inflammation with fecal microbiome composition, child development, and growth metrics (Aim 3). This chapter summarizes the results of these 3 aims, discusses the potential impact of the research, and future directions for the research community regarding this topic. Understanding the complex interactions between infant diet, gut microbiome, GI inflammation, and development is essential in vulnerable populations facing high rates of malnutrition, growth faltering, and heightened exposure to environmental pathogens.

Review of Findings and Potential Impact

Across all 3 aims, the type of milk consumed by the infant played an undeniable role in modifying the relationship between gut microbial composition, infant development, and

localized GI inflammation. **Aim 1** explored how an egg intervention in comparison to a multiple micronutrients powder control, as well as diet diversity and milk type, influenced the gut microbiome and localized gastrointestinal (GI) inflammation in rural Guatemalan infants. The key results highlighted important bacterial differences between breastfed and non-breastfed infants in response to an egg intervention. Co-abundance groups containing beneficial bacterial ASVs decreased in relative abundance in the non-breastmilk group following egg intervention, whereas co-abundance groups with ASV members often associated with dysbiosis or pathogenicity, increased. The resilient microbiota composition observed in breastfed infants suggests that breastmilk plays a dominant role in stabilizing microbial composition, potentially minimizing the impact of complementary food introduction on gut microbiota composition. Additionally, the observed relationship between fecal calprotectin and milk type indicates that non-breastmilk feeding may be linked to lower fecal calprotectin levels, though the biological significance of this result is uncertain. Lower fecal calprotectin levels in the non-breastmilk group could indicate lower levels of GI inflammation but could also signal a lack of appropriate immune response, given reported findings of temporary immune responses to complimentary food introduction in healthy infants. These findings underscore the need to develop complementary feeding strategies that account for both milk type and the type of infant formula consumed, to bolster gut microbiome health and overall gastrointestinal well-being in malnourished infants.

In **Aim 2**, associations between microbial co-abundance groups with ECD and growth scores lost significance when stratified by milk type, which once again underscores the driving influence of milk type in this population. However, the positive relationship between co-abundance group 18 and language scores in non-breastfed (fluid animal milk, powdered animal

milk, infant formula) infants suggests microbial composition and neurodevelopment may be connected. The finding is interesting given the association is specific to non-breastfed infants. When comparing language scores between breastfed and non-breastfed infants, language scores were slightly higher in the non-breastmilk fed infants. The specific amplicon sequence variants (ASVs) present in the co-abundance group in question contain bacteria previously associated with higher neurodevelopment scores and commensal bacteria known to produce short chain fatty acids (SCFA) that strengthen the gut barrier and protect against external pathogens (*Faecalibacterium, Lachnospira, Ruminococcus*, and *Roseburia*). These findings strengthen the validity of the results. What remains in question is the mechanism behind the relationship between this bacterial profile and neurodevelopment, and what bacterial byproducts may be contributing to higher language scores. However, the absence of this relationship in breastfed infants raises important questions about whether maternal nutritional status or milk composition might differentially influence neurodevelopment in a malnourished population.

Aim 3 results further underscored the need to interpret fecal calprotectin findings with caution, given that milk type once again modified the relationship with the co-abundance groups. While lower fecal calprotectin levels in non-breastfed infants may suggest reduced GI inflammation, it is equally plausible that they indicate a blunted and potentially healthy immune response. This highlights the importance of considering additional inflammatory markers and immune competence when assessing gut health in malnourished populations. While the borderline significant associations between fecal calprotectin and Z-BMI/Z-weight-for-length suggest a possible link between GI inflammation and growth, the effect sizes were minimal. This may indicate that the impact of GI inflammation on growth is only relevant in cases of severe or prolonged inflammation, as seen in conditions like environmental enteropathy. The lack of

association with stunting further supports the idea that GI inflammation in this cohort may not be severe enough to cause significant growth impairments. However, the cross-sectional nature of the data may be limiting the effect. Low-grade, chronic GI inflammation may have larger effects on growth at later stages of development, as the child ages.

Future Directions

Longitudinal studies are critically needed to assess the directionality of the relationship between gut microbiota composition and infant development. It is unclear in infants whether microbial composition, or poor development itself, is the instigating factor in this process. Particularly in populations experiencing malnutrition, it is possible that low birthweight and lack of nutrients leads to neurodevelopmental deficits which impact gut motility, mucus production, and intestinal barrier integrity, ultimately creating a feedback loop of suboptimal development outcomes. It is critical to ascertain the causal point at which this process begins after birth, to circumvent any disruptions in ECD and growth. Given the population in this study was malnourished, longitudinal research is needed to determine whether lower fecal calprotectin levels in non-breastfed infants indicates improved or impaired gut health over time. Future studies should also explore other gastrointestinal and systemic inflammatory biomarkers in combination with fecal calprotectin.

These findings also have implications for infant feeding practices in resource-limited settings. The role that breastmilk appears to play in buffering threats to the infant host speaks to the importance of promoting breastfeeding in these settings, as well as infant formulas which closely mimic the composition of breastmilk (e.g. HMO containing formulas) when breastfeeding is not possible. While our study did not investigate the impact of human milk oligosaccharides (HMO), given what appears to be a lack of resilience in non-breastfed infants to

complimentary feeding, infant formula containing HMO should be investigated in malnourished infant populations to determine whether HMO consumption mirrors the benefits of breastfeeding and confers greater gut microbiome resilience. If HMO-enriched formulas can replicate the microbiome and immune-modulating benefits of breastmilk, they may serve as a valuable intervention to support gut health and development in non-breastfed infants. Future studies should also collect and analyze the contents of mothers' breastmilk to determine whether maternal malnutrition influences its protective effects and how it differs from nutrient composition in infant formula. The micro- and macronutrient contents of breastmilk is likely depleted in a malnourished context, which may be a factor in infant ECD and growth.

Additionally, the observed relationships between microbiome composition and language raise important considerations for early infant nutrition programs. If specific microbial taxa contribute to neurodevelopment, interventions aimed at modulating the gut microbiome—through diet, probiotics, or prebiotics—could become a novel strategy for improving ECD outcomes in malnourished populations. While there were observed associations between microbial profiles and neurodevelopmental outcomes, it remains uncertain whether specific bacterial metabolites, host-microbe interactions, or nutritional factors drive these relationships. Whole-metagenome sequencing and transcriptomics should be used versus 16S rRNA sequencing to capture fungi and protists, resolve microbial diversity at the strain level, and provide functional annotation to infer metabolic potential.

Conclusions

This research lays the groundwork for optimizing early-life nutrition strategies to improve health outcomes in rural low-middle income populations. Aim 1 investigated how an egg intervention, diet diversity, and milk type influenced the gut microbiome and localized GI

inflammation in rural Guatemalan infants, revealing that milk type played a dominant role in shaping microbial composition, responses to complimentary feeding, and localized GI inflammation. Findings from Aim 2 demonstrated that microbial co-abundance groups were associated with language scores and that milk type plays a modifying role in co-abundance group associations in relation to ECD and growth. Finally, Aim 3 highlighted that while fecal calprotectin was weakly associated with weight-related growth metrics, the effect sizes were not clinically meaningful, and the biological significance of lower fecal calprotectin levels in non-breastfed infants remains uncertain. Together, these findings emphasize the critical role of milk type in infant gut microbiome development, immune response, and potential growth and developmental outcomes.

APPENDIX A: Chapter 3 Supplemental Materials

Supplemental Tables

Table S3.1. Kruskal-Wallis test comparisons of Faith's Phylogenetic Diversity and Shannon Diversity between treatment (A = egg, B = standard care controls), minimum diet diversity (MDD), and most frequently consumed milk type.

Variable	Alpha Diversity Metric	Group 1 (n)	Group 2 (n)	H Value	p-value	q-value
Treatment	Faith's PD	A (106)	B (93)	0.000152	0.990158	0.990158
Treatment	Shannon Diversity	A (106)	B (93)	0.788801	0.374463	0.374463
MDD	Faith's PD	No (52)	Yes (114)	1.484884	0.223012	0.223012
MDD	Shannon Diversity	No (52)	Yes (114)	2.509455	0.113165	0.113165
Milk Type	Shannon Diversity	Breastmilk (154)	Non-breastmilk (34)	36.924164	1.228144E-09	1.228144E-09
Milk Type	Faith's PD	Breastmilk (154)	Non-breastmilk (34)	22.001637	0.000003	0.000003

Table S3.2. PERMANOVA test comparisons of unweighted and weighted Unifrac between treatment, minimum diet diversity (MDD), and most frequently consumed milk type.

Variable	Beta-Diversity Metric	Group 1	Group 2	Sample Size	pseudo-F	p-value	q-value
Treatment	Unweighted	A	B	199	0.742936	0.791	0.791
Treatment	Weighted	A	B	199	1.581753	0.166	0.166
MDD	Unweighted	No	Yes	166	0.949575	0.467	0.467
MDD	Weighted	No	Yes	166	0.974871	0.373	0.373
Milk Type	Unweighted	Breastmilk	Non-breastmilk	188	6.8483	0.001	0.001
Milk Type	Weighted	Breastmilk	Non-breastmilk	188	39.324471	0.001	0.001

Table S3.3. Top 10 most relatively abundant ASVs stratified by treatment group.

ASV (Group: Egg Treatment)	%	ASV (Group: Control)	%
ASV000J Bifidobacterium	35.0	ASV000J Bifidobacterium	32.6
ASV0004 Bifidobacterium	6.29	ASV0004 Bifidobacterium	5.34
ASV0010 Enterobacteriaceae	3.36	ASV0010 Enterobacteriaceae	3.52
ASV099S Megasphaera	2.71	ASV099S Megasphaera	2.95
ASV00AT Bifidobacterium	2.53	ASV0002 Bacteroides	2.94
ASV000H Blautia	2.48	ASV000H Blautia	2.79
ASV0279 Bifidobacterium	2.38	ASV009K Prevotella	2.45
ASV009K Prevotella	2.13	ASV00AT Bifidobacterium	2.05
ASV0014 Faecalibacterium	2.09	ASV005L Bacteroides	2.05
ASV00A0 Prevotella	2.00	ASV02XU Ruminococcus	1.74

Table S3.4. Top 10 most relatively abundant ASVs stratified by milk type.

ASV (Group: Breastmilk)	%	ASV (Group: Not breastmilk)	%
ASV000J Bifidobacterium	40.7	ASV000J Bifidobacterium	11.4
ASV0004 Bifidobacterium	4.73	ASV0004 Bifidobacterium	10.9
ASV0010 Enterobacteriaceae	3.68	ASV000H Blautia	5.87
ASV099S Megasphaera	3.60	ASV0002 Bacteroides	5.13
ASV00AT Bifidobacterium	2.78	ASV005L Bacteroides	3.60
ASV0279 Bifidobacterium	2.46	ASV0014 Faecalibacterium	3.35
ASV059A Lactobacillus	2.09	ASV00A0 Prevotella	3.03
ASV0065 Veillonella	1.84	ASV009K Prevotella	2.67
ASV009K Prevotella	1.81	ASV0010 Enterobacteriaceae	2.49
ASV000H Blautia	1.77	ASV00LE Roseburia	2.21

Table S3.5. Amplicon sequence variant (ASV) module membership for guild co-abundance grouping method.

ASV	CAG
ASV0065 Veillonella	1
ASV006E Actinomyces	1
ASV007C Streptococcus	1
ASV00MT Haemophilus	1
ASV00T3 Streptococcus	1
ASV00WZ Clostridium_sensu_stricto_1	1
ASV032P Neisseria	1
ASV0337 Rothia	1
ASV066P Haemophilus	1
ASV081D Paeniclostridium	1
ASV087G Streptococcus	1
ASV0DFK Rothia	1
ASV0UN1 Streptococcus	1
ASV004F Streptococcus	2
ASV005A Intestinibacter	2
ASV0089 Romboutsia	2
ASV0313 Clostridium_sensu_stricto_1	2
ASV032I Turicibacter	2
ASV04LS Clostridium_sensu_stricto_1	2
ASV07DK Terrisporobacter	2
ASV0FY6 Sutterella	2
ASV000J Bifidobacterium	3
ASV0010 Escherichia-Shigella	3
ASV00BD Veillonella	3
ASV00GX Enterococcus	3
ASV00QW Veillonella	3
ASV01AN Enterococcus	3
ASV02CE Clostridioides	3
ASV058T uncultured	3
ASV05P3 Enterococcus	3
ASV06V4 Clostridium_sensu_stricto_1	3
ASV0AKK Clostridium_sensu_stricto_1	3
ASV0BGP Veillonella	3
ASV0YGW Clostridium_sensu_stricto_1	3
ASV005U Sutterella	4
ASV005Z Lactobacillus	4

ASV	CAG
ASV02ZZ Lactobacillus	4
ASV034A Bifidobacterium	4
ASV034R Megasphaera	4
ASV059A Lactobacillus	4
ASV0A7W Lachnoclostridium	4
ASV0C9G Lactobacillus	4
ASV0YFD Lactobacillus	4
ASV0YFE Lactobacillus	4
ASV0YI4 Lactobacillus	4
ASV001D Bifidobacterium	5
ASV005K Lactobacillus	5
ASV00AT Bifidobacterium	5
ASV00U8 Lactobacillus	5
ASV00UH Lactobacillus	5
ASV04LU Lachnospira	5
ASV07L8 Bifidobacterium	5
ASV0AKT Lactobacillus	5
ASV000U Bacteroides	6
ASV0013 Bacteroides	6
ASV001Q Megamonas	6
ASV05PT Bacteroides	6
ASV0CWA Sutterella	6
ASV0E3Q Fusobacterium	6
ASV0EJD Phascolarctobacterium	6
ASV0GBT Fusobacterium	6
ASV0YFC Campylobacter	6
ASV0YFG Megasphaera	6
ASV0008 Parasutterella	7
ASV002N UBA1819	7
ASV003H Incertae_Sedis	7
ASV02D0 Anaerostipes	7
ASV02XR Tyzzerella	7
ASV02ZL	7
ASV03JF Lachnospiraceae_UCG-004	7
ASV05R6	7
ASV05TD Leuconostoc	7
ASV09TB Blautia	7
ASV0A7B Lachnoclostridium	7

ASV	CAG
ASV0018 Lachnoclostridium	8
ASV001S Flavonifractor	8
ASV003K Lachnoclostridium	8
ASV005L Bacteroides	8
ASV005V Lachnoclostridium	8
ASV0092 Eggerthella	8
ASV0097 Clostridium	8
ASV00E0 Erysipelatoclostridium	8
ASV00KY Hungatella	8
ASV02XU Ruminococcus	8
ASV0344 Bacteroides	8
ASV03JI Tyzzerella	8
ASV0DH2 Blautia	8
ASV0279 Bifidobacterium	9
ASV0ATZ Bifidobacterium	9
ASV00UN Lactococcus	10
ASV01P9	10
ASV0BF4 Streptococcus	10
ASV0D5J Lactococcus	10
ASV036J Lactobacillus	11
ASV03MM	11
ASV04DB Staphylococcus	11
ASV05A6 Erysipelatoclostridium	11
ASV05PN Weissella	11
ASV05XL Veillonella	11
ASV06LA Corynebacterium	11
ASV07DC Epulopiscium	11
ASV0DFL Bifidobacterium	11
ASV0YFH Veillonella	11
ASV00DH Streptococcus	12
ASV00IQ uncultured	12
ASV02Y7 Blautia	12
ASV034D Streptococcus	12
ASV0367 Bifidobacterium	12
ASV03YT NK4A214_group	12
ASV065U Actinomyces	12
ASV0DH0 Ruminococcus	12
ASV0023 Anaerotruncus	13

ASV	CAG
ASV00JI Peptostreptococcus	13
ASV02ZI uncultured	13
ASV030Q Campylobacter	13
ASV03VD Saccharimonadales	13
ASV0672 Fusobacterium	13
ASV0674 Veillonella	13
ASV09A2 Acidaminococcus	13
ASV09AJ Peptostreptococcus	13
ASV0YGZ Tuzzerella	13
ASV0YHS uncultured	13
ASV024G Parabacteroides	14
ASV0317 Sellimonas	14
ASV032K Eubacterium	14
ASV0ATU Lachnoclostridium	14
ASV0UUL Acidaminococcus	14
ASV0YGD Butyricicoccus	14
ASV000G Bacteroides	15
ASV008P Granulicatella	15
ASV04VN Lachnoclostridium	15
ASV05U4 Veillonella	15
ASV067C Veillonella	15
ASV07P0 Finegoldia	15
ASV0GNW Anaerococcus	15
ASV0002 Bacteroides	16
ASV000H Blautia	16
ASV0014 Faecalibacterium	16
ASV001P Anaerostipes	16
ASV0028 Dorea	16
ASV002J Anaerostipes	16
ASV003G Dorea	16
ASV003J Eubacterium	16
ASV003L Lachnospiraceae_UCG-004	16
ASV003U Eubacterium	16
ASV0049 Lachnoclostridium	16
ASV004S	16
ASV000R CAG-352	17
ASV000T	17
ASV001A Blautia	17

ASV	CAG
ASV001E Faecalibacterium	17
ASV001R Fusicatenibacter	17
ASV0027 Blautia	17
ASV002F Blautia	17
ASV002M Subdoligranulum	17
ASV0046 Lachnospiraceae_ND3007_group	17
ASV0050 Monoglobus	17
ASV00LE Agathobacter	17
ASV000F Ruminococcus	18
ASV000Y Faecalibacterium	18
ASV0024 Erysipelotrichaceae_UCG-003	18
ASV004E Coprococcus	18
ASV006D Faecalibacterium	18
ASV00PA Roseburia	18
ASV01S6 Lachnospira	18
ASV05Q6 Ruminococcus	18
ASV000M Lachnospiraceae_NK4A136_group	19
ASV000P Roseburia	19
ASV001N Eubacterium	19
ASV001V Odoribacter	19
ASV002E Lachnospiraceae_UCG-004	19
ASV002H Lachnospiraceae_NK4A136_group	19
ASV0034 Lachnoclostridium	19
ASV003I Lachnospiraceae_UCG-008	19
ASV004P Colidextribacter	19
ASV01S0 Lachnospiraceae_NK4A136_group	19
ASV003T Agathobacter	20
ASV01JE Roseburia	20
ASV02Z2 Subdoligranulum	20
ASV0B1Y Subdoligranulum	20
ASV000A UCG-002	21
ASV001F Christensenellaceae_R-7_group	21
ASV001G CAG-56	21
ASV0035 Lachnospiraceae_UCG-010	21
ASV004A Coprococcus	21
ASV005W Eubacterium	21
ASV006P Lachnospiraceae_FCS020_group	21
ASV008R UCG-003	21

ASV	CAG
ASV01RN Ruminococcus	21
ASV026G Ruminococcus	21
ASV0004 Bifidobacterium	22
ASV004W Collinsella	22
ASV009K Prevotella	22
ASV00A0 Prevotella	22
ASV00H6 Dialister	22
ASV01XP Sutterella	22
ASV02WS Catenibacterium	22
ASV02WV Prevotella	22
ASV02XN Bifidobacterium	22
ASV05NN Lactobacillus	22
ASV099S Megasphaera	22
ASV0B52 Sutterella	22
ASV004N Eubacterium	23
ASV0063 Eubacterium	23
ASV01UN Eubacterium	23
ASV02ZX Eubacterium	23
ASV000B Bacteroides	24
ASV000N Bacteroides	24
ASV0016 Bacteroides	24
ASV004R Bacteroides	24
ASV0051 Bacteroides	24
ASV0078 Bacteroides	24
ASV00AD Colidextribacter	24
ASV02WR Bacteroides	24
ASV02ZR Ruminococcus	24
ASV05RM Flavonifractor	24
ASV0007 Akkermansia	25
ASV002L Lachnospiraceae_NK4A136_group	25
ASV004I Lachnoclostridium	25
ASV0081 GCA-900066575	25
ASV02X6 Ruminococcus	25
ASV02XI Ruminococcus	25
ASV0DGK Sarcina	25
ASV000C UCG-002	26
ASV000V Parabacteroides	26
ASV0047	26

ASV	CAG
ASV01K1 Alistipes	26
ASV01NR Phascolarctobacterium	26
ASV028C Parabacteroides	26
ASV05TP	26
ASV000E Alistipes	27
ASV000I Parabacteroides	27
ASV0015 Alistipes	27
ASV0022 Bilophila	27
ASV0029	27
ASV003C Incertae_Sedis	27
ASV004C	27
ASV005P Family_XIII_AD3011_group	27
ASV008D UCG-010	27
ASV030G Senegalimassilia	27
ASV05Q4 UCG-002	27
ASV0DU5 Holdemanella	27
ASV0GHF Bifidobacterium	27
ASV000K Subdoligranulum	28
ASV0011 Ruminococcus	28
ASV001B Alistipes	28
ASV002Y UCG-003	28
ASV005D Eubacterium	28
ASV005S Lachnospira	28
ASV0067 Butyricicoccus	28
ASV01VK Bacteroides	28
ASV02X0 Roseburia	28
ASV02ZY Dialister	28
ASV05RB Blautia	28
ASV0UQG Sutterella	28
ASV0019 Bacteroides	29
ASV002A Bacteroides	29
ASV01YL Lachnospiraceae_UCG-001	29
ASV0DF6 Bacteroides	29
ASV000Q Parabacteroides	30
ASV001K Tyzzerella	30
ASV0060 Butyricicoccus	30
ASV02WT NK4A214_group	30
ASV02Y2 Peptococcus	30

ASV	CAG
ASV02Y8 Olsenella	30
ASV0329 Slackia	30
ASV05QI Lachnospiraceae_FCS020_group	30
ASV05T8 uncultured	30
ASV002K	31
ASV004U Intestinimonas	31
ASV00H9 Rothia	31
ASV05BL Butyricicoccus	31
ASV09XR Gastranaerophilales	31
ASV0DHZ Prevotella	31
ASV0YGM Butyricicoccus	31
ASV001X Ruminococcus	32
ASV002X uncultured	32
ASV005I Lachnospiraceae_UCG-010	32
ASV025Z Holdemanella	32
ASV02WX Desulfovibrio	32
ASV02XM Allisonella	32
ASV02ZC Blautia	32
ASV05NR Holdemanella	32
ASV05R4 Blautia	32
ASV0DIJ	32
ASV0FVA Muribaculaceae	32
ASV0G9T Prevotella	32
ASV003E Bacteroides	33
ASV003P	33
ASV005Y Ruminococcus	33
ASV01VM Sutterella	33
ASV04TB Blautia	33
ASV05R1 Lachnospiraceae_UCG-003	33
ASV080M Erysipelotrichaceae_UCG-003	33
ASV0AJD Anaerosporobacter	33

Table S3.6. The top 10 co-abundance group relative abundance table stratified by treatment group.

Control group (Treatment B)	%	Egg group (Treatment A)	%
CAG3	40.0	CAG3	41.9
CAG22	14.8	CAG22	15.2
CAG16	10.5	CAG16	8.60
CAG17	5.50	CAG17	4.99
CAG8	4.63	CAG5	4.27
CAG5	3.61	CAG8	4.08
CAG1	3.07	CAG1	3.77
CAG18	2.67	CAG9	3.02
CAG4	2.33	CAG4	2.47
CAG6	2.23	CAG18	1.99

Table S3.7. The top 10 co-abundance group relative abundance table stratified by milk type.

Breastmilk Group	%	Non-breastmilk Group	%
CAG3	48.7	CAG16	20.1
CAG22	13.6	CAG22	19.3
CAG16	6.89	CAG3	15.1
CAG5	4.81	CAG17	10.4
CAG1	3.76	CAG8	6.92
CAG8	3.57	CAG18	5.12
CAG17	3.29	CAG6	3.43
CAG4	3.06	CAG24	2.99
CAG9	2.78	CAG1	2.53
CAG6	1.91	CAG2	1.43

Table S3.8. Per subject relative abundance of all co-abundance groups by treatment and milk type interaction subgroups.

Egg + Breastmilk group	%	Control + Breastmilk group	%	Egg + Non-breastmilk group	%	Control + Non-breastmilk group	%
CAG3	48.5	CAG3	48.9	CAG16	22.4	CAG3	19.9
CAG22	13.6	CAG22	13.5	CAG22	22.2	CAG16	18.4
CAG16	5.97	CAG16	8.08	CAG17	10.2	CAG22	17.2
CAG5	5.06	CAG5	4.49	CAG3	8.22	CAG17	10.5
CAG8	4.12	CAG1	3.47	CAG6	6.88	CAG8	8.95
CAG1	3.99	CAG4	3.11	CAG18	5.37	CAG18	4.95
CAG17	3.55	CAG17	2.95	CAG8	4.05	CAG24	3.07
CAG9	3.34	CAG8	2.86	CAG1	3.10	CAG1	2.12
CAG4	3.02	CAG6	2.69	CAG24	2.88	CAG10	1.75
CAG6	1.32	CAG9	2.04	CAG9	2.19	CAG21	1.40
CAG18	1.20	CAG18	1.69	CAG25	1.96	CAG2	1.35
CAG24	1.15	CAG24	1.11	CAG2	1.53	CAG5	1.28
CAG10	0.862	CAG15	0.601	CAG33	1.40	CAG19	1.19
CAG2	0.659	CAG2	0.548	CAG31	1.18	CAG6	0.992
CAG23	0.497	CAG32	0.467	CAG19	1.07	CAG25	0.987
CAG11	0.358	CAG25	0.465	CAG5	0.768	CAG33	0.761
CAG15	0.316	CAG28	0.421	CAG23	0.758	CAG32	0.759
CAG32	0.275	CAG23	0.397	CAG28	0.610	CAG28	0.693
CAG21	0.258	CAG27	0.310	CAG32	0.572	CAG20	0.637
CAG25	0.245	CAG10	0.268	CAG7	0.549	CAG27	0.526
CAG20	0.226	CAG19	0.208	CAG10	0.478	CAG7	0.513
CAG28	0.201	CAG7	0.184	CAG26	0.317	CAG23	0.360
CAG19	0.192	CAG29	0.165	CAG21	0.315	CAG31	0.340
CAG33	0.174	CAG21	0.164	CAG27	0.266	CAG12	0.294
CAG7	0.173	CAG11	0.164	CAG20	0.219	CAG26	0.280
CAG14	0.154	CAG33	0.140	CAG29	0.144	CAG4	0.251
CAG27	0.153	CAG20	0.127	CAG15	0.141	CAG30	0.185
CAG29	0.0977	CAG26	0.122	CAG12	0.132	CAG14	0.0916
CAG13	0.0949	CAG13	0.113	CAG14	0.0723	CAG11	0.0860
CAG26	0.0899	CAG14	0.0822	CAG30	0.0597	CAG15	0.0798
CAG30	0.0570	CAG30	0.0755	CAG13	0.0227	CAG29	0.0611
CAG12	0.0424	CAG31	0.0388	CAG11	0.0151	CAG13	0.0473
CAG31	0.0338	CAG12	0.0353	CAG4	0.0104	CAG9	-

Table S3.9. MaAsLin2 Analysis Results: Associations of Gut Microbiome ASV (10% prevalence in samples) with most frequently consumed milk type and module assignment.

ASV Feature	Coefficient	N.not. 0**	P value	Q value (FDR adjusted p value)	Module
ASV000R.Ruminococcus	2.77275354	32	0.00055358	0.0838341	2
ASV0011.Ruminococcus	1.62856911	18	0.00092259	0.09954789	2
ASV0018.Clostridium	2.32303146	18	1.96E-10	4.24E-07	1
ASV001N.Ruminococcaceae	1.78938377	22	5.77E-05	0.02075766	2
ASV001P.Coprococcus	3.70590471	84	0.00027458	0.05925331	2
ASV0022.Bilophila	1.28541101	18	2.80E-05	0.01704995	1
ASV0027.Blautia	3.52441593	50	0.00058272	0.0838341	2
ASV0067.Ruminococcaceae	2.06610522	34	0.00013286	0.0409578	2
ASV0089.Peptostreptococcacee	1.60648625	43	0.00323547	0.21819188	1
ASV008R.Oscillospira	2.82539621	28	3.16E-05	0.01704995	2
ASV05Q6.Ruminococcus	3.37144306	18	1.51E-08	1.63E-05	2
ASV05R6.Lachnospiraceae	1.29989572	22	4.74E-05	0.02047739	1
ASV07DK.SMB53	1.64595074	24	0.00015806	0.04263781	1
ASV03VD.CW040	-0.6573799	13	0.00020631	0.08897246	NA

*The model was adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, and diarrhea in the past 14 days, poverty score, medication, and treatment. Breastmilk was the reference group.
**Number of participants with non-zero feature.

Table S3.10. MaAsLin2 association of infant gut microbiome co-abundance groupings with treatment group (egg versus multiple micronutrients powder controls).

Feature	Coefficient	N not 0**	P value	Q value (FDR adjusted p value)
CAG13	0.70420139	96	0.04601735	0.35606019
CAG9	0.51335153	15	0.09512617	0.5369897
CAG15	0.7142335	81	0.1104768	0.5369897
CAG12	-0.4749944	53	0.17823968	0.70334905
CAG18	0.7788749	130	0.19881571	0.72230293
CAG3	-0.2788821	186	0.25653344	0.82401467
CAG33	0.40433497	72	0.40456162	0.92604857
CAG29	-0.2920406	37	0.41336124	0.93537744
CAG5	0.32635999	156	0.47311659	0.95803749
CAG17	0.40860395	145	0.50006089	0.97077407
CAG1	0.18024837	186	0.56227488	0.97679694
CAG2	-0.1384692	164	0.7364368	0.97679694
CAG4	-0.3696759	130	0.56309604	0.97679694
CAG6	-0.1548847	96	0.82397033	0.97679694
CAG7	0.06439564	108	0.88897538	0.97679694
CAG8	-0.0564521	181	0.89472434	0.97679694
CAG11	-0.124682	83	0.77035027	0.97679694
CAG14	0.20124167	60	0.63447825	0.97679694
CAG16	-0.2770424	178	0.56766853	0.97679694
CAG19	0.15653912	105	0.75055178	0.97679694
CAG20	-0.0809215	54	0.86400326	0.97679694
CAG21	-0.249646	79	0.62156037	0.97679694
CAG22	0.20572328	167	0.75768554	0.97679694
CAG23	-0.131696	99	0.81299343	0.97679694
CAG24	-0.3622772	117	0.54838473	0.97679694
CAG25	0.10057217	93	0.84741184	0.97679694
CAG26	0.20665593	57	0.60337543	0.97679694
CAG27	0.21250077	95	0.65719454	0.97679694
CAG28	-0.3302181	91	0.52547469	0.97679694
CAG30	-0.2410865	61	0.533149	0.97679694

Feature	Coefficient	N not 0**	P value	Q value (FDR adjusted p value)
CAG31	0.10848785	52	0.75806878	0.97679694
CAG32	-0.0252612	73	0.96040129	0.98528215
CAG10	0.02158674	80	0.96766517	0.9873938

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, animal contact, and diarrhea in the past 14 days.
**Number of participants with non-zero feature.

Table S3.11. MaAsLin2 association of infant gut microbiome co-abundance groupings with Minimum Diet Diversity (MDD).

Feature	Coefficient	N not 0 **	P value	Q value (FDR adjusted p value)
CAG26	0.84773649	44	0.05322606	0.52657971
CAG13	0.74999552	86	0.06366854	0.5481031
CAG22	1.18227309	146	0.11448093	0.69745304
CAG5	0.74229333	146	0.21668415	0.83307693
CAG7	-0.5182814	90	0.30606729	0.86206282
CAG12	-0.4225334	48	0.24900987	0.86206282
CAG16	0.58584502	156	0.29997872	0.86206282
CAG23	-0.6415681	88	0.30285293	0.86206282
CAG27	0.57293396	83	0.27141199	0.86206282
CAG33	0.53207575	63	0.31040404	0.86206282
CAG10	-0.5572995	70	0.34468991	0.87792499
CAG29	-0.3547852	33	0.36255128	0.88079943
CAG32	0.48020338	67	0.38311102	0.8819925
CAG30	0.30267649	54	0.47294649	0.94806957
CAG3	-0.213907	165	0.50626004	0.95253523
CAG20	0.34058063	47	0.50828624	0.95253523
CAG4	-0.4607366	120	0.54291142	0.95675846
CAG11	0.20208084	73	0.6683828	0.9610487
CAG14	-0.2014398	53	0.66536589	0.9610487
CAG18	0.32518141	113	0.63409553	0.9610487
CAG21	0.25347652	65	0.65490003	0.9610487
CAG25	-0.2945442	79	0.62063861	0.9610487
CAG31	0.17559026	40	0.64241352	0.9610487
CAG15	-0.1804491	76	0.7172001	0.96274997
CAG1	-0.1157511	165	0.73905109	0.96527936
CAG6	-0.2546502	88	0.74054527	0.96527936
CAG19	-0.1396585	89	0.79918894	0.98285347
CAG2	-0.0903663	144	0.84751056	0.98596258
CAG9	-0.0620618	13	0.85833101	0.98596258
CAG17	0.1071245	127	0.8863704	0.98596258
CAG24	0.10589436	98	0.87677554	0.98596258
CAG8	-0.0445582	159	0.92839018	0.99335745
CAG28	0.02853403	80	0.96083989	0.99335745

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, and diarrhea in the past 14 days.

**Number of participants with non-zero feature.

Table S3.12. MaAsLin2 association of infant gut microbiome co-abundance groupings with treatment group stratified by breastmilk (egg versus multiple micronutrients powder controls).

Feature	Coefficient	N not 0 **	P value	Q value (FDR adjusted p value)
CAG1	0.089289	153	0.347565	0.79763
CAG10	0.318108	60	0.564039	0.573662
CAG11	0.202441	69	0.480548	0.674198
CAG12	-0.31491	42	0.344109	0.361676
CAG13	0.904612	80	0.396123	0.023886
CAG14	0.178397	51	0.469994	0.704834
CAG15	0.747191	67	0.504364	0.140718
CAG16	-0.22014	145	0.577871	0.70382
CAG17	0.379602	114	0.69442	0.585486
CAG18	0.882174	100	0.672965	0.19203
CAG19	0.130457	78	0.524589	0.803967
CAG2	-0.24879	131	0.475719	0.601816
CAG20	-0.0377	37	0.480665	0.937593
CAG21	0.03242	56	0.53388	0.951664
CAG22	-0.09705	134	0.750415	0.897287
CAG23	-0.06421	76	0.618839	0.917504
CAG24	-0.59259	87	0.683496	0.387419
CAG25	-0.48619	66	0.562183	0.3886
CAG26	-0.04685	38	0.41416	0.910094
CAG27	0.074872	75	0.525775	0.886964
CAG28	-0.37292	66	0.570524	0.514403
CAG29	-0.38878	30	0.38915	0.31948
CAG3	-0.09498	153	0.269808	0.72535
CAG30	-0.00212	48	0.408786	0.995866
CAG31	-0.22213	34	0.32783	0.499159
CAG32	0.179838	59	0.538787	0.739039
CAG33	0.310351	54	0.489802	0.527352
CAG4	-0.01622	119	0.74681	0.982705
CAG5	0.556851	138	0.481852	0.24978
CAG6	-0.60622	77	0.763853	0.428738
CAG7	0.01906	82	0.503891	0.969881
CAG8	-0.05435	149	0.460657	0.906245
CAG9	0.549264	13	0.35499	0.124041

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, animal contact, and diarrhea in the past 14 days.

**Number of participants with non-zero feature.

Table S3.13. MaAsLin2 association of infant gut microbiome co-abundance groupings with treatment group stratified by non-breastmilk (egg versus multiple micronutrients powder controls).

Feature	Coefficient	N not 0 **	P value	Q value (FDR adjusted p value)
CAG1	0.602359	33	0.529192	0.956533
CAG10	-0.77783	20	0.699214	0.956533
CAG11	-1.56129	14	0.254704	0.906447
CAG12	-1.22574	11	0.434421	0.956533
CAG13	0.703094	16	0.525912	0.956533
CAG14	0.885442	9	0.341834	0.956533
CAG15	0.577901	14	0.570562	0.956533
CAG16	0.089631	33	0.865295	0.963171
CAG17	2.647739	31	0.050637	0.671578
CAG18	1.260833	30	0.413091	0.956533
CAG19	0.610149	27	0.663218	0.956533
CAG2	0.080788	33	0.926669	0.963171
CAG20	0.644688	17	0.591083	0.956533
CAG21	-1.18409	23	0.510241	0.956533
CAG22	3.228204	33	0.052676	0.671578
CAG23	0.385846	23	0.782622	0.959499
CAG24	0.601709	30	0.71529	0.956533
CAG25	4.438152	27	0.021716	0.60924
CAG26	1.506373	19	0.396454	0.956533
CAG27	-0.01627	20	0.989799	0.989799
CAG28	0.173957	25	0.914934	0.963171
CAG29	0.925098	7	0.214485	0.85231
CAG3	-2.16687	33	0.025553	0.60924
CAG30	-0.37233	13	0.693281	0.956533
CAG31	3.414639	18	0.0616	0.74536
CAG32	-0.7983	14	0.649262	0.956533
CAG33	2.691062	18	0.183343	0.829634
CAG4	-1.85944	11	0.065006	0.753094
CAG5	-1.01031	18	0.574532	0.956533
CAG6	2.034151	19	0.424951	0.956533
CAG7	-0.47917	26	0.79781	0.959499
CAG8	-0.75497	32	0.613557	0.956533
CAG9	1.231946	2	0.023853	0.60924

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, animal contact, and diarrhea in the past 14 days.

**Number of participants with non-zero feature.

Table S3.14. MaAsLin2 association of infant gut microbiome co-abundance groupings with MDD group stratified by breastmilk (egg versus multiple micronutrients powder controls).

Feature	Coefficient	N not 0 **	P value	Q value (FDR adjusted p value)
CAG1	-0.27959	148	0.441904	0.997963
CAG10	-0.47385	59	0.426977	0.997963
CAG11	0.206045	66	0.685262	0.997963
CAG12	-0.17687	41	0.632566	0.997963
CAG13	0.629091	77	0.139221	0.739215
CAG14	-0.27262	48	0.582548	0.997963
CAG15	-0.43029	66	0.422314	0.997963
CAG16	0.430698	140	0.478819	0.997963
CAG17	0.106833	110	0.884566	0.997963
CAG18	0.390883	98	0.582105	0.997963
CAG19	0.153315	76	0.780779	0.997963
CAG2	-0.25923	127	0.615982	0.997963
CAG20	0.167357	36	0.748896	0.997963
CAG21	0.32025	54	0.5747	0.997963
CAG22	1.14266	129	0.153277	0.739215
CAG23	-0.56576	75	0.392581	0.997963
CAG24	-0.15874	85	0.82702	0.997963
CAG25	-0.532	65	0.37935	0.997963
CAG26	0.331853	35	0.448173	0.997963
CAG27	0.682915	73	0.214279	0.883902
CAG28	0.067893	65	0.911031	0.997963
CAG29	-0.26601	30	0.527293	0.997963
CAG3	-0.14815	148	0.607358	0.997963
CAG30	0.052964	46	0.902934	0.997963
CAG31	-0.20305	32	0.563272	0.997963
CAG32	0.584608	58	0.306762	0.997963
CAG33	0.264815	52	0.609967	0.997963
CAG4	-0.43096	115	0.586135	0.997963
CAG5	0.811705	133	0.111946	0.708922
CAG6	-0.67059	76	0.412461	0.997963
CAG7	-0.89214	79	0.102333	0.708922
CAG8	-0.06306	144	0.899107	0.997963
CAG9	-0.0571	13	0.88157	0.997963

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, and diarrhea in the past 14 days.

**Number of participants with non-zero feature.

Table S3.15. Linear regression results displaying the association between treatment group and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	**SE**	**P Value**	**β (95% CI)**	**SE**	**P Value**
Intercept	360 (286, 435)	37.8	8.70E-18	512 (-103, 1127)	312	0.102
Treatment (Egg)	8.76 (-93.3, 110)	51.7	0.866	-10.2 (-119, 98.5)	55.1	0.852
Age				-20.7 (-68.3, 26.8)	24.1	0.390
Sex (Girl)				10.0 (-92.7, 113)	52.0	0.847
Milk Type (Non-breastmilk)				-169 (-311, -28.2)	71.5	0.0189
Antibiotics ever (Yes)				88.787 (-37.0, 215)	63.7	0.165
Water source (Borehole/tubewell)				89.1 (-98.7, 277)	95.2	0.350
Water source (Other)				142 (-35.2, 319)	89.7	0.116
Water source (Piped water)				87.8 (-52.4, 228)	71.0	0.218
Water source (Protected dug well)				42.5 (-148, 233)	96.5	0.660
Antibiotics 14 days (Yes)				-77.6 (-212, 56.8)	68.1	0.256
Diarrhea 14 days (Yes)				15.0 (-96.7, 126.6)	56.6	0.792
Animal contact (Yes)				29.3 (-84.0, 142.6)	57.4	0.610
Model Summary						
R^2	0.000156			0.0781		
Adjusted R^2	-0.00528			0.0134		
Model P value	0.866			0.281		
Sample size (n)	186			184		

Table S3.16. Linear regression results displaying the association between minimum diet diversity (MDD) and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	SE	P Value	β (95% CI)	SE	P Value
Intercept	423 (319, 527)	52.8	2.39E-13	750 (29.4, 1470)	364	0.0415
MDD (No)	-59.3 (-184, 65.6)	63.2	0.350	-57.8 (-187, 71.3)	65.3	0.377
Age				-37.8 (-92.6, 17.0)	27.7	0.175
Sex (Girl)				5.24 (-112, 122)	59.1	0.930
Milk Type (Non-breastmilk)				-143 (-383, 97.7)	122	0.243
Antibiotics ever (Yes)				107 (-31.3, 245)	69.8	0.129
Water source (Borehole/tube well)				114 (-93.0, 322)	105	0.278
Water source (Other)				158 (-48.9, 366)	105	0.133
Water source (Piped water)				114 (-46.1, 274)	80.9	0.162
Water source (Protected dug well)				17.1 (-205, 239)	112	0.879
Antibiotics 14 days (Yes)				-119 (-275, 37.0)	79.0	0.134
Diarrhea 14 days (Yes)				16.3 (-113, 145)	65.2	0.802
Animal contact (Yes)				22.4 (-106, 151)	65.2	0.731
Model Summary						
R^2	0.00560			0.0864		
Adjusted R^2	-0.000771			0.0103		
Model P value	0.350			0.336		
Sample size (n)	158			157		

Supplemental Figures

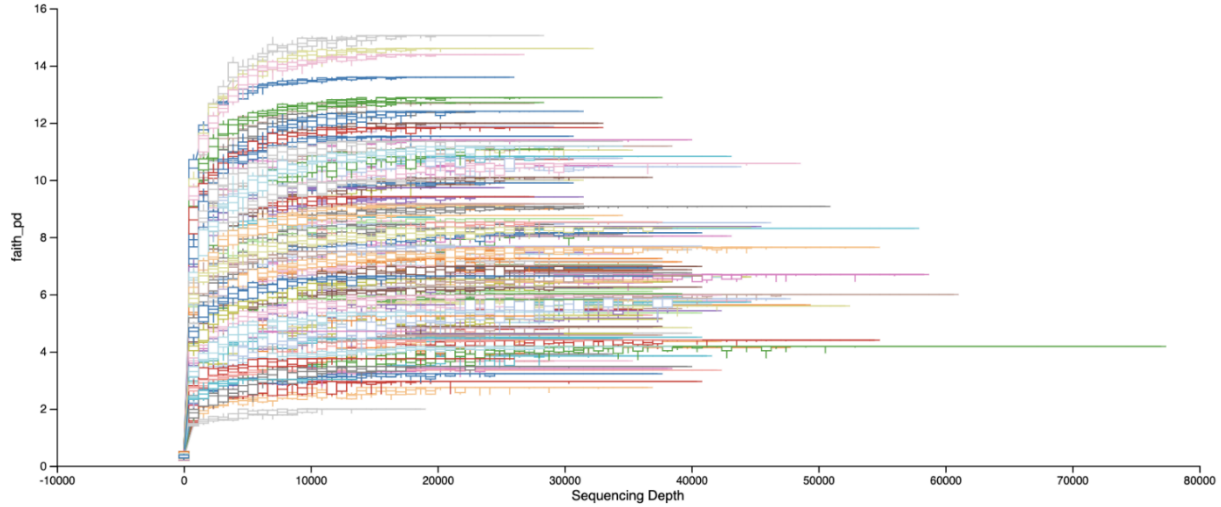

Figure S3.1. Plot of alpha rarefaction curves using Faith's phylogenetic diversity.

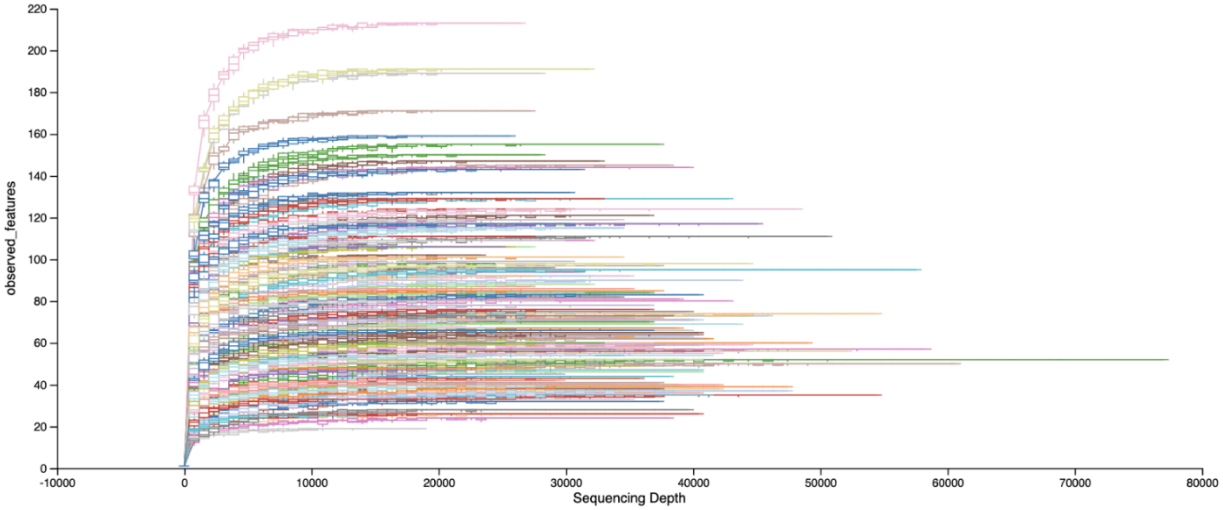

Figure S3.2. Plot of alpha rarefaction curves using observed features.

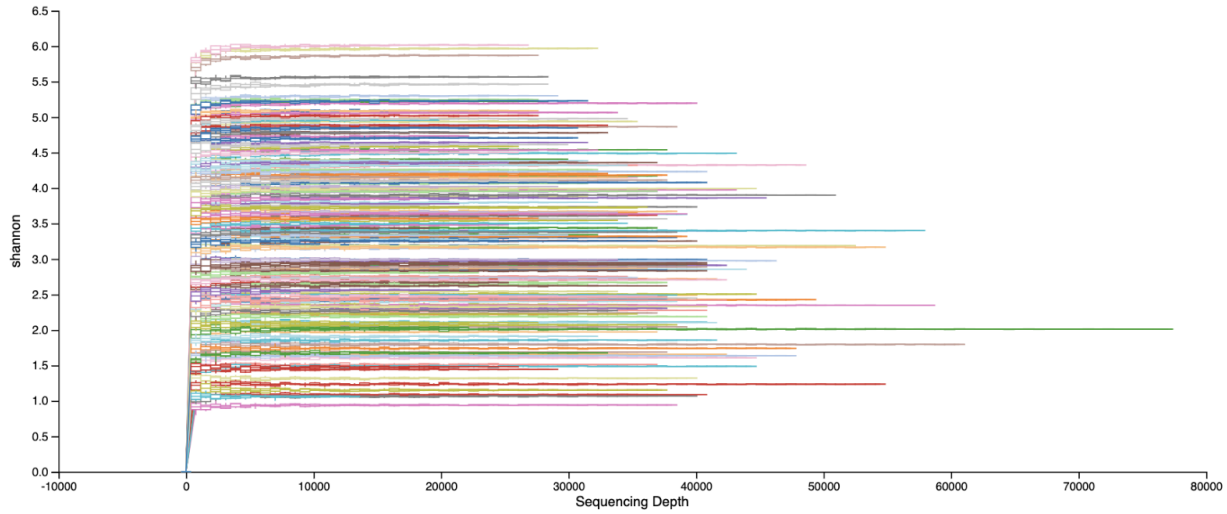

Figure S3.3. Plot of alpha rarefaction curves using Shannon diversity.

Alpha-Diversity Comparisons

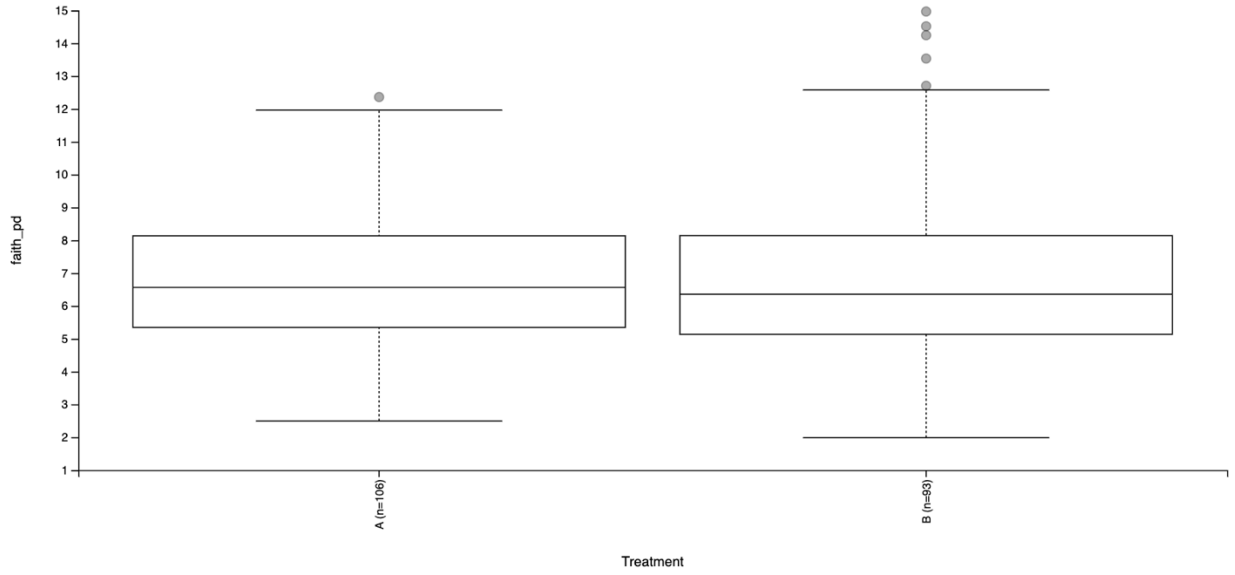

Figure S3.4. Faith's Phylogenetic Diversity (Faith's PD) boxplot of treatment groups (A = egg group, B = standard care controls).

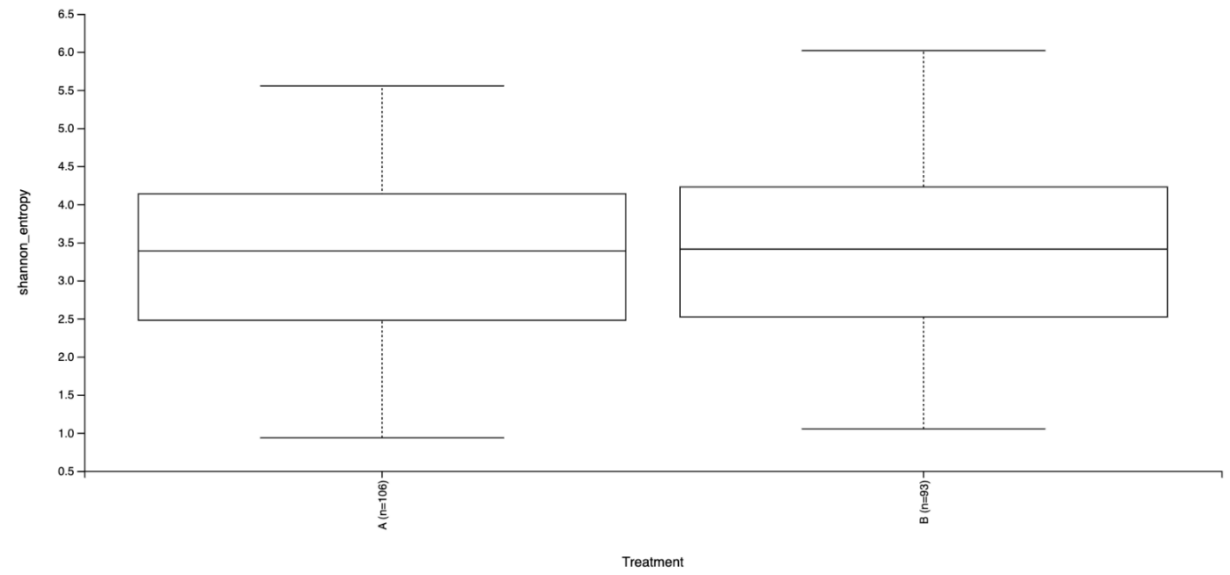

Figure S3.5. Shannon Diversity (Faith's PD) boxplot of treatment groups (A = egg group, B = standard care controls).

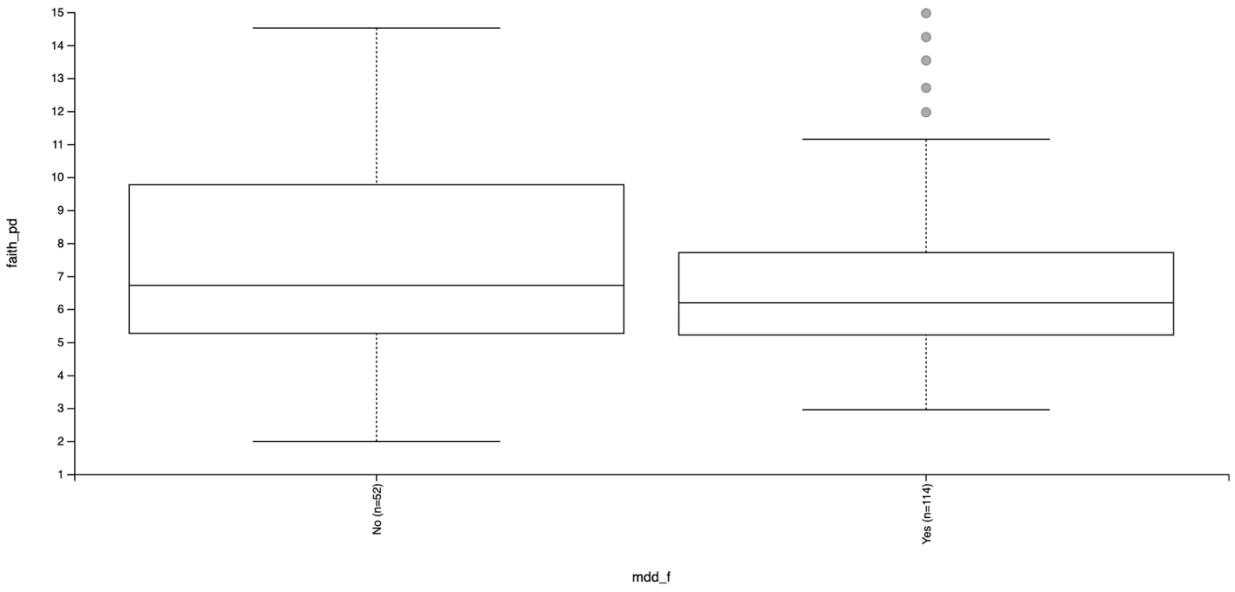

Figure S3.6. Faith's Phylogenetic Diversity (Faith's PD) boxplot of minimum diet diversity (MDD) groups.

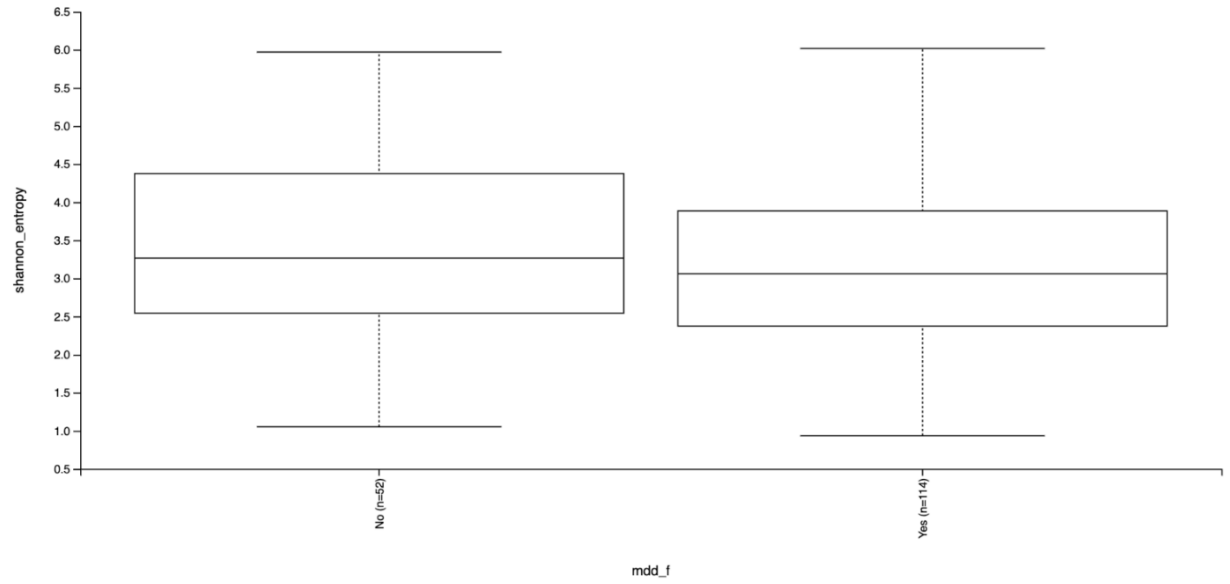

Figure S3.7. Shannon Diversity (Faith's PD) boxplot of minimum diet diversity (MDD) groups.

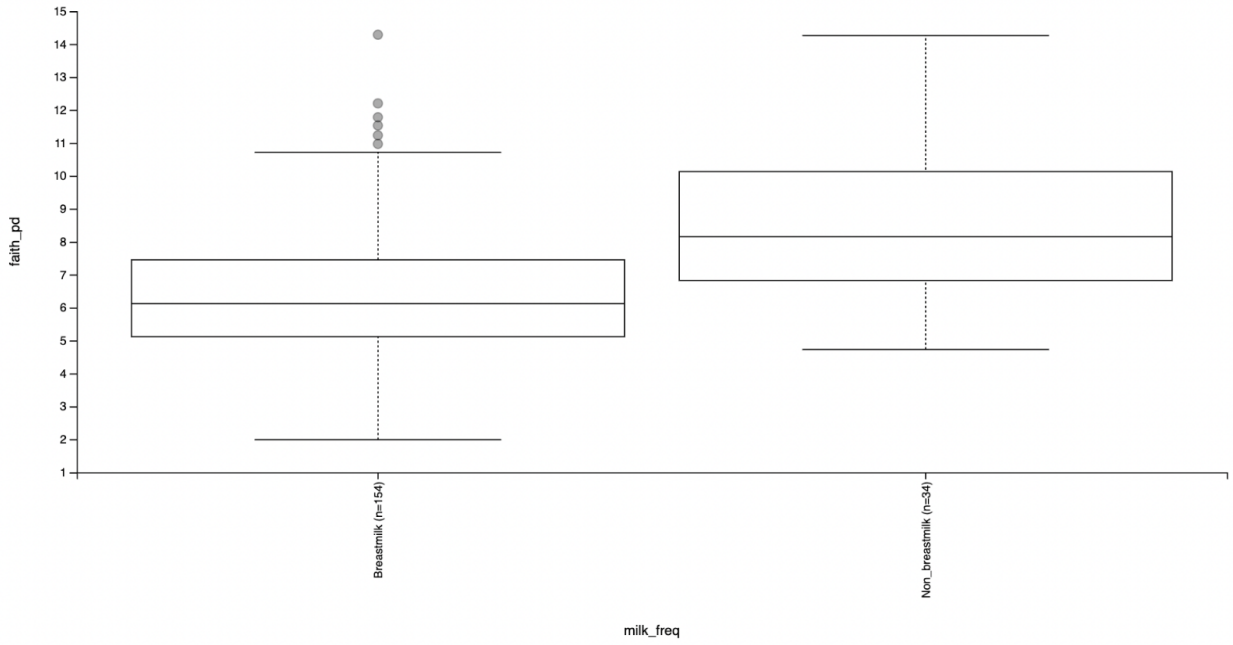

Figure S3.8. Faith's Phylogenetic Diversity (Faith's PD) boxplot of milk type most frequently consumed groups.

Beta-Diversity Comparisons

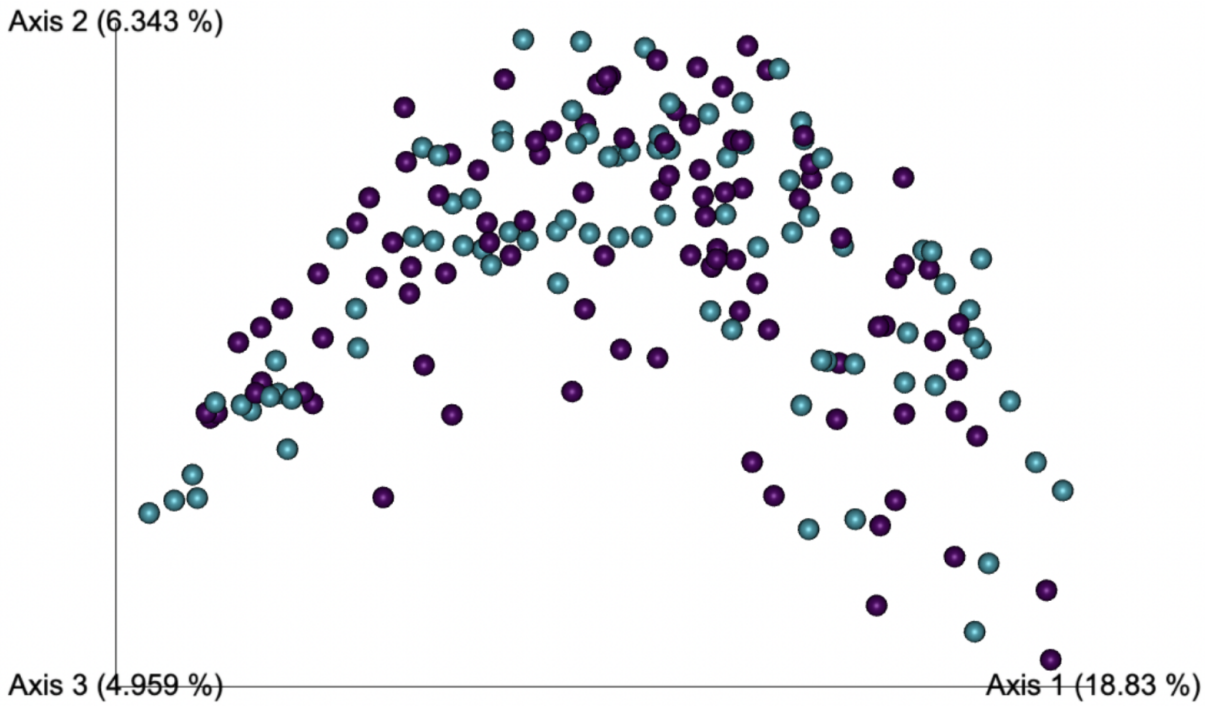

Figure S3.9. Unweighted Unifrac PCoA plot of treatment groups (purple = egg group, blue = standard care controls, n = 199).

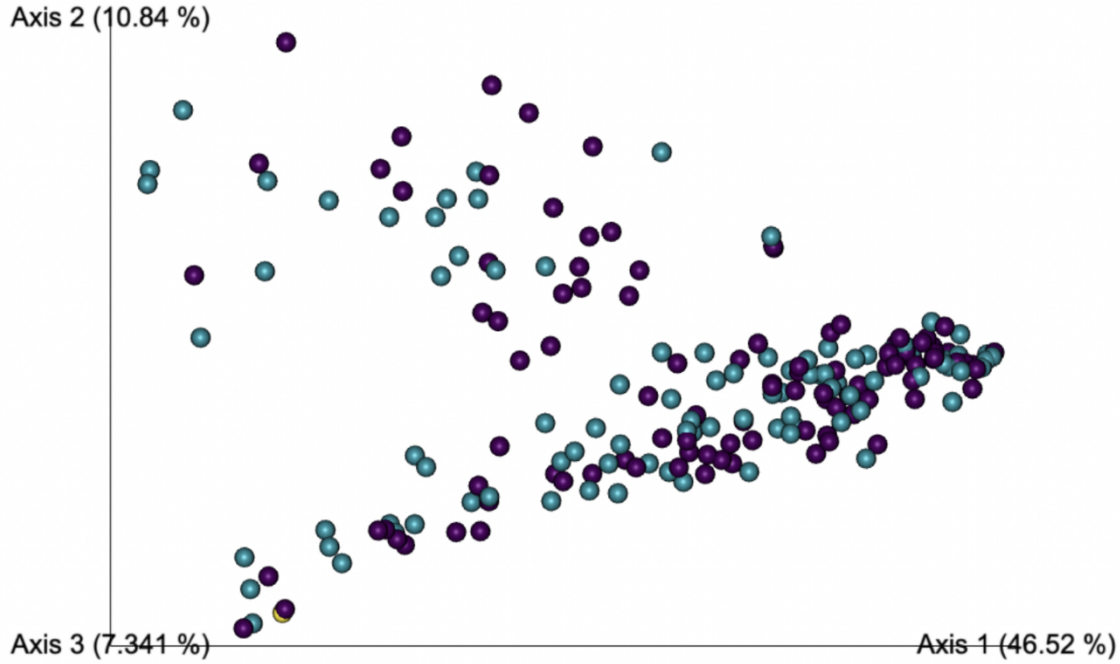

Figure S3.10. Weighted Unifrac PCoA plot of treatment groups (purple = egg group, blue = standard care controls, n = 199).

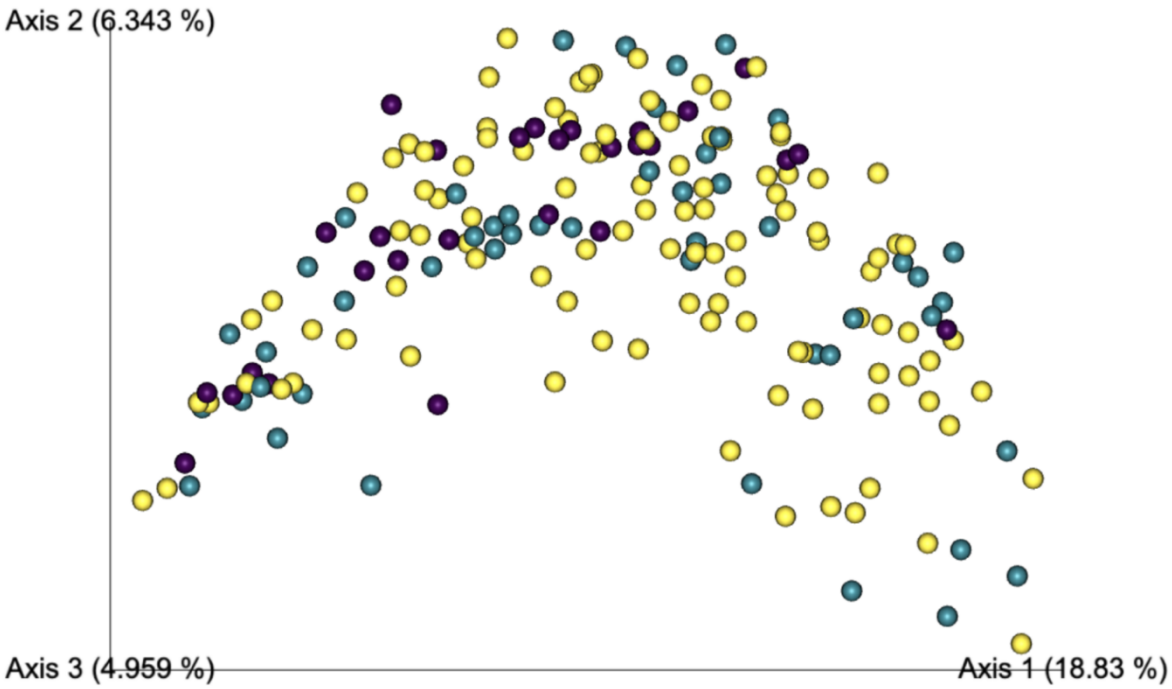

Figure S3.11. Unweighted Unifrac PCoA plot of minimum diet diversity groups (MDD) (yellow = meets MDD, teal = does not meet MDD, purple = NA).

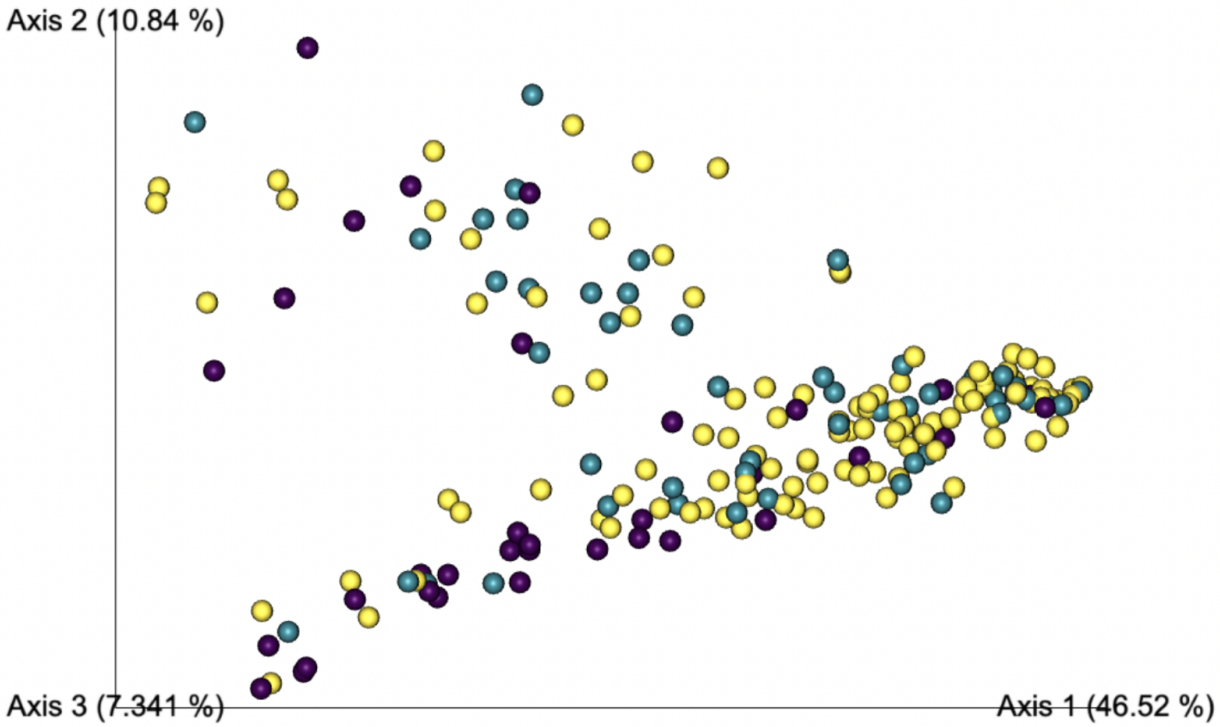

Figure S3.12. Weighted Unifrac PCoA plot of minimum diet diversity groups (MDD) (yellow = meets MDD, teal = does not meet MDD, purple = NA).

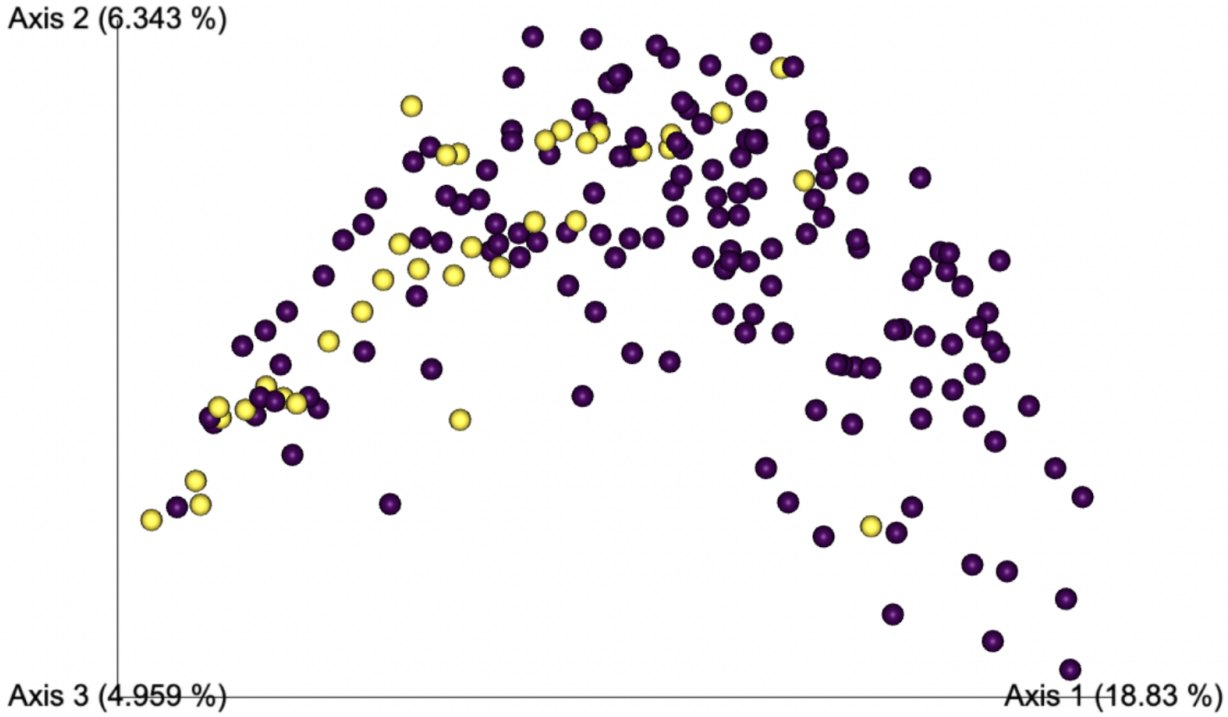

Figure S3.13. Unweighted Unifrac PCoA plot of most frequently consumed milk type groups (purple = breastmilk, yellow = non-breastmilk).

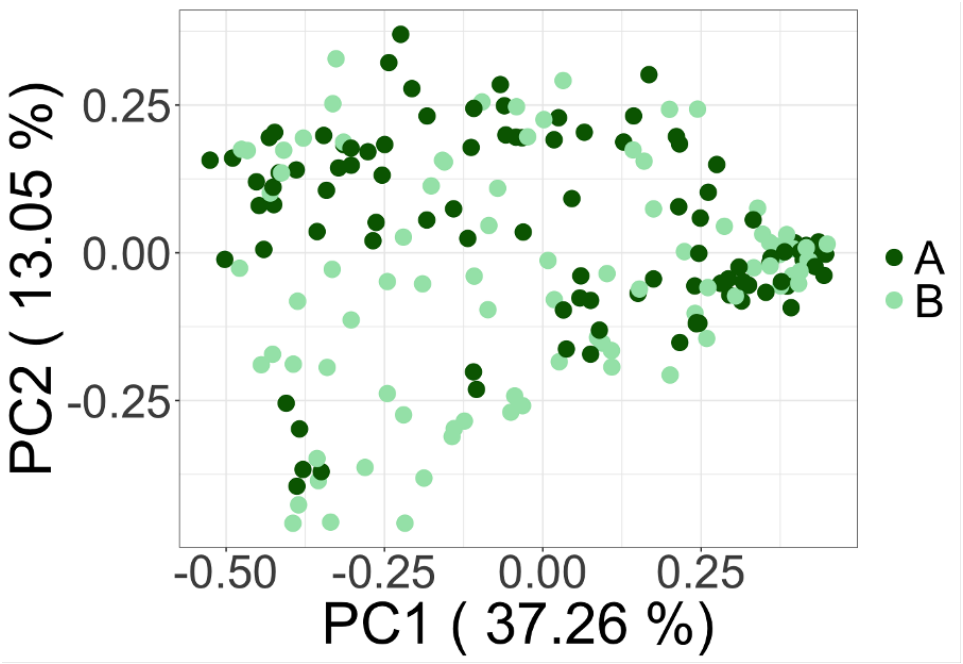

Figure S3.14. Individual point PCoA plot using Bray-Curtis distance colored by treatment group (A = egg group, B = control group, n = 199).

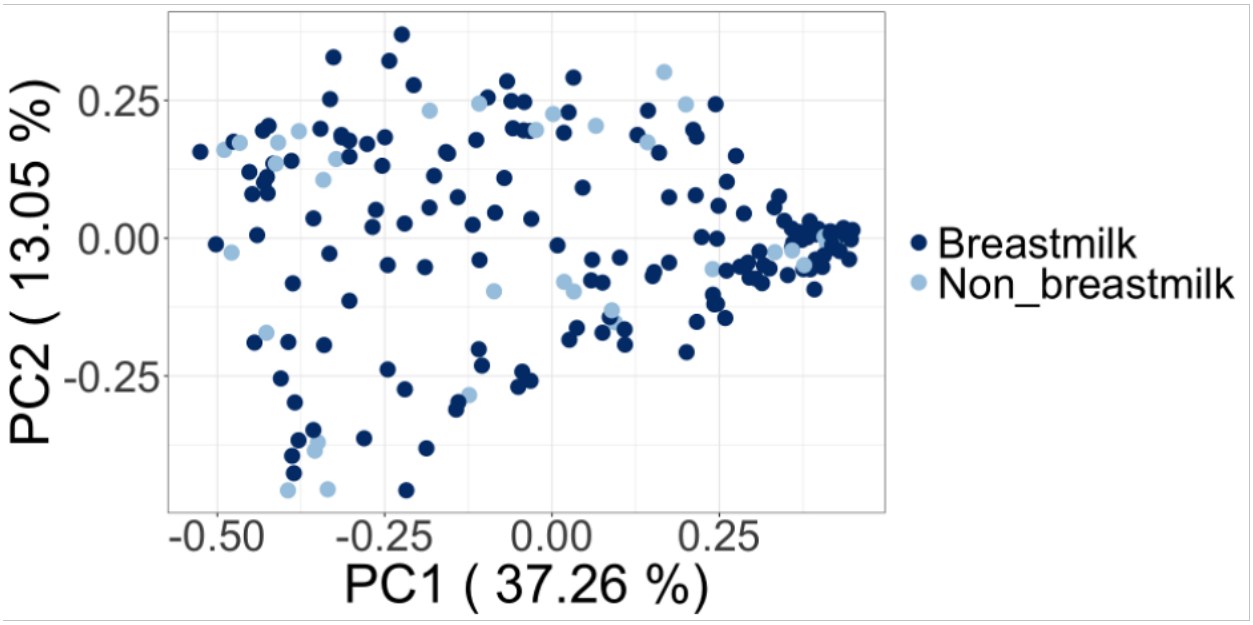

Figure S3.15. Individual point PCoA plot using Bray-Curtis distance colored by milk type (breastmilk, non-breastmilk, n = 188).

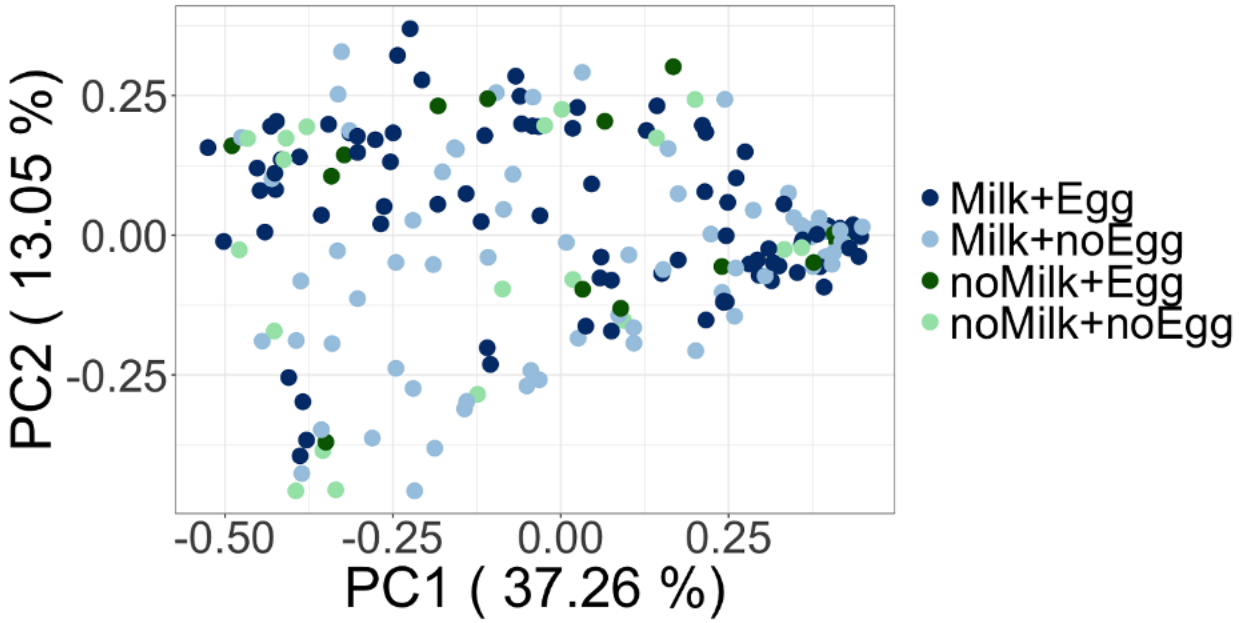

Figure S3.16. Individual point PCoA plot using Bray-Curtis distance colored by treatment and milk type interactions (breastmilk + egg [n = 87], breastmilk + control [n = 67], non-breastmilk + egg [n = 14], non-breastmilk + control [n = 20]).

APPENDIX B: Chapter 4 Supplemental Materials

Supplemental Tables

Table S4.1. Z-score mean and SD stratified by treatment and breastmilk subgroups.

Egg + Breastmilk group			**Control + Breastmilk group**		
Variable	**Mean**	**SD**	**Variable**	**Mean**	**SD**
Z-length-for-age	-1.90149425	0.9668115	Z-length-for-age	-1.87484848	1.1522583
Z-BMI	0.28337209	0.6486286	Z-BMI	0.22261538	0.8370396
Z-weight-for-length	0.03604651	0.6394359	Z-weight-for-length	-0.01276923	0.8079014
Z-weight-for-age	-0.9305814	0.7411205	Z-weight-for-age	-0.95292308	0.8954743
Z-head-circumference	-0.72954023	0.859697	Z-head-circumference	-0.94515152	0.9060434

Table S4.2. Z-score mean and SD stratified by treatment and non-breastmilk subgroups.

Egg + Non-Breastmilk group			**Control + Non-Breastmilk group**		
Variable	**Mean**	**SD**	**Variable**	**Mean**	**SD**
Z-length-for-age	-2.31928571	1.2439853	Z-BMI	0.5611111	0.8220698
Z-BMI	0.4	0.6916758	Z-head-circumference	-0.8035	0.6595475
Z-weight-for-length	0.08785714	0.6530894	Z-length-for-age	-1.6405	1.1623229
Z-weight-for-age	-1.10571429	0.7782793	Z-weight-for-age	-0.5772222	1.0598213
Z-head-circumference	-1.04214286	0.713056	Z-weight-for-length	0.3238889	0.8791917

Table S4.3. Kruskal-Wallis test comparisons of Faith's Phylogenetic Diversity, Shannon Diversity, and Pielou's Evenness between high and low CREDI scores, anthropometric z-scores, and stunting.

Variable	Alpha Diversity Metric	H Value	P value	Q value
CREDI Overall	Faith's PD	4.041698	0.044389	0.044389
CREDI Overall	Shannon Diversity	0.773197	0.37923	0.37923
CREDI Overall	Pielou's Evenness	0.100018	0.751808	0.751808
CREDI Cognitive	Faith's PD	4.534594	0.033216	0.033216
CREDI Cognitive	Shannon Diversity	2.00951	0.156316	0.156316
CREDI Cognitive	Pielou's Evenness	0.725958	0.394197	0.394197
CREDI Motor	Faith's PD	3.873929	0.049042	0.049042
CREDI Motor	Shannon Diversity	1.541225	0.214436	0.214436
CREDI Motor	Pielou's Evenness	0.66394	0.415173	0.415173
CREDI Socioemotional	Faith's PD	3.728851	0.05348	0.05348
CREDI Socioemotional	Shannon Diversity	1.210166	0.271299	0.271299
CREDI Socioemotional	Pielou's Evenness	0.386076	0.534369	0.534369
CREDI Language	Faith's PD	3.39196	0.065515	0.065515
CREDI Language	Shannon Diversity	2.32328	0.127451	0.127451
CREDI Language	Pielou's Evenness	0.962389	0.326586	0.326586
Z-Length-for-Age	Faith's PD	0.03987	0.841734	0.841734
Z-Length-for-Age	Shannon Diversity	0.570469	0.450072	0.450072
Z-Length-for-Age	Pielou's Evenness	0.470379	0.492813	0.492813
Z-Weight-for-Age	Faith's PD	0.553458	0.456908	0.456908
Z-Weight-for-Age	Shannon Diversity	0.04821	0.826208	0.826208
Z-Weight-for-Age	Pielou's Evenness	0.00353	0.952624	0.952624
Z-Head-Circumference-for-Age	Faith's PD	0.436971	0.508588	0.508588
Z-Head-Circumference-for-Age	Shannon Diversity	0.144969	0.70339	0.70339
Z-Head-Circumference-for-Age	Pielou's Evenness	0.174686	0.675981	0.675981
Z-BMI	Faith's PD	0.097254	0.75515	0.75515
Z-BMI	Shannon Diversity	0.026364	0.871014	0.871014
Z-BMI	Pielou's Evenness	0.036375	0.848744	0.848744
Z-Weight-for-Length	Faith's PD	0.176487	0.67441	0.67441
Z-Weight-for-Length	Shannon Diversity	0.006384	0.936319	0.936319
Z-Weight-for-Length	Pielou's Evenness	0.026364	0.871014	0.871014
Stunting	Faith's PD	0.108933	0.741362	0.741362

Variable	Alpha Diversity Metric	H Value	P value	Q value
Stunting	Shannon Diversity	0.32013	0.571529	0.571529
Stunting	Pielou's Evenness	0.208493	0.647951	0.647951

Table S4.4. PERMANOVA test comparisons of unweighted and weighted Unifrac between high and low CREDI scores, anthropometric z-scores, and stunting.

Variable	Beta-Diversity Metric	Sample Size	pseudo-F	P value	Q value
CREDI Overall	Unweighted UniFrac	198	1.687082	0.041	0.041
CREDI Overall	Weighted UniFrac	198	0.795129	0.452	0.452
CREDI Cognitive	Unweighted UniFrac	198	1.949878	0.021	0.021
CREDI Cognitive	Weighted UniFrac	198	1.182587	0.28	0.28
CREDI Motor	Unweighted UniFrac	198	1.875089	0.035	0.035
CREDI Motor	Weighted UniFrac	198	1.513873	0.164	0.164
CREDI Socioemotional	Unweighted UniFrac	198	1.328749	0.137	0.137
CREDI Socioemotional	Weighted UniFrac	198	1.229807	0.227	0.227
CREDI Language	Unweighted UniFrac	198	1.766257	0.022	0.022
CREDI Language	Weighted UniFrac	198	2.042639	0.089	0.089
Z-Length-for-Age	Unweighted UniFrac	198	1.100959	0.298	0.298
Z-Length-for-Age	Weighted UniFrac	198	0.752151	0.506	0.506
Z-Weight-for-Age	Unweighted UniFrac	193	1.016905	0.402	0.402
Z-Weight-for-Age	Weighted UniFrac	193	0.965295	0.368	0.368
Z-Head-Circumference-for-Age	Unweighted UniFrac	198	0.660029	0.898	0.898
Z-Head-Circumference-for-Age	Weighted UniFrac	198	0.684057	0.537	0.537
Z-BMI	Unweighted UniFrac	193	0.67678	0.897	0.897
Z-BMI	Weighted UniFrac	193	1.029505	0.331	0.331
Z-Weight-for-Length	Unweighted UniFrac	193	0.635836	0.953	0.953
Z-Weight-for-Length	Weighted UniFrac	193	0.566393	0.641	0.641
Stunting	Unweighted UniFrac	198	1.095215	0.267	0.267
Stunting	Weighted UniFrac	198	0.831921	0.464	0.464

Table S4.5. Significant ASV associations with CREDI score modules using MaAsLin2.

Feature	Variable*	Coefficient	N.not.0**	P value	Q value
ASV02CE. Clostridioides	CREDI Overall	-0.47	29	0.00136	0.074
ASV0013.Bacteroides	CREDI Overall	0.775	31	0.00177	0.0875
ASV001Q.Megamonas	CREDI Overall	1.08	53	0.00437	0.174
ASV002E. Lachnospiraceae_UCG-004	CREDI Overall	0.55	60	0.00902	0.243
ASV02CE. Clostridioides	CREDI Cognitive	-0.478	30	0.000223	0.0204
ASV002E. Lachnospiraceae_UCG-004	CREDI Cognitive	0.558	64	0.00394	0.147
ASV0067.Butyricicoccus	CREDI Language	0.423	43	0.00418	0.186
ASV02CE. Clostridioides	CREDI Language	-0.38	29	0.00528	0.215
ASV0013.Bacteroides	CREDI Motor	0.728	31	0.00368	0.147
ASV001Q.Megamonas	CREDI Motor	1.06	53	0.00565	0.189
ASV02CE. Clostridioides	CREDI Socioemotional	-0.583	29	4.56E-05	0.00524
ASV0013.Bacteroides	CREDI Socioemotional	0.65	31	0.0079	0.226

Table S4.6. Significant ASV associations with growth scores using MaAsLin2.

Feature	Variable*	Coefficient	N.not .0**	P value	Q value
ASV0065.Veillonella	Z-BMI	-0.685	142	0.00933	0.228
ASV099S.Megasphaera	Z-Head-Circumference	-0.977	67	0.00742	0.205
ASV004I.Lachnoclostridium	Z-Length-for-Age	0.365	39	0.00447	0.161
ASV0DFK.Rothia	Z-Length-for-Age	-0.553	72	0.00470	0.167
ASV003E.Bacteroides	Z-Length-for-Age	0.539	25	0.00547	0.174
ASV00E0.Clostridium	Z-Length-for-Age	0.548	82	0.00624	0.185
ASV01S6.Lachnospira	Z-Weight-for-Length	-0.619	45	0.00252	0.107
ASV0065.Veillonella	Z-Weight-for-Length	-0.693	142	0.00871	0.220
ASV087G.Streptococcus	Z-Weight-for-Length	-0.698	67	0.00919	0.229

Table S4.7. MaAslin2 associations between co-abundance groupings and CREDI overall scores.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.22260127	96	0.0028023	0.09247601
CAG28	0.88727419	91	0.00366008	0.11061698
CAG16	0.51829484	178	0.07169746	0.57242325
CAG19	0.48539139	105	0.09213955	0.65155825
CAG1	-0.3003861	186	0.10055098	0.65559918
CAG10	-0.4933455	80	0.11980053	0.65559918
CAG23	0.5263669	99	0.11135462	0.65559918
CAG24	0.55019169	117	0.12531423	0.66699509
CAG12	-0.284354	53	0.16934876	0.77314951
CAG17	0.4122562	145	0.25258887	0.84217002
CAG27	0.33679752	95	0.23922506	0.84217002
CAG7	0.31240435	108	0.26053235	0.84289879
CAG14	0.23406685	60	0.35345777	0.90597218
CAG18	0.34817003	130	0.3323778	0.90597218
CAG13	-0.1816474	96	0.37839688	0.92007769
CAG21	0.26604998	79	0.37424656	0.92007769
CAG3	0.12086984	186	0.40805865	0.92605514
CAG2	-0.1875679	164	0.45068461	0.93401715
CAG20	0.20950916	54	0.45157157	0.93401715
CAG32	0.22640639	73	0.45097066	0.93401715
CAG5	-0.1786005	156	0.50992015	0.95329553
CAG8	-0.1683401	181	0.51035013	0.95329553
CAG4	0.08558883	130	0.81858504	0.9750255
CAG9	0.08027356	15	0.65853	0.9750255
CAG11	-0.1127374	83	0.65764469	0.9750255
CAG22	0.19557416	167	0.62219538	0.9750255
CAG25	0.08598953	93	0.78222903	0.9750255
CAG29	-0.114107	37	0.58352852	0.9750255
CAG31	-0.1066221	52	0.61016266	0.9750255
CAG33	0.15874108	72	0.57903689	0.9750255
CAG30	-0.0318346	61	0.89021193	0.97823757
CAG15	0.02739518	81	0.91777825	0.97947663
CAG26	0.0271881	57	0.90876052	0.97947663

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.
**Number of participants with non-zero feature.

Table S4.8. MaAslin2 associations between co-abundance groupings and CREDI overall scores stratified by breastmilk consumption.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.30286942	77	0.00320754	0.76873471
CAG28	0.85616571	66	0.00941404	0.8010473
CAG1	-0.2747762	153	0.17402577	0.99823744
CAG2	-0.0852266	131	0.76669658	0.99823744
CAG3	0.11330514	153	0.47643164	0.99823744
CAG4	0.14315642	119	0.73895494	0.99823744
CAG5	-0.1068611	138	0.70701466	0.99823744
CAG7	0.42857046	82	0.15459669	0.99823744
CAG8	-0.1200827	149	0.65936609	0.99823744
CAG9	0.11931399	13	0.56762555	0.99823744
CAG10	-0.568466	60	0.08595105	0.99823744
CAG11	-0.127606	69	0.65319022	0.99823744
CAG12	-0.2320271	42	0.24416682	0.99823744
CAG13	-0.1947828	80	0.39411978	0.99823744
CAG14	0.15477581	51	0.57655386	0.99823744
CAG15	-0.0545403	67	0.85312154	0.99823744
CAG16	0.61245305	145	0.07037822	0.99823744
CAG17	0.35775463	114	0.38151919	0.99823744
CAG18	0.19652901	100	0.62183742	0.99823744
CAG19	0.41873746	78	0.17029297	0.99823744
CAG20	-0.0192652	37	0.94646417	0.99823744
CAG21	0.00601021	56	0.98479489	0.99823744
CAG22	0.16923348	134	0.70187845	0.99823744
CAG23	0.30541665	76	0.4010552	0.99823744
CAG24	0.65632782	87	0.10202999	0.99823744
CAG25	0.09216948	66	0.78071266	0.99823744
CAG26	-0.1274203	38	0.60260853	0.99823744
CAG27	0.36499395	75	0.23704542	0.99823744
CAG29	-0.1260816	30	0.57352866	0.99823744
CAG30	-0.1565495	48	0.51429273	0.99823744
CAG31	-0.1375111	34	0.47868451	0.99823744
CAG32	0.19319321	59	0.5413753	0.99823744
CAG33	-0.0554713	54	0.84619538	0.99823744

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.9. MaAslin2 associations between co-abundance groupings and CREDI overall scores stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG18	2.22337728	30	0.00083672	0.21634535
CAG2	-0.8282377	33	0.0778994	0.93943023
CAG12	-1.2161309	11	0.13243582	0.93943023
CAG19	1.21741359	27	0.07828385	0.93943023
CAG20	1.08537609	17	0.10540378	0.93943023
CAG21	1.84692455	23	0.04700293	0.93943023
CAG22	1.0994709	33	0.20306876	0.93943023
CAG23	1.35367804	23	0.05620993	0.93943023
CAG26	1.10819805	19	0.20609107	0.94458405
CAG15	0.6726719	14	0.22118787	0.95464869
CAG6	1.19658975	19	0.28799377	0.95496528
CAG28	0.9887284	25	0.2352167	0.95496528
CAG31	1.04966747	18	0.2763479	0.95496528
CAG33	1.33131295	18	0.25248387	0.95496528
CAG5	-0.7948682	18	0.38863284	0.95560963
CAG11	-0.6983182	14	0.37794337	0.95560963
CAG17	0.66951726	31	0.34779768	0.95560963
CAG24	0.8144976	30	0.36288357	0.95560963
CAG10	-0.7401686	20	0.51380567	0.9726234
CAG16	0.20634129	33	0.47442986	0.9726234
CAG1	-0.1399543	33	0.77444701	0.9862591
CAG3	-0.2125876	33	0.67095351	0.9862591
CAG4	0.14370159	11	0.80141648	0.9862591
CAG7	-0.2816874	26	0.74301973	0.9862591
CAG8	-0.2236903	32	0.78441442	0.9862591
CAG9	-0.092357	2	0.75943328	0.9862591
CAG14	0.1153843	9	0.81253022	0.9862591
CAG25	0.43558176	27	0.66093532	0.9862591
CAG29	0.12905649	7	0.75434673	0.9862591
CAG30	0.14887644	13	0.77370116	0.9862591
CAG32	0.33391436	14	0.74460166	0.9862591
CAG27	-0.1526975	20	0.84065975	0.99001029
CAG13	-0.115988	16	0.84856951	0.99053647

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.10. MaAslin2 associations between co-abundance groupings and CREDI language scores.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.17494196	96	0.00264373	0.0997065
CAG28	0.68566232	91	0.01938933	0.3566146
CAG18	0.74333374	130	0.02981951	0.44984855
CAG21	0.52891197	79	0.06416748	0.5635907
CAG19	0.49187304	105	0.07344254	0.5702597
CAG10	-0.4732421	80	0.11840931	0.68703421
CAG16	0.41237785	178	0.13252533	0.71741695
CAG24	0.48077725	117	0.16219977	0.73828862
CAG25	0.37921926	93	0.20223468	0.80285647
CAG17	0.43089278	145	0.21095643	0.80784
CAG1	-0.2090839	186	0.2317905	0.81744769
CAG9	0.20522787	15	0.23760202	0.81744769
CAG23	0.36656225	99	0.24584358	0.81744769
CAG29	-0.2216832	37	0.26343434	0.81744769
CAG31	0.21970199	52	0.27067358	0.81744769
CAG33	0.29903224	72	0.27254573	0.81744769
CAG5	-0.246785	156	0.34158707	0.87562004
CAG15	0.23456552	81	0.34562007	0.87562004
CAG13	-0.152822	96	0.43937189	0.94712344
CAG30	-0.1778415	61	0.42005468	0.94712344
CAG3	-0.0411126	186	0.76915093	0.97724639
CAG4	-0.2023102	130	0.57106468	0.97724639
CAG7	0.14134084	108	0.59028752	0.97724639
CAG8	-0.1361645	181	0.57843794	0.97724639
CAG12	-0.089924	53	0.6508733	0.97724639
CAG20	0.15931492	54	0.55023548	0.97724639
CAG22	0.1277081	167	0.73650434	0.97724639
CAG27	0.14393475	95	0.59844208	0.97724639
CAG32	0.10670355	73	0.71090125	0.97724639
CAG2	0.04131476	164	0.86249784	0.9825587
CAG11	0.02848725	83	0.90696037	0.9825587
CAG14	0.02205553	60	0.92735731	0.98816433
CAG26	0.00648251	57	0.97725674	0.99797335

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.11. MaAslin2 associations between co-abundance groupings and CREDI language scores stratified by breastmilk consumption.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.13554746	77	0.00979462	0.84008705
CAG1	-0.1081833	153	0.59061619	0.99895488
CAG2	0.06646154	131	0.81582112	0.99895488
CAG3	0.00229235	153	0.9884363	0.99895488
CAG4	-0.0711944	119	0.86747049	0.99895488
CAG5	-0.1528694	138	0.58807297	0.99895488
CAG7	0.13109365	82	0.6581805	0.99895488
CAG8	-0.0967329	149	0.72048756	0.99895488
CAG9	0.29117771	13	0.15897072	0.99895488
CAG10	-0.2716468	60	0.40974709	0.99895488
CAG11	0.1490685	69	0.59644589	0.99895488
CAG12	-0.1564106	42	0.42988534	0.99895488
CAG13	-0.1459426	80	0.52056259	0.99895488
CAG14	-0.1790884	51	0.51527191	0.99895488
CAG15	0.20331503	67	0.47857563	0.99895488
CAG16	0.46855607	145	0.16058749	0.99895488
CAG17	0.19139671	114	0.63638162	0.99895488
CAG18	0.42620515	100	0.28031659	0.99895488
CAG19	0.33186388	78	0.27264752	0.99895488
CAG20	-0.17897	37	0.52827939	0.99895488
CAG21	0.24142007	56	0.44033135	0.99895488
CAG22	-0.037968	134	0.93081622	0.99895488
CAG23	0.24689895	76	0.49370195	0.99895488
CAG24	0.33724111	87	0.39895179	0.99895488
CAG25	0.26017768	66	0.42793268	0.99895488
CAG26	-0.2219873	38	0.36063898	0.99895488
CAG27	0.07071584	75	0.81630917	0.99895488
CAG28	0.56070202	66	0.08773298	0.99895488
CAG29	-0.2173936	30	0.32552754	0.99895488
CAG30	-0.258619	48	0.27735835	0.99895488
CAG31	0.04021999	34	0.83485447	0.99895488
CAG32	0.09921957	59	0.75224609	0.99895488
CAG33	-0.077461	54	0.78320131	0.99895488

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.12. MaAslin2 associations between co-abundance groupings and CREDI language scores stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG18	1.95287947	30	4.57E-05	0.02260745
CAG1	-0.5281046	33	0.14940549	0.91148184
CAG4	-0.5069937	11	0.24161864	0.91148184
CAG5	-0.629867	18	0.37651263	0.91148184
CAG6	1.81484176	19	0.02762275	0.91148184
CAG7	0.61531581	26	0.34853479	0.91148184
CAG9	-0.2409913	2	0.29387502	0.91148184
CAG10	-1.5956145	20	0.05605189	0.91148184
CAG11	-0.9007927	14	0.13243441	0.91148184
CAG13	-0.4135532	16	0.37294582	0.91148184
CAG14	0.36859301	9	0.32017587	0.91148184
CAG16	0.19776097	33	0.3730808	0.91148184
CAG17	0.88154478	31	0.10029893	0.91148184
CAG19	0.98489923	27	0.06404937	0.91148184
CAG20	0.92809358	17	0.0705243	0.91148184
CAG21	1.53451984	23	0.03099034	0.91148184
CAG22	1.03160862	33	0.11775142	0.91148184
CAG23	0.43478288	23	0.44847212	0.91148184
CAG24	1.07540297	30	0.11079501	0.91148184
CAG25	0.72282382	27	0.34114002	0.91148184
CAG26	1.13497215	19	0.08753542	0.91148184
CAG27	0.5376788	20	0.35374893	0.91148184
CAG28	0.7695344	25	0.23183862	0.91148184
CAG29	-0.2671846	7	0.39754704	0.91148184
CAG31	1.05048218	18	0.15342566	0.91148184
CAG33	1.37009309	18	0.12148108	0.91148184
CAG3	-0.2695208	33	0.48385181	0.93125287
CAG15	0.28609397	14	0.50764009	0.94187988
CAG2	-0.1919752	33	0.61242673	0.95308568
CAG8	-0.2575752	32	0.68318701	0.97739183
CAG12	-0.1670182	11	0.7959104	0.99108937
CAG30	0.02184044	13	0.95653085	0.99193601
CAG32	0.02929636	14	0.97054991	0.99260786

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.13. MaAslin2 associations between co-abundance groupings and CREDI cognitive scores.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.36676193	96	0.00063617	0.02583839
CAG28	0.78611389	91	0.00897074	0.21529784
CAG10	-0.7382659	80	0.01724893	0.30358112
CAG1	-0.3281102	186	0.06711766	0.5806977
CAG19	0.44215047	105	0.11786448	0.69734331
CAG23	0.50071278	99	0.12242391	0.69734331
CAG16	0.42017649	178	0.13593803	0.70783772
CAG22	0.57235545	167	0.14110292	0.70783772
CAG24	0.49633036	117	0.1603474	0.72985711
CAG12	-0.2741973	53	0.17837754	0.75008827
CAG14	0.33153186	60	0.1812656	0.75008827
CAG17	0.47714337	145	0.1775877	0.75008827
CAG25	0.37665398	93	0.21791124	0.80225775
CAG2	-0.296064	164	0.22623906	0.80752153
CAG21	0.34314215	79	0.24409995	0.80785624
CAG18	0.38249556	130	0.27928402	0.82070555
CAG20	0.29470218	54	0.28172386	0.82070555
CAG32	0.31815533	73	0.28154192	0.82070555
CAG29	-0.2117907	37	0.29870109	0.83269875
CAG27	0.23764734	95	0.39725113	0.89934898
CAG11	-0.1657411	83	0.50797822	0.94738189
CAG15	0.1692303	81	0.50825132	0.94738189
CAG3	0.05303823	186	0.71257832	0.96913406
CAG5	-0.0883575	156	0.74072375	0.96913406
CAG8	-0.1347066	181	0.59277446	0.96913406
CAG9	0.09307363	15	0.60290056	0.96913406
CAG13	0.08561366	96	0.6736075	0.96913406
CAG30	0.12589704	61	0.57884077	0.96913406
CAG31	0.12352957	52	0.54730001	0.96913406
CAG33	0.15551928	72	0.57926964	0.96913406
CAG7	0.04792348	108	0.85907766	0.97212609
CAG26	0.03303018	57	0.88760949	0.97212609
CAG4	0.03053499	130	0.93373456	0.97240996

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.14. MaAslin2 associations between co-abundance groupings and CREDI cognitive scores stratified by breastmilk consumption.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.41619663	77	0.00190055	0.76312948
CAG10	-0.8846325	60	0.00925005	0.76312948
CAG28	0.91461733	66	0.00711055	0.76312948
CAG1	-0.3581379	153	0.08620176	0.99712052
CAG2	-0.1782207	131	0.54904128	0.99712052
CAG3	0.10147049	153	0.53819478	0.99712052
CAG4	0.04660937	119	0.9165797	0.99712052
CAG5	0.11103956	138	0.70609242	0.99712052
CAG7	0.34396481	82	0.26478548	0.99712052
CAG8	-0.0875279	149	0.75625581	0.99712052
CAG9	0.13322806	13	0.53782099	0.99712052
CAG11	-0.0808061	69	0.78323326	0.99712052
CAG12	-0.135127	42	0.51334463	0.99712052
CAG13	0.14660443	80	0.53607016	0.99712052
CAG14	0.27731625	51	0.33358507	0.99712052
CAG15	-0.0050311	67	0.98659899	0.99712052
CAG16	0.5150804	145	0.13901728	0.99712052
CAG17	0.46998295	114	0.26478966	0.99712052
CAG18	0.099504	100	0.80941855	0.99712052
CAG19	0.36201219	78	0.25114353	0.99712052
CAG20	0.01098674	37	0.97041179	0.99712052
CAG21	0.11424259	56	0.72649072	0.99712052
CAG22	0.62834173	134	0.16685476	0.99712052
CAG23	0.40948408	76	0.27580056	0.99712052
CAG24	0.58013083	87	0.16320941	0.99712052
CAG25	0.21227974	66	0.53534496	0.99712052
CAG26	-0.137518	38	0.58752169	0.99712052
CAG27	0.3645721	75	0.249767	0.99712052
CAG29	-0.2110361	30	0.36026023	0.99712052
CAG30	0.02019513	48	0.93531813	0.99712052
CAG31	-0.0837945	34	0.67700713	0.99712052
CAG32	0.20074771	59	0.5400769	0.99712052
CAG33	-0.1657259	54	0.57229654	0.99712052

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.15. MaAslin2 associations between co-abundance groupings and CREDI cognitive scores stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG18	1.88491512	30	0.00053998	0.26729005
CAG2	-0.5920759	33	0.13252317	0.97553972
CAG5	-1.0906457	18	0.14438382	0.97553972
CAG6	0.88095037	19	0.34620639	0.97553972
CAG11	-0.8392036	14	0.19423446	0.97553972
CAG12	-1.0255879	11	0.12463177	0.97553972
CAG15	0.47528313	14	0.29921473	0.97553972
CAG19	0.97394339	27	0.08972418	0.97553972
CAG20	0.65633092	17	0.2456032	0.97553972
CAG21	1.11334642	23	0.15983474	0.97553972
CAG23	0.79011885	23	0.190549	0.97553972
CAG24	0.71184716	30	0.33580035	0.97553972
CAG25	0.95560579	27	0.23624661	0.97553972
CAG26	0.91487317	19	0.20717304	0.97553972
CAG31	1.3389313	18	0.08482365	0.97553972
CAG33	1.23220005	18	0.19819763	0.97553972
CAG1	-0.1018268	33	0.8011116	0.98989022
CAG3	-0.28135	33	0.49494101	0.98989022
CAG4	-0.1659491	11	0.72532317	0.98989022
CAG7	-0.4048169	26	0.56777289	0.98989022
CAG8	-0.18326	32	0.78649564	0.98989022
CAG9	-0.029068	2	0.90738353	0.98989022
CAG10	-0.3260627	20	0.7293389	0.98989022
CAG14	0.14443772	9	0.71944785	0.98989022
CAG16	0.16297951	33	0.49502797	0.98989022
CAG17	0.47925876	31	0.4182449	0.98989022
CAG22	0.45910825	33	0.52827238	0.98989022
CAG28	0.25511925	25	0.71653259	0.98989022
CAG30	0.08719087	13	0.83882764	0.98989022
CAG32	0.57359573	14	0.49651842	0.98989022
CAG13	-0.0153959	16	0.97557654	0.99336858
CAG27	-0.0440697	20	0.94412318	0.99336858
CAG29	0.01392275	7	0.96750448	0.99336858

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.16. MaAslin2 associations between co-abundance groupings and CREDI motor scores.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.09961464	96	0.00989874	0.23332733
CAG28	0.71488039	91	0.0247472	0.41869073
CAG4	0.58063562	130	0.14009831	0.75325652
CAG12	-0.320759	53	0.13604683	0.75325652
CAG3	0.21238963	186	0.16073474	0.80367368
CAG27	0.3963476	95	0.18004597	0.80986798
CAG23	0.44364348	99	0.19479209	0.83984554
CAG2	-0.3100847	164	0.23267923	0.85624152
CAG1	-0.204497	186	0.28621692	0.90404133
CAG7	0.25099421	108	0.37735871	0.90404133
CAG8	-0.2561627	181	0.33425078	0.90404133
CAG10	-0.3221177	80	0.32772173	0.90404133
CAG13	-0.2360864	96	0.27544588	0.90404133
CAG16	0.33957501	178	0.25391482	0.90404133
CAG19	0.29754767	105	0.31919463	0.90404133
CAG24	0.37638629	117	0.31665443	0.90404133
CAG31	-0.2286517	52	0.29049407	0.90404133
CAG32	0.29678718	73	0.34295933	0.90404133
CAG5	-0.2379012	156	0.39779227	0.90463699
CAG17	0.28302571	145	0.44911067	0.92629293
CAG15	0.17113557	81	0.52598233	0.97149721
CAG9	-0.009933	15	0.95803487	0.99074825
CAG11	-0.0435477	83	0.86909268	0.99074825
CAG14	0.12779003	60	0.62582893	0.99074825
CAG18	0.08154776	130	0.82746643	0.99074825
CAG20	0.05128228	54	0.85938209	0.99074825
CAG21	0.1282619	79	0.68015534	0.99074825
CAG22	0.10568279	167	0.79729861	0.99074825
CAG25	-0.0564327	93	0.86162828	0.99074825
CAG26	-0.1001214	57	0.68440129	0.99074825
CAG29	-0.0933772	37	0.66789263	0.99074825
CAG30	0.01697357	61	0.94343501	0.99074825
CAG33	-0.1480762	72	0.61986715	0.99074825

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.
**Number of participants with non-zero feature.

Table S4.17. MaAslin2 associations between co-abundance groupings and CREDI motor scores stratified by breastmilk consumption.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.32798417	77	0.00347869	0.80800597
CAG28	0.72241918	66	0.03341721	0.83353421
CAG1	-0.2231584	153	0.28398071	0.99088598
CAG2	-0.1423762	131	0.6303882	0.99088598
CAG3	0.13853848	153	0.39782689	0.99088598
CAG4	0.52503751	119	0.23429218	0.99088598
CAG5	-0.2960906	138	0.31120786	0.99088598
CAG7	0.36248861	82	0.23722842	0.99088598
CAG8	-0.2136169	149	0.44584692	0.99088598
CAG9	0.05366889	13	0.80305253	0.99088598
CAG10	-0.4279837	60	0.20981264	0.99088598
CAG11	-0.0787865	69	0.78739676	0.99088598
CAG12	-0.2279366	42	0.26690283	0.99088598
CAG13	-0.2282575	80	0.33222836	0.99088598
CAG14	0.25519828	51	0.37109332	0.99088598
CAG15	0.12035273	67	0.68612983	0.99088598
CAG16	0.42900297	145	0.21592097	0.99088598
CAG17	0.28764648	114	0.49318325	0.99088598
CAG18	0.10566703	100	0.79678232	0.99088598
CAG19	0.26184139	78	0.40450125	0.99088598
CAG20	-0.1628036	37	0.58030647	0.99088598
CAG22	0.05672006	134	0.90049801	0.99088598
CAG23	0.26748051	76	0.47469162	0.99088598
CAG24	0.555668	87	0.17943682	0.99088598
CAG25	0.1257381	66	0.71213165	0.99088598
CAG26	-0.1779682	38	0.48014228	0.99088598
CAG27	0.40246966	75	0.20117865	0.99088598
CAG29	-0.1117852	30	0.62646447	0.99088598
CAG30	-0.0984487	48	0.690711	0.99088598
CAG31	-0.2270382	34	0.25549322	0.99088598
CAG32	0.26422265	59	0.41722703	0.99088598
CAG33	-0.164792	54	0.57237961	0.99088598
CAG21	-0.0037036	56	0.99090494	0.99910339

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.18. MaAslin2 associations between co-abundance groupings and CREDI motor scores stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0	P Value	Q Value
CAG1	0.21186095	33	0.71610213	0.99366058
CAG2	-1.2108173	33	0.02664066	0.99366058
CAG3	0.33344442	33	0.57592689	0.99366058
CAG4	0.6658126	11	0.32224137	0.99366058
CAG5	-0.271834	18	0.80689443	0.99366058
CAG6	0.59696456	19	0.66152254	0.99366058
CAG7	-0.7425027	26	0.46623507	0.99366058
CAG8	-0.455499	32	0.64000841	0.99366058
CAG9	-0.0736658	2	0.8380011	0.99366058
CAG10	0.1857212	20	0.89150872	0.99366058
CAG11	-0.4973131	14	0.60167113	0.99366058
CAG12	-1.5982298	11	0.09506036	0.99366058
CAG13	-0.2644105	16	0.71497713	0.99366058
CAG14	-0.5046399	9	0.37973376	0.99366058
CAG15	0.76091841	14	0.2474678	0.99366058
CAG16	0.09821301	33	0.77691122	0.99366058
CAG17	0.21354559	31	0.80431842	0.99366058
CAG18	1.24542085	30	0.16940526	0.99366058
CAG19	1.26808627	27	0.12890692	0.99366058
CAG20	1.02337427	17	0.20789733	0.99366058
CAG21	1.31024846	23	0.25642745	0.99366058
CAG22	1.44859649	33	0.15771847	0.99366058
CAG23	1.97159032	23	0.01628247	0.99366058
CAG24	-0.1180263	30	0.91312808	0.99366058
CAG25	-0.0439641	27	0.97050186	0.99366058
CAG26	0.45166097	19	0.67262969	0.99366058
CAG27	-0.1781416	20	0.84423384	0.99366058
CAG28	0.48641239	25	0.63081751	0.99366058
CAG29	-0.0209891	7	0.96604473	0.99366058
CAG30	0.11390214	13	0.85389534	0.99366058
CAG31	0.2779883	18	0.81233397	0.99366058
CAG32	0.28047287	14	0.81884161	0.99366058
CAG33	0.54211402	18	0.70130408	0.99366058

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.19. MaAslin2 associations between co-abundance groupings and CREDI socioemotional scores.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.25732145	96	0.00236299	0.07453043
CAG28	0.87194369	91	0.00473306	0.13781568
CAG10	-0.752698	80	0.01804272	0.31896951
CAG2	-0.4536759	164	0.07216993	0.60068793
CAG1	-0.3082468	186	0.09796866	0.63531887
CAG12	-0.2907965	53	0.16535587	0.7458196
CAG22	0.47022858	167	0.23962401	0.84724204
CAG24	0.42230826	117	0.2483351	0.85415339
CAG16	0.31706976	178	0.27402054	0.86835634
CAG29	-0.2284245	37	0.28039938	0.87592624
CAG3	0.14250746	186	0.33453272	0.91194942
CAG32	0.26860703	73	0.37816323	0.92668713
CAG14	0.21428829	60	0.40074264	0.93561291
CAG4	0.28838655	130	0.45277993	0.93888586
CAG15	0.20322157	81	0.43917474	0.93888586
CAG19	0.20383979	105	0.48374233	0.94555064
CAG23	0.23740097	99	0.47691019	0.94555064
CAG11	-0.1635955	83	0.52464392	0.94958351
CAG17	0.22856843	145	0.5302736	0.94958351
CAG9	0.10274471	15	0.57600749	0.96980853
CAG25	0.17762523	93	0.57298191	0.96980853
CAG5	-0.0489355	156	0.85837468	0.98123939
CAG7	0.0499863	108	0.85687662	0.98123939
CAG8	-0.1094189	181	0.67221946	0.98123939
CAG13	0.06781602	96	0.74804766	0.98123939
CAG18	0.09944678	130	0.78487445	0.98123939
CAG20	0.05858174	54	0.83535295	0.98123939
CAG27	0.06707152	95	0.81623781	0.98123939
CAG26	-0.0258805	57	0.9140927	0.98423979
CAG31	0.01834878	52	0.93069461	0.98449864
CAG33	-0.0183653	72	0.94963984	0.98449864
CAG21	0.00314881	79	0.99170921	0.99444792
CAG30	-0.0022507	61	0.99229069	0.99444792

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.20. MaAslin2 associations between co-abundance groupings and CREDI socioemotional scores stratified by breastmilk consumption.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG6	1.57414533	77	0.00065327	0.32336779
CAG10	-0.9338433	60	0.0068082	0.61686713
CAG28	0.98609124	66	0.00423277	0.61686713
CAG1	-0.3122964	153	0.14154515	0.99668099
CAG2	-0.2942794	131	0.32976708	0.99668099
CAG3	0.17131273	153	0.3059249	0.99668099
CAG4	0.08410043	119	0.85246707	0.99668099
CAG5	0.04716108	138	0.87480029	0.99668099
CAG7	0.31695826	82	0.31215924	0.99668099
CAG8	-0.0678067	149	0.81297546	0.99668099
CAG9	0.11208604	13	0.6101113	0.99668099
CAG11	-0.0794109	69	0.79022449	0.99668099
CAG12	-0.1404117	42	0.50392211	0.99668099
CAG13	0.14234847	80	0.55443289	0.99668099
CAG14	0.17256373	51	0.55425734	0.99668099
CAG15	0.07168647	67	0.81381047	0.99668099
CAG16	0.38571718	145	0.27654215	0.99668099
CAG17	0.13501428	114	0.75310044	0.99668099
CAG18	-0.1390825	100	0.74009291	0.99668099
CAG19	0.15232594	78	0.63537298	0.99668099
CAG20	-0.0918709	37	0.76018517	0.99668099
CAG21	-0.204278	56	0.53807118	0.99668099
CAG22	0.6268041	134	0.17487506	0.99668099
CAG23	0.17250578	76	0.65204024	0.99668099
CAG24	0.40866983	87	0.33469399	0.99668099
CAG25	0.02839776	66	0.93501538	0.99668099
CAG26	-0.1658466	38	0.51969195	0.99668099
CAG27	0.24044343	75	0.45577035	0.99668099
CAG29	-0.3450572	30	0.14008151	0.99668099
CAG30	-0.0691019	48	0.78464471	0.99668099
CAG31	-0.1704935	34	0.4038613	0.99668099
CAG32	0.08354933	59	0.80199588	0.99668099
CAG33	-0.237178	54	0.42625023	0.99668099

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.21. MaAslin2 associations between co-abundance groupings and CREDI socioemotional scores stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0*	P Value	Q Value
CAG2	-0.8102119	33	0.04418124	0.9889216
CAG3	-0.1396342	33	0.74839041	0.9889216
CAG4	0.23844053	11	0.63014985	0.9889216
CAG5	-0.9202707	18	0.24641677	0.9889216
CAG6	-0.1821763	19	0.85471305	0.9889216
CAG7	-0.6603643	26	0.37139974	0.9889216
CAG8	-0.2678103	32	0.70593408	0.9889216
CAG9	0.06784094	2	0.79581796	0.9889216
CAG10	0.25440985	20	0.79728238	0.9889216
CAG11	-0.8240462	14	0.22635366	0.9889216
CAG12	-1.3296673	11	0.05299506	0.9889216
CAG13	0.37466392	16	0.47463521	0.9889216
CAG14	0.08546192	9	0.83983064	0.9889216
CAG15	0.62267677	14	0.19106647	0.9889216
CAG16	0.21665074	33	0.38568439	0.9889216
CAG17	0.33736691	31	0.58946236	0.9889216
CAG18	1.57368425	30	0.01073055	0.9889216
CAG19	0.60759853	27	0.32729884	0.9889216
CAG20	0.33788616	17	0.5752189	0.9889216
CAG21	0.42698184	23	0.61695374	0.9889216
CAG22	-0.1276895	33	0.8680902	0.9889216
CAG23	0.73611711	23	0.24823152	0.9889216
CAG24	0.50769126	30	0.51653701	0.9889216
CAG25	0.76336101	27	0.3719754	0.9889216
CAG26	0.66507138	19	0.38862848	0.9889216
CAG27	-0.614045	20	0.34595425	0.9889216
CAG29	0.15334534	7	0.66832624	0.9889216
CAG31	0.87073042	18	0.29949863	0.9889216
CAG32	0.5349101	14	0.54655703	0.9889216
CAG33	1.21252673	18	0.2294096	0.9889216
CAG28	0.08041042	25	0.91337786	0.9914957
CAG1	0.01785822	33	0.96647493	0.99390269
CAG30	-0.0245654	13	0.95650829	0.99390269

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, stunting, play activities score, and poverty score.

**Number of participants with non-zero feature.

Table S4.22. MaAslin2 associations between co-abundance groupings and stunting.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG15	1.38608703	79	0.00312443	0.08019356
CAG7	-1.041129	100	0.02678339	0.34079071
CAG33	-0.8917476	67	0.07178883	0.44924363
CAG29	-0.616396	30	0.07679781	0.45487936
CAG16	-0.7012827	163	0.16327324	0.67956968
CAG1	0.40909539	170	0.19937416	0.75500707
CAG6	0.91272068	88	0.20355895	0.75504339
CAG30	-0.4919871	51	0.20428663	0.75504339
CAG31	0.4262024	48	0.22265416	0.77928957
CAG23	-0.6401291	89	0.25785586	0.82177575
CAG22	0.72287546	154	0.27996527	0.83016239
CAG19	-0.4847626	95	0.34009179	0.87625219
CAG10	-0.4799429	73	0.38681966	0.89781875
CAG8	-0.3342644	166	0.41405704	0.91092549
CAG32	0.35877994	64	0.4811064	0.95806533
CAG13	0.25669625	90	0.48645186	0.9604306
CAG3	-0.0583457	170	0.80198287	0.98747577
CAG4	0.29687949	119	0.64318263	0.98747577
CAG5	0.23746161	142	0.61473468	0.98747577
CAG9	-0.0958027	14	0.7713715	0.98747577
CAG11	0.17159016	77	0.69148803	0.98747577
CAG12	-0.0718381	50	0.84062628	0.98747577
CAG14	-0.250375	57	0.57113303	0.98747577
CAG17	-0.3618722	133	0.56024878	0.98747577
CAG18	-0.3178426	118	0.6089873	0.98747577
CAG20	-0.182582	50	0.70764653	0.98747577
CAG21	-0.1449491	71	0.77820955	0.98747577
CAG24	-0.1769378	105	0.77641011	0.98747577
CAG25	0.27379426	83	0.60740063	0.98747577
CAG26	-0.1341503	51	0.73858659	0.98747577
CAG27	-0.1348262	88	0.77678678	0.98747577
CAG28	-0.1565432	83	0.77117595	0.98747577
CAG2	0.02924086	149	0.94748444	0.99764803

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.
**Number of participants with non-zero feature.

Table S4.23. MaAslin2 associations between co-abundance groupings and stunting stratified by breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG15	1.45535253	65	0.00657235	0.5793454
CAG7	-1.1727811	76	0.02564114	0.61913895
CAG33	-0.9879918	51	0.05059674	0.64795793
CAG16	-0.8964384	133	0.13693562	0.86107821
CAG29	-0.5614413	24	0.14171546	0.86107821
CAG1	0.30538399	140	0.39681345	0.95931082
CAG6	0.78894987	70	0.31705206	0.95931082
CAG8	-0.2900299	137	0.50102942	0.95931082
CAG11	0.42515386	64	0.391293	0.95931082
CAG12	0.2230581	39	0.51126675	0.95931082
CAG17	-0.6528315	105	0.36443328	0.95931082
CAG19	-0.370769	71	0.49447086	0.95931082
CAG20	-0.4212765	34	0.40528698	0.95931082
CAG22	0.92185274	124	0.22423592	0.95931082
CAG23	-0.701338	69	0.27188	0.95931082
CAG27	-0.3116941	69	0.55265149	0.95931082
CAG28	-0.4737986	61	0.42798683	0.95931082
CAG30	-0.406196	40	0.3352434	0.95931082
CAG18	-0.4017179	91	0.57085326	0.9740563
CAG2	0.15204459	119	0.77239772	0.98015132
CAG5	0.073299	126	0.88406203	0.98015132
CAG9	0.09431445	12	0.80577054	0.98015132
CAG10	-0.1931463	55	0.73880944	0.98015132
CAG13	0.14703104	75	0.72692398	0.98015132
CAG14	-0.0743577	48	0.87933309	0.98015132
CAG21	-0.1446099	51	0.79054086	0.98015132
CAG24	-0.2410644	78	0.73401125	0.98015132
CAG25	0.22150621	59	0.70528735	0.98015132
CAG26	-0.1147402	34	0.78080464	0.98015132
CAG31	0.11602378	31	0.72774445	0.98015132
CAG32	0.17474853	52	0.75257263	0.98015132
CAG3	-0.0303719	140	0.90298097	0.9838302
CAG4	-0.0502103	109	0.94654363	0.9838302

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.24. MaAslin2 associations between co-abundance groupings and stunting stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG1	1.54365746	30	0.12195563	0.89707901
CAG2	-1.5349527	30	0.06795842	0.89707901
CAG4	1.31006787	10	0.22243807	0.89707901
CAG5	2.04756851	16	0.28558716	0.89707901
CAG7	-1.6861413	24	0.35669562	0.89707901
CAG12	-1.8097047	11	0.26439235	0.89707901
CAG15	1.17242069	14	0.29933282	0.89707901
CAG22	-1.6386945	30	0.35784593	0.89707901
CAG29	-0.9981463	6	0.21163003	0.89707901
CAG30	-0.9831395	11	0.2800257	0.89707901
CAG18	-0.971477	27	0.38769393	0.90468488
CAG27	-1.1306905	19	0.3826815	0.90468488
CAG32	1.72449947	12	0.37151223	0.90468488
CAG19	-1.2024233	24	0.41047371	0.90688604
CAG13	0.92664268	15	0.4254957	0.91629483
CAG11	1.02557867	13	0.43864606	0.91686827
CAG16	-0.4491533	30	0.45399643	0.92788139
CAG24	-1.2720183	27	0.44914782	0.92788139
CAG17	0.97320461	28	0.47051103	0.93448719
CAG21	-1.457573	20	0.46790908	0.93448719
CAG10	1.2708484	18	0.55441001	0.96416674
CAG8	-0.7241374	29	0.61940404	0.97121845
CAG14	-0.3855575	9	0.69376333	0.97121845
CAG20	0.47281661	16	0.75094681	0.97121845
CAG23	-0.4447066	20	0.73737277	0.97121845
CAG25	0.65730094	24	0.72844271	0.97121845
CAG28	0.68606386	22	0.6871965	0.97121845
CAG33	-1.0451314	16	0.65157234	0.97121845
CAG26	0.53197391	17	0.77712227	0.97357018
CAG31	0.45902725	17	0.78521044	0.97357018
CAG6	0.47776093	18	0.86483663	0.99248865
CAG9	0.07029382	2	0.90186665	0.99693873
CAG3	-0.1014905	30	0.91884781	0.99792946

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.25. MaAslin2 associations between co-abundance groupings and Z-BMI.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG10	0.59195351	71	0.03108365	0.30990089
CAG13	-0.3492123	89	0.05601328	0.4188708
CAG1	-0.2322668	166	0.14358525	0.64404257
CAG21	-0.3344753	69	0.19105885	0.74804399
CAG2	-0.2881333	145	0.19361731	0.75049664
CAG3	0.12754218	166	0.26620642	0.84222891
CAG6	-0.2946453	86	0.40714566	0.90715631
CAG18	-0.2500811	115	0.4208864	0.90864259
CAG14	-0.1558229	56	0.47805525	0.94489206
CAG22	0.23801054	150	0.47145369	0.94489206
CAG4	-0.1925308	117	0.54614615	0.9767135
CAG5	0.14701665	139	0.53013376	0.9767135
CAG11	-0.13129	76	0.54032688	0.9767135
CAG24	0.19807059	102	0.52411273	0.9767135
CAG33	-0.156281	67	0.52741295	0.9767135
CAG8	-0.1074148	162	0.6000132	0.9854285
CAG29	-0.0923212	30	0.59756761	0.9854285
CAG7	-0.0519909	97	0.82414926	0.99028844
CAG9	0.01576216	14	0.92390114	0.99028844
CAG12	0.01477275	49	0.93295834	0.99028844
CAG15	0.02995474	76	0.89972484	0.99028844
CAG17	-0.0714138	130	0.8174901	0.99028844
CAG19	-0.1123029	92	0.65620613	0.99028844
CAG20	0.08502152	47	0.72142794	0.99028844
CAG23	-0.0904361	87	0.74545342	0.99028844
CAG25	-0.0876055	80	0.74056979	0.99028844
CAG26	-0.040604	50	0.83636567	0.99028844
CAG27	-0.0433421	86	0.8548803	0.99028844
CAG28	-0.042957	80	0.87033034	0.99028844
CAG30	-0.0841722	49	0.66177738	0.99028844
CAG31	0.06622088	48	0.70054963	0.99028844
CAG32	-0.0031532	61	0.98998386	0.99428814
CAG16	0.00091029	159	0.99711628	0.99711628

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.
**Number of participants with non-zero feature.

Table S4.26. MaAslin2 associations between co-abundance groupings and Z-BMI stratified by breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG13	-0.4363973	75	0.03690635	0.71967376
CAG1	-0.2747586	138	0.12644686	0.9320964
CAG10	0.41203409	54	0.15516255	0.9320964
CAG2	-0.3116025	117	0.23768326	0.94104355
CAG3	0.1589114	138	0.20356756	0.94104355
CAG21	-0.3122721	51	0.25168467	0.94104355
CAG30	-0.2670887	39	0.2076009	0.94104355
CAG4	-0.2975368	107	0.42791931	0.97416634
CAG11	-0.2181004	64	0.37844066	0.97416634
CAG14	-0.1885838	48	0.44364172	0.97416634
CAG22	0.27894785	122	0.45941482	0.97416634
CAG24	0.3258249	77	0.35882158	0.97416634
CAG31	-0.1316501	31	0.42897317	0.97416634
CAG26	0.13459113	34	0.51555222	0.9813989
CAG32	-0.1611163	50	0.56014411	0.9813989
CAG8	-0.1161545	135	0.59152899	0.98453705
CAG5	0.06428304	124	0.79818401	0.986356
CAG6	-0.1349271	69	0.73152333	0.986356
CAG7	-0.0925078	75	0.72675939	0.986356
CAG9	0.06933771	12	0.71916692	0.986356
CAG15	0.0825314	64	0.76244306	0.986356
CAG16	-0.1068185	131	0.72479017	0.986356
CAG18	-0.0813274	90	0.81962181	0.986356
CAG20	0.08246795	33	0.74585941	0.986356
CAG23	0.11858643	68	0.70932974	0.986356
CAG25	0.04258979	58	0.88484758	0.986356
CAG27	-0.1331396	68	0.61408248	0.986356
CAG28	-0.1439569	59	0.62679028	0.986356
CAG29	-0.0676711	24	0.72555748	0.986356
CAG12	-0.0191063	39	0.91051715	0.99001689
CAG17	-0.0265649	104	0.94118293	0.99023846
CAG19	0.02076581	70	0.93881255	0.99023846
CAG33	0.0017596	51	0.99448154	0.99913953

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.27. MaAslin2 associations between co-abundance groupings and Z-BMI stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG6	-2.618183	17	0.01635238	0.65139859
CAG18	-1.0248548	25	0.02781133	0.65139859
CAG22	-1.4062612	28	0.04565518	0.65139859
CAG23	-0.9857481	19	0.04519291	0.65139859
CAG10	1.4667919	17	0.10423059	0.69867068
CAG19	-0.9271584	22	0.15631576	0.77756934
CAG11	0.79419346	12	0.18114	0.79748579
CAG21	-0.8987764	18	0.27725127	0.88903563
CAG30	0.39170761	10	0.27670715	0.88903563
CAG5	0.81149898	15	0.35383012	0.89864681
CAG14	-0.3134626	8	0.41046584	0.89864681
CAG24	-0.7194808	25	0.35037139	0.89864681
CAG25	-0.6683714	22	0.40532905	0.89864681
CAG26	-0.7524674	16	0.32899161	0.89864681
CAG31	0.66024388	17	0.32416905	0.89864681
CAG33	-0.9180701	16	0.36534368	0.89864681
CAG1	0.28139823	28	0.54948164	0.91871404
CAG2	0.23632592	28	0.55623499	0.91871404
CAG20	-0.3841525	14	0.55944591	0.91871404
CAG8	-0.2979232	27	0.63712312	0.95427047
CAG3	-0.0973007	28	0.81259203	0.95615726
CAG12	0.21466503	10	0.75625501	0.95615726
CAG13	-0.1723236	14	0.72671795	0.95615726
CAG15	-0.1520029	12	0.76924694	0.95615726
CAG16	-0.0763147	28	0.77924691	0.95615726
CAG17	-0.1411062	26	0.81882088	0.95615726
CAG28	0.25461302	21	0.70517761	0.95615726
CAG32	0.17980708	11	0.82358106	0.95615726
CAG27	0.07986647	18	0.8604541	0.96379845
CAG7	-0.1156387	22	0.87628705	0.96639368
CAG4	0.06163219	10	0.9009847	0.97053726
CAG9	0.03599019	2	0.88503934	0.97053726
CAG29	0.01515346	6	0.96795569	0.9886976

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.28. MaAslin2 associations between co-abundance groupings and Z-head-circumference.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG20	0.59666642	50	0.01299794	0.23096341
CAG7	0.57399785	100	0.01408486	0.23840284
CAG6	-0.5216299	88	0.1443722	0.64219544
CAG1	-0.2179114	170	0.16988803	0.68913733
CAG11	0.28851372	77	0.17933636	0.70270213
CAG16	0.32160581	163	0.19962321	0.74375746
CAG32	-0.3258665	64	0.19818381	0.74375746
CAG13	-0.2128588	90	0.24607561	0.85732407
CAG10	0.31523537	73	0.25350236	0.85764737
CAG12	0.1822989	50	0.30493745	0.85764737
CAG22	-0.3445576	154	0.30140137	0.85764737
CAG21	-0.229481	71	0.37023459	0.86625953
CAG24	0.26102053	105	0.39992651	0.90571591
CAG14	0.17227869	57	0.43387633	0.93200399
CAG28	-0.2030649	83	0.4485602	0.93200399
CAG19	-0.1836916	95	0.4683872	0.93325606
CAG31	-0.1177949	48	0.49944381	0.94697584
CAG27	-0.1521463	88	0.52056358	0.9620015
CAG5	-0.1403637	142	0.55034245	0.97045118
CAG30	-0.1190202	51	0.53845571	0.97045118
CAG3	-0.0676298	170	0.55942138	0.97898742
CAG2	0.06351229	149	0.77397434	0.98105024
CAG4	-0.1581362	119	0.62039368	0.98105024
CAG8	0.04013578	166	0.84410906	0.98105024
CAG15	-0.0451876	79	0.84877695	0.98105024
CAG17	0.09867816	133	0.74997527	0.98105024
CAG18	0.06367844	118	0.83709032	0.98105024
CAG25	0.11329041	83	0.66963924	0.98105024
CAG26	-0.1082377	51	0.58878512	0.98105024
CAG33	0.1294042	67	0.60175021	0.98105024
CAG9	0.00951507	14	0.9538138	0.99899131
CAG23	0.0068432	89	0.9806712	0.99899131
CAG29	-0.0211279	30	0.90358876	0.99899131

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.29. MaAslin2 associations between co-abundance groupings and Z-head-circumference stratified by breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG7	0.60258031	76	0.0222288	0.62186964
CAG20	0.47482632	34	0.06009021	0.73720976
CAG6	-0.6849036	70	0.08218463	0.80271126
CAG32	-0.4320775	52	0.11856684	0.80271126
CAG13	-0.3182422	75	0.13026795	0.83410373
CAG11	0.29355277	64	0.23748163	0.95682417
CAG1	-0.1505881	140	0.40499666	0.96992447
CAG10	0.22943964	55	0.42938425	0.96992447
CAG12	0.15561494	39	0.36066756	0.96992447
CAG15	-0.1668261	65	0.54030507	0.96992447
CAG16	0.33437179	133	0.2698019	0.96992447
CAG19	-0.235798	71	0.3862136	0.96992447
CAG21	-0.193337	51	0.47871319	0.96992447
CAG22	-0.3633609	124	0.34023171	0.96992447
CAG24	0.26953913	78	0.44848873	0.96992447
CAG25	0.21139898	59	0.47157266	0.96992447
CAG27	-0.1717481	69	0.51426196	0.96992447
CAG28	-0.2530697	61	0.39868665	0.96992447
CAG33	0.14540958	51	0.56903593	0.96992447
CAG2	0.119869	119	0.64939543	0.99183273
CAG3	-0.0338536	140	0.78651739	0.99183273
CAG4	-0.0712044	109	0.84968478	0.99183273
CAG5	-0.0607038	126	0.80977607	0.99183273
CAG14	0.05022952	48	0.83803608	0.99183273
CAG17	0.13386771	105	0.71125929	0.99183273
CAG18	0.08880951	91	0.80288484	0.99183273
CAG26	0.01728454	34	0.93342447	0.99183273
CAG30	-0.076561	40	0.71778566	0.99183273
CAG31	-0.0336926	31	0.84033943	0.99183273
CAG8	0.00395121	137	0.98543621	0.99397869
CAG9	0.00688799	12	0.97145204	0.99397869
CAG23	0.00353588	69	0.99121211	0.99397869
CAG29	-0.0089089	24	0.96309183	0.99397869

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.30. MaAslin2 associations between co-abundance groupings and Z-head-circumference stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG14	0.90931802	9	0.07409814	0.98036008
CAG1	-0.5932559	30	0.28559936	0.9888884
CAG3	-0.4265459	30	0.42801475	0.9888884
CAG4	-0.2589381	10	0.66516664	0.9888884
CAG5	-0.3343588	16	0.7533341	0.9888884
CAG6	1.79564256	18	0.23041274	0.9888884
CAG7	0.99069753	24	0.31991034	0.9888884
CAG8	0.95260951	29	0.22185492	0.9888884
CAG10	-1.0848132	18	0.3510525	0.9888884
CAG11	0.2968847	13	0.68327512	0.9888884
CAG15	0.42895789	14	0.49070734	0.9888884
CAG16	0.12878055	30	0.69590861	0.9888884
CAG18	0.60680496	27	0.32091487	0.9888884
CAG19	-0.4085497	24	0.61072118	0.9888884
CAG20	0.65803054	16	0.41396811	0.9888884
CAG21	-0.5410486	20	0.62305986	0.9888884
CAG22	-0.2270461	30	0.8177022	0.9888884
CAG23	-0.4068062	20	0.57273383	0.9888884
CAG24	0.58382434	27	0.52552933	0.9888884
CAG25	0.4017777	24	0.69718906	0.9888884
CAG27	-0.5176238	19	0.46558977	0.9888884
CAG28	-0.7776506	22	0.39873807	0.9888884
CAG29	0.23015728	6	0.60516258	0.9888884
CAG30	-0.3058214	11	0.54320691	0.9888884
CAG31	-0.2791521	17	0.76126443	0.9888884
CAG32	0.41679381	12	0.69526815	0.9888884
CAG2	-0.044739	30	0.92632023	0.99347845
CAG12	0.16897111	11	0.8514378	0.99347845
CAG13	0.11340568	15	0.85950931	0.99347845
CAG17	-0.0705117	28	0.92429181	0.99347845
CAG9	0.00495652	2	0.9872908	0.99588054
CAG26	-0.0629656	17	0.95110049	0.99588054
CAG33	-0.021546	16	0.98643302	0.99588054

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.31. MaAslin2 associations between co-abundance groupings and Z-length-for-age.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG33	0.51369576	67	0.03882726	0.33845651
CAG31	-0.3276173	48	0.06144244	0.42359801
CAG8	0.37642348	166	0.06618017	0.42465609
CAG7	0.40802791	100	0.08530962	0.49267553
CAG15	-0.4056863	79	0.08836991	0.49267553
CAG6	-0.5759086	88	0.11006573	0.52969135
CAG14	0.30119809	57	0.1743542	0.70659333
CAG2	0.28542239	149	0.19985211	0.73008364
CAG29	0.22376957	30	0.20250572	0.73091909
CAG30	0.24575885	51	0.20730577	0.73926915
CAG10	0.32476153	73	0.24373479	0.79601194
CAG12	0.20779017	50	0.24638465	0.79601194
CAG16	0.29513486	163	0.24372878	0.79601194
CAG32	-0.2830586	64	0.26843075	0.80317374
CAG22	-0.3649103	154	0.27808706	0.81314064
CAG25	-0.2542508	83	0.34220794	0.87240947
CAG1	-0.1354677	170	0.39875576	0.88347719
CAG26	-0.1708545	51	0.3975019	0.88347719
CAG5	-0.1789726	142	0.45036785	0.8966438
CAG3	0.05516995	170	0.63713827	0.97221774
CAG11	0.09438918	77	0.66418595	0.97221774
CAG13	-0.0949977	90	0.60863333	0.97221774
CAG27	0.09301665	88	0.69726222	0.97221774
CAG4	0.11917865	119	0.71154224	0.97327917
CAG9	0.06152233	14	0.71053958	0.97327917
CAG17	0.0596389	133	0.84867337	0.99159247
CAG18	0.02508007	118	0.93604987	0.99159247
CAG19	-0.0401445	95	0.87533496	0.99159247
CAG20	0.04415723	50	0.85688683	0.99159247
CAG21	-0.0220067	71	0.93223757	0.99159247
CAG23	0.03378972	89	0.9056453	0.99159247
CAG24	-0.0901408	105	0.77355143	0.99159247
CAG28	-0.0016665	83	0.99508786	0.99700237

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.
**Number of participants with non-zero feature.

Table S4.32. MaAslin2 associations between co-abundance groupings and Z-length-for-age stratified by breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG33	0.5184318	51	0.04081295	0.68225996
CAG7	0.41862951	76	0.11423653	0.85627476
CAG1	-0.1107011	140	0.54079908	0.96401036
CAG2	0.23402219	119	0.37437111	0.96401036
CAG4	0.37517613	109	0.31719797	0.96401036
CAG6	-0.5253244	70	0.18368661	0.96401036
CAG8	0.29402067	137	0.17288459	0.96401036
CAG10	0.16887376	55	0.56113831	0.96401036
CAG13	-0.1602629	75	0.44776148	0.96401036
CAG14	0.20076409	48	0.41347344	0.96401036
CAG15	-0.3059253	65	0.26075266	0.96401036
CAG16	0.39418925	133	0.19298031	0.96401036
CAG22	-0.4307224	124	0.25804072	0.96401036
CAG25	-0.1699405	59	0.56301731	0.96401036
CAG26	-0.2474363	34	0.23054451	0.96401036
CAG27	0.23317392	69	0.37568566	0.96401036
CAG28	0.17037258	61	0.57031738	0.96401036
CAG29	0.17601117	24	0.35993113	0.96401036
CAG30	0.27028976	40	0.200744	0.96401036
CAG31	-0.103453	31	0.53602645	0.96401036
CAG32	-0.227174	52	0.41362576	0.96401036
CAG20	0.13043303	34	0.60796785	0.96622068
CAG3	0.05128994	140	0.68158923	0.97398535
CAG9	-0.0283523	12	0.88291662	0.97398535
CAG11	-0.0751553	64	0.76297942	0.97398535
CAG12	-0.0289259	39	0.8653294	0.97398535
CAG17	0.16823877	105	0.64182156	0.97398535
CAG18	0.1060566	91	0.76568339	0.97398535
CAG19	-0.0523119	71	0.84781373	0.97398535
CAG21	0.06893848	51	0.80081167	0.97398535
CAG23	0.14700703	69	0.64695906	0.97398535
CAG24	-0.0139696	78	0.96870969	0.99105744
CAG5	0.0061832	126	0.9804463	0.99521221

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.33. MaAslin2 associations between co-abundance groupings and Z-length-for-age stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG2	1.00807676	30	0.00930966	0.68711354
CAG12	1.58312155	11	0.03452426	0.68711354
CAG8	1.31760258	29	0.0487796	0.70757473
CAG5	-1.6556663	16	0.06552393	0.75790298
CAG14	0.75496063	9	0.0982013	0.75790298
CAG15	-0.8895096	14	0.09592459	0.75790298
CAG4	-0.6638386	10	0.20219332	0.9218335
CAG7	0.89130063	24	0.31469325	0.9218335
CAG17	-0.6132767	28	0.34641991	0.9218335
CAG20	-0.6928468	16	0.33186119	0.9218335
CAG23	-0.8098622	20	0.19753554	0.9218335
CAG29	0.39432208	6	0.31399523	0.9218335
CAG32	-0.9344668	12	0.31732983	0.9218335
CAG1	-0.3781742	30	0.44867928	0.92241501
CAG28	-0.6303262	22	0.44324421	0.92241501
CAG31	-0.625716	17	0.44063276	0.92241501
CAG9	0.16956274	2	0.53800733	0.97239482
CAG3	-0.1463813	30	0.76198806	0.9771965
CAG6	-0.4065465	18	0.76537176	0.9771965
CAG10	-0.5798747	18	0.57900102	0.9771965
CAG13	0.22919931	15	0.68721288	0.9771965
CAG16	0.15489727	30	0.59673025	0.9771965
CAG18	0.10650895	27	0.84718219	0.9771965
CAG19	-0.1227376	24	0.86412288	0.9771965
CAG22	0.39922656	30	0.64812327	0.9771965
CAG24	0.38523279	27	0.63899137	0.9771965
CAG25	-0.2067502	24	0.82234278	0.9771965
CAG27	-0.0929708	19	0.88392927	0.9771965
CAG30	0.06036949	11	0.89337668	0.9771965
CAG33	0.12456107	16	0.91206345	0.97953876
CAG11	0.0315963	13	0.96120551	0.99809636
CAG21	-0.0093359	20	0.99242916	0.99809636
CAG26	0.02378124	17	0.97924675	0.99809636

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.34. MaAslin2 associations between co-abundance groupings and Z-weight-for-age.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG10	0.7152359	71	0.01172667	0.22573847
CAG6	-0.7128224	86	0.05181798	0.38340096
CAG13	-0.3078947	89	0.10465297	0.56220551
CAG1	-0.2654302	166	0.10658831	0.5660207
CAG15	-0.3323792	76	0.1760617	0.70730875
CAG33	0.33914376	67	0.18482243	0.71754592
CAG3	0.15019247	166	0.20626771	0.75727755
CAG8	0.24920034	162	0.23966338	0.79796522
CAG31	-0.2123902	48	0.2330666	0.79796522
CAG12	0.20774928	49	0.25271011	0.79995939
CAG7	0.27097412	97	0.26289992	0.80436928
CAG21	-0.2817609	69	0.28844436	0.82032826
CAG25	-0.2794846	80	0.30742761	0.82032826
CAG26	-0.2075537	50	0.30760713	0.82032826
CAG32	-0.2704578	61	0.29803241	0.82032826
CAG16	0.25080404	159	0.33596278	0.84647059
CAG29	0.13837177	30	0.44503622	0.89103392
CAG30	0.13200433	49	0.50788098	0.92036554
CAG18	-0.17102	115	0.59562673	0.95279683
CAG19	-0.1392594	92	0.59429264	0.95279683
CAG22	-0.1780926	150	0.60332802	0.95279683
CAG14	0.10403898	56	0.64785137	0.95689886
CAG9	0.06934751	14	0.68507711	0.96555044
CAG2	0.03662087	145	0.87370386	0.98916467
CAG4	-0.0483613	117	0.88378302	0.98916467
CAG5	-0.0300772	139	0.901448	0.98916467
CAG11	-0.0246264	76	0.91182379	0.98916467
CAG17	-0.0309456	130	0.9231485	0.98916467
CAG20	0.06546307	47	0.79114305	0.98916467
CAG23	-0.089413	87	0.75682885	0.98916467
CAG24	0.08499532	102	0.79211034	0.98916467
CAG27	0.02719367	86	0.9118522	0.98916467
CAG28	-0.0446145	80	0.87007692	0.98916467

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.
**Number of participants with non-zero feature.

Table S4.35. MaAslin2 associations between co-abundance groupings and Z-weight-for-age stratified by breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG13	-0.4299812	75	0.04484952	0.73717698
CAG33	0.4405752	51	0.08901084	0.88803838
CAG1	-0.2765624	138	0.13319963	0.88912122
CAG10	0.44700768	54	0.13196323	0.88912122
CAG3	0.15408513	138	0.22889292	0.97390538
CAG6	-0.508963	69	0.20502529	0.97390538
CAG7	0.26623015	75	0.32535543	0.97390538
CAG8	0.16834721	135	0.44739626	0.97390538
CAG11	-0.2112839	64	0.40490682	0.97390538
CAG15	-0.1992122	64	0.47588579	0.97390538
CAG16	0.2516287	131	0.41761333	0.97390538
CAG20	0.16435065	33	0.52805521	0.97390538
CAG24	0.22919289	77	0.52895094	0.97390538
CAG31	-0.1722745	31	0.31186909	0.97390538
CAG32	-0.3020048	50	0.28553837	0.97390538
CAG4	0.06718943	107	0.86142045	0.97454174
CAG5	0.04855385	124	0.85044975	0.97454174
CAG12	-0.0303054	39	0.86183146	0.97454174
CAG17	0.09560439	104	0.79541984	0.97454174
CAG19	-0.0492924	70	0.85879757	0.97454174
CAG21	-0.1637585	51	0.55806443	0.97454174
CAG22	-0.1733586	122	0.65374628	0.97454174
CAG23	0.17928633	68	0.58205477	0.97454174
CAG25	-0.123906	58	0.68071828	0.97454174
CAG26	-0.114404	34	0.58957503	0.97454174
CAG27	0.08827066	68	0.7441988	0.97454174
CAG28	0.05749276	59	0.84968727	0.97454174
CAG29	0.10487593	24	0.59519725	0.97454174
CAG30	0.03665569	39	0.86633118	0.97454174
CAG2	-0.024521	117	0.92791765	0.98314755
CAG9	0.02523443	12	0.89836179	0.98314755
CAG18	0.02918793	90	0.93632264	0.98314755
CAG14	0.01381869	48	0.95634641	0.98386717

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.36. MaAslin2 associations between co-abundance groupings and Z-weight-for-age stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG2	1.03959633	28	0.00864236	0.65019886
CAG12	1.44309284	10	0.0427342	0.65019886
CAG23	-1.419086	19	0.00552865	0.65019886
CAG6	-2.2831006	17	0.07033529	0.68610659
CAG15	-0.8458712	12	0.12430502	0.70141805
CAG8	0.85049606	27	0.21269796	0.82576436
CAG18	-0.6851784	25	0.21029291	0.82576436
CAG19	-0.8370953	22	0.25390933	0.84681009
CAG20	-0.8152206	14	0.25498601	0.84681009
CAG14	0.45490071	8	0.27627866	0.85268739
CAG4	-0.5462067	10	0.31033378	0.88167676
CAG30	0.40623701	10	0.30970457	0.88167676
CAG11	0.66871207	12	0.31620864	0.89002674
CAG7	0.75201691	22	0.35315577	0.90762368
CAG17	-0.5973357	26	0.374745	0.90762368
CAG22	-0.6869909	28	0.40623014	0.90762368
CAG29	0.33956066	6	0.41044055	0.90762368
CAG5	-0.7936118	15	0.4145754	0.90764713
CAG25	-0.6859517	22	0.44143887	0.90764713
CAG3	-0.2762459	28	0.54093079	0.933152
CAG9	0.17228421	2	0.52911792	0.933152
CAG10	0.61251354	17	0.55687495	0.933152
CAG21	-0.6084571	18	0.51242328	0.933152
CAG26	-0.4501978	16	0.60252472	0.933152
CAG32	-0.558029	11	0.52947725	0.933152
CAG33	-0.6030278	16	0.59499305	0.933152
CAG28	-0.2686518	21	0.71855592	0.96203724
CAG1	-0.1058482	28	0.8398335	0.9739225
CAG16	0.06139562	28	0.83875312	0.9739225
CAG24	-0.215815	25	0.8032166	0.9739225
CAG27	0.071857	18	0.8864513	0.98858496
CAG31	0.04264467	17	0.95495825	0.98976102
CAG13	-0.0160325	14	0.97662881	0.9904817

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.37. MaAslin2 associations between co-abundance groupings and Z-weight-for-length.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG10	0.63244161	71	0.02131298	0.30178426
CAG13	-0.3883881	89	0.03358121	0.31689532
CAG1	-0.2663114	166	0.09369482	0.53440748
CAG21	-0.3249445	69	0.20478928	0.80180208
CAG2	-0.2575589	145	0.2462304	0.82646482
CAG3	0.13281059	166	0.24760964	0.82646482
CAG6	-0.3635434	86	0.30688225	0.85934411
CAG18	-0.2378228	115	0.44476647	0.93577193
CAG4	-0.1622924	117	0.61153816	0.98686156
CAG5	0.14231215	139	0.54401781	0.98686156
CAG11	-0.1276406	76	0.55227239	0.98686156
CAG14	-0.1164573	56	0.59667019	0.98686156
CAG19	-0.1233776	92	0.625313	0.98686156
CAG22	0.18603176	150	0.57425712	0.98686156
CAG24	0.19078458	102	0.54013542	0.98686156
CAG8	-0.0674063	162	0.74255817	0.98812378
CAG9	0.03319291	14	0.84081307	0.98812378
CAG12	0.05585721	49	0.75076002	0.98812378
CAG15	-0.017518	76	0.94134709	0.98812378
CAG16	0.05711686	159	0.82085325	0.98812378
CAG17	-0.0384685	130	0.90122683	0.98812378
CAG20	0.08418342	47	0.72447086	0.98812378
CAG23	-0.1001454	87	0.7196302	0.98812378
CAG25	-0.1167479	80	0.65948669	0.98812378
CAG26	-0.0703236	50	0.72092646	0.98812378
CAG27	-0.0222787	86	0.92521834	0.98812378
CAG28	-0.0222986	80	0.93257746	0.98812378
CAG29	-0.0605497	30	0.72964076	0.98812378
CAG30	-0.0645923	49	0.73752094	0.98812378
CAG31	0.02429501	48	0.88799167	0.98812378
CAG32	-0.0444205	61	0.85983548	0.98812378
CAG33	-0.0720774	67	0.77115792	0.98812378
CAG7	-0.0029579	97	0.98992973	0.99566952

*Models were adjusted for age, sex, milk type, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.
**Number of participants with non-zero feature.

Table S4.38. MaAslin2 associations between co-abundance groupings and Z-weight-for-length stratified by breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG13	-0.4876495	75	0.01953342	0.56905441
CAG1	-0.3058189	138	0.08882947	0.89849212
CAG10	0.42698369	54	0.14091836	0.90954529
CAG2	-0.2882108	117	0.27534277	0.95858207
CAG3	0.16338515	138	0.19139896	0.95858207
CAG30	-0.2402713	39	0.25760911	0.95858207
CAG21	-0.2855779	51	0.29510435	0.96424409
CAG4	-0.2320094	107	0.53699211	0.98438127
CAG11	-0.2337436	64	0.34546727	0.98438127
CAG14	-0.1573325	48	0.52323071	0.98438127
CAG22	0.22057446	122	0.55902545	0.98438127
CAG24	0.33127915	77	0.3511883	0.98438127
CAG31	-0.1504304	31	0.36630577	0.98438127
CAG32	-0.1941766	50	0.48277755	0.98438127
CAG5	0.07125797	124	0.77700324	0.98881297
CAG6	-0.1947782	69	0.62058006	0.98881297
CAG7	-0.0397621	75	0.88072641	0.98881297
CAG8	-0.0875699	135	0.68610858	0.98881297
CAG9	0.07889316	12	0.68266228	0.98881297
CAG12	-0.0218918	39	0.89761891	0.98881297
CAG15	0.04530228	64	0.8683391	0.98881297
CAG16	-0.0361045	131	0.90537303	0.98881297
CAG18	-0.0485683	90	0.89177216	0.98881297
CAG20	0.09531669	33	0.70815846	0.98881297
CAG23	0.12809707	68	0.68742517	0.98881297
CAG26	0.10837755	34	0.60097049	0.98881297
CAG27	-0.1014973	68	0.70095989	0.98881297
CAG28	-0.1036899	59	0.72642862	0.98881297
CAG29	-0.0388357	24	0.8404912	0.98881297
CAG33	0.09154059	51	0.71912698	0.98881297
CAG17	0.0205756	104	0.95446338	0.99164984
CAG25	0.01967606	58	0.94669841	0.99164984
CAG19	0.00838257	70	0.97529984	0.99638167

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Table S4.39. MaAslin2 associations between co-abundance groupings and Z-weight-for-length stratified by non-breastmilk consumption.

Feature	Coefficient	N.not.0**	P Value	Q Value
CAG6	-2.6885021	17	0.01511955	0.65654064
CAG18	-1.049263	25	0.02661101	0.65654064
CAG22	-1.4100748	28	0.04935344	0.65654064
CAG23	-1.0822652	19	0.02836869	0.65654064
CAG10	1.44412017	17	0.1170374	0.74617739
CAG19	-0.936661	22	0.15939183	0.76913028
CAG11	0.82817713	12	0.16965169	0.78490022
CAG30	0.40662085	10	0.26652398	0.87605002
CAG21	-0.8974786	18	0.28655184	0.88709796
CAG2	0.35181432	28	0.38556709	0.89495004
CAG5	0.73734985	15	0.40908954	0.89495004
CAG14	-0.2653136	8	0.49512374	0.89495004
CAG20	-0.471447	14	0.48003649	0.89495004
CAG24	-0.7232231	25	0.35627456	0.89495004
CAG25	-0.6814801	22	0.40419045	0.89495004
CAG26	-0.8125251	16	0.29897856	0.89495004
CAG31	0.60446911	17	0.37655108	0.89495004
CAG33	-0.957596	16	0.35290215	0.89495004
CAG12	0.44723475	10	0.52252815	0.89665831
CAG1	0.23015662	28	0.63113847	0.96403246
CAG3	-0.1123647	28	0.78774624	0.96403246
CAG8	-0.2106	27	0.74358064	0.96403246
CAG13	-0.1888607	14	0.70640201	0.96403246
CAG15	-0.2251261	12	0.66871937	0.96403246
CAG17	-0.1897096	26	0.76191039	0.96403246
CAG28	0.23528089	21	0.73119283	0.96403246
CAG16	-0.0594504	28	0.8301298	0.96436987
CAG27	0.09285174	18	0.8407327	0.96436987
CAG9	0.03305747	2	0.89613423	0.97910915
CAG29	0.03218686	6	0.93314107	0.97910915
CAG32	0.0738073	11	0.92837039	0.97910915
CAG7	-0.0508074	22	0.94640873	0.98307348
CAG4	-0.0042309	10	0.99330251	0.99594143

*Models were adjusted for age, sex, antibiotics ever, antibiotics in the past 14 days, water source, treatment, diarrhea in the past 14 days, low birthweight, and poverty score.

**Number of participants with non-zero feature.

Supplemental Figures

Alpha-Diversity

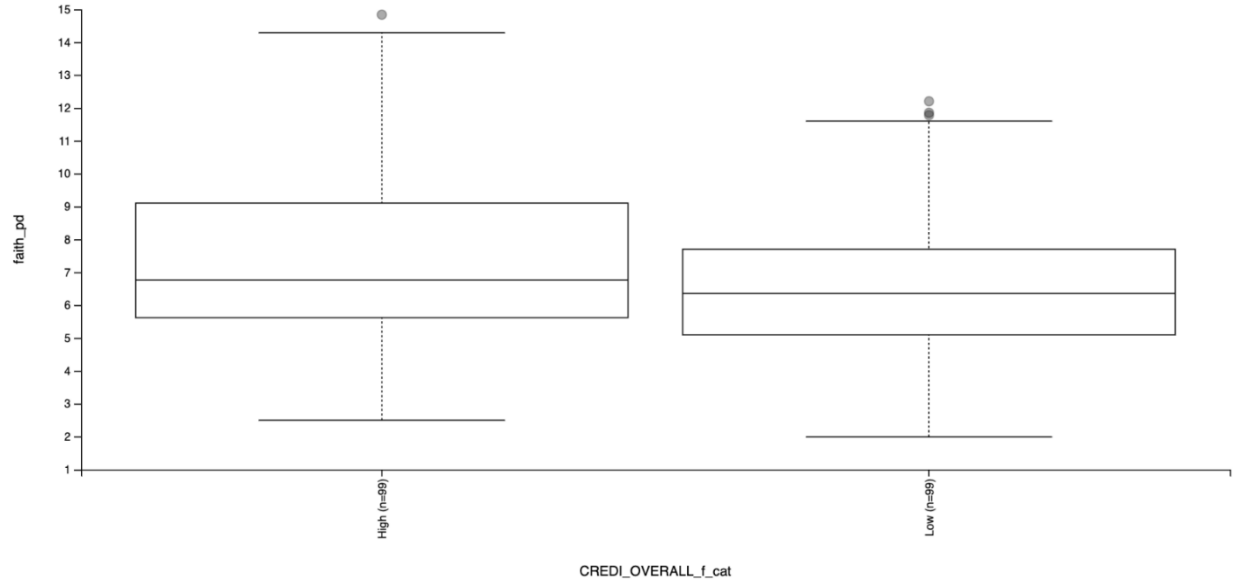

Figure S4.1. Faith's Phylogenetic Diversity (Faith's PD) boxplot of CREDI overall score high and low groups.

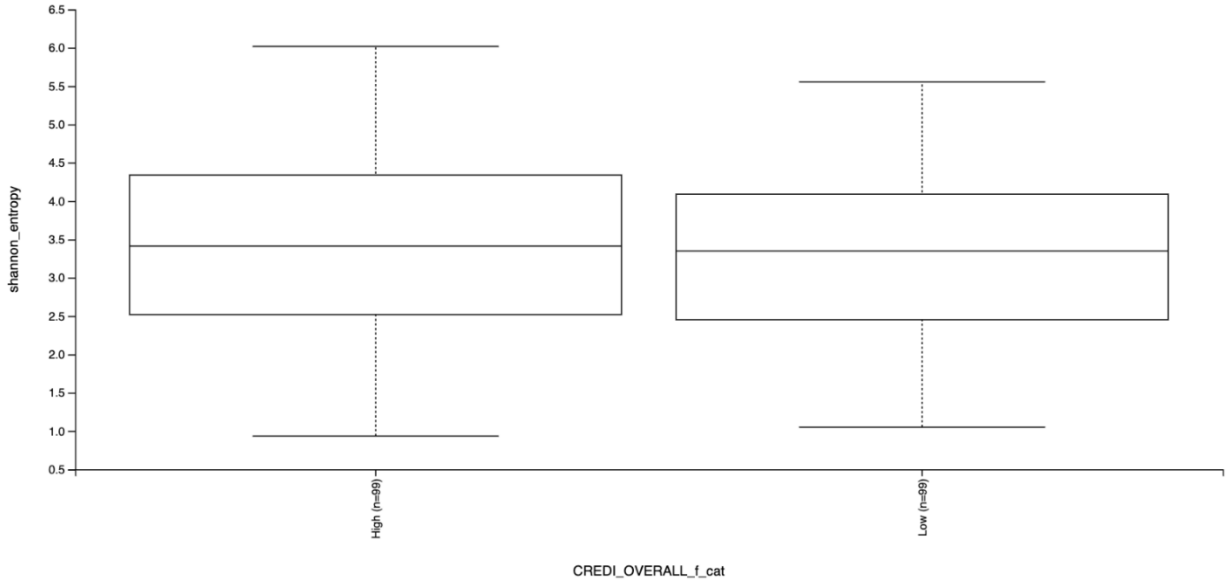

Figure S4.2. Shannon diversity boxplot of CREDI overall score high and low groups.

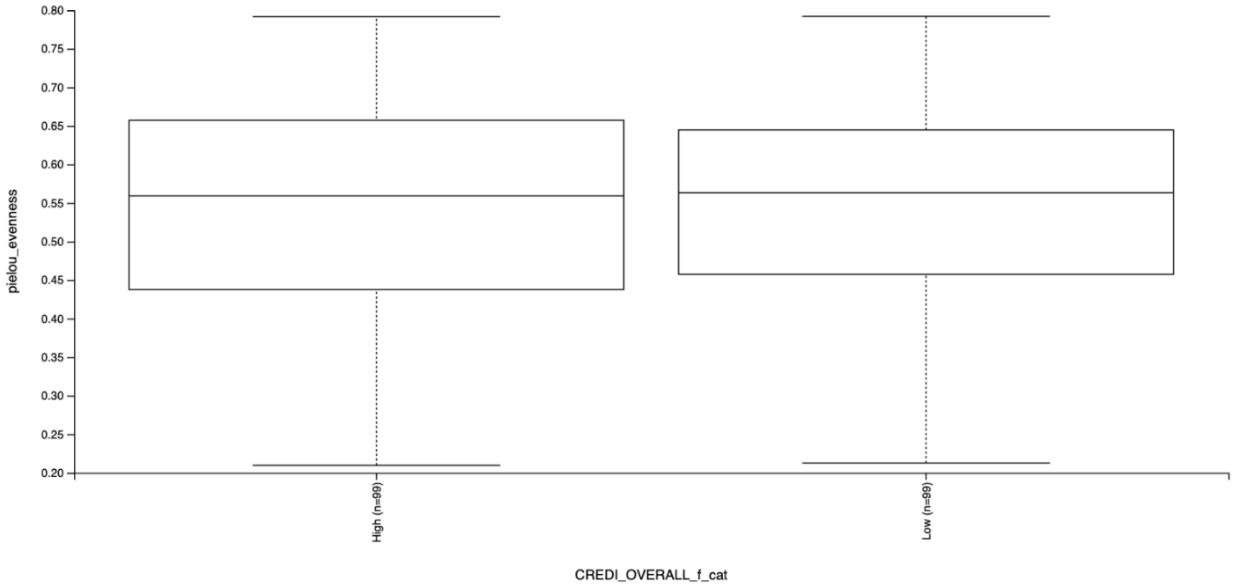

Figure S4.3. Pielou's Evenness alpha-diversity boxplot of CREDI overall score high and low groups.

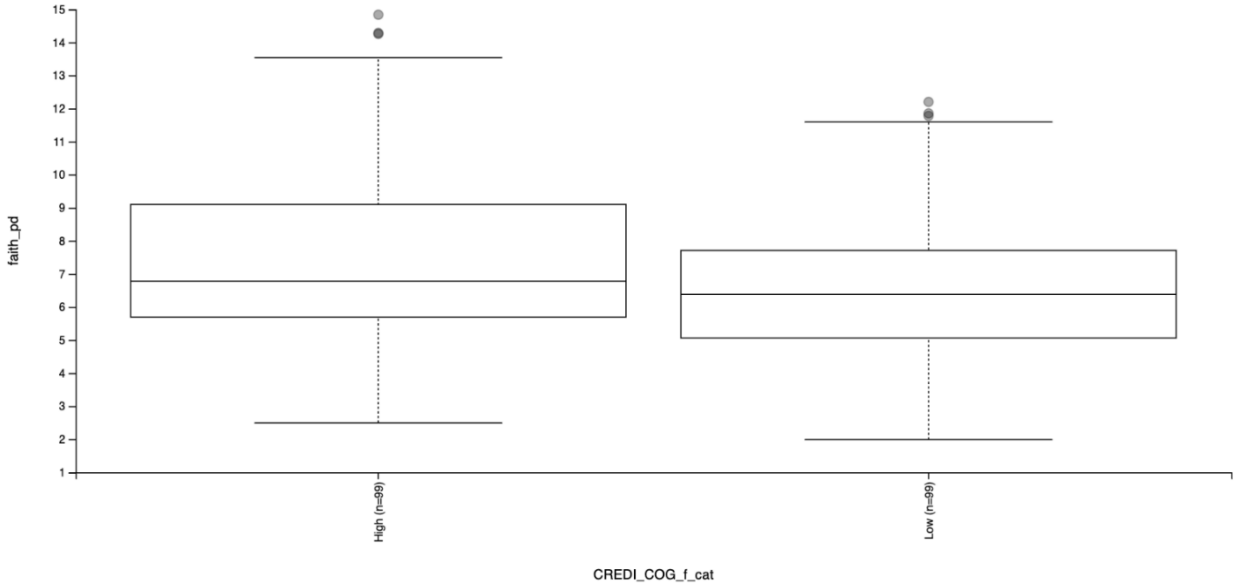

Figure S4.4. Faith's PD boxplot of CREDI cognitive score high and low groups.

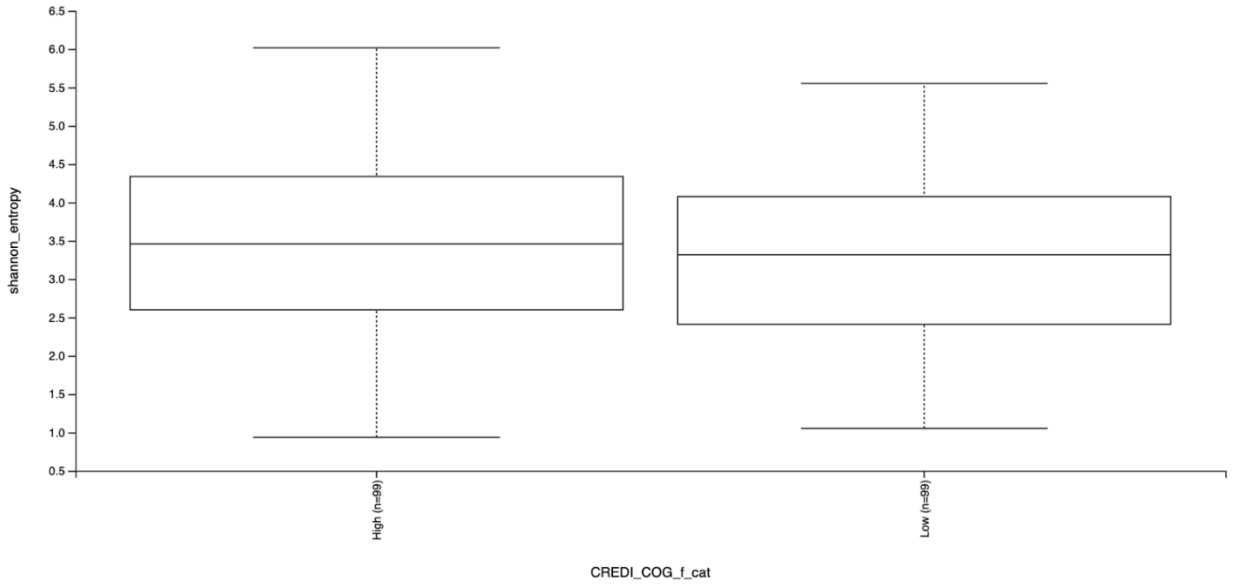

Figure S4.5. Shannon diversity boxplot of CREDI cognitive score high and low groups.

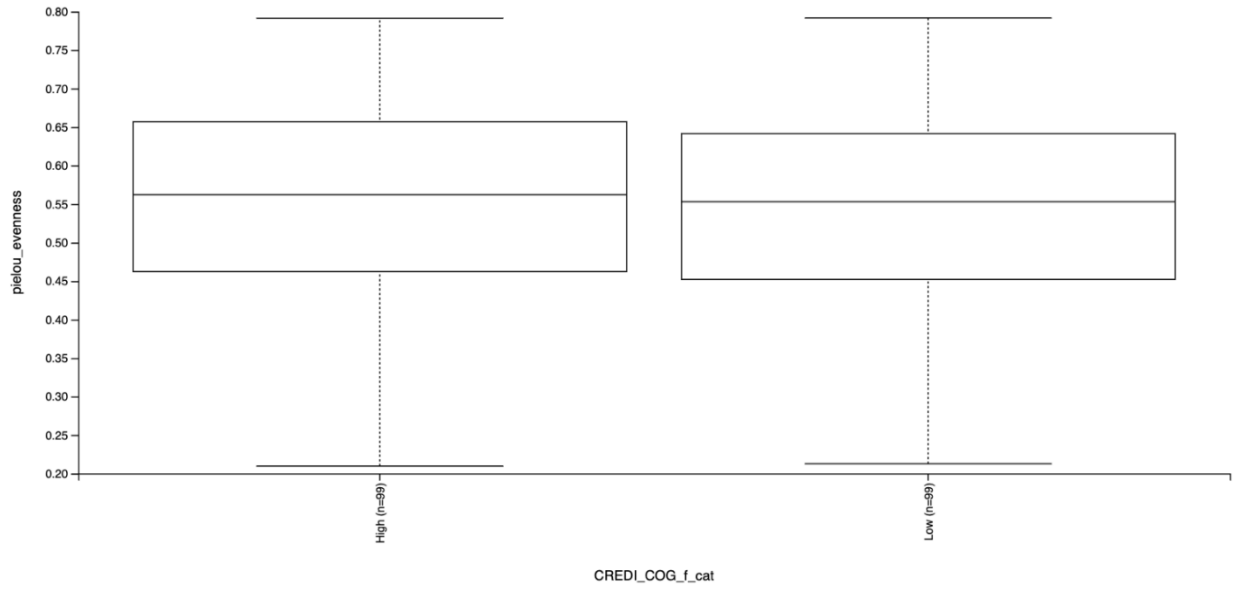

Figure S4.6. Pielou's Evenness alpha-diversity boxplot of CREDI cognitive score high and low groups.

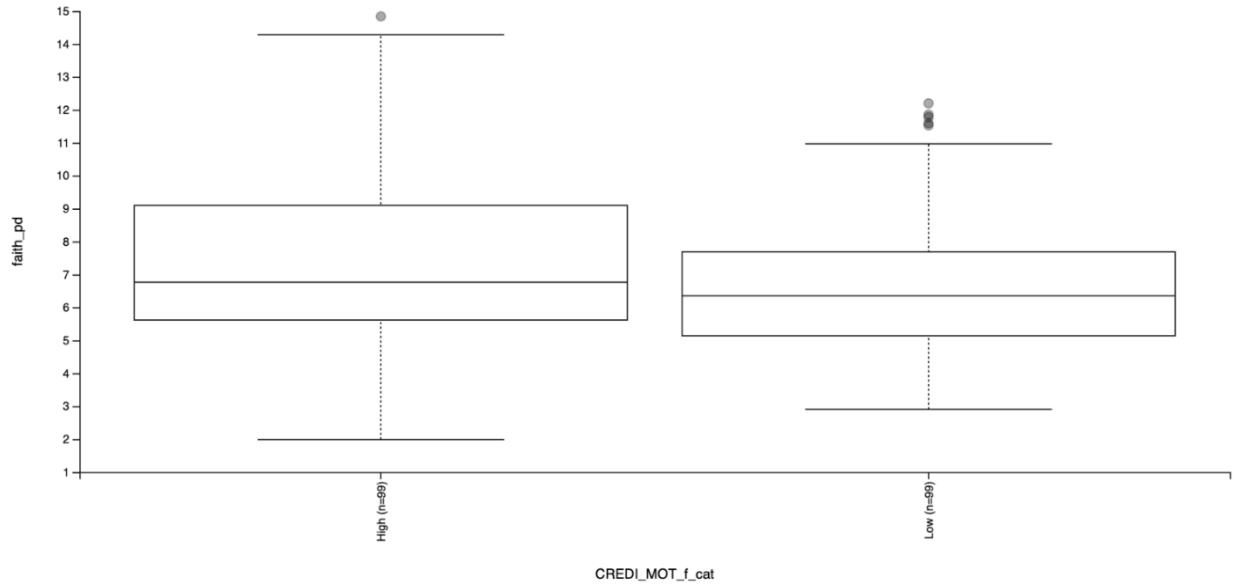

Figure S4.7. Faith's PD boxplot of CREDI motor score high and low groups.

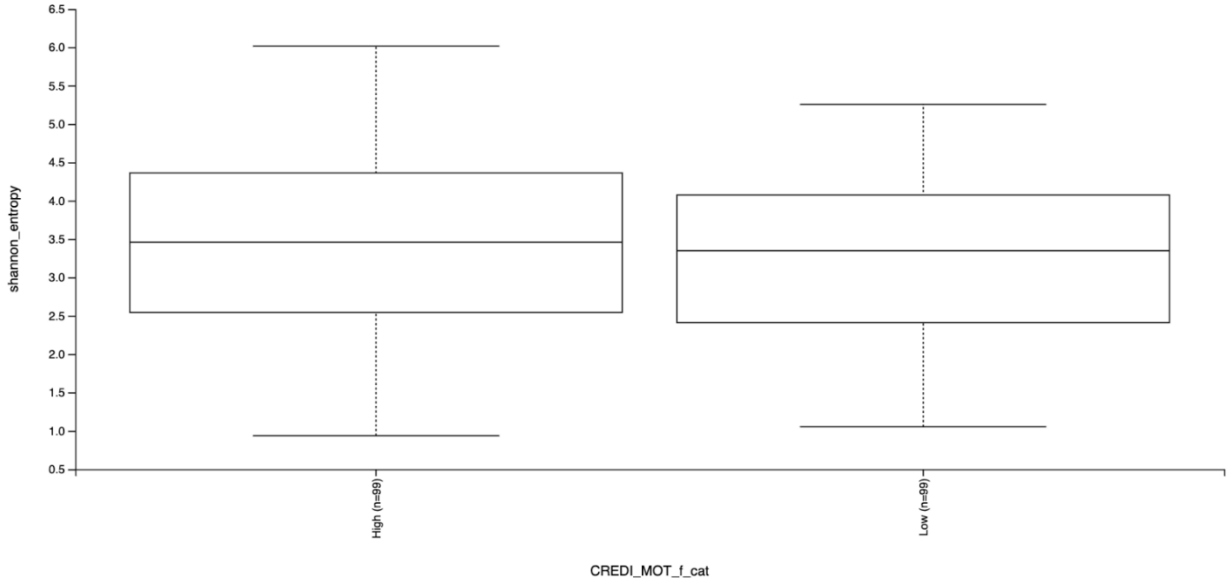

Figure S4.8. Shannon diversity boxplot of CREDI motor score high and low groups.

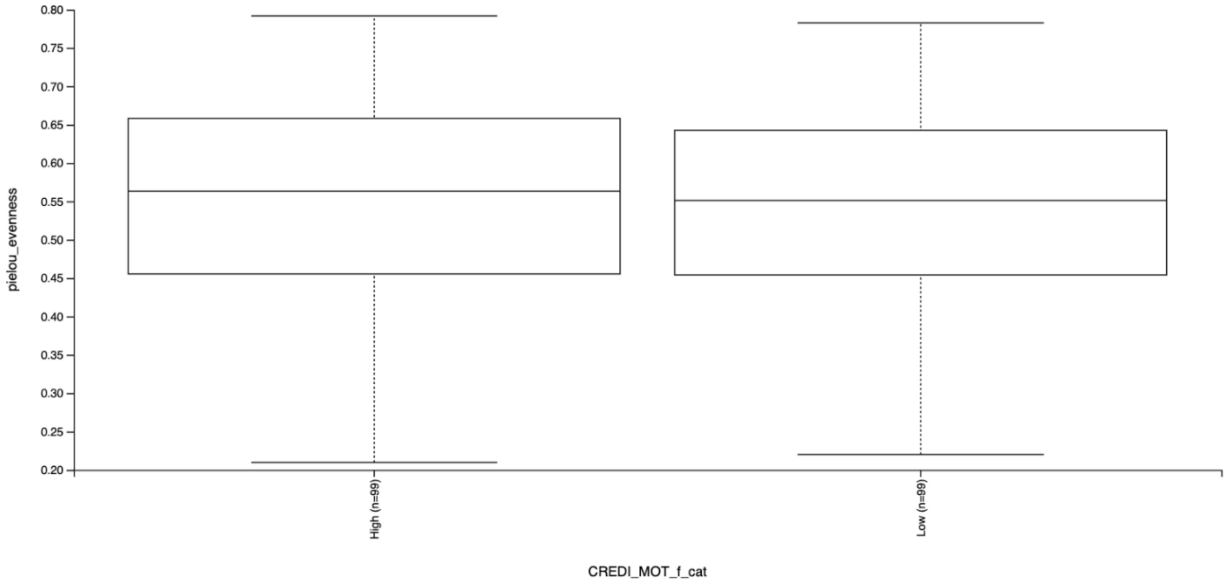

Figure S4.9. Pielou's Evenness alpha-diversity boxplot of CREDI motor score high and low groups.

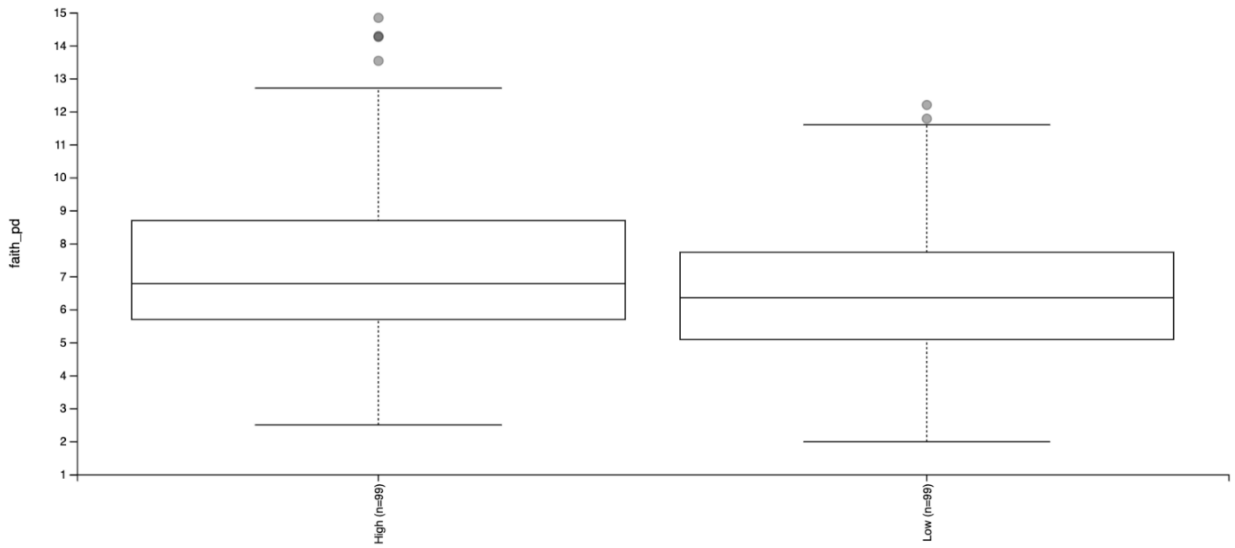

Figure S4.10. Faith's PD boxplot of CREDI socioemotional score high and low groups.

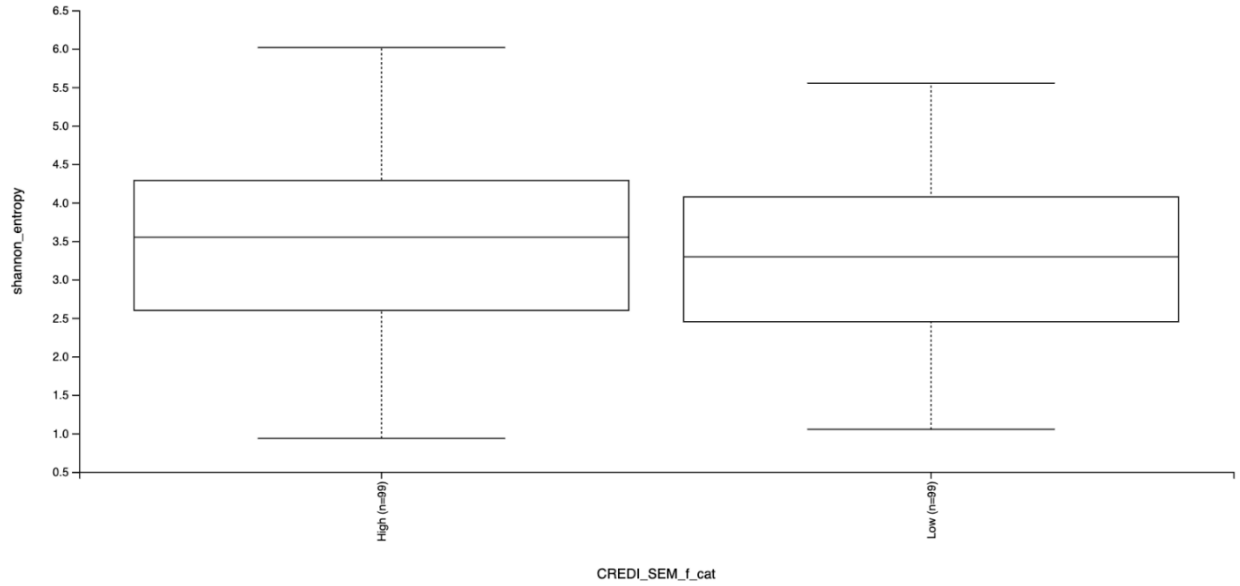

Figure S4.11. Shannon diversity boxplot of CREDI socioemotional score high and low groups.

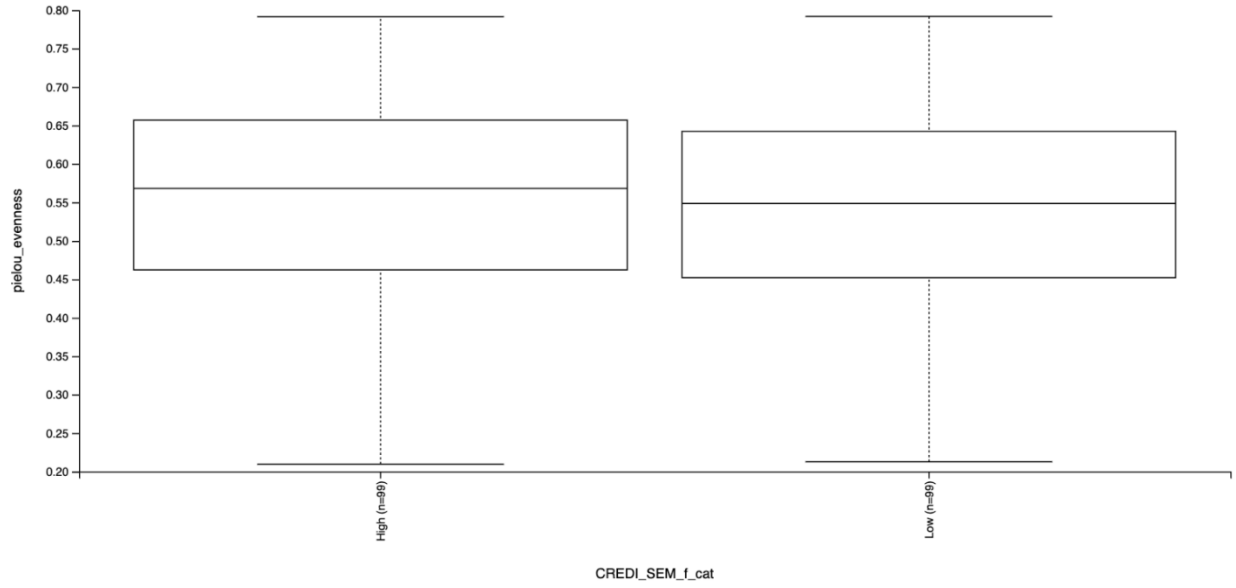

Figure S4.12. Pielou's Evenness alpha-diversity boxplot of CREDI socioemotional score high and low groups.

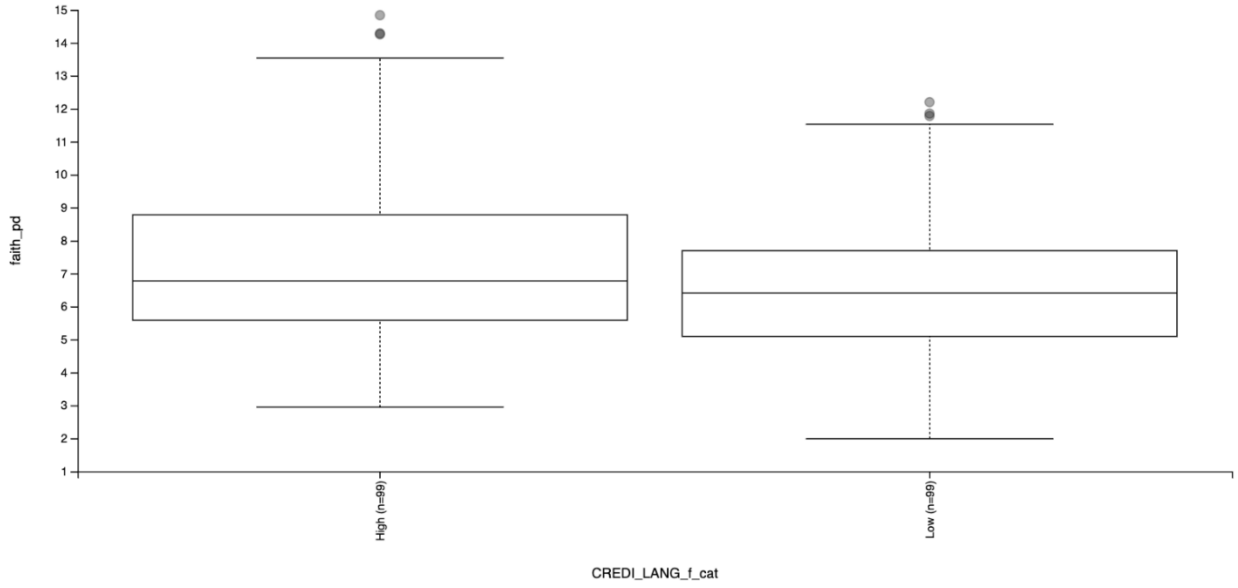

Figure S4.13. Faith's PD boxplot of CREDI language score high and low groups.

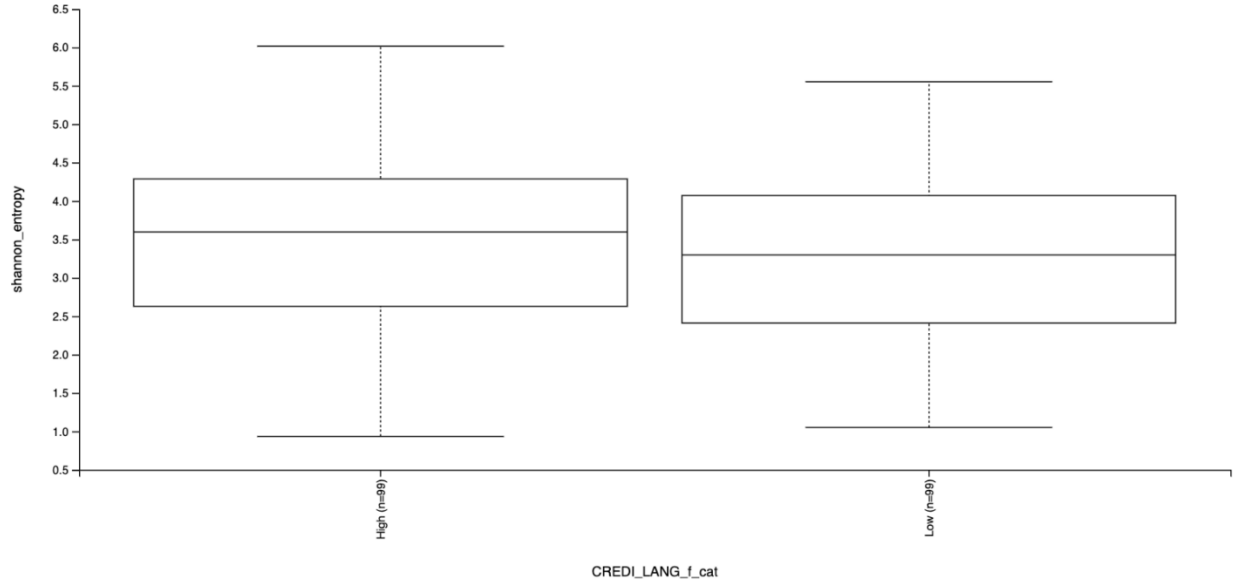

Figure S4.14. Shannon diversity boxplot of CREDI language score high and low groups.

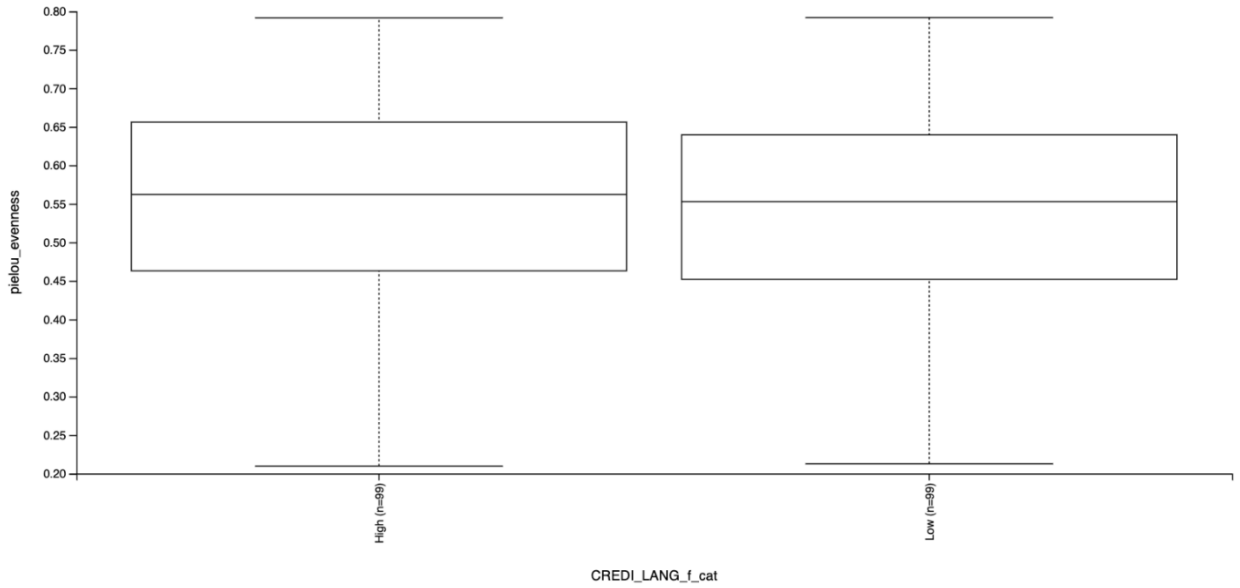

Figure S4.15. Pielou's Evenness alpha-diversity boxplot of CREDI language score high and low groups.

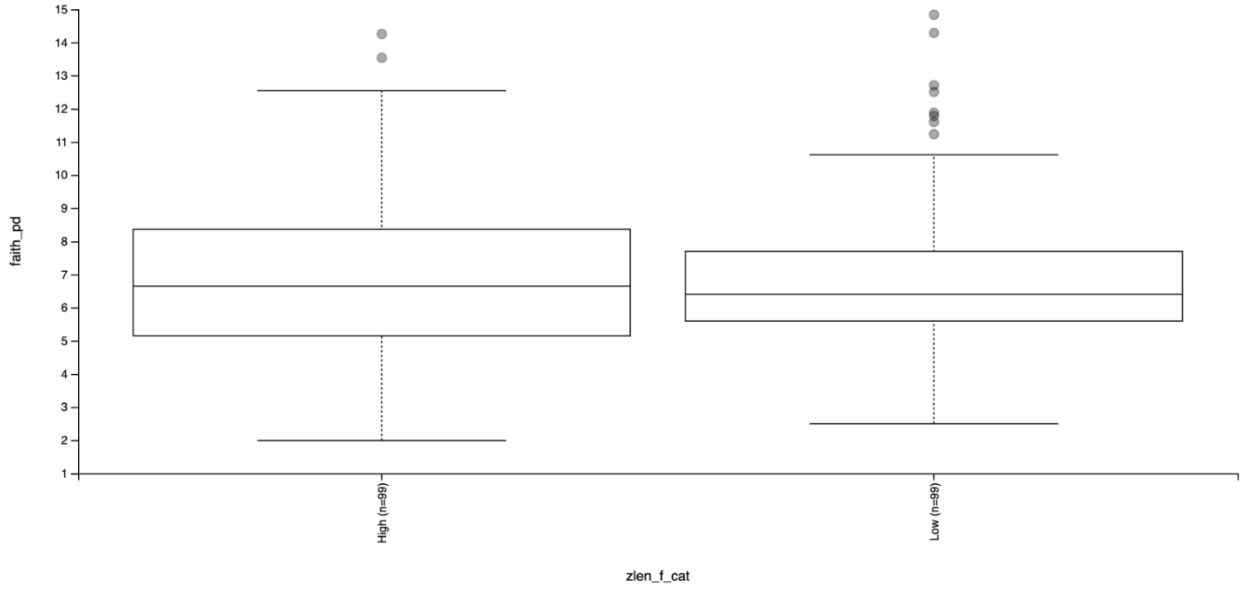

Figure S4.16. Faith's PD boxplot of Z-length-for-age high and low groups.

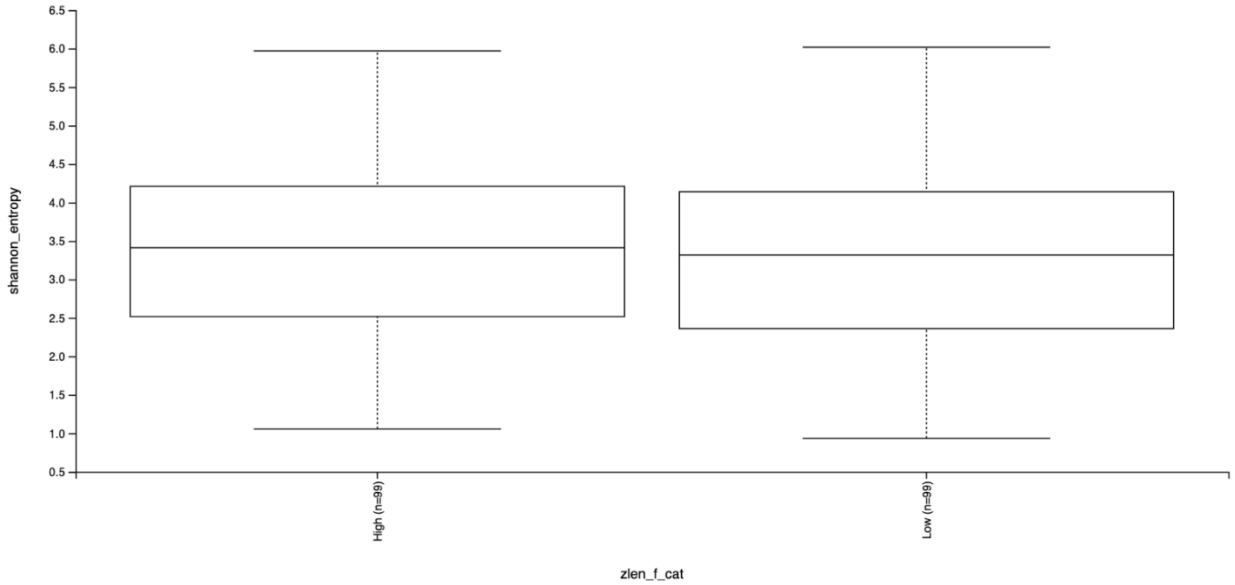

Figure S4.17. Shannon diversity boxplot of Z-length-for-age high and low groups.

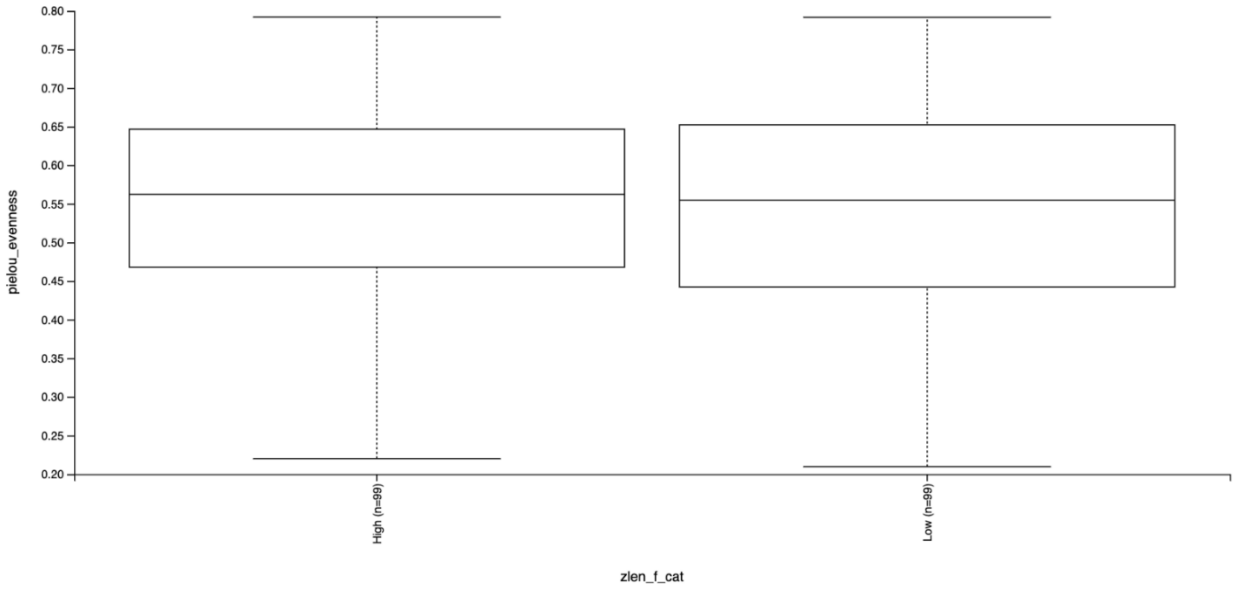

Figure S4.18. Pielou's Evenness alpha-diversity boxplot of Z-length-for-age high and low groups.

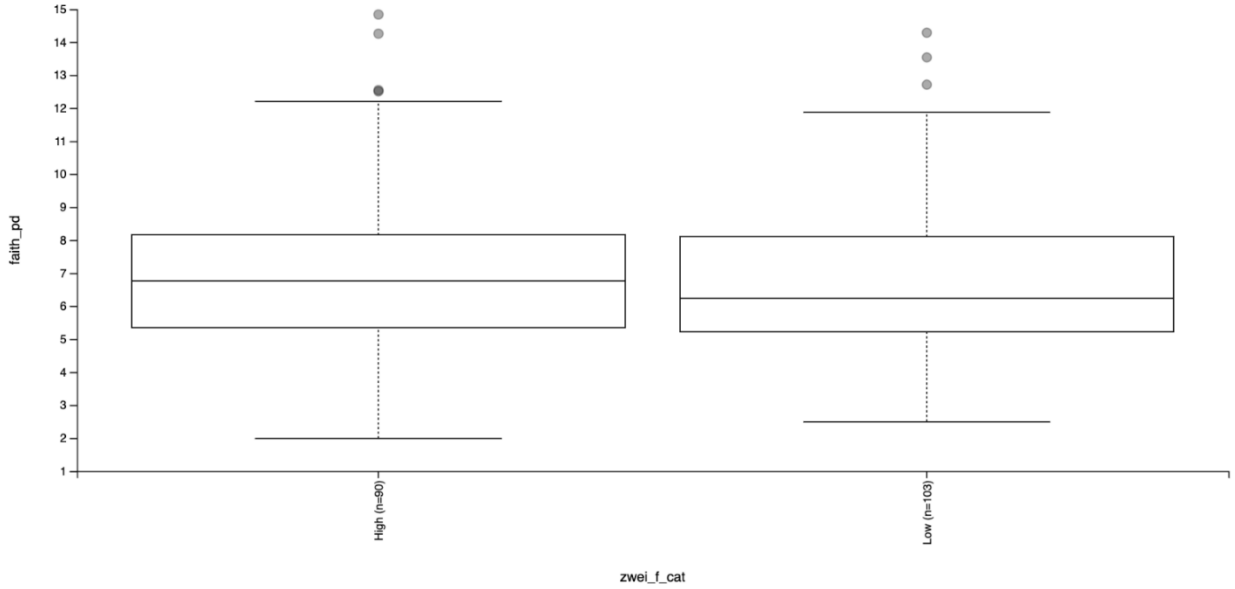

Figure S4.19. Faith's PD boxplot of Z-weight-for-age high and low groups.

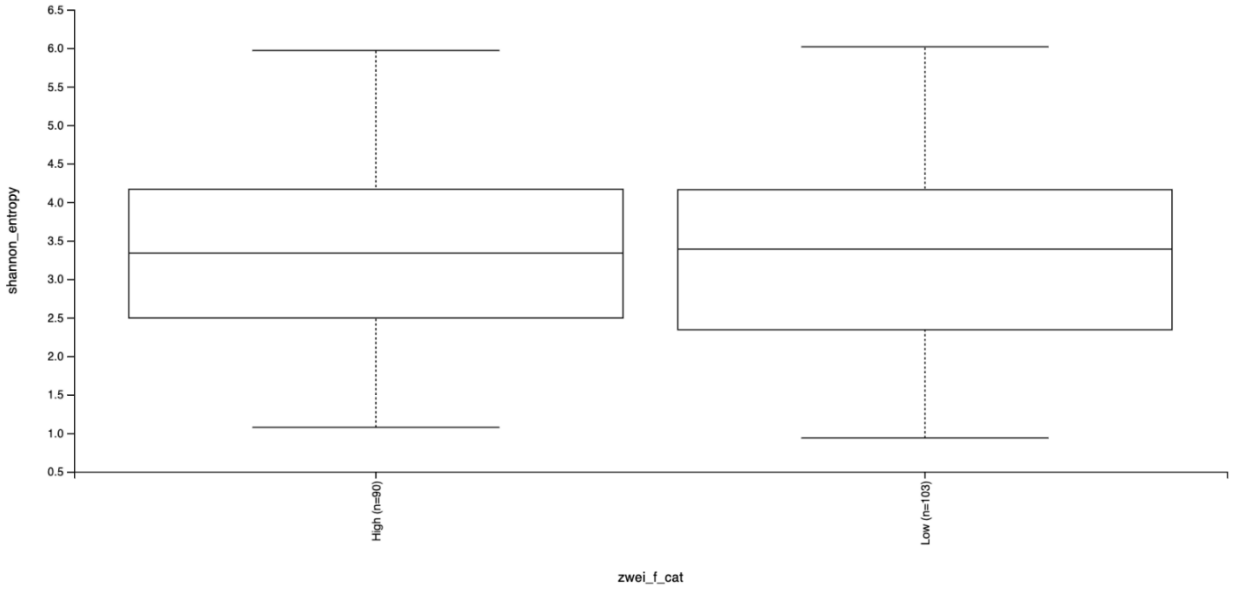

Figure S4.20. Shannon diversity boxplot of Z-weight-for-age high and low groups.

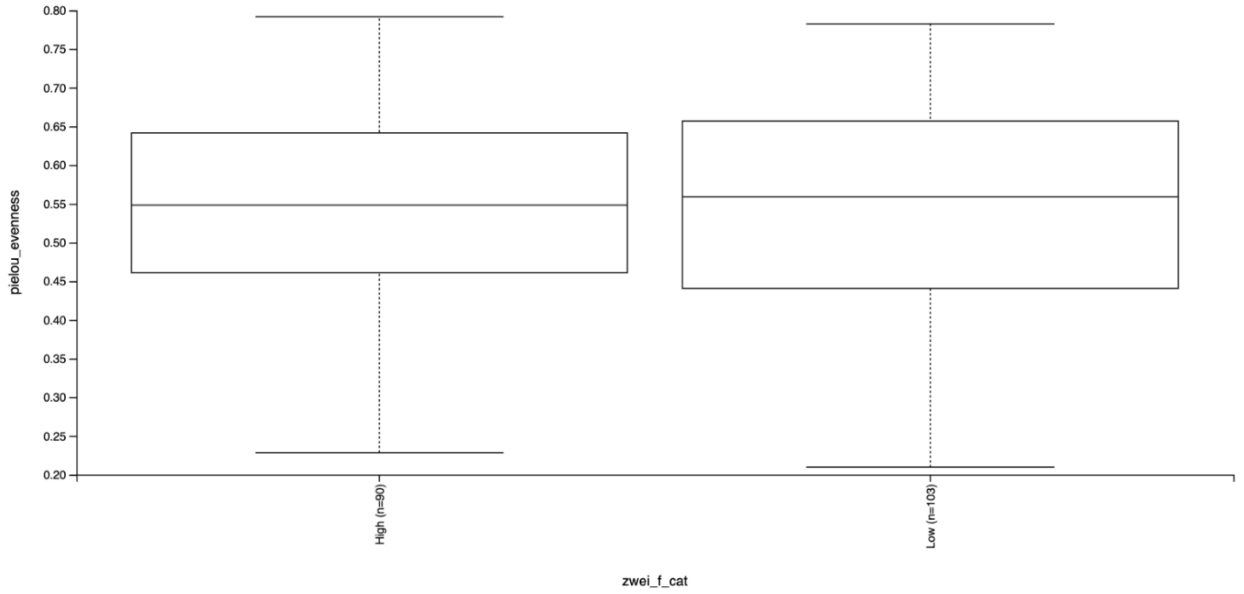

Figure S4.21. Pielou's Evenness alpha-diversity boxplot of Z-weight-for-age high and low groups.

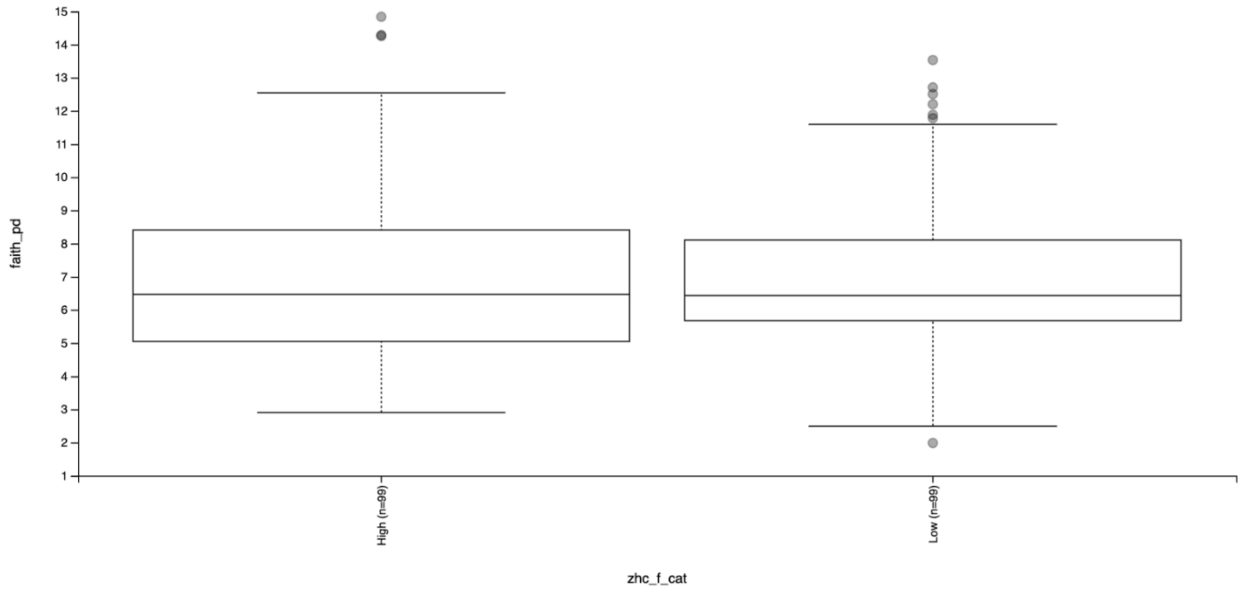

Figure S4.22. Faith's PD boxplot of Z-head-circumference high and low groups.

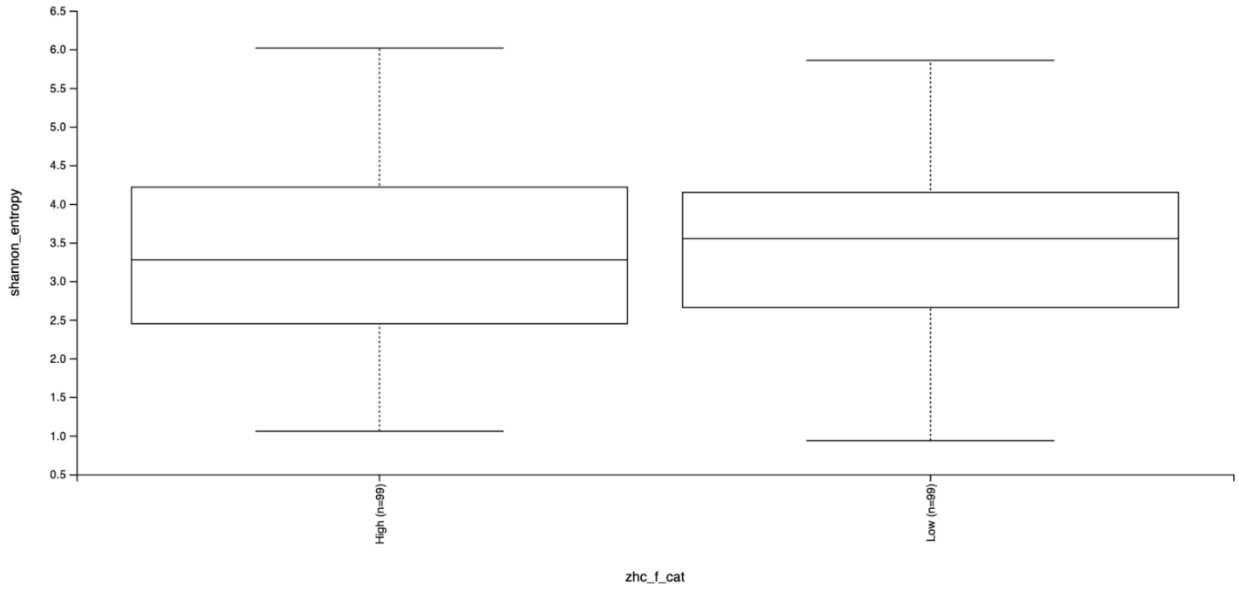

Figure S4.23. Shannon diversity boxplot of Z-head-circumference high and low groups.

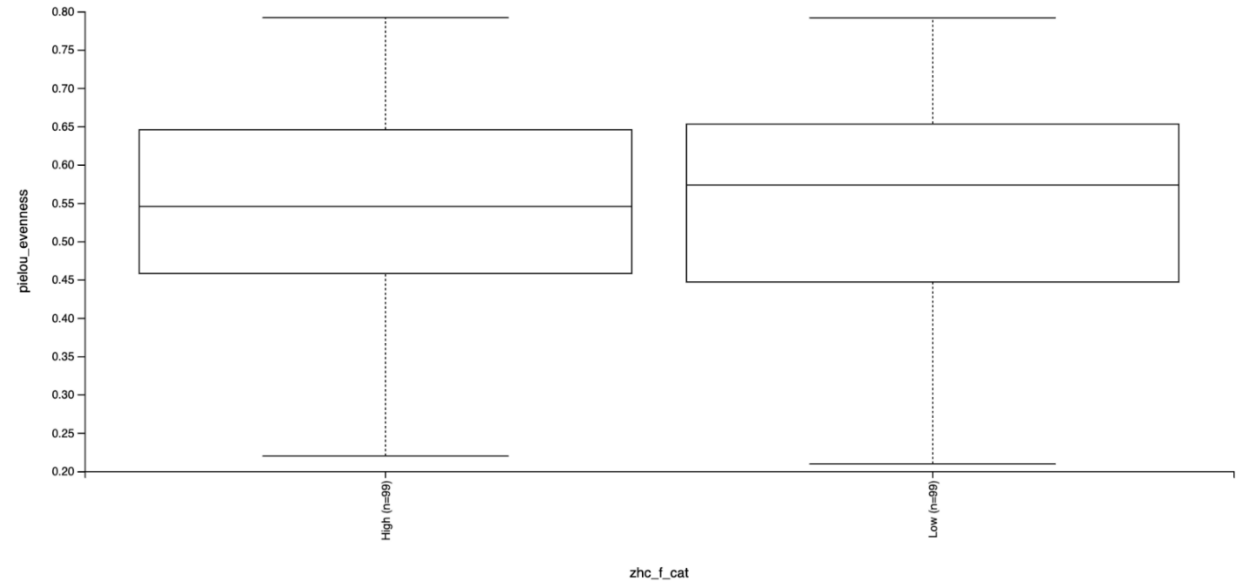

Figure S4.24. Pielou's Evenness alpha-diversity boxplot of Z-head-circumference high and low groups.

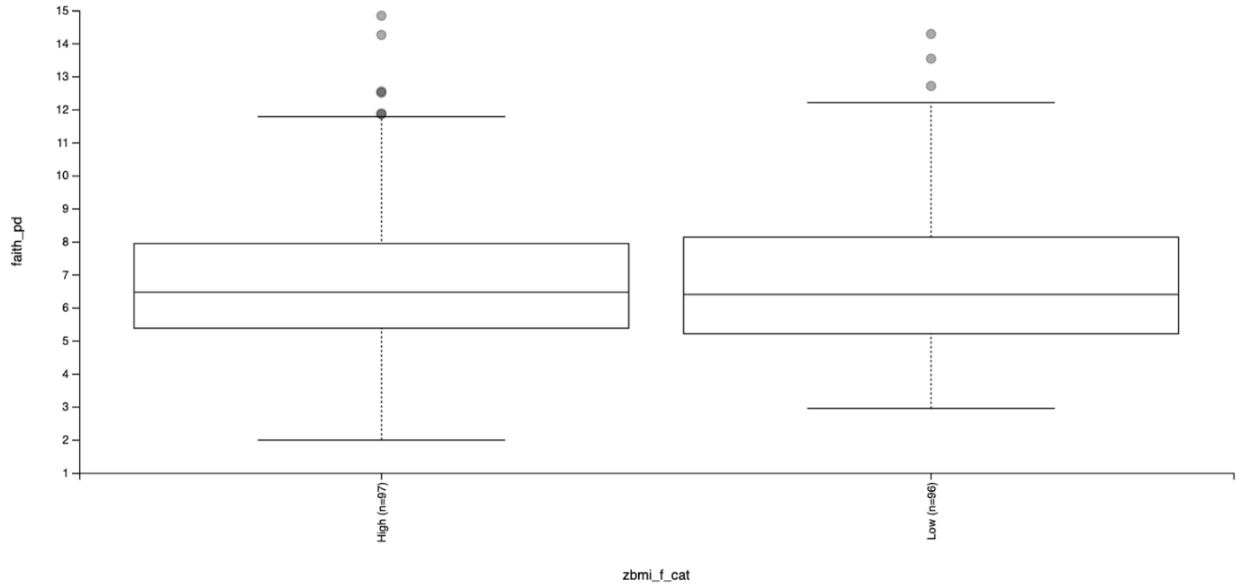

Figure S4.25. Faith's PD boxplot of Z-BMI high and low groups.

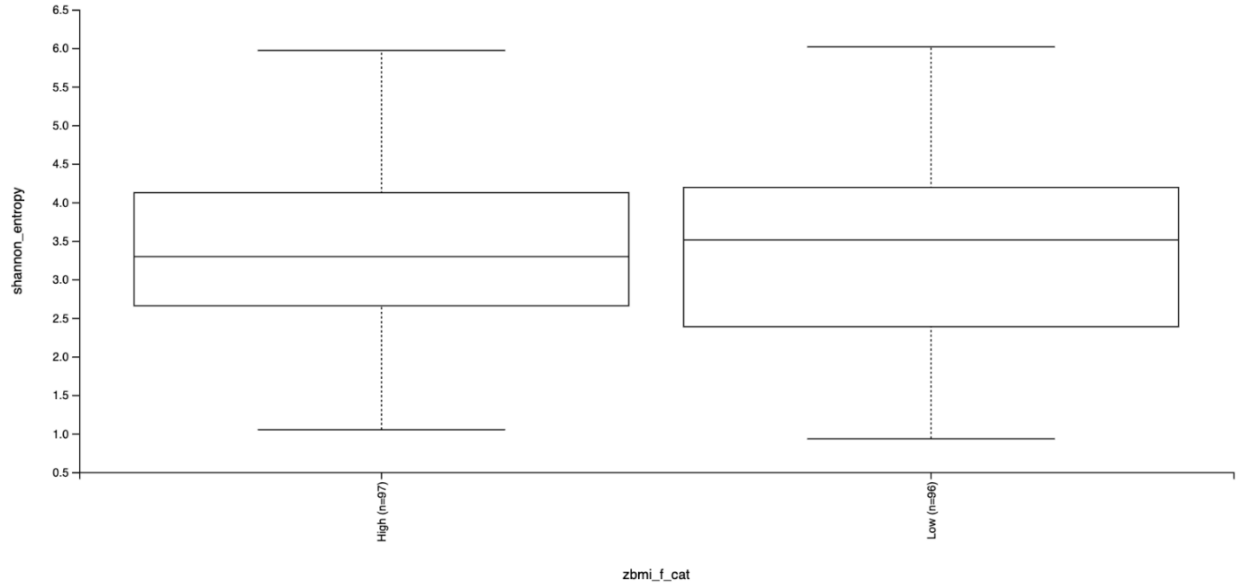

Figure S4.26. Shannon diversity boxplot of Z-BMI high and low groups.

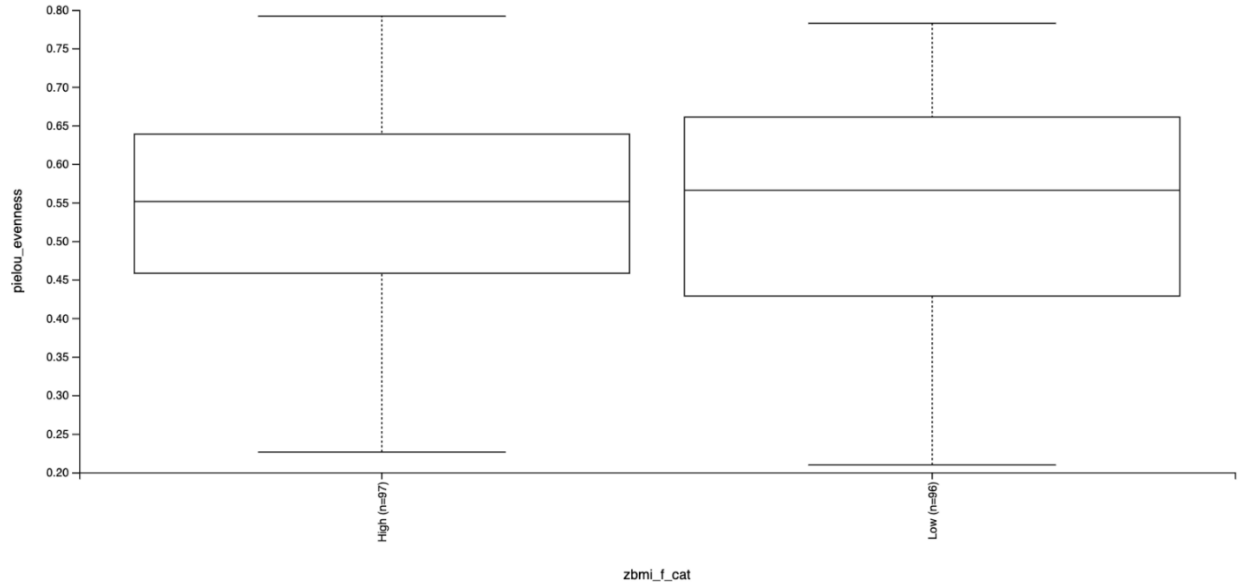

Figure S4.27. Pielou's Evenness alpha-diversity boxplot of Z-BMI high and low groups.

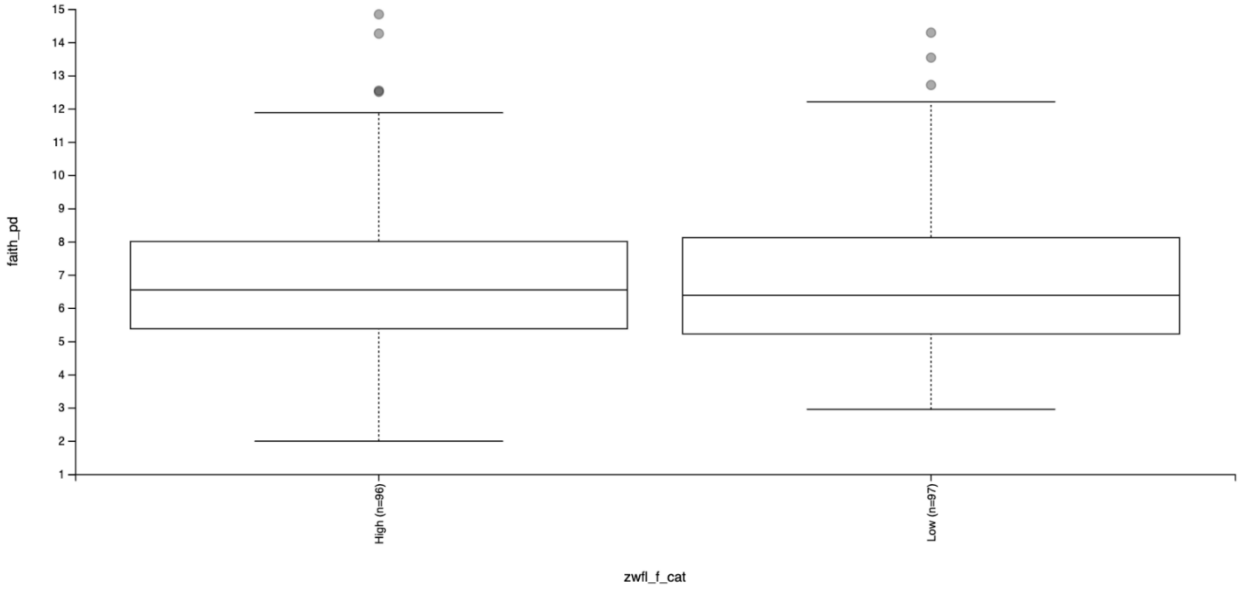

Figure S4.28. Faith's PD boxplot of Z-weight-for-length high and low groups.

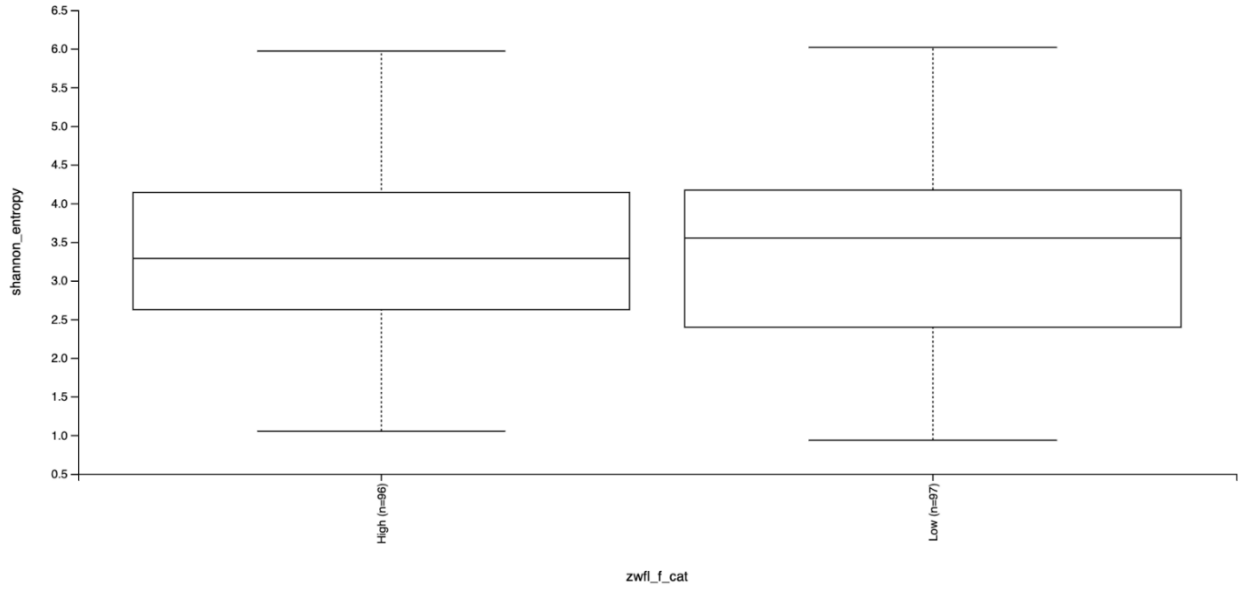

Figure S4.29. Shannon diversity boxplot of Z-weight-for-length score high and low groups.

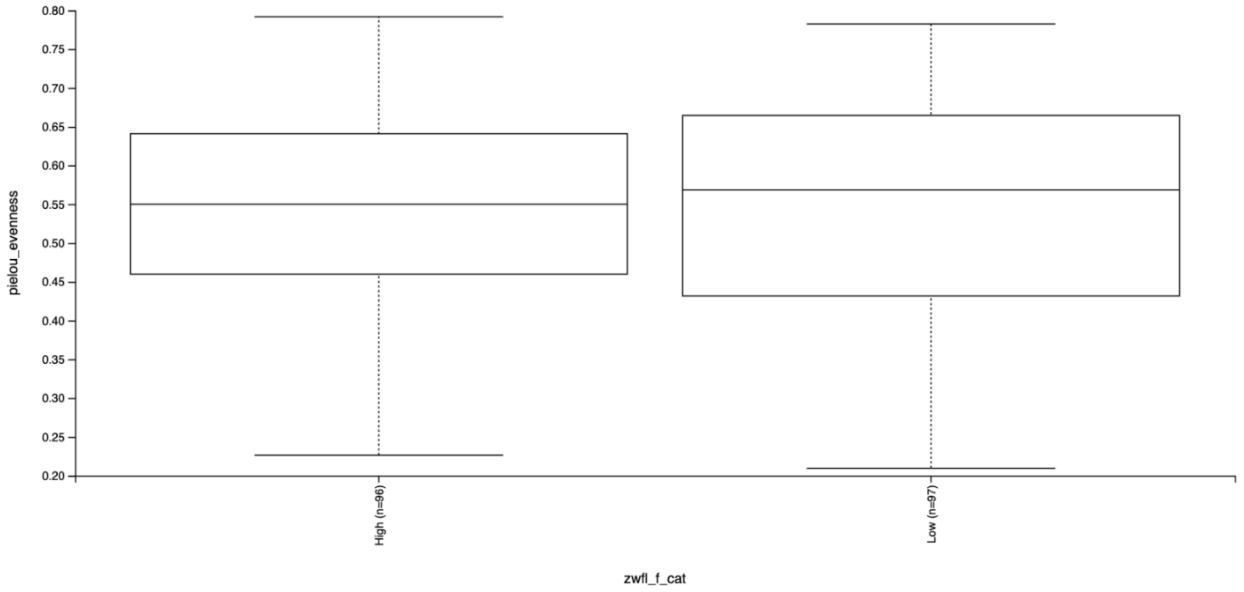

Figure S4.30. Pielou's Evenness alpha-diversity boxplot of Z-weight-for-length high and low groups.

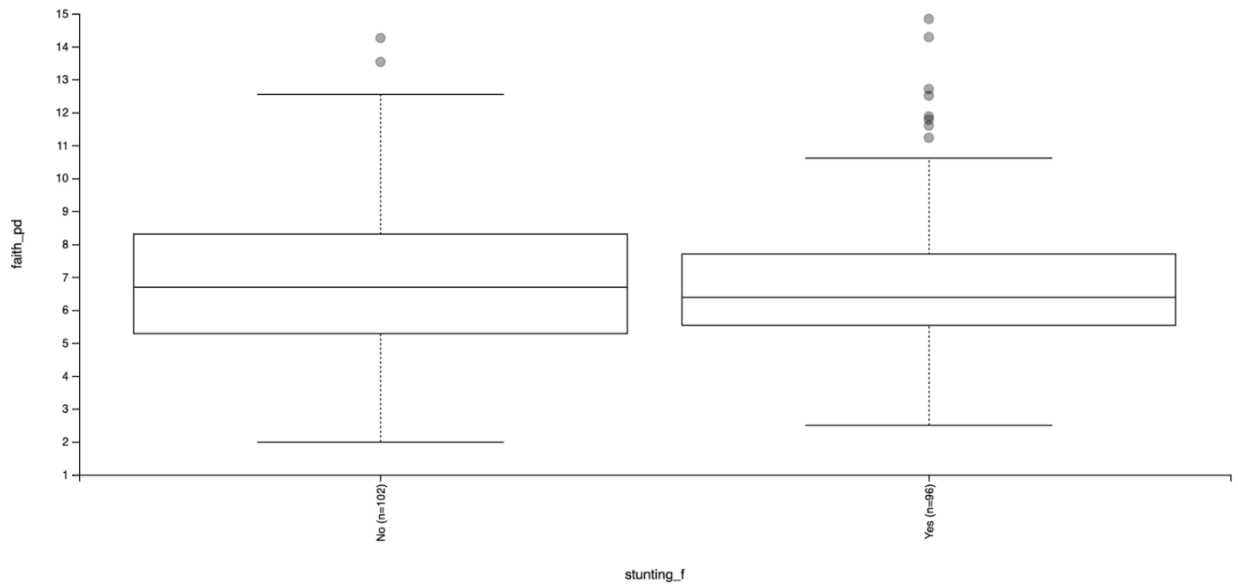

Figure S4.31. Faith's PD boxplot of stunting (yes, no).

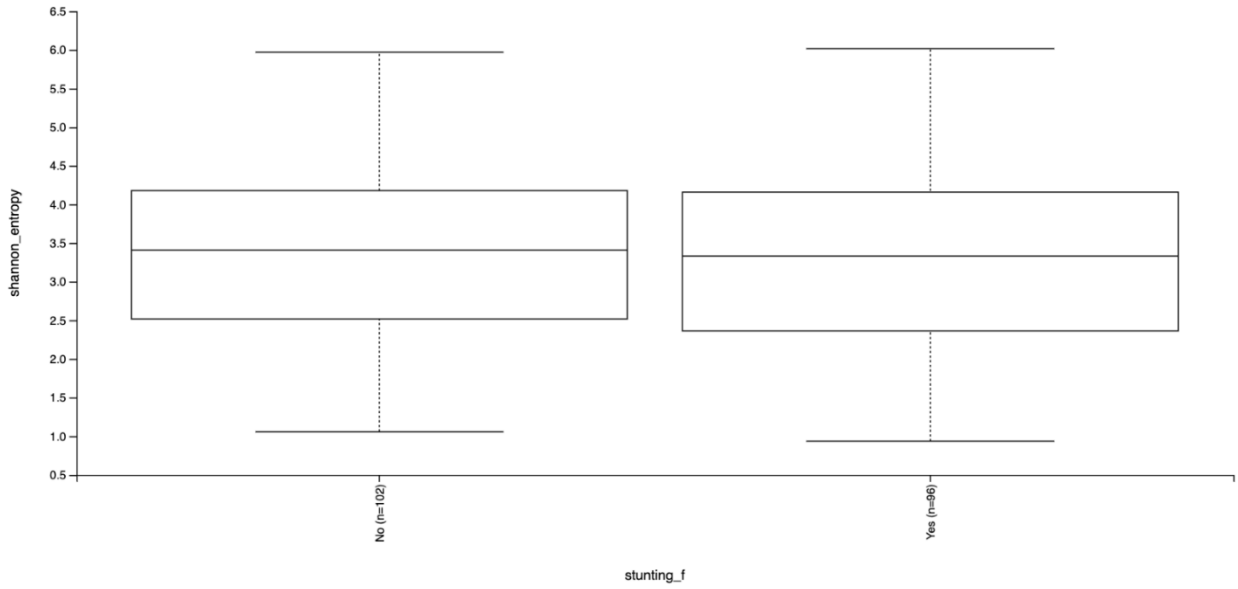

Figure S4.32. Shannon diversity boxplot of stunting (yes, no) score high and low groups.

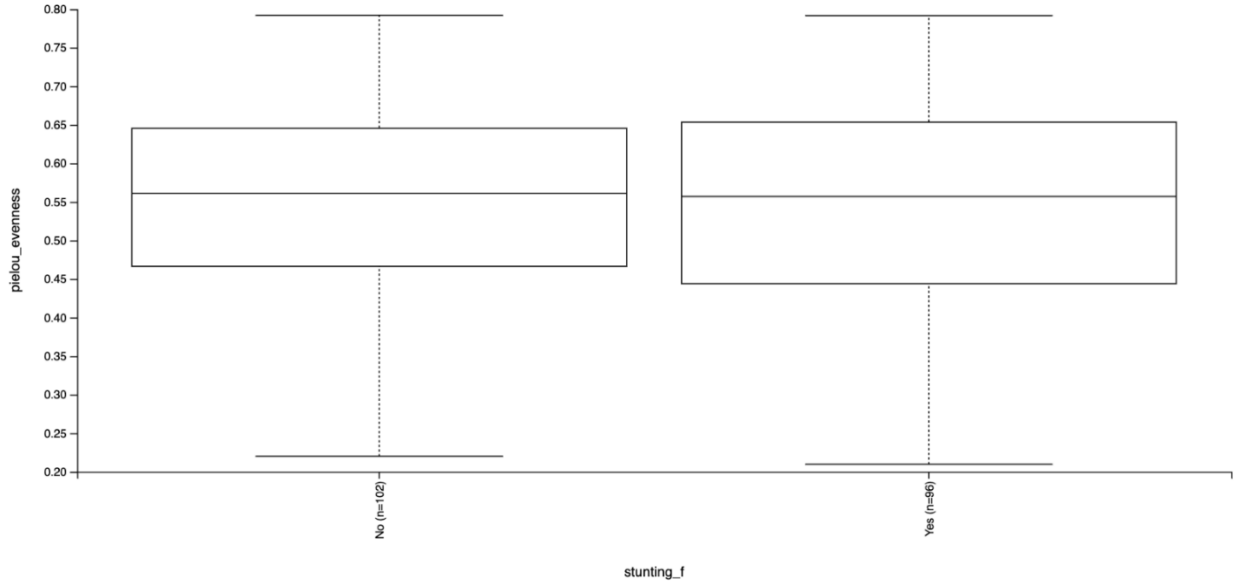

Figure S4.33. Pielou's Evenness alpha-diversity boxplot of stunting (yes, no).

Beta-Diversity

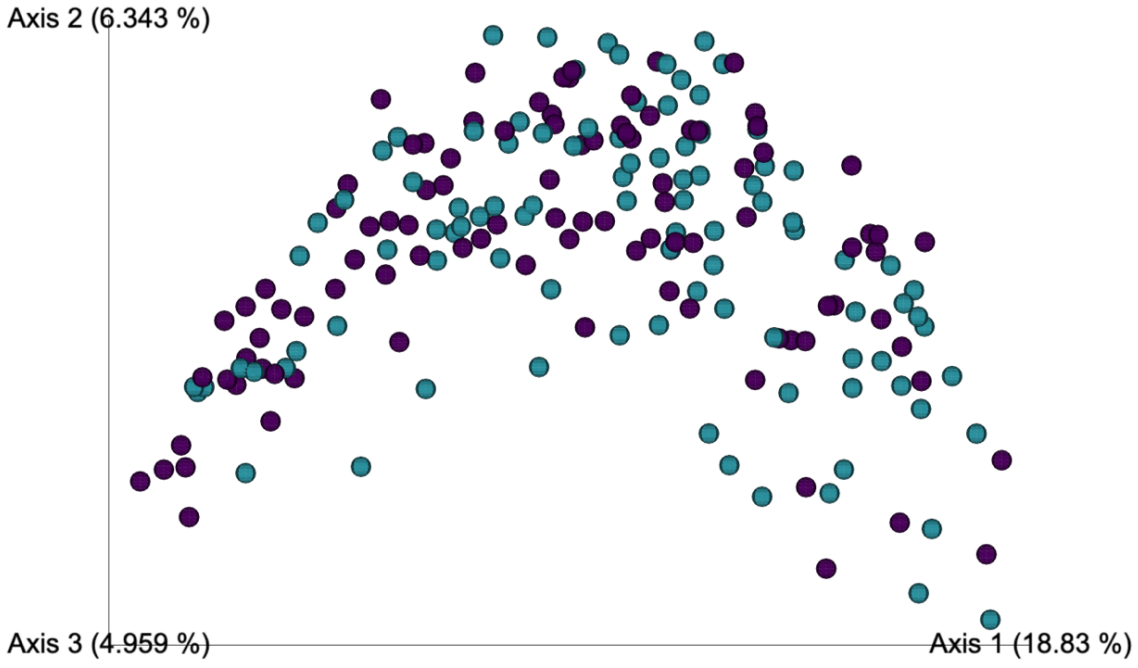

Figure S4.34. PCoA plot of high and low CREDI Overall scores unweighted Unifrac beta diversity. Purple = high scores, green = low scores.

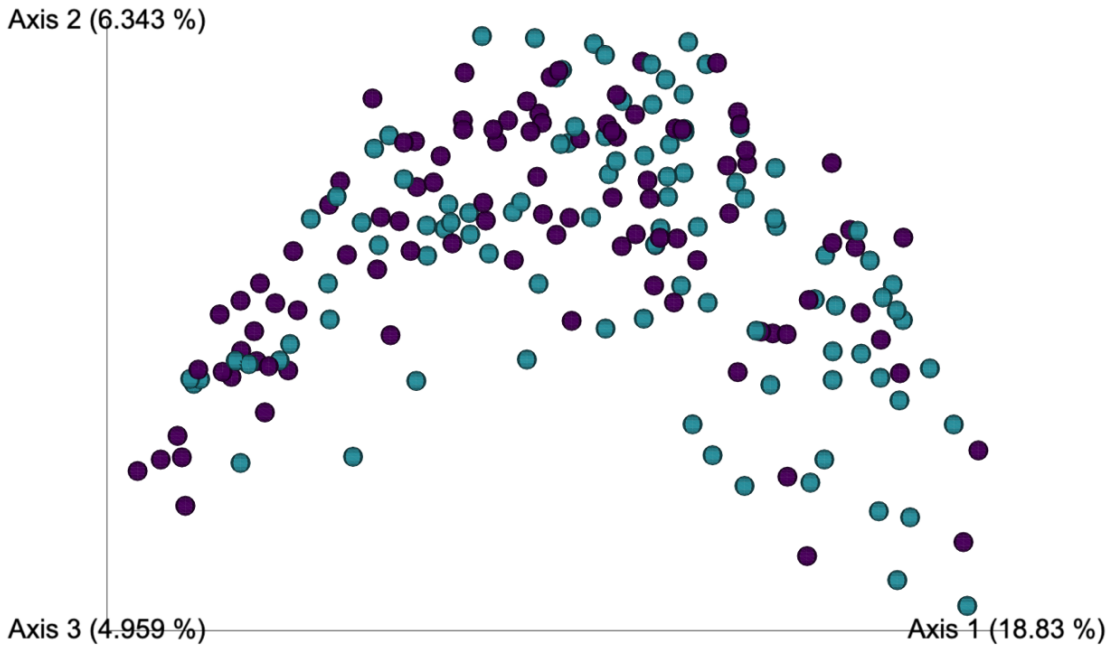

Figure S4.35. PCoA plot of high and low CREDI Cognitive scores unweighted Unifrac beta diversity. Purple = high scores, green = low scores.

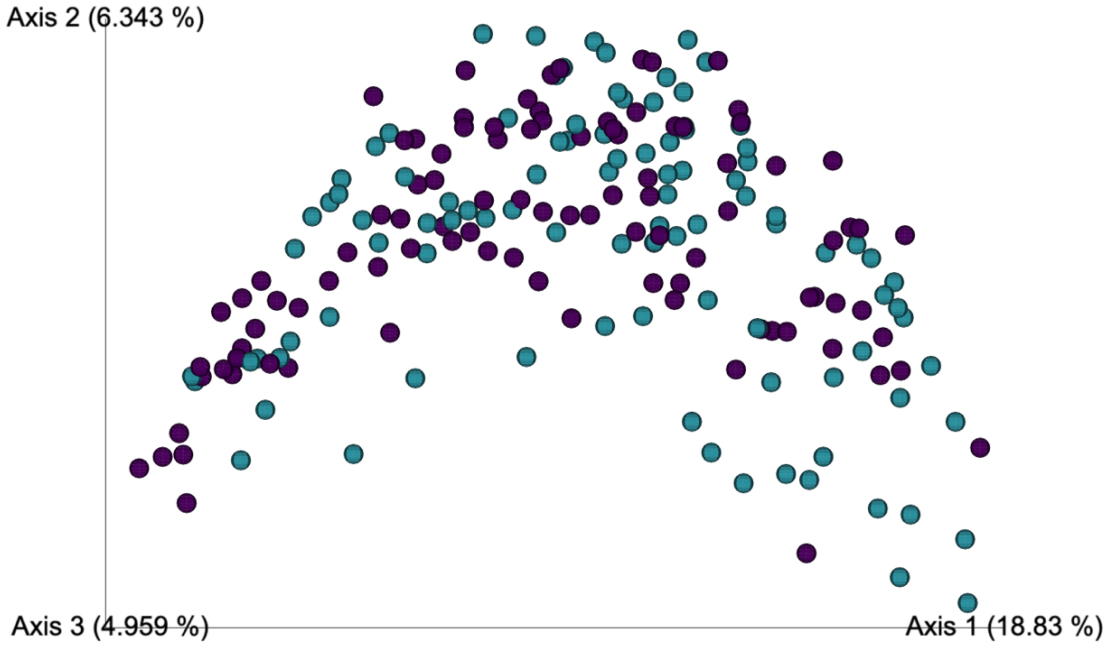

Figure S4.36. PCoA plot of high and low CREDI Language scores unweighted Unifrac beta diversity. Purple = high scores, green = low scores.

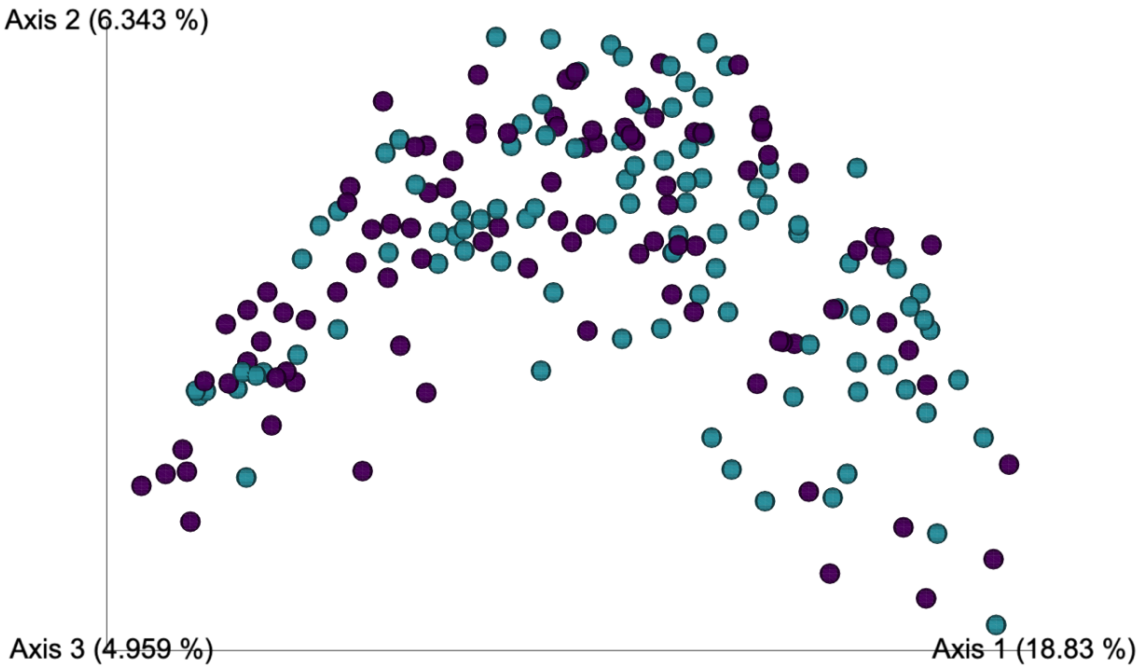

Figure S4.37. PCoA plot of high and low CREDI Motor scores unweighted Unifrac beta diversity. Purple = high scores, green = low scores.

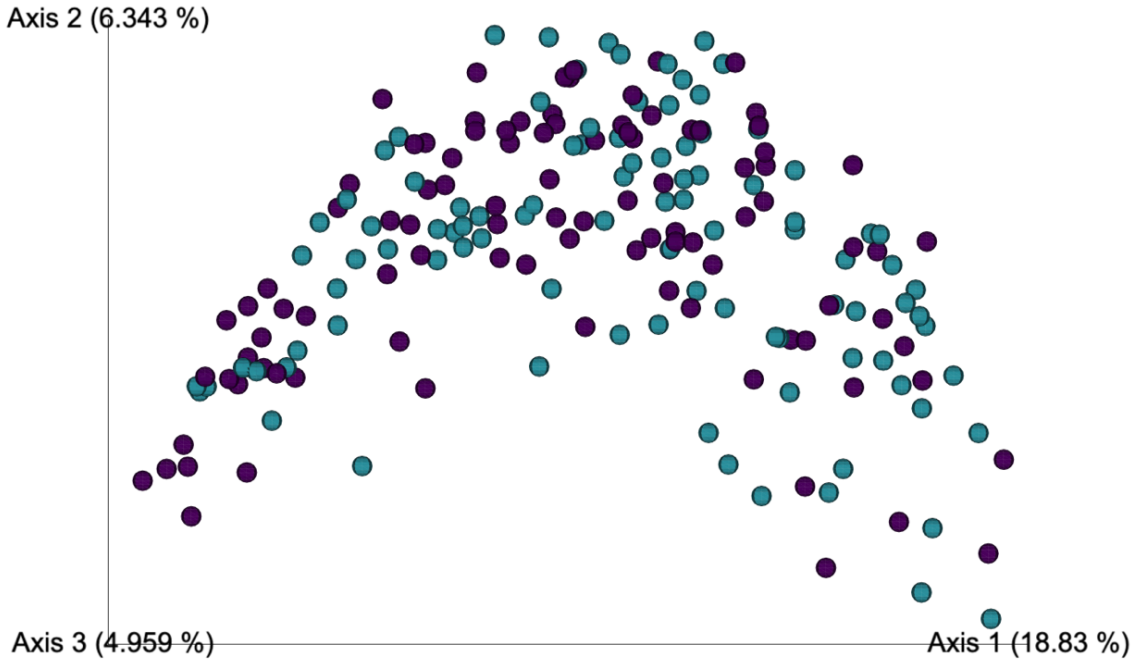

Figure S4.38. PCoA plot of high and low CREDI Socioemotional scores unweighted Unifrac beta diversity. Purple = high scores, green = low scores.

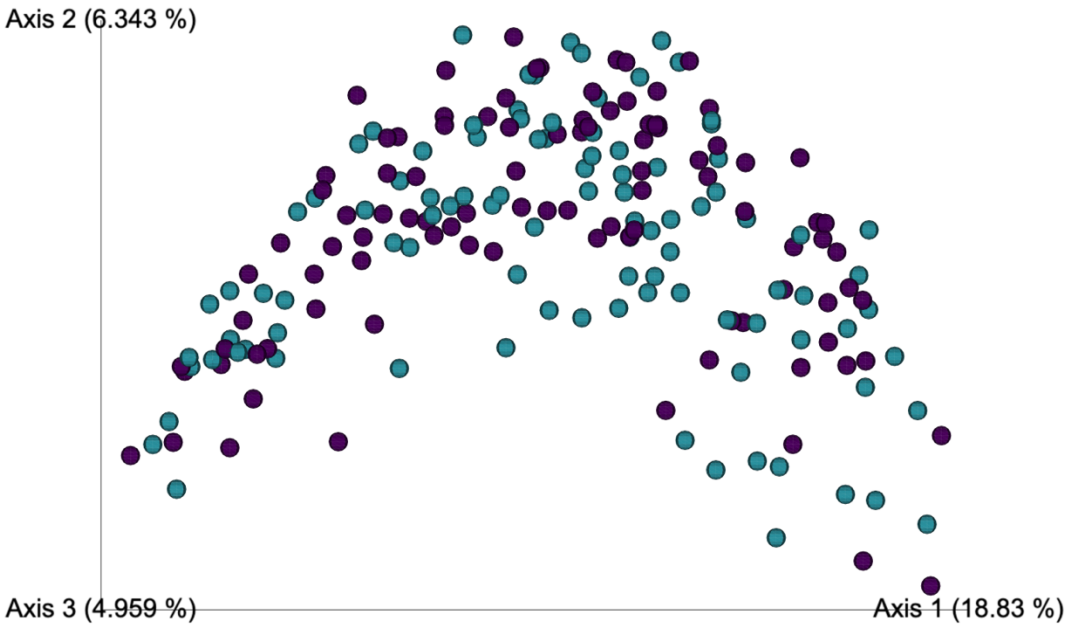

Figure S4.39. PCoA plot of high and low Z-length-for-age scores unweighted Unifrac beta diversity. Purple = high scores, green = low scores.

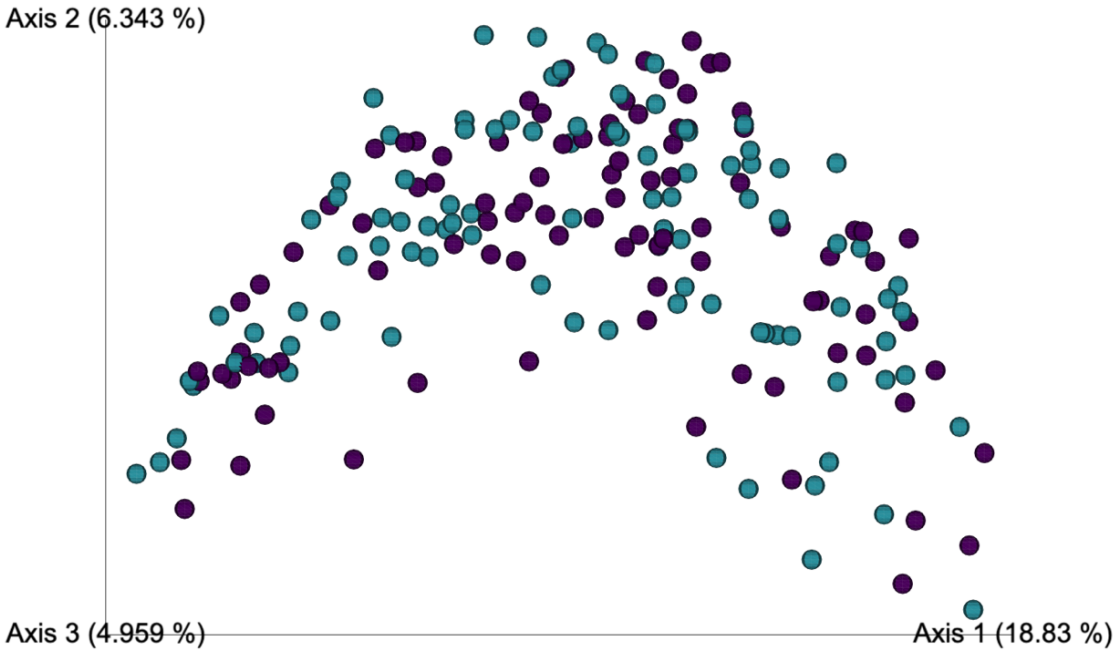

Figure S4.40. PCoA plot of high and low Z-BMI scores unweighted Unifrac beta diversity. Purple = high scores, green = low scores.

APPENDIX C: Chapter 5 Supplemental Materials

Supplemental Tables

Table S5.1. Linear regression results displaying the association between Z-Length-for-Age and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	SE	P Value	β (95% CI)	SE	P Value
Intercept	-1.83 (-2.06, -1.60)	0.115	1.7E-36	-0.792 (-2.789, 1.204)	1.01	0.434
Fecal calprotectin	-0.000197 (-0.000646, 0.000252)	0.000227	0.387	-0.000130 (-0.000597, 0.000337)	0.000236	0.584
Age				-0.0652 (-0.218, 0.0872)	0.0771	0.399
Sex (Girl)				0.431 (0.104, 0.758)	0.165	0.0100
Milk Type (Non-breastmilk)				-0.195 (-0.650, 0.261)	0.231	0.400
Antibiotics ever (Yes)				0.0160 (-0.376, 0.408)	0.198	0.936
Water source (Borehole/tubewell)				-0.335 (-0.924, 0.255)	0.298	0.264
Water source (Other)				-0.223 (-0.784, 0.337)	0.284	0.433
Water source (Piped water)				-0.274 (-0.711, 0.163)	0.221	0.217
Water source (Protected dug well)				-0.203 (-0.790, 0.384)	0.297	0.496
Low birthweight (Yes)				-0.492 (-0.919, -0.0639)	0.216	0.0245
Antibiotics 14 days (Yes)				0.0901 (-0.344, 0.524)	0.220	0.682
Diarrhea 14 days (Yes)				-0.179 (-0.530, 0.172)	0.178	0.316
Treatment (Egg)				-0.153 (-0.487, 0.181)	0.169	0.367
Poverty score				-0.0000463 (-0.00564, 0.00555)	0.00283	0.987
Model Summary						
R^2	0.00409			0.113		
Adjusted R^2	-0.00135			0.0326		

Model P value	0.387	0.157
Sample size (n)	185	169

Table S5.2. Linear regression results displaying the association between Z-Head-Circumference and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	**SE**	**P Value**	**β (95% CI)**	**SE**	**P Value**
Intercept	-0.760 (-0.938, -0.582)	0.0903	1.07E-14	-0.582 (-2.219, 1.056)	0.829	0.484
Fecal calprotectin	-0.000209 (-0.000563, 0.000145)	0.000179	0.245	-0.000176 (-0.000559, 0.000207)	0.000194	0.364
Age				0.000194 (-0.125, 0.125)	0.0633	0.998
Sex (Girl)				0.289 (0.0215, 0.557)	0.136	0.0344
Milk Type (Non-breastmilk)				-0.193 (-0.566, 0.181)	0.189	0.310
Antibiotics ever (Yes)				-0.0743 (-0.396, 0.247)	0.163	0.649
Water source (Borehole/tubewell)				-0.141 (-0.625, 0.342)	0.245	0.564
Water source (Other)				-0.589 (-1.048, -0.129)	0.233	0.0124
Water source (Piped water)				-0.190 (-0.549, 0.168)	0.182	0.296
Water source (Protected dug well)				-0.275 (-0.756, 0.206)	0.244	0.261
Low birthweight (Yes)				-0.0979 (-0.449, 0.253)	0.178	0.582
Antibiotics 14 days (Yes)				-0.0408 (-0.397, 0.315)	0.180	0.821
Diarrhea 14 days (Yes)				-0.184 (-0.472, 0.104)	0.146	0.208
Treatment (Egg)				0.125 (-0.149, 0.400)	0.139	0.369
Poverty score				-0.0000345 (-0.00465, 0.00456)	0.00232	0.988
Model Summary						
R^2	0.00737			0.101		
Adjusted R^2	0.00195			0.0197		
Model P value	0.245			0.251		
Sample size (n)	185			169		

Table S5.3. Linear regression results displaying the association between CREDI overall score and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	SE	P Value	β (95% CI)	SE	P Value
Intercept	47.5 (47.2, 47.8)	0.134	1.01E-261	39.7 (37.8, 41.5)	0.918	1.48E-92
Fecal Calprotectin	-0.000123 (-0.000647, 0.000400)	0.000265	0.642	0.000211 (-0.000225, 0.000647)	0.000221	0.340
Age				0.569 (0.432, 0.706)	0.069	6.35E-14
Sex (Girl)				-0.135 (-0.435, 0.166)	0.152	0.377
Milk Type (Non-breastmilk)				0.276 (-0.134, 0.686)	0.208	0.186
Antibiotics ever (Yes)				0.036 (-0.332, 0.405)	0.186	0.845
Water source (Borehole/tubewell)				-0.0706 (-0.618, 0.477)	0.277	0.799
Water source (Other)				-0.199 (-0.715, 0.318)	0.262	0.449
Water source (Piped water (inside home or land))				0.257 (-0.144, 0.659)	0.203	0.208
Water source (Protected dug well)				0.325 (-0.231, 0.881)	0.282	0.250
Antibiotics 14 days (Yes)				-0.125 (-0.514, 0.263)	0.197	0.525
Diarrhea 14 days (Yes)				0.220 (-0.101, 0.542)	0.163	0.178
Treatment (Egg)				0.0142 (-0.297, 0.325)	0.157	0.928
Poverty score				0.00278 (-0.00236, 0.00791)	0.00260	0.288

Play activities score (2-3)		0.546 (0.162, 0.930)	0.194	0.00555
Play activities score (4-6)		0.860 (0.423, 1.30)	0.221	0.000146
Stunting (Yes)		-0.514 (-0.820, -0.209)	0.155	0.00110
Model Summary				
R^2	0.00118	0.422		
Adjusted R^2	-0.00428	0.366		
Model P value	0.642	2.88E-13		
Sample size (n)	185	184		

Table S5.4. Linear regression results displaying the association between CREDI cognitive score and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	SE	P Value	β (95% CI)	SE	P Value
Intercept	48.9 (48.8, 49.1)	0.0903	2.89E-295	44.4 (43.1, 45.7)	0.646	1.92E-124
Fecal Calprotectin	0.0000150 (-0.000339, 0.000369)	0.000179	0.934	0.000241 (-0.0000660, 0.000547)	0.000155	0.123
Age				0.305 (0.209, 0.401)	0.0488	3.36E-09
Sex (Girl)				-0.0169 (-0.228, 0.195)	0.107	0.875
Milk Type (Non-breastmilk)				0.156 (-0.133, 0.444)	0.146	0.289
Antibiotics ever (Yes)				0.107 (-0.152, 0.366)	0.131	0.415
Water source (Borehole/tubewell)				-0.0594 (-0.445, 0.326)	0.195	0.761
Water source (Other)				-0.266 (-0.629, 0.0975)	0.184	0.151
Water source (Piped water (inside home or land))				0.0213 (-0.261, 0.304)	0.143	0.882
Water source (Protected dug well)				0.240 (-0.151, 0.631)	0.198	0.228
Antibiotics 14 days (Yes)				-0.0859 (-0.359, 0.187)	0.138	0.536
Diarrhea 14 days (Yes)				0.0354 (-0.191, 0.262)	0.115	0.758
Treatment (Egg)				0.0837 (-0.135, 0.302)	0.111	0.451
Poverty score				0.00302 (-0.000597, 0.00663)	0.00183	0.101

Play activities score (2-3)		0.553 (0.283, 0.823)	0.137	0.0000810
Play activities score (4-6)		0.770 (0.463, 1.08)	0.156	1.82E-06
Stunting (Yes)		-0.268 (-0.483, -0.0526)	0.109	0.0150
Model Summary				
R^2	0.0000380	0.371		
Adjusted R^2	-0.00543	0.310		
Model P value	0.934	1.39E-10		
Sample size (n)	185	184		

Table S5.5. Linear regression results displaying the association between CREDI motor score and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	SE	P Value	β (95% CI)	SE	P Value
Intercept	48.1 (48.0, 48.2)	0.0860	9.52E-298	42.6 (41.5, 43.8)	0.588	3.75E-128
Fecal Calprotectin	-0.000128 (-0.000465, 0.000209)	0.000171	0.454	0.0000519 (-0.000227, 0.000331)	0.000141	0.714
Age				0.410 (0.322, 0.497)	0.0444	1.25E-16
Sex (Girl)				-0.086 (-0.279, 0.106)	0.0975	0.377
Milk Type (Non-breastmilk)				0.115 (-0.148, 0.377)	0.133	0.389
Antibiotics ever (Yes)				0.052 (-0.184, 0.287)	0.119	0.667
Water source (Borehole/tubewell)				-0.0453 (-0.396, 0.305)	0.178	0.799
Water source (Other)				0.0200(-0.311, 0.350)	0.167	0.905
Water source (Piped water (inside home or land))				0.140 (-0.117, 0.397)	0.130	0.285
Water source (Protected dug well)				0.200 (-0.156, 0.556)	0.180	0.268
Antibiotics 14 days (Yes)				-0.116 (-0.365, 0.132)	0.126	0.357
Diarrhea 14 days (Yes)				0.122 (-0.0844, 0.328)	0.104	0.246
Treatment (Egg)				0.0119 (-0.187, 0.211)	0.101	0.906
Poverty score				0.00231 (-0.000978, 0.00560)	0.00167	0.167

Play activities score (2-3)		0.228 (-0.0176, 0.474)	0.124	0.0686
Play activities score (4-6)		0.360 (0.0809, 0.640)	0.142	0.0118
Stunting (Yes)		-0.385 (-0.581, -0.189)	0.099	0.000148
Model Summary				
R^2	0.00306	0.430		
Adjusted R^2	-0.00238	0.375		
Model P value	0.454	1.01E-13		
Sample size (n)	185	184		

Table S5.6. Linear regression results displaying the association between CREDI language score and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	SE	P Value	β (95% CI)	SE	P Value
Intercept	48.8 (48.6, 49.0)	0.0797	5.40E-305	45.2 (44.1, 46.4)	0.589	3.27E-132
Fecal Calprotectin	-0.0000449 (-0.000357, 0.000267)	0.000158	0.777	0.000143 (-0.000136, 0.000423)	0.000142	0.313
Age				0.245 (0.157, 0.333)	0.0445	1.41E-07
Sex (Girl)				0.0131 (-0.180, 0.206)	0.0977	0.893
Milk Type (Non-breastmilk)				0.171 (-0.0916, 0.434)	0.133	0.200
Antibiotics ever (Yes)				-0.0251 (-0.261, 0.211)	0.120	0.834
Water source (Borehole/tubewell)				0.0179 (-0.333, 0.369)	0.178	0.920
Water source (Other)				-0.264 (-0.595, 0.0670)	0.168	0.117
Water source (Piped water (inside home or land))				0.203 (-0.0545, 0.461)	0.131	0.122
Water source (Protected dug well)				0.300 (-0.0564, 0.657)	0.181	0.0983
Antibiotics 14 days (Yes)				-0.0102 (-0.259, 0.239)	0.126	0.936
Diarrhea 14 days (Yes)				0.120 (-0.0867, 0.326)	0.105	0.254
Treatment (Egg)				-0.0684 (-0.268, 0.131)	0.101	0.499
Poverty score				0.00171 (-0.00159, 0.00501)	0.00167	0.307

Play activities score (2-3)		0.381 (0.135, 0.627)	0.125	0.003
Play activities score (4-6)		0.568 (0.288, 0.849)	0.142	0.0000927
Stunting (Yes)		-0.204 (-0.400, -0.00801)	0.0994	0.0415
Model Summary				
R^2	0.000439	0.333		
Adjusted R^2	-0.00502	0.269		
Model P value	0.777	8.62E-09		
Sample size (n)	185	184		

Table S5.7. Linear regression results displaying the association between CREDI socioemotional score and fecal calprotectin unadjusted and adjusted models.

Variable	Unadjusted			Adjusted		
	β (95% CI)	SE	P Value	β (95% CI)	SE	P Value
Intercept	48.8 (48.6, 49.0)	0.0887	1.80E-296	43.7 (42.5, 44.9)	0.612	4.58E-127
Fecal Calprotectin	-0.0000190 (-0.000367, 0.000329)	0.000176	0.914	0.000207 (-0.0000839, 0.000497)	0.000147	0.162
Age				0.352 (0.260, 0.443)	0.0463	2.00E-12
Sex (Girl)				-0.0365 (-0.237, 0.164)	0.101	0.720
Milk Type (Non-breastmilk)				0.182 (-0.0909, 0.456)	0.138	0.189
Antibiotics ever (Yes)				0.122 (-0.124, 0.367)	0.124	0.329
Water source (Borehole/tubewell)				0.0565 (-0.309, 0.422)	0.185	0.760
Water source (Other)				-0.189 (-0.533, 0.155)	0.174	0.279
Water source (Piped water (inside home or land))				0.0830 (-0.185, 0.351)	0.136	0.542
Water source (Protected dug well)				0.189 (-0.182, 0.559)	0.188	0.317
Antibiotics 14 days (Yes)				-0.0338 (-0.293, 0.225)	0.131	0.797
Diarrhea 14 days (Yes)				0.0572 (-0.157, 0.272)	0.109	0.599
Treatment (Egg)				0.0746 (-0.133, 0.282)	0.105	0.478
Poverty score				0.00319 (-0.000231, 0.00662)	0.00173	0.0674
Play activities score (2-3)				0.515 (0.259, 0.771)	0.130	0.000105

Play activities score (4-6)		0.685 (0.394, 0.976)	0.147	6.77E-06
Stunting (Yes)		-0.265 (-0.468, -0.0607)	0.103	0.011
Model Summary				
R^2	0.0000636	0.413		
Adjusted R^2	-0.00540	0.357		
Model P value	0.914	8.30E-13		
Sample size (n)	185	184		

Supplemental Figures

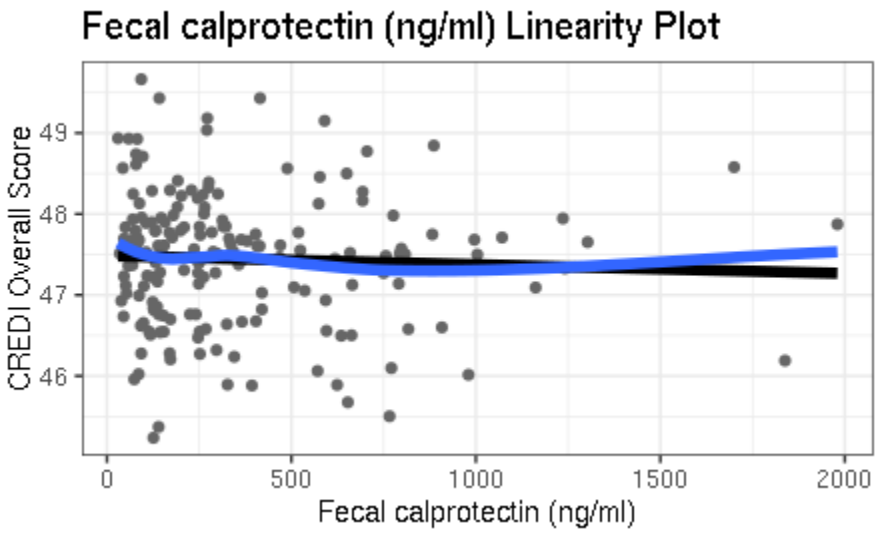

Figure S5.1. Scatterplot assessing linearity between fecal calprotectin and CREDI Overall score from the adjusted model.

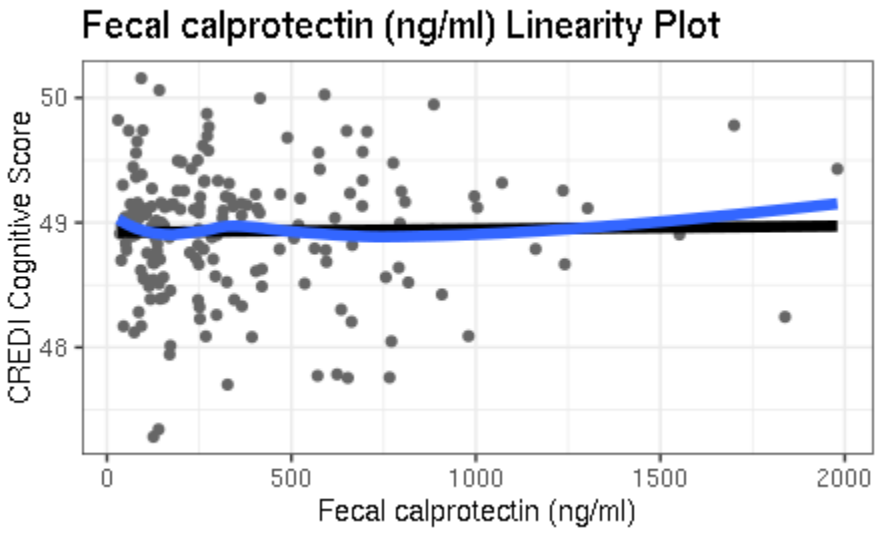

Figure S5.2. Scatterplot assessing linearity between fecal calprotectin and CREDI Cognitive score from the adjusted model.

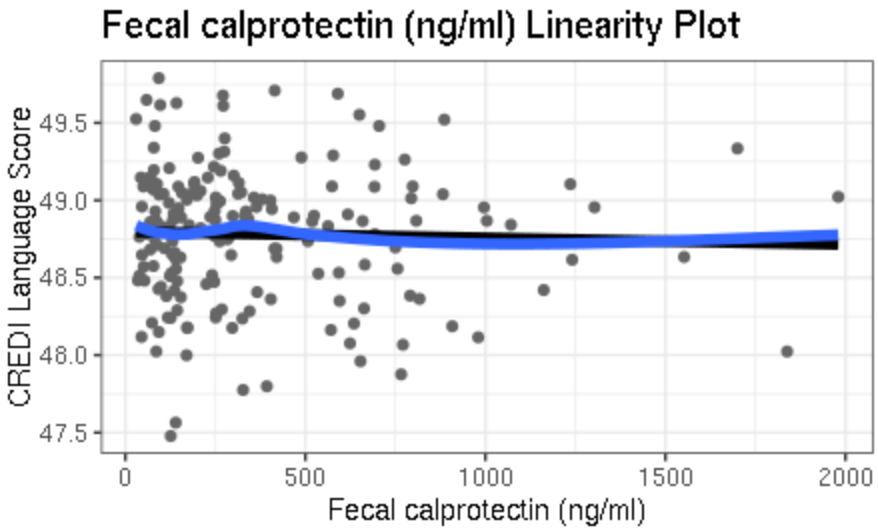

Figure S5.3. Scatterplot assessing linearity between fecal calprotectin and CREDI Language score from the adjusted model.

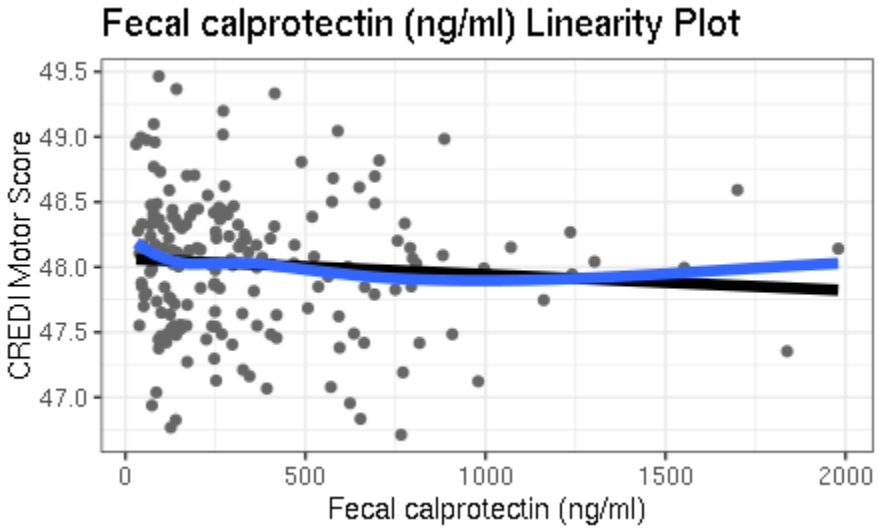

Figure S5.4. Scatterplot assessing linearity between fecal calprotectin and CREDI Motor score from the adjusted model.

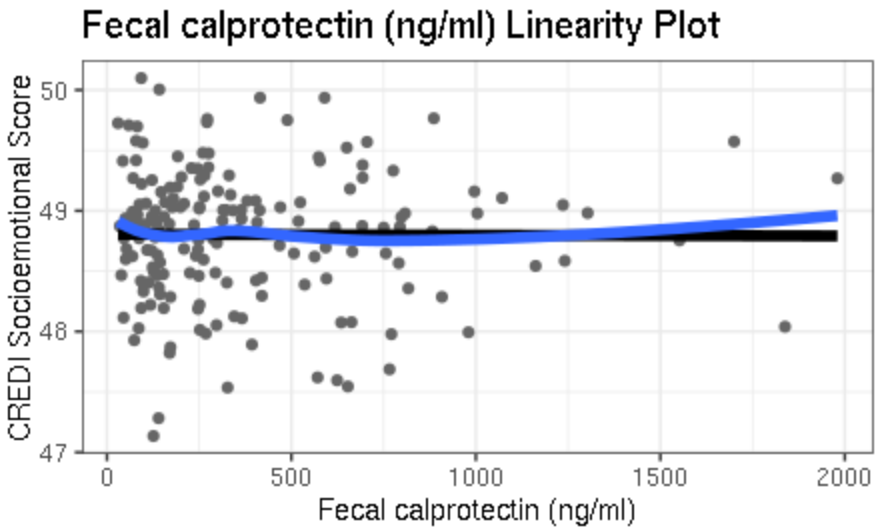

Figure S5.5. Scatterplot assessing linearity between fecal calprotectin and CREDI Socioemotional score from the adjusted model.

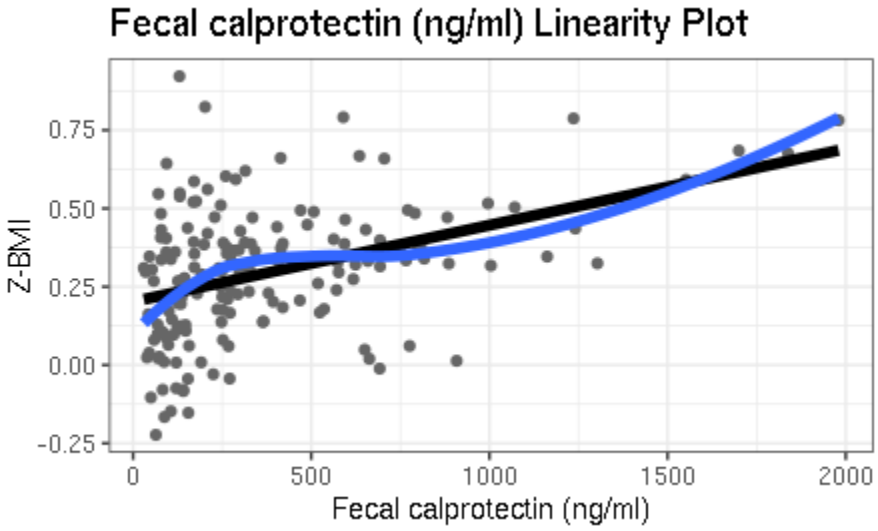

Figure S5.6. Scatterplot assessing linearity between fecal calprotectin and Z-BMI from the adjusted model.

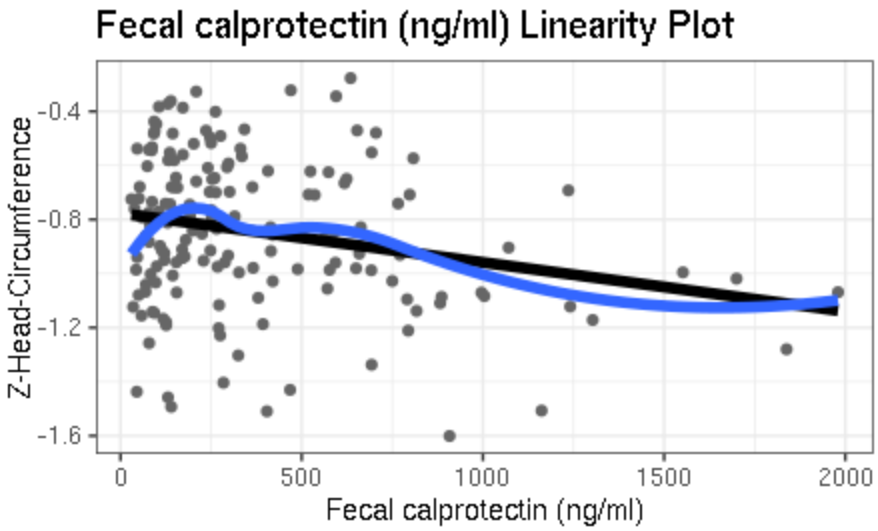

Figure S5.7. Scatterplot assessing linearity between fecal calprotectin and Z-Head-Circumference from the adjusted model.

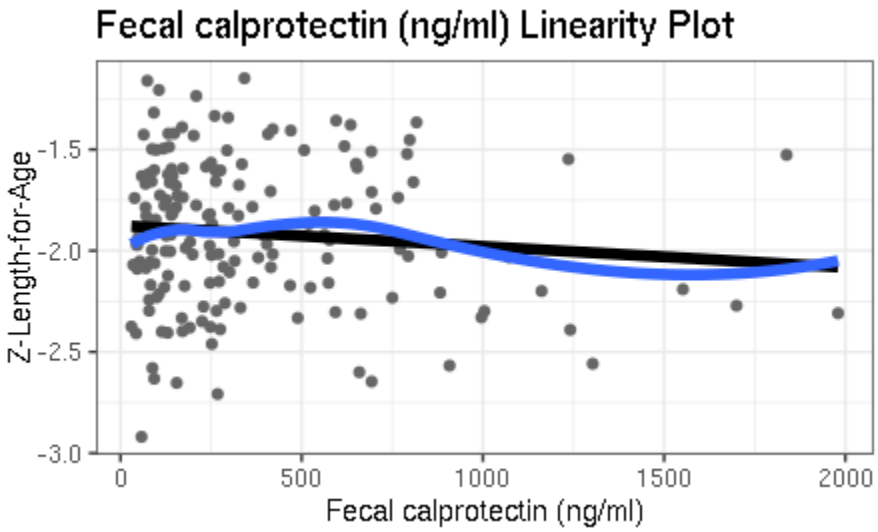

Figure S5.8. Scatterplot assessing linearity between fecal calprotectin and Z-Length-for-Age from the adjusted model.

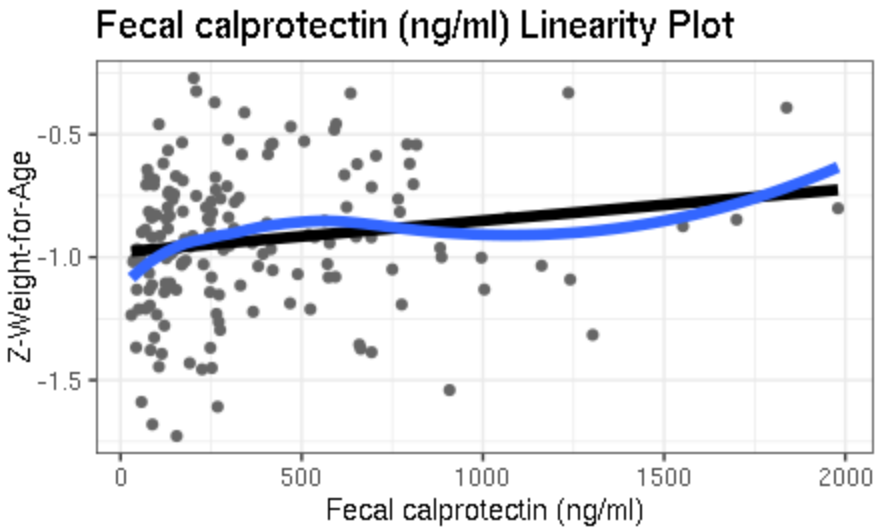

Figure S5.9. Scatterplot assessing linearity between fecal calprotectin and Z-Weight-for-Age from the adjusted model.

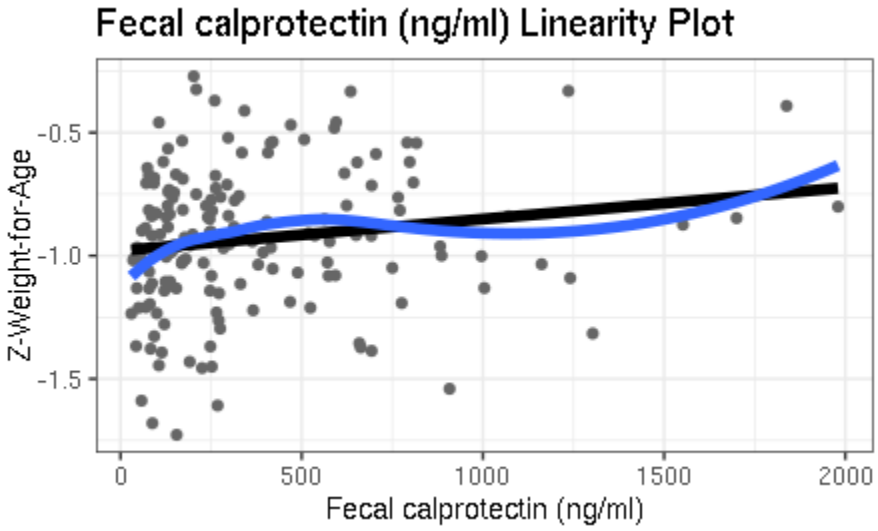

Figure S5.10. Scatterplot assessing linearity between fecal calprotectin and Z-Weight-for-Age from the adjusted model.

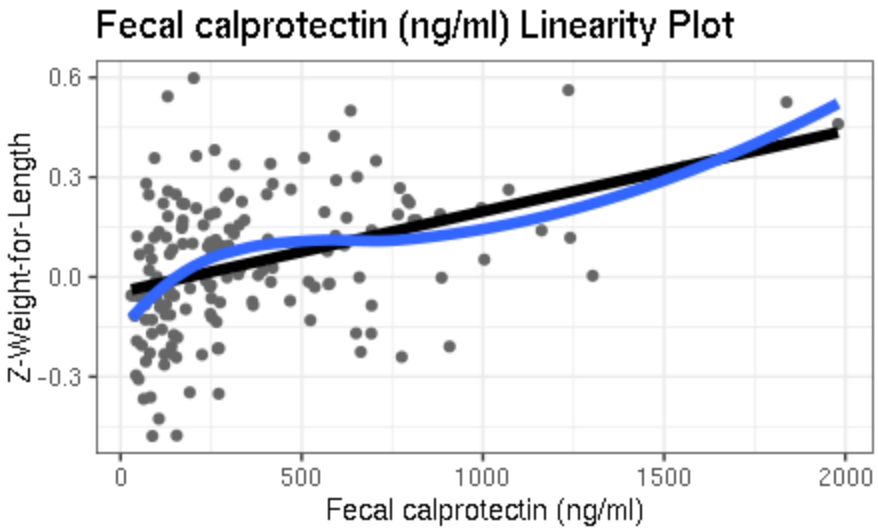

Figure S5.11. Scatterplot assessing linearity between fecal calprotectin and Z-Weight-for-Length from the adjusted model.

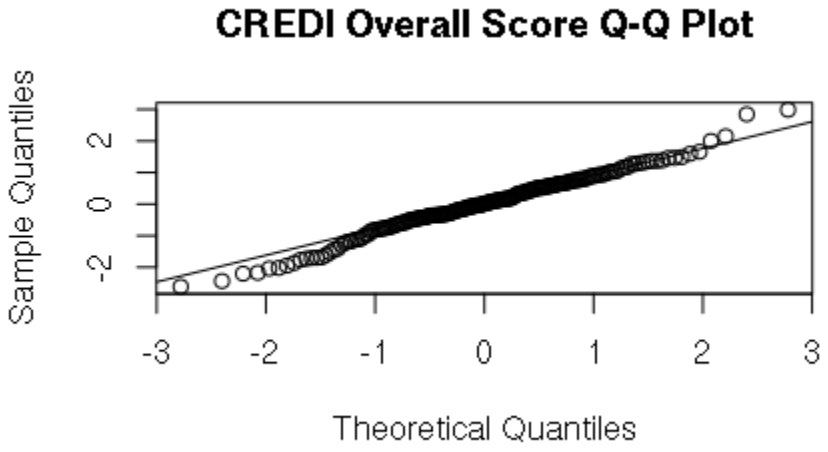

Figure S5.12. Q-Q Plot of Residuals: Adjusted Model of Fecal Calprotectin and CREDI Overall Score.

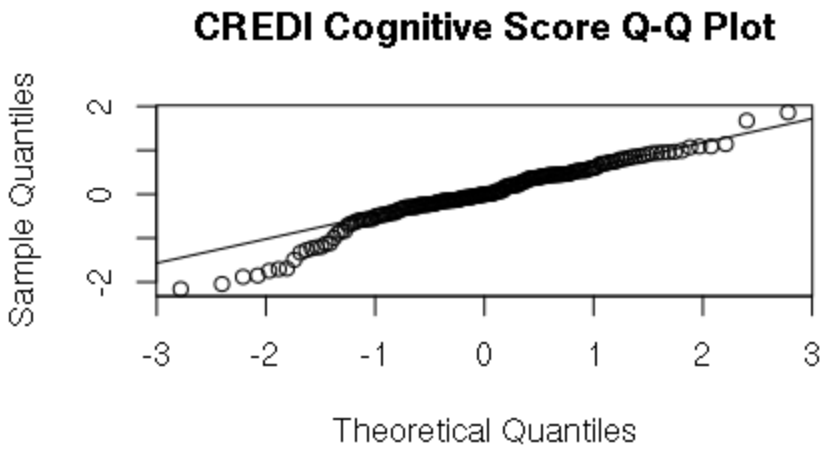

Figure S5.13. Q-Q Plot of Residuals: adjusted model of fecal calprotectin and CREDI Cognitive score.

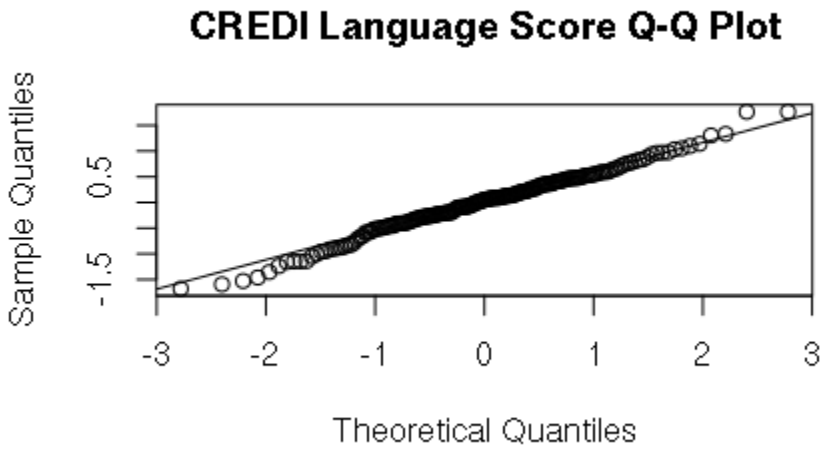

Figure S5.14. Q-Q Plot of Residuals: adjusted model of fecal calprotectin and CREDI Language score.

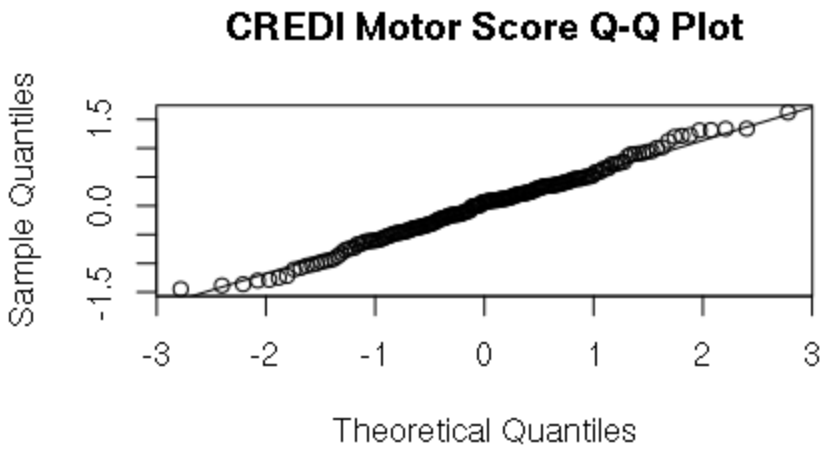

Figure S5.15. Q-Q Plot of Residuals: adjusted model of fecal calprotectin and CREDI Motor score.

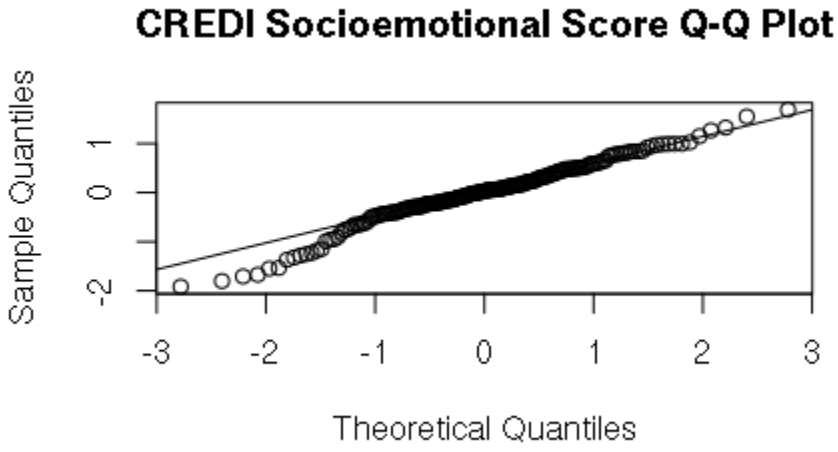

Figure S5.16. Q-Q Plot of Residuals: adjusted model of fecal calprotectin and CREDI Socioemotional score.

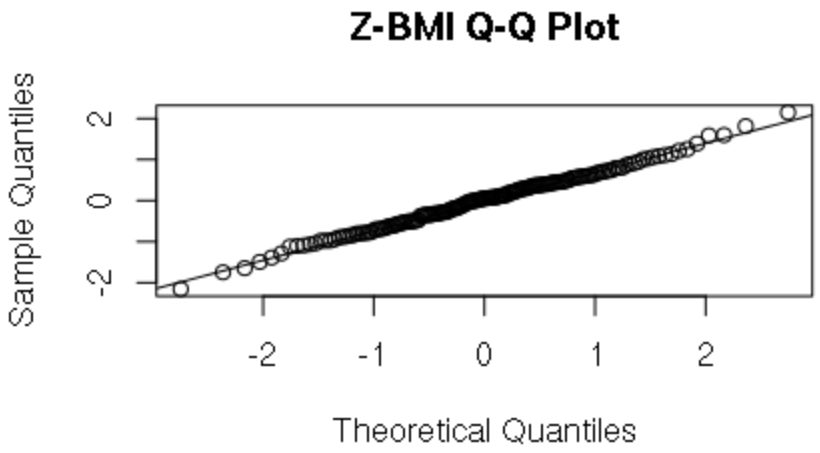

Figure S5.17. Q-Q Plot of Residuals: adjusted model of fecal calprotectin and Z-BMI score.

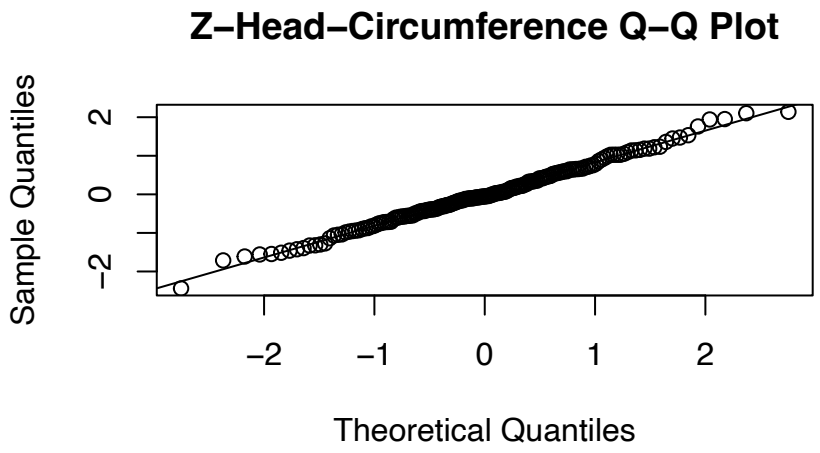

Figure S5.18. Q-Q Plot of Residuals: adjusted model of fecal calprotectin and Z-Head-Circumference score.

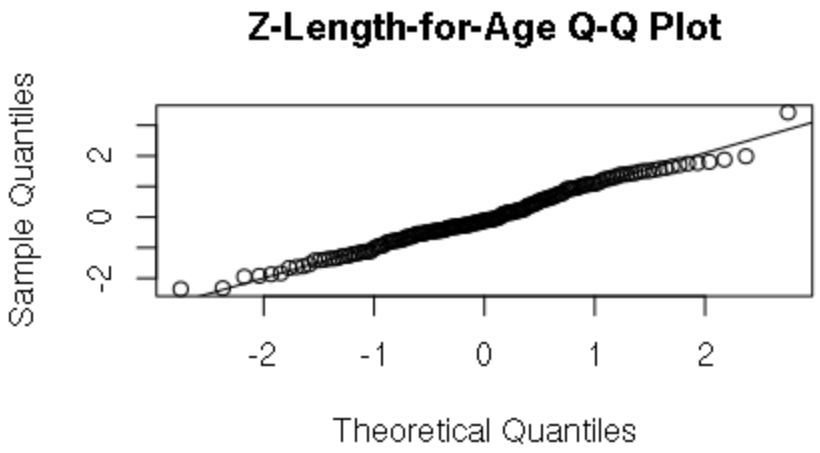

Figure S5.19. Q-Q Plot of Residuals: adjusted model of fecal calprotectin and Z-BMI score.

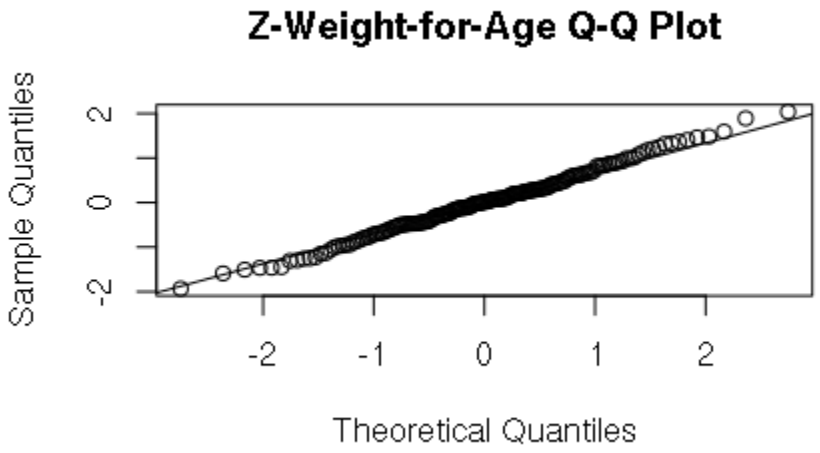

Figure S5.20. Q-Q Plot of Residuals: adjusted model of fecal calprotectin and Z-Weight-for-Age score.

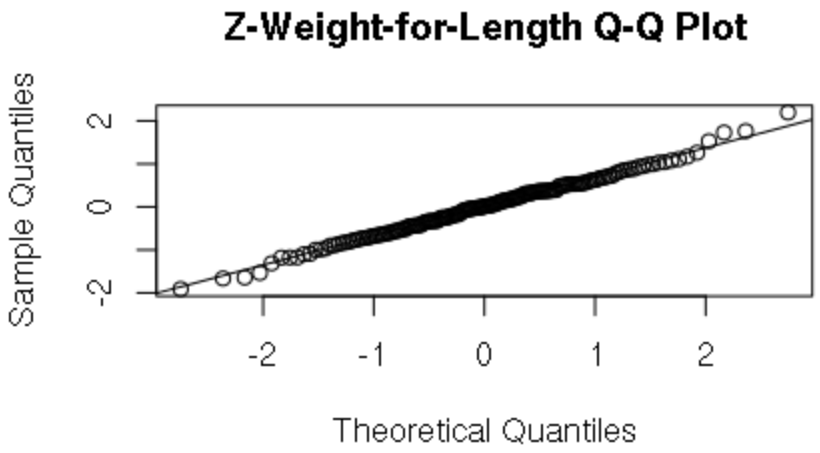

Figure S5.21. Q-Q Plot of Residuals: adjusted model of fecal calprotectin and Z-Weight-for-Length score.

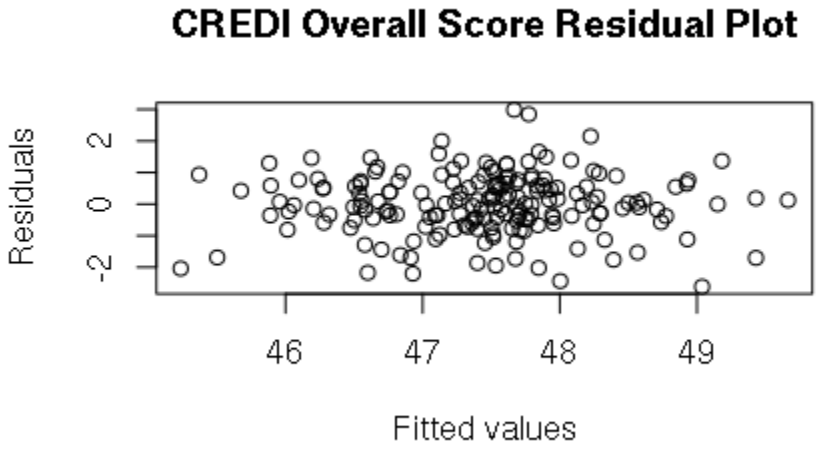

Figure S5.22. Residuals Plot: adjusted model of fecal calprotectin and CREDI Overall Score.

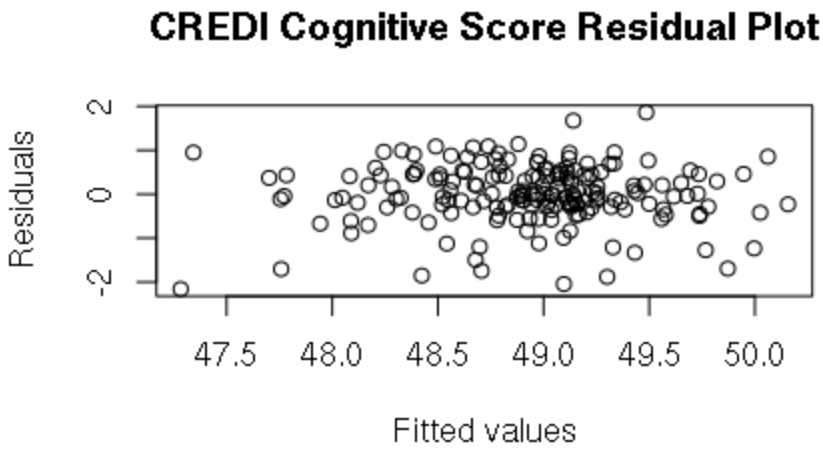

Figure S5.23. Residuals Plot: adjusted model of fecal calprotectin and CREDI Cognitive Score.

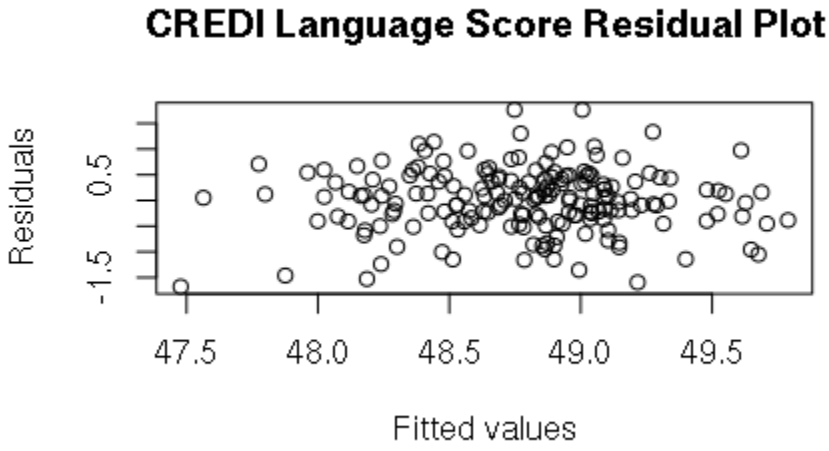

Figure S5.24. Residuals Plot: adjusted model of fecal calprotectin and CREDI Language Score.

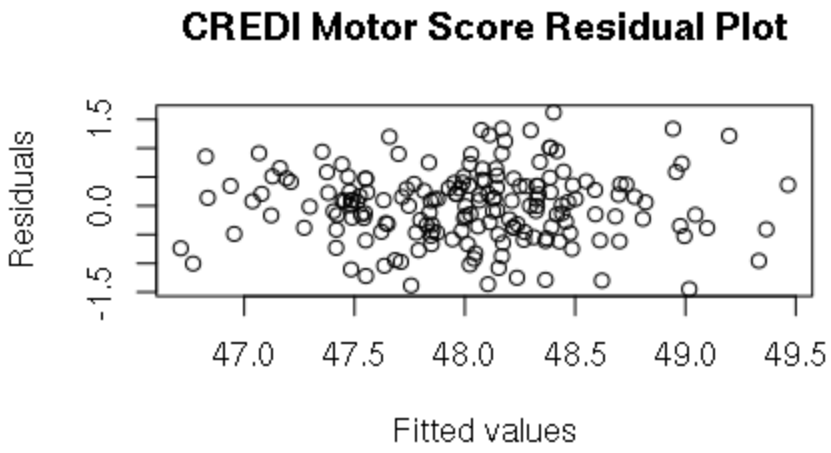

Figure S5.25. Residuals Plot: adjusted model of fecal calprotectin and CREDI Motor Score.

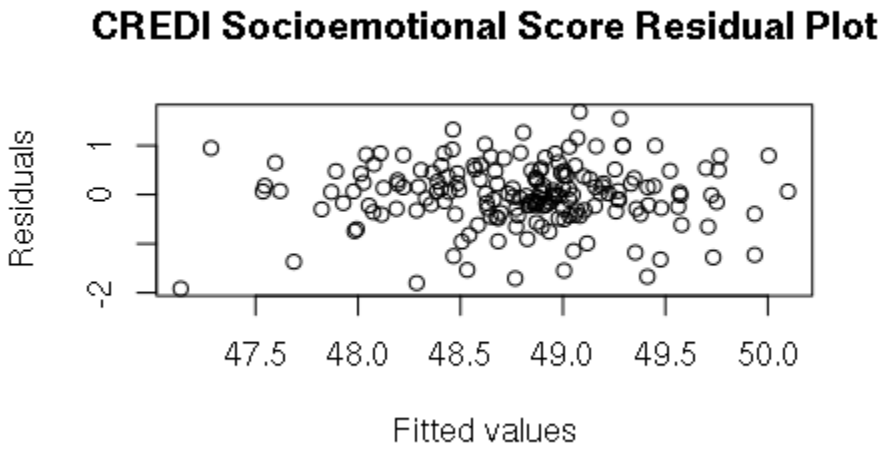

Figure S5.26. Residuals Plot: adjusted model of fecal calprotectin and CREDI Socioemotional Score.

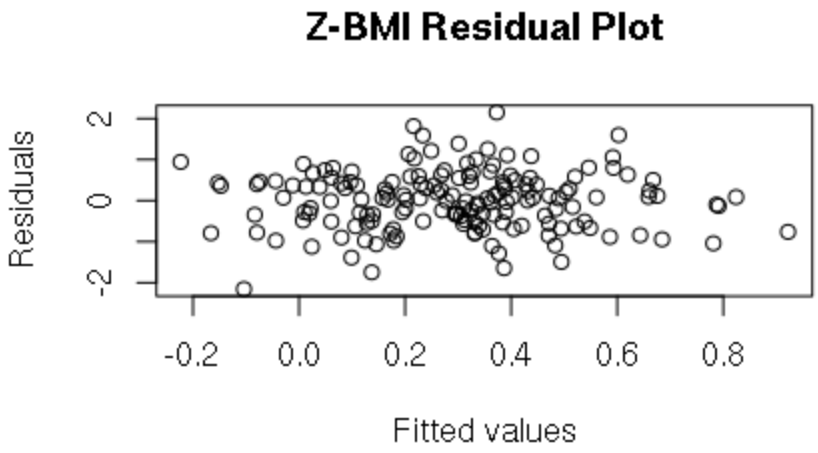

Figure S5.27. Residuals Plot: adjusted model of fecal calprotectin and Z-BMI Score.

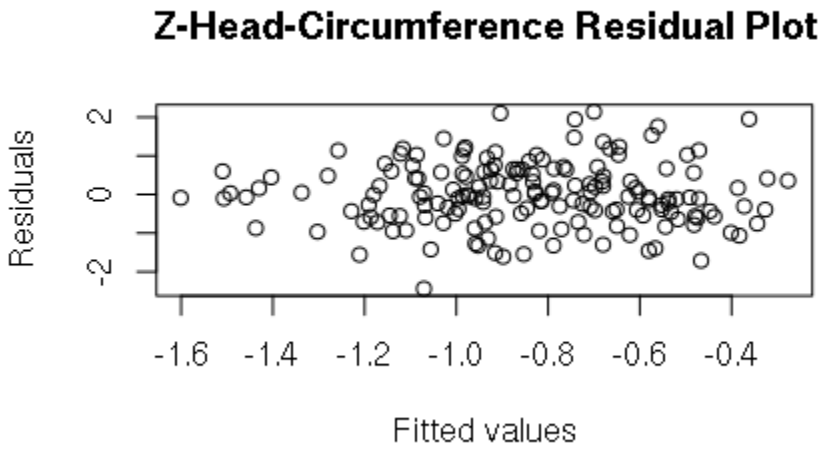

Figure S5.28. Residuals Plot: adjusted model of fecal calprotectin and Z-Head-Circumference Score.

Figure S5.29. Residuals Plot: adjusted model of fecal calprotectin and Z-Length-for-Age Score.

Figure S5.30. Residuals Plot: adjusted model of fecal calprotectin and Z-Weight-for-Age Score.

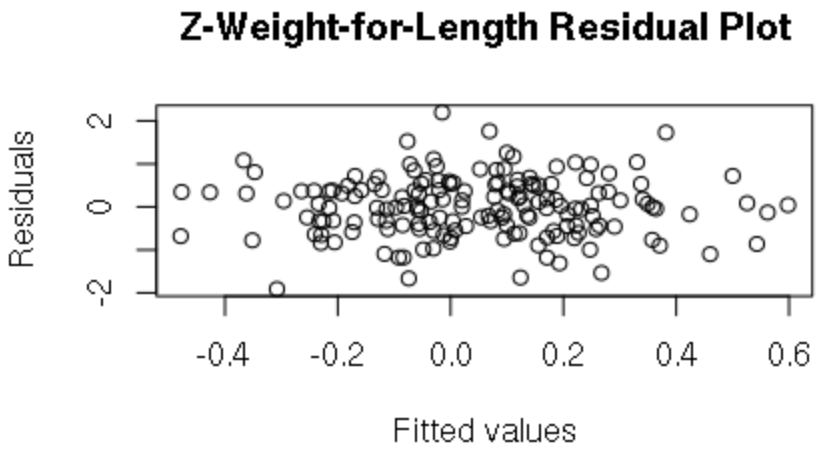

Figure S5.31. Residuals Plot: adjusted model of fecal calprotectin and Z-Weight-for-Length Score.