*Article*

# The Macronutrient Composition of Infant Formula Produces Differences in Gut Microbiota Maturation That Associate with Weight Gain Velocity and Weight Status

Julie A. Mennella [1,†], Yun Li [2], Kyle Bittinger [3], Elliot S. Friedman [4], Chunyu Zhao [3], Hongzhe Li [2], Gary D. Wu [4] and Jillian C. Trabulsi [5,*,†]

1    Monell Chemical Senses Center, Philadelphia, PA 19104, USA; mennella@monell.org
2    Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; yunli1@pennmedicine.upenn.edu (Y.L.); hongzhe@pennmedicine.upenn.edu (H.L.)
3    Division of Gastroenterology, Hepatology, and Nutrition, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; bittingerk@chop.edu (K.B.); chunyu.zhao@czbiohub.org (C.Z.)
4    Division of Gastroenterology and Hepatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; elliotf@pennmedicine.upenn.edu (E.S.F.); gdwu@pennmedicine.upenn.edu (G.D.W.)
5    Department of Behavioral Health and Nutrition, University of Delaware, Newark, DE 19713, USA
*    Correspondence: trabulsi@udel.edu; Tel.: +1-302-831-4991
†    These authors contributed equally to this work.

**Abstract:** This proof-of-principle study analyzed fecal samples from 30 infants who participated in a randomized controlled trial on the effects of the macronutrient composition of infant formula on growth and energy balance. In that study, infants randomized to be fed cow milk formula (CMF) had faster weight-gain velocity during the first 4 months and higher weight-for-length Z scores up to 11.5 months than those randomized to an isocaloric extensive protein hydrolysate formula (EHF). Here we examined associations among infant formula composition, gut microbial composition and maturation, and children's weight status. Fecal samples collected before and monthly up to 4.5 months after randomization were analyzed by shotgun metagenomic sequencing and targeted metabolomics. The EHF group had faster maturation of gut microbiota than the CMF group, and increased alpha diversity driven by Clostridia taxa. Abundance of *Ruminococcus gnavus* distinguished the two groups after exclusive feeding of the assigned formula for 3 months. Abundance of Clostridia at 3–4 months negatively correlated with prior weight-gain velocity and body weight phenotypes when they became toddlers. Macronutrient differences between the formulas likely led to the observed divergence in gut microbiota composition that was associated with differences in transient rapid weight gain, a well-established predictor of childhood obesity and other comorbidities.

**Keywords:** infant; microbiota; metabolome; diet; gene; infant formula; rapid weight gain; randomized controlled trial

## 1. Introduction

There are sensitive periods during early life when diet interacts with the gut microbiota in shaping the structure and function of the mucosal immune system and, in turn, modifying metabolic and immune health [1–3]. While the initial colonization pattern of newborns' gut microbiome is often considered chaotic, with low species diversity, a growing body of evidence—the vast majority from observational and cross-sectional research—suggests that maturation and diversification of the gut bacterial environment during infancy is shaped by early diet [4–9].

Although the early diet is unique in typically consisting of a sole source of nutrition in liquid form—breast milk or breast milk substitute—the type of diet is important in

shaping the gut microbiome. During the first 4 months, gut microbial alpha diversity of exclusively breastfed infants, as expressed by the standardized Shannon index, was consistently lower than in breastfed infants whose diets were supplemented with infant formula and/or solid foods [9,10]. Supplementation with relatively small amounts of infant formula can shift the more "stable" and uniform microbiome of exclusively breastfed infants to one with a broader and more diversified spectrum of bacteria [8,11,12]. Prominence of Enterobacteriaceae, especially *E. coli* and *K. pneumoniae*, and abundance of Clostridiaceae have been observed among formula-fed infants. In particular, the relative abundance of Clostridia, a class of bacteria involved in amino acid fermentation [13–15], is greater in formula-fed than in breastfed infants younger than 6 months [5,11,16].

Based on widespread evidence that diet plays a significant role in the composition of the gut microbiota and long-term health, we hypothesized that there would be significant differences in the gut microbiome over time in infants exclusively fed different infant formulas that are isocaloric yet with different macronutrient composition. Specifically, we posited differences between infants fed cow milk formula (CMF), the most commonly ingested type of infant formula [17], and infants fed an extensive protein hydrolysate formula (EHF), a type of formula for infants with intolerance to cow's milk proteins. Unlike the intact proteins found in CMF, the proteins in EHF have been hydrolyzed to reduce the burden of digestion and allergenicity; consequently, EHF is abundant in free amino acids and small peptides [18]. While the carbohydrate source in CMF is lactose, EHF carbohydrate comprises non-lactose sources that, along with free amino acids, have been shown to impact the composition and function of the adult gut microbiome [19,20]. Randomized controlled trials (RCTs) have consistently revealed that infants fed EHF exhibit more normative weight gain during early infancy [21–23] and decreased risks of allergic diseases during childhood when compared to those fed CMF. For example, in the German Infant Nutritional Intervention RCT, breastfed infants with a family history of atopy who were randomized at birth to be fed EHF, only if human milk was insufficient, remained at decreased risk for atopy 20 years later compared to those randomized to CMF [24].

In the present proof-of-principle study, we randomly selected a subsample of 30 infants from our RCT of exclusively formula-fed infants with no family history of atopy. The RCT demonstrated that the accelerated weight gain pattern among CMF-fed infants compared to EHF-fed infants was due to both energy intake and energy loss mechanisms. In this subsample, we focused on group-dependent differences in the maturation of the microbiota from 0.5 to 4.5 months, when formula feeding was exclusive and when the microbiota is highly plastic and responsive to external factors, such as diet [25]. We posited that there would be time-dependent differences in gut microbiome composition that were related to differences in the macronutrient composition of the randomized formulas that would, in turn, be associated with clinical outcomes. We used an integrated analysis of shotgun metagenomic sequencing and metabolomics on fecal samples collected before (baseline: 0.5 and 0.75 months) and at fixed intervals after randomization (1.5, 2.5, 3.5, and 4.5 months) to determine the effects of infant formula diet on the composition and metabolic function of the infant gut microbiome and its relationship to clinical outcomes.

## 2. Materials and Methods

### 2.1. Participants and Trial Design

Study participants were enrolled in a longitudinal, double-blind RCT on healthy, term infants designed to determine the effects of feeding isocaloric commercial infant formulas (CMF, EHF) that differed in macronutrient composition (see Section 2.1.2) on early weight gain and energy balance. From the intent-to-treat cohort of 113 infants, we randomly selected 30 infants while balancing for sex and race/ethnicity. The study design, inclusion and exclusion criteria, and CONSORT table for the 12-month-long RCT have been published previously [26]. In brief, infants who were born at term and whose mothers chose to exclusively formula feed when infants were 2 weeks of age were randomized to be fed either CMF or EHF for the first year. The Office of Regulatory Affairs at the University

of Pennsylvania approved the study, and the trial was registered online at clinicaltrials.gov (NCT01700205) prior to its start.

### 2.1.1. Study Cohort

The cohort of 30 exclusively formula-fed infants (15 CMF, 15 EHF) was racially diverse (57% black, 27% white, 17% more than one race). There were no significant differences between the formula treatment groups in sex ratio of infants or any baseline infant or maternal characteristics measured [26], including infant anthropometric measurements taken at baseline (Table 1).

**Table 1.** Baseline characteristics of study cohort at study entry (0.5 months).

| Characteristics | Infant Formula Treatment Group | | |
| --- | --- | --- | --- |
| | CMF (*n* = 15) | EHF (*n* = 15) | *p* Value |
| Age in months | 0.39 ± 0.02 | 0.40 ± 0.02 | 0.90 |
| Female, *n* (%) | 6 (40%) | 6 (40%) | 1.00 |
| Race/ethnicity, *n* (%) | | | |
|     Black | 9 (60%) | 8 (53%) | |
|     White | 4 (27%) | 4 (27%) | 0.88 |
|     More than one race/ethnicity | 2 (13%) | 3 (20%) | |
| Anthropometry, Z score; 0.5 mos | | | |
|     Weight for age (WAZ) | −0.54 ± 0.21 | −0.37 ± 0.21 | 0.57 |
|     Length for age (LAZ) | −0.73 ± 0.27 | −0.65 ± 0.67 | 0.85 |
|     Weight for length (WLZ) | −0.25 ± 0.24 | −0.06 ± 0.24 | 0.59 |
| Body composition [1], 0.75 mos | | | |
|     Fat mass (kg) | 0.43 ± 0.08 | 0.49 ± 0.08 | 0.60 |
|     Percent body fat (%) | 11.6 ± 2.0 | 12.8 ± 2.2 | 0.70 |

Mean ± standard error of the mean (SEM) or *n* (%). *n* = 30 unless otherwise indicated; *p* values for main effect of infant formula treatment group (CMF, cow milk formula; EHF, extensive protein hydrolysate formula). [1] CMF, mboxemphn = 15, EHF, *n* = 12.

### 2.1.2. Composition of Infant Formula Diets

The two infant formulas (Enfamil [CMF], Nutramigen [EHF]; Mead Johnson Nutrition) were isocaloric (67 kcal/100 mL) and contained no added prebiotics or probiotics. While the percent energy provided from fat (48% [5.3 g/100 kcal]) and type of fat blend (i.e., palm olein, soy, coconut, high-oleic sunflower oils) were similar, the percent energy from carbohydrate (EHF: 41% [10.3 g/100 kcal], CMF: 44% [11.3 g/100 kcal]) and the percent energy from protein (EHF: 11% [2.8 g/100 kcal], CMF: 8% [2.1 g/100 kcal]) slightly differed [27]. The major differentiators were the form of the protein and the source of carbohydrate. CMF contained intact cow milk proteins and its source of carbohydrate was lactose, whereas EHF contained small peptides [28] and substantially higher concentrations of free amino acids (80,375 μmol/L) because of its extensive hydrolyzation than CMF (864 μmol/L) [18], and its source of carbohydrate was primarily corn syrup solids with some modified cornstarch.

### 2.1.3. Clinical Phenotypes

We tested measures of anthropometry and body composition from the RCT [26] for associations with infants' microbiota identified in this proof-of-concept study. At baseline visits (0.5, 0.75 months) and each month thereafter until 12.5 months, infants were weighed and measured in triplicate by research personnel certified in standard anthropometric techniques and with a calibrated pediatric scale (Scale-Tronix, White Plains, NY, USA) and an infantometer (Harpenden Infantometer 702; Crymych, Dyfed, UK) that were accurate to 0.001 kg and 0.1 cm, respectively. Anthropometric data were converted to weight-for length Z (WLZ) and length-for age Z scores using World Health Organization growth standards [29]. Velocities of early weight (g/day) and length (cm/day) accretion were calculated by dividing the change of weight in grams or length in centimeters by the change of age in days from 0.5 to 4.5 months, when infant formula accounted for 93% (±3) of daily energy intake. Of the 30 children in this cohort, 26 completed the final RCT visit at 12.5 months. At this age, only 24% (±5) of their daily energy intake was

from infant formula. We determined children's weight status at 1 year based on age- and sex-specific WLZ percentiles, which defines overweight as values greater than the 85th percentile [30]. There was no difference between the groups in the age at which mothers began complementing their diets with solids ($p = 0.56$).

Measures of body composition were obtained by the double-labeled water method at the baseline visit and at 3.5 and 12.5 months [26,31]. After obtaining a baseline urine sample, infants were dosed based on body weight with 0.3 g $H_2^{18}O$ and 0.15 g $^2H_2O$ per kilogram estimated total body water (TBW). Mothers were given special study diapers to collect morning urine samples each day thereafter for one week. The elimination rates of $^{18}O$ and $^2H$ were used to determine isotope dilution spaces (kg) and TBW, from which we determined fat-free mass (kg) for each infant at each age by dividing TBW derived from $^{18}O$ and $^2H$ dilution by age-specific hydration constants [32,33]. From these data, we determined the fat mass (kg) and percent body fat for each infant at each age. Of the 30 infants, body composition data were available for 27 infants at 0.75 months, 28 infants at 3.5 months, and 25 infants at 12.5 months.

*2.2. Methodology*

During the trial, convenient stool samples were collected in 76 × 20 mm feces containers (Sarstedt, Inc., Newton, NC) and frozen at −70 °C until transport to the Children's Hospital of Philadelphia Microbiome Center for analysis. For the present shotgun metagenomics and bacterial gene analyses, we analyzed 148 fecal samples. For each infant, we had at least one baseline (0.5, 0.75) sample and at least three post-randomization stool samples (1.5, 2.5, 3.5, 4.5 months); sample size at each time point ranged from 19 to 29. For the metabolomics analysis, we analyzed 137 fecal samples from 29 infants; the sample size at each time point ranged from 19 to 26. Only one baseline sample was available for the remaining infant, which we thus eliminated from this analysis.

2.2.1. Shotgun Metagenomics

The gut microbiome was characterized by shotgun metagenomic sequencing using established methods [34]. Genomic DNA was extracted using the MO BIO PowerSoil kit and prepared for sequencing with the Nextera XT kit (Illumina, Inc., San Diego, CA, USA). Libraries were sequenced on Illumina HiSeq 2500 using the paired-end 125 bp sequencing protocol. Sequence reads were processed to remove adapter sequences and low-quality reads with Trimmomatic v. 0.33 [35]. Taxonomic assignments were generated with MetaPhlAn2 [36]. Gene orthologs were assigned by aligning reads to the KEGG database [37] using DIAMOND v. 2.0.4 [38]. Negative (mock purification of DNA-free water) and positive (standard fecal and pond-sediment samples) controls were included in each assay. Samples of each formula (CMF, EHF) were assayed for DNA content, which yielded no detectable bacterial DNA. Shotgun libraries were generated from 1 ng of DNA using the NexteraXT kit (Illumina, Inc.). Libraries were sequenced on the Illumina HiSeq using 2 × 125 bp chemistry in high-output mode.

2.2.2. Targeted Metabolomics

Targeted metabolomics of 17 fecal amino acids were characterized using established methods [39] with a limit of detection of 5 nmol/g. Briefly, methanol-extracted fecal samples (5 µL/mg stool) were prepared using the Waters AccQ-Tag Ultra Amino Acid Derivatization and Chemistry Kit (Waters Corporation, Milford, MA, USA). Samples were analyzed on Waters Acquity uPLC System with an AccQ-Tag Ultra C18 1.7 µm 2.1 × 10 m column and a photodiode-array detector (Waters Corporation, Milford, MA, USA).

*2.3. Statistical Analyses*

The abundance of genes was normalized across samples using quantile normalization (up to 95% percentile). Genes absent from more than 50% of samples were considered low-occurrence genes and were removed from the analyses. Linear mixed-effects modeling

was used to determine the effect of treatment group (CMF, EHF), time (0.5, 0.75, 1.5, 2.5, 3.5, 4.5 months), and group × time interaction on (a) Shannon diversity indices, (b) relative abundance of each bacteria at the species levels, (c) the quantile normalized abundance of the top 1000 most prevalent genes, and (d) concentration of each fecal amino acid separately. The species for which more than 90% of the samples were undetected were excluded from the linear mixed effects modeling. The relative abundance of species, and the quantile normalized abundance of genes, were log10 transformed in the linear mixed effects model. A permutational multivariate analysis of variance (PERMANOVA) test on Bray–Curtis distances was used to examine differences in beta diversity between formula groups at each time point separately. A Bayesian logistic regression model was applied to determine whether specific species (using log10 transformed relative abundance in the model) were associated with formula group at each time point separately. Receiver operator characteristic (ROC) curves [40] were plotted and area under the curve values were calculated to evaluate the model.

To probe the relationships between continuous outcomes related to gut microbiota or metabolome and formula-induced differences in clinical phenotypes reported previously [26,41], simple and partial Pearson correlation analyses were conducted. Statistical analyses of data on gut microbiota, fecal amino acids, and bacterial genes were implemented in R (version 4.1.1) and higher [42] or Graphpad Prism (version 9.2.0; Graphpad Software Inc., San Diego, CA, USA); associations with clinical outcomes were conducted using Statistica version 14.0 (StatSoft Inc., Tulsa, OK, USA). The Benjamini and Hochberg false discovery rate (FDR) [43]; $p < 0.05$ was considered significant.

## 3. Results

### 3.1. Outgrowth of Ruminococcus Gnavus and Other Clostridia Species Driven by Formula-Induced Differences in Gut Microbiota

The most abundant classes of bacterial taxa found in the microbiota of the 30 infants in this study, aged 0.5 to 4.5 months, were Verrucomicrobiae, Bacilli, Actinobacteria, Bacteroidia, Clostridia, Negativiutes, and Gammaproteobacteria. The EHF group had greater Shannon diversity in the gut microbiota over time than the CMF group (group × time, $p = 0.004$; linear mixed effects model). The most substantial group difference was the increased prevalence of obligately anaerobic communities of Clostridia in the EHF group (Figure 1A,B). Figure 1C,D depict the relative abundance of the top 15 Clostridia species in CMF and EHF groups, respectively.

To determine whether the significant group × time interaction ($p = 0.004$) in the Shannon diversity index of the complete microbial community (Figure 2A) was due primarily to differences in Clostridia, we compared Shannon indexes of the complete community without Clostridia and to that of Clostridia only. As shown in Figure 2B, removal of Clostridia species from the Shannon diversity index resulted in no significant group, time, or group × time effect ($p > 0.05$). However, there was a significant group × time interaction in the Shannon diversity index of Clostridia only ($p = 0.027$; Figure 2C), thus providing further evidence that the greater Shannon diversity among the EHF group was driven primarily by Clostridia.

While there were no group differences before randomization to the study formulas before 1.5 months (PERMANOVA tests on Bray–Curtis distances, $p > 0.05$; Figure A1), there were significant group differences in the beta diversity index at 2.5, 3.5, and 4.5 months ($p = 0.027$, 0.024, and 0.031, respectively). By 4.5 months, the outgrowth of the most abundant taxa of Clostridia—*Ruminoccocus gnavus*—accounted for approximately one-fifth (mean relative abundance, 19.1%) of the community in the EHF group (Figure 1D), but only 6.6% in the CMF group (Figure 1C). To evaluate the predictive value of *R. gnavus* relative abundance by group, we applied logistic regression models to each time point separately (the relative abundances were log10 transformed) and constructed ROC curves (Figure 3). The area under the curve was greater than 0.8 at both 3.5 and 4.5 months, demonstrating that relative abundance of *R. gnavus* alone could distinguish the microbiome of infants

fed EHF from those fed CMF after the infants had been fed the formula for approximately 3 months.
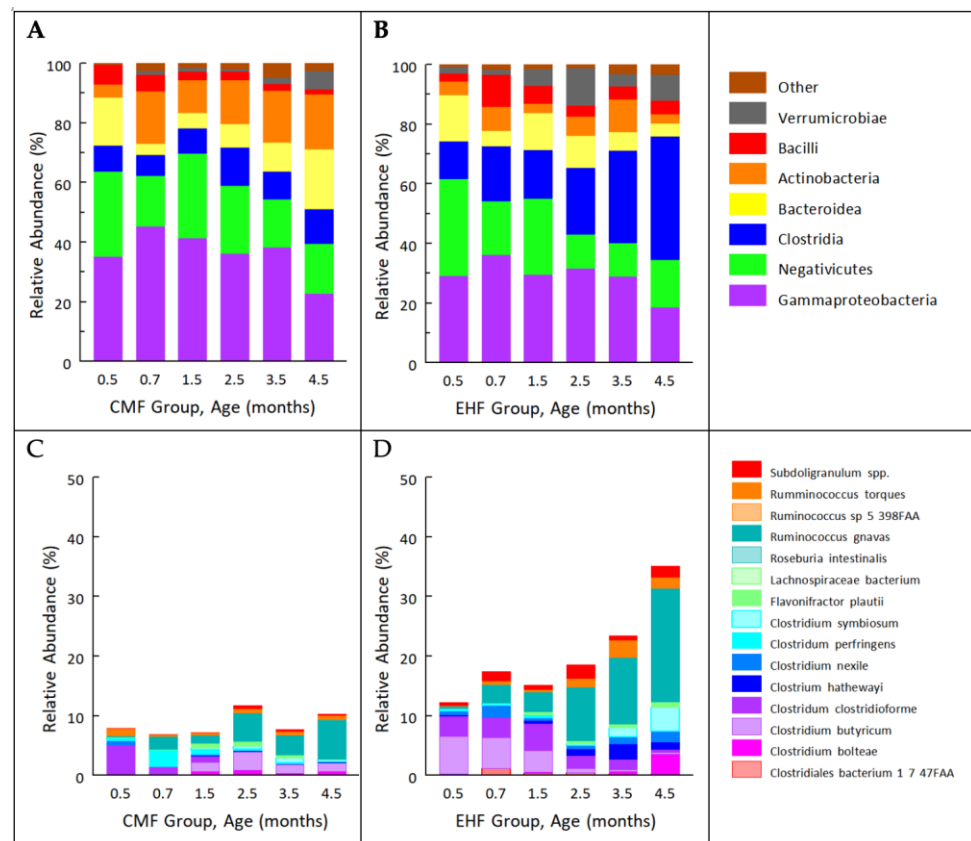


**Figure 1.** Relative abundance of microbial classes in (**A**) CMF (cow milk formula) and (**B**) EHF (extensive protein hydrolysate formula) groups and the top 15 species of Clostridia in (**C**) CMF and (**D**) EHF groups from 0.5 to 4.5 months.
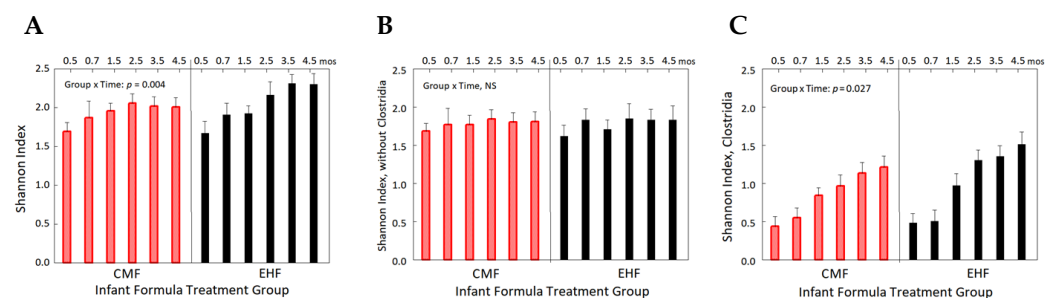


**Figure 2.** Shannon diversity indexes of (**A**) complete microbial community, (**B**) microbial community without Clostridia, and (**C**) Clostridia only. Red bars depict the CMF (cow milk formula) group and black bars depict the EHF (extensive protein hydrolysate formula) group from 0.5 to 4.5 months (linear mixed-effect modeling).

Although *R. gnavus* was the most prominent signature distinguishing the microbiome of EHF- from CMF-fed infants, other microbial taxa differed between the groups. Using a linear mixed-effect model ($p$'s < 0.05), we identified seven taxa that had significant group × time interactions (Figure 4). The taxa *Streptococcus thermophilus* had significant group × time interaction ($p < 0.001$), group ($p < 0.001$) and time ($p = 0.005$) effects with relative abundance remaining high in CMF but decreasing in the EHF group over time. Conversely, the remaining six taxa, each within the Clostridia class of Firmicutes, had significant group × time interaction (all $p$'s < 0.04) and time (all $p$'s < 0.001) effects with relative abundance being higher in EHF than CMF group.
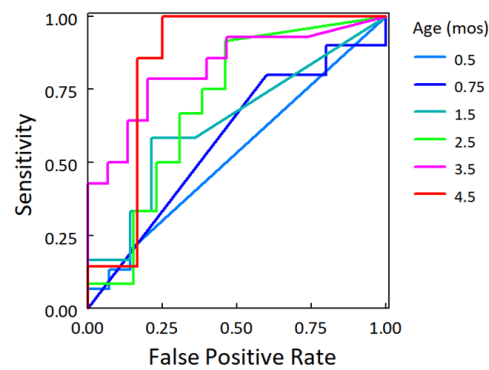
**Figure 3.** Receiver operating characteristic curves revealed discriminatory power of *Ruminococcous gnavus* to distinguish CMF group from EHF group by 3.5–4.5 months; logistic regression model.
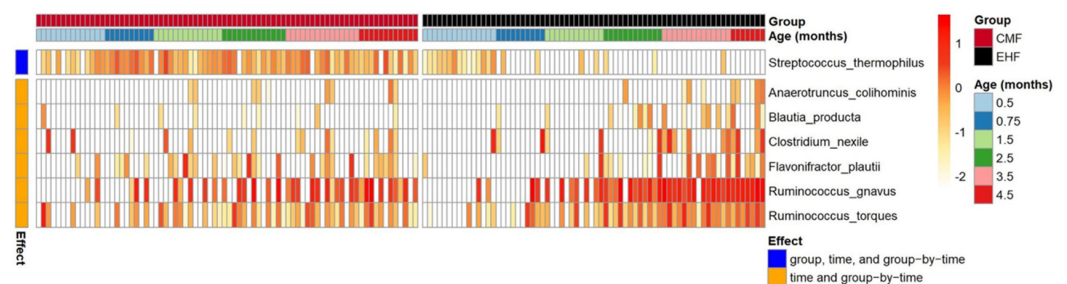


**Figure 4.** Heatmap of bacterial species from 0.5 to 4.5 months in CMF (red bar) and EHF (black bar) groups. Red signifies higher abundance. Linear mixed effects models revealed significant group × time, group and time effects (blue) for *Steptococcus thermophiles* and significant group × time and time effects (orange) for *Ruminococcus gnavus*, *Blautia producta*, *Ruminococcus torques*, *Clostridium nexile*, *Flavonifractor plautii*, and *Anaerotruncus colihominis*.

Table A1 lists the 47 bacterial taxa that were not affected by the randomization but significantly changed from 0.5 to 4.5 months (time effect only, $p < 0.05$). In general, some species in the Veillonella, Lactobacillus, Streptococcus genera are in higher relative abundance during the first month and then decreased over time, whereas some Bifidobacteria and Clostridium species increased over time [44,45].

### 3.2. Formula-Induced Differences in Expression of Genes Related to Carbohydrate Metabolism

Of the top 1000 most abundant genes, we found significant group × time interactions for six genes (all FDR, $p$'s $< 0.05$): the type IV secretion system protein VirD4, carbonic anhydrase, tRNA nucleotidyltransferase, and three genes associated with carbohydrate metabolism: beta-fructofuranosidase, glucose-1-phosphate adenylyltransferase and lactase permease (Figure 5A–F).
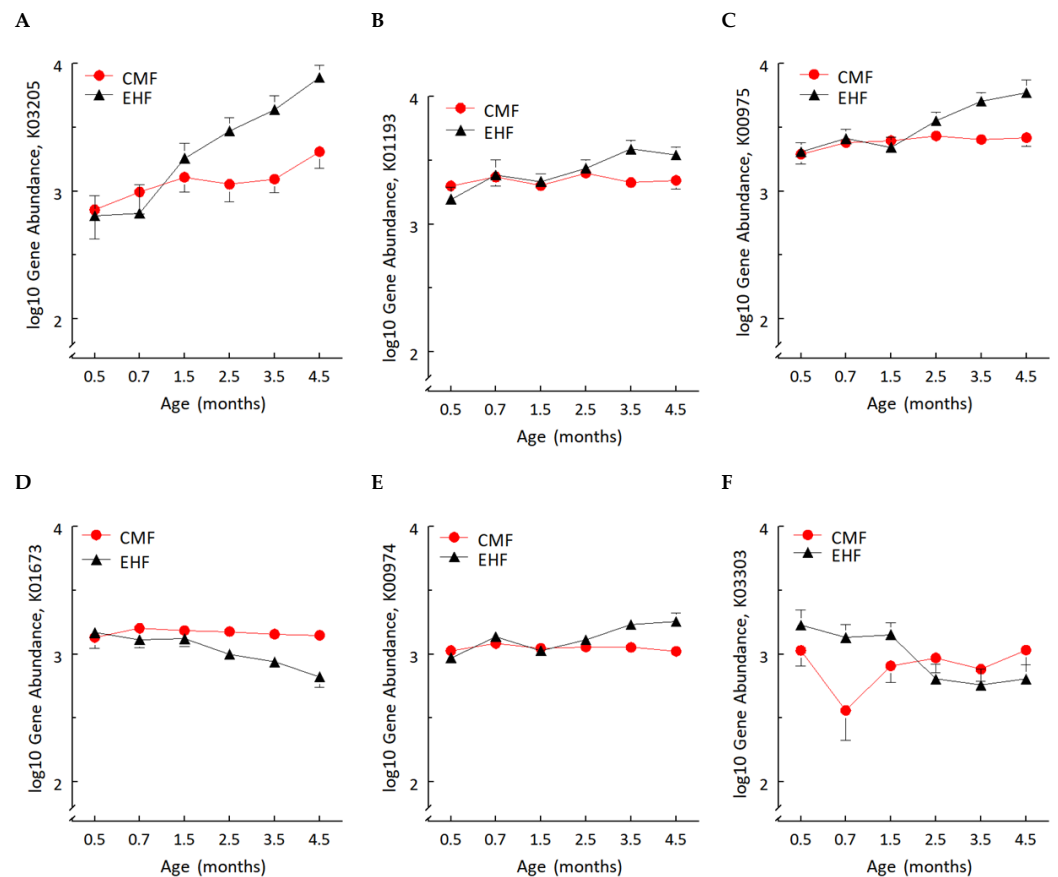
**Figure 5.** Log$_{10}$ gene abundance over time in CMF (red circles) and EHF (black triangles) groups for gene K03205, type IV secretion system protein VirD4 (**A**); K01193, beta-fructofuranosidase (**B**); K00975, glucose 1-phosphate adenylyltransferase (**C**); K01673, carbonic anhydrase (**D**); K00974, tRNA nucleotidyltransferase (**E**); and K03303, lactase permease (**F**). Significant group × time effect for each gene; linear mixed models.

### 3.3. The Two Groups Shared Similar Fecal Amino Acid Concentrations despite Different Free Amino Acid Concentrations in CMF and EHF

The concentrations of fecal amino acids over time by formula group are shown in Figure 6. Fecal concentrations of histidine and cysteine were below the limit of detection in 78% and 96% of the samples, respectively, so they were not included in the analyses. There were no significant group or group × time effect identified for any of the remaining amino acids ($p$'s > 0.05).
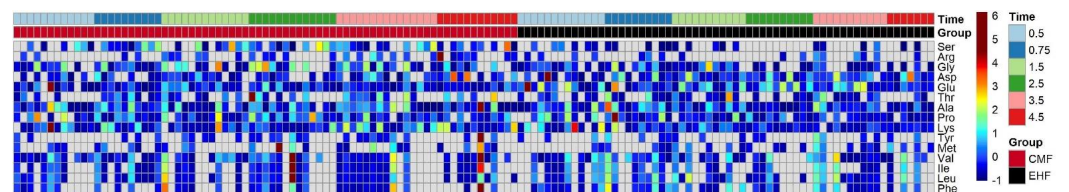


**Figure 6.** Heatmap of fecal amino acid concentrations (standardized by mean and standard deviation per amino acid) from 0.5 to 4.5 months in CMF (cow milk formula; red bar) and EHF (extensive protein hydrolysate formula; black bar) groups. Red signifies higher abundance. No significant group or group × time effect for any of the amino acids identified; linear mixed-effects model.

### 3.4. Increases in the Relative Abundance of Clostridia Related to Leaner Phenotypes

While no measured infant phenotypes differed at baseline (Table 1), the two randomized groups significantly differed in subsequent velocities of weight gain from 0.5 to 4.5 months, in fat mass and percent body fat at 3.5 months, and in WLZ scores at 4.5 months,

but did not differ in velocities of linear growth (Table 2). The two groups also significantly differed in relative abundance of Clostridia at both 3.5 and 4.5 months (Figure 1B); we refer to these data, when combined, as occurring at 3–4 months.

**Table 2.** Growth and body composition in study cohort at 3.5–4.5 months.

| | Infant Formula Treatment Group | | |
|---|---|---|---|
| **Characteristics** | **CMF** (*n* = 15) | **EHF** (*n* = 15) | *p* **Value** |
| Anthropometry, Z scores; 4.5 months | | | |
|     Weight for age (WAZ) | 0.30 ± 0.23 | −0.55 ± 0.23 | 0.02 |
|     Length for age (LAZ) | −0.47 ± 0.27 | −0.66 ± 0.27 | 0.61 |
|     Weight for length (WLZ) | 0.88 ± 0.20 | −0.07 ± 0.20 | <0.001 |
| Weight-gain velocity (g/day), 0.5–4.5 months | 31.77 ± 1.49 | 25.49 ± 1.49 | <0.001 |
| Length-gain velocity (cm/day), 0.5–4.5 months | 0.104 ± 0.003 | 0.094 ± 0.003 | 0.18 |
| Body composition, 3.5 months [1] | | | |
|     Fat mass (kg) | 1.61 ± 0.08 | 1.15 ± 0.12 | 0.01 |
|     Percent body fat (%) | 24.3 ± 1.5 | 19.0 ± 1.5 | 0.02 |

Mean ± standard error of mean (SEM). *n* = 30 unless otherwise indicated; *p* values for main effect of infant formula treatment group (CMF, cow milk formula; EHF, extensive protein hydrolysate formula). [1] CMF, *n* = 14; EHF, *n* = 14.

As shown in Figure 7, the three types of outcomes that differed between the randomized groups—gut Clostridia, velocities of growth and anthropometric phenotypes—were related to each other. The relative abundance of Clostridia at 3–4 months was significantly and inversely correlated with prior weight-gain velocities (Figure 7A; r = −0.51, *p* = 0.004) and length gain velocities (Figure 7B; r = −0.53, *p* = 0.003), WLZ at 4.5 months (r = −0.41; *p* = 0.02; Figure 7C), and fat mass at 3.5 months (r = −0.41; *p* = 0.03; Figure 7D).
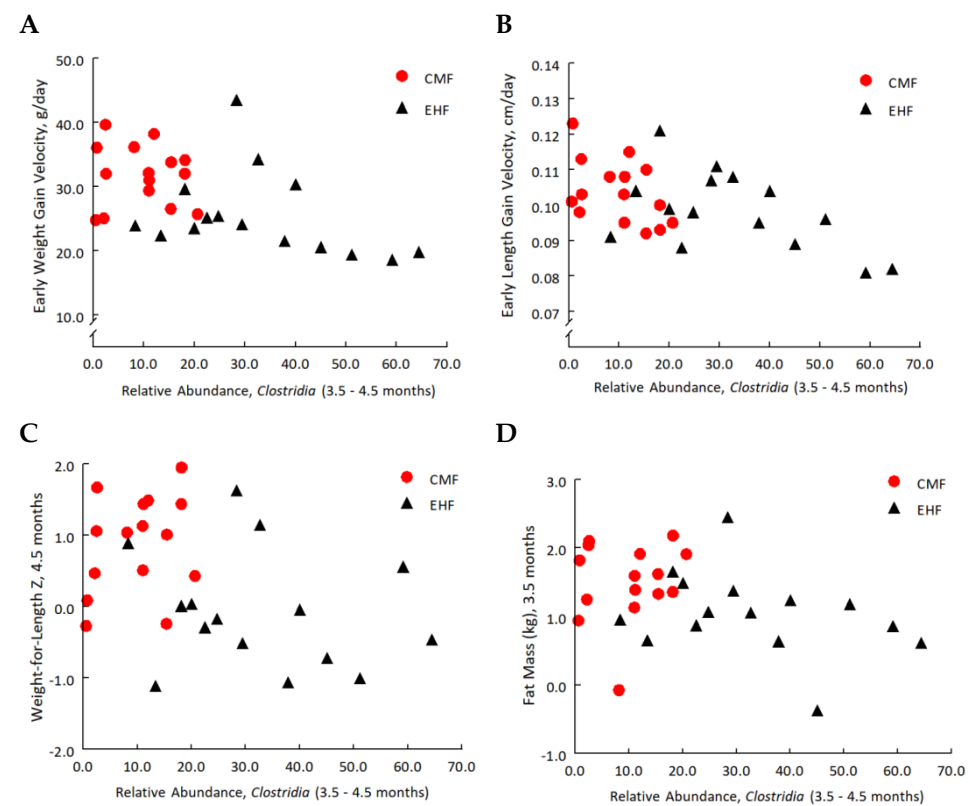


**Figure 7.** Pearson correlations between relative abundance of *Clostridia* (3–4 months) and (**A**) early weight gain velocity (g/day; *p* = 0.004), (**B**) early length gain velocity (cm/day; *p* = 0.003), (**C**) WLZ at 4.5 months (*p* = 0.02), and (**D**) fat mass (kg) at 3.5 months (*p* = 0.03). Red circles denote CMF (cow milk formula) and black triangles denote EHF (extensive protein hydrolysate) groups.

As expected [41,46], greater velocity of weight gain predicted weight status (r = 0.59; *p* < 0.001) and fat mass (r = 0.45; *p* = 0.02) at 1 year. Children who were with overweight at 1 year tended to have lower relative abundance of Clostridia at 3–4 months (15.7 ± 5.3) compared to 1-year-olds who were a healthy weight (27.8 ± 4.2; *p* = 0.08).

## 4. Discussion

While the human gut gradually changes from a predominantly aerobic to anaerobic environment after birth [47,48], in this proof-of-concept study the composition of the gut microbiome underwent striking time-dependent separation based on the type of formula infants were fed, at the age when formula was their sole source of nutrition. Exclusively feeding of EHF, a formula rich in free amino acids, was associated with a more rapid maturation of the infant gut microbiota, with increased alpha diversity dominated by classes of bacteria that degrade carbohydrates (e.g., Lachnospiraceae, Negativicutes, Clostridia). In general, some Veillonella, Lactobacillus, and Streptococcus species within these classes, which are in high relative abundance in the first month [49] and are more commonly associated with the proximal gastrointestinal tract [44], decreased over time, whereas some Bifidobacteria and Clostridium species, which are normally found in a more mature distal gastrointestinal tract community, increased over time [44,45]. However, there were two notable exceptions.

While *Streptococcus thermophilus* steadily decreased over time in the EHF group, its relative abundance remained high in the CMF group. Conversely, a number of Clostridia species decreased over time in the CMF group, as expected, but significantly increased over time in the EHF group. Indeed, of the seven taxa that exhibited group-dependent or group × time dependent changes, Streptococcus accounted for one taxa while Clostridia accounted for six, including taxa that are associated with the induction of immune tolerance in murine model systems [50,51], that utilize proteins and free amino acids as primary nutrient sources [52,53], or that produce a large array of metabolites by utilizing simple and complex carbohydrates [54]. The carbohydrate-fermentative species *R. gnavus* [55] exhibited the greatest increase in the EHF group over time and its outgrowth distinguished the two infant formula groups approximately 3 months after infants began feeding their assigned formulas.

These two key compositional differences between CMF and EHF formed the basis of our hypothesis on how feeding CMF or EHF shaped the microbiota in different ways (i.e., increased abundance of Streptococcus among CMF infants and increased abundance of Clostridia among EHF infants). These differences may be acting alone or in concert in causing the observed effects on the infants' gut microbiota. First, the formulas differ in the form of the protein. The free amino acid concentration is substantially higher in EHF (>9200%) than CMF [18]. Because EHF transits the gastrointestinal tract at a faster rate [56], there may be increased concentrations of amino acids in the colon, where they act as substrates for colonic bacterial growth [57]. Certain Clostridia species may have bioenergetic growth advantage when infants feed EHF because of Strickland fermentation—the coupled oxidation and reduction of amino acids to organic acids—a well-known reaction in Clostridia [58]. However, despite the drastic differences in free amino acids between the infant formulas, the overall concentrations of amino acids in their feces were similar between the two groups, suggesting that, in addition to Strickland fermentation by the gut microbiota, they also likely are being absorbed in the small intestine of the EHF-fed infants.

Second, differences in carbohydrate source between the infant formulas and, in turn, the availability of carbon sources for metabolism may contribute to group differences in bacterial taxa. Unlike CMF, in EHF one of the carbohydrate sources is modified cornstarch, composed of glucose monomers connected by alpha-1-4 glycosidic linkages and alpha 1–6 glycosidic branch linkages, the latter of which leads to slower enzymatic degradation in the gastrointestinal tract [59]. Not only is the digestion of cornstarch slow in the small intestine, but its fermentation in the colon continues several hours after its ingestion [20]. Indeed, we observed a very robust induction of two Ruminococcus species, *R. gnavus* and

*R. torques*, known to be mucinophilic [60–62]. Ruminococcus, a genus of bacteria in the Clostridia class, require fermentable carbohydrates for growth [20], and are stimulated by diets high in resistant starches, including cornstarch which contains some resistant starch [20,63]. In addition, *Streptococcus thermophilus*, which was the only taxa that was higher in the microbiota of the CMF than EHF group, has been shown to have increased growth in lactose-containing media compared to media containing the monomers of lactose, namely glucose and galactose [64,65]. In this regard, it is interesting that three of the six bacterial genes showing a significant group by time effect are involved with carbohydrate metabolism [66].

Building on the findings from the entire intent-to-treat cohort of 113 infants [26,41], we found in this subsample of 30 infants that the CMF group had faster weight-gain velocity during the first 4 months, greater fat mass and higher WLZ scores than the EHF group. The present study also revealed that the EHF group had greater abundance of Clostridia class bacteria in the gut than the CMF group, a difference that emerged after feeding the assigned formulas for 3 months. Both of the formula-induced outcomes—relative abundance of Clostridia and clinical phenotypes related to weight —were associated with each other. Greater relative abundance of Clostridia was associated with more normative trajectories of weight and length gain during preceding months and with leaner phenotypes. However, after adjustment for formula treatment group, only the partial correlation with gains in length remained significant. Nevertheless, how fast infants gained weight during the sensitive period of the first 4 months predicted their weight status at one year [67], and overweight 1-year-olds tended to have lower abundance of Clostridia in their guts than those at a healthy weight. Whether or not there is a cause-and-effect relationship between the leaner phenotypes among EHF-fed infants and the greater abundance of will require additional investigation.

Although the present proof-of-principle study focused on only 30 of the 113 infants enrolled in the RCT and only on fecal samples collected during the first 4.5 months of the year-long trial [26], associations between the abundance of gastrointestinal Clostridia and phenotypes related to obesity have been previously reported in children [68] and adults [69]. Taken together, the data suggest that the formula-induced differences in early weight gain and later weight status [46,70] may be mediated by early priming of the gut microbiota, particularly as it relates to the Clostridia class of bacteria. The impact of Clostridia species on lipid absorption [71] and mucosal immune regulation [50,51] could play a role in the more normative early rapid weight gain among EHF-fed infants [26], which is, in turn, associated with normative weigh status at one year [46]. Additionally, since Clostridia spp. have been shown to enhance immune tolerance in murine models [50], our finding that the consumption of EHF led to increased abundance of certain Clostridia species might be relevant to the observed association between the ingestion of EHF during infancy and the decreased risk for developing atopic disease during childhood and adolescence [72]. Further, a potential energy mechanism for the association between the loss of Clostridia abundance and diversity and obesity and metabolic syndrome comes from T-Myd88-knockout mice [71], increases in Clostridia colonization were associated with a leaner phenotype and down-regulation of CD36, an enterocyte receptor involved in lipid absorption, suggesting decreased energy absorption from the gut. In contrast, lack of colonization with the protective Clostridia in germ-free mice resulted in elevated CD36 expression, weight gain, fatty liver disease, adipose tissue inflammation, and insulin resistance [71].

Advantages of this proof-of-principle study include the randomized design, which minimizes selection bias and enables determination of effects of early infant formula diet while keeping other variables constant, thus providing insights into mechanisms of long-term effects of feeding infant formulas that differ in macronutrient composition. While the small sample size and increased likelihood of type II error are limitations, the findings from the proof-of-principle approach hold promise that similarly designed RCTs could serve as

a model system for determining mechanisms underlying the divergent clinical outcomes of infants fed different formulas.

In conclusion, diet has strong selective pressure on shaping the composition and functioning of the microbiome [73,74]. What infants are fed is more important than the mode of feeding, since formula-fed infants are not a homogeneous group. Although isocaloric, infant formulas can have different macronutrient composition that have significantly different impacts, in the short term, on infant satiety [75], early weight gain, energy intake, and fecal energy loss [26] and, in the longer term, risks for obesity [21,76] and atopic disease [24]. The present proof-of-principle study revealed distinct compositional changes in the gut microbiota of infants fed formulas differing in the source of and form of protein and carbohydrate for only 3–4 months. The long-lasting consequences of the formula-induced changes to the early bacterial environment as the child transitions to a diet devoid of infant formula remain unknown. Whether the compositional changes in early gut bacterial environment have independent effects on clinical outcomes of growth trajectories, body weight, and fat mass, and whether such changes are biomarkers for later obesity and immune functioning, are important areas of research, for which data are needed from larger, longitudinal RCT cohorts [77] that experimentally manipulate early diet.

## Appendix A

**A**



**Figure A1.** Multidimensional scaling plot of beta diversity based on Bray–Curtis distance metric for CMF and EHF groups showing (**A**) all ages combined and (**B**–**G**) at each time point. There were no significant differences between CMF (cow milk formula, red symbols) and EHF (extensive protein hydrolysate formula, black symbols) groups at 0.5, 0.75, and 1.5 months ($p > 0.05$, Permutational Multivariate Analysis of Variance (PERMANOVA) tests on Bray–Curtis distances), but there were significant group differences at 2.5 ($p = 0.03$), 3.5 ($p = 0.02$), and 4.5 months ($p = 0.03$); PERMANOVA tests on Bray–Curtis distances).

**Table A1.** Bacterial taxa (N = 47) with significant time effects [1].

| Species Name | Time, β Coefficient | Time, *p*-Value |
|---|---|---|
| *Actinomyces odontolyticus* | 0.12 | $3.4 \times 10^{-5}$ |
| *Akkermansia muciniphila* | 0.28 | $3.1 \times 10^{-5}$ |
| *Anaerostipes, unclassified* | 0.25 | $6.0 \times 10^{-8}$ |
| *Bacteroides uniformis* | 0.07 | 0.02 |
| *Bifidobacterium bifidum* | 0.35 | $1.0 \times 10^{-5}$ |
| *Bifidobacterium breve* | 0.26 | $1.5 \times 10^{-4}$ |
| *Bifidobacterium longum* | 0.39 | $3.7 \times 10^{-5}$ |
| *Bifidobacterium pseudocatenulatum* | 0.11 | 0.02 |
| *Bilophila, unclassified* | 0.06 | 0.01 |
| *Clostridiales bacterium 1 7 47FAA* | 0.23 | $9.0 \times 10^{-6}$ |
| *Clostridium bartlettii* | 0.20 | 0.02 |
| *Clostridium bolteae* | 0.22 | $3.7 \times 10^{-5}$ |
| *Clostridium difficile* | 0.28 | $8.3 \times 10^{-8}$ |
| *Clostridium hathewayi* | 0.11 | 0.03 |
| *Clostridium perfringens* | −0.26 | $4.5 \times 10^{-5}$ |
| *Clostridium ramosum* | 0.36 | $1.4 \times 10^{-9}$ |
| *Clostridium symbiosum* | 0.13 | $3.7 \times 10^{-5}$ |
| *Collinsella aerofaciens* | 0.22 | $4.4 \times 10^{-4}$ |
| *Coprobacillus, unclassified* | 0.43 | $5.4 \times 10^{-13}$ |
| *Eggerthella, unclassified* | 0.12 | $7.8 \times 10^{-5}$ |
| *Enterococcus avium* | 0.11 | $1.6 \times 10^{-5}$ |
| *Enterococcus faecalis* | −0.12 | 0.02 |
| *Erysipelotrichaceae bacterium 21 3* | 0.14 | $3.6 \times 10^{-5}$ |
| *Erysipelotrichaceae bacterium 2 2 44A* | 0.12 | $3.0 \times 10^{-3}$ |
| *Erysipelotrichaceae bacterium 6 1 45* | 0.09 | 0.02 |
| *Eubacterium limosum* | 0.06 | 0.01 |
| *Granulicatella, unclassified* | 0.08 | $1.2 \times 10^{-3}$ |
| *Klebsiella pneumoniae* | 0.09 | 0.01 |
| *Lachnospiraceae bacterium 2 1 58FAA* | 0.23 | $6.0 \times 10^{-8}$ |
| *Lachnospiraceae bacterium 7 1 58FAA* | 0.06 | 0.01 |
| *Lachnospiraceae bacterium 9 1 43BFAA* | 0.10 | $1.3 \times 10^{-5}$ |
| *Lactobacillus fermentum* | −0.13 | $13.5 \times 10^{-4}$ |
| *Lactobacillus gasseri* | −0.24 | $3.3 \times 10^{-5}$ |
| *Lactococcus lactis* | 0.19 | $3.8 \times 10^{-3}$ |
| *Megasphaera, unclassified* | 0.19 | 0.04 |
| *Parabacteroides distasonis* | 0.58 | 0.03 |
| *Solobacterium moorei* | 0.06 | $7.1 \times 10^{-5}$ |
| *Staphylococcus hominis* | −0.11 | $1.2 \times 10^{-4}$ |
| *Streptococcus mitis oralis pneumoniae* | 0.08 | 0.01 |
| *Streptococcus peroris* | 0.09 | 0.01 |
| *Streptococcus salivarius* | −0.33 | $8.3 \times 10^{-8}$ |
| *Streptococcus vestibularis* | −0.41 | $8.3 \times 10^{-8}$ |
| *Subdoligranulum, unclassified* | 0.16 | $9.9 \times 10^{-6}$ |
| *Veillonella atypica* | −0.18 | $1.2 \times 10^{-4}$ |
| *Veillonella parvula* | −0.13 | 0.02 |
| *Veillonella ratti* | 0.30 | $4.3 \times 10^{-4}$ |
| *Veillonella, unclassified* | −0.15 | 0.01 |

[1] Linear mixed models determined the effect of treatment group (CMF (cow milk formula), EHF (extensive protein hydrolysate formula)), time (0.5, 0.75, 1.5, 2.5, 3.5, 4.5 months), and group × time interaction on fecal bacteria taxa. Table lists the 47 taxa that had only significant time effects.

## References

1. Koleva, P.T.; Kim, J.-S.; Scott, J.; Kozyrskyj, A.L. Microbial programming of health and disease starts during fetal life. *Birth Defects Res. Part C Embryo Today Rev.* **2015**, *105*, 265–277. [CrossRef] [PubMed]
2. Bäckhed, F.; Roswall, J.; Peng, Y.; Feng, Q.; Jia, H.; Kovatcheva-Datchary, P.; Li, Y.; Xia, Y.; Xie, H.; Zhong, H.; et al. Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe* **2015**, *17*, 852. [CrossRef] [PubMed]
3. Tremaroli, V.; Bäckhed, F. Functional interactions between the gut microbiota and host metabolism. *Nature* **2012**, *489*, 242–249. [CrossRef] [PubMed]
4. Fanaro, S.; Chierici, R.; Guerrini, P.; Vigi, V. Intestinal microflora in early infancy: Composition and development. *Acta Paediatr.* **2007**, *92*, 48–55. [CrossRef] [PubMed]
5. He, X.; Sotelo-Orozco, J.; Rudolph, C.; Lönnerdal, B.; Slupsky, C.M. The role of protein and free amino acids on intake, metabolism, and gut microbiome: A comparison between breast-fed and formula-fed rhesus monkey infants. *Front. Pediatr.* **2020**, *7*, 563. [CrossRef]

6. Buddington, R.K.; Sangild, P.T. Companion Animals Symposium: Development of the mammalian gastrointestinal tract, the resident microbiota, and the role of diet in early life. *J. Anim. Sci.* **2011**, *89*, 1506–1519. [CrossRef] [PubMed]

7. Ding, T.; Schloss, P.D. Dynamics and associations of microbial community types across the human body. *Nat.* **2014**, *509*, 357–360. [CrossRef]

8. Ho, N.T.; Li, F.; Lee-Sarwar, K.A.; Tun, H.M.; Brown, B.; Pannaraj, P.S.; Bender, J.M.; Azad, M.B.; Thompson, A.L.; Weiss, S.T.; et al. Meta-analysis of effects of exclusive breastfeeding on infant gut microbiota across populations. *Nat. Commun.* **2018**, *9*, 4169. [CrossRef]

9. Ma, J.; Li, Z.; Zhang, W.; Zhang, C.; Zhang, Y.; Mei, H.; Zhuo, N.; Wang, H.; Wang, L.; Wu, D. Comparison of gut microbiota in exclusively breast-fed and formula-fed babies: A study of 91 term infants. *Sci. Rep.* **2020**, *10*, 15792. [CrossRef]

10. Bokulich, N.A.; Chung, J.; Battaglia, T.; Henderson, N.; Jay, M.; Li, H.; Lieber, A.D.; Wu, F.; Perez-Perez, G.I.; Chen, Y.; et al. Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **2016**, *8*, 343ra82. [CrossRef]

11. Guaraldi, F.; Salvatori, G. Effect of breast and formula feeding on gut microbiota shaping in newborns. *Front. Cell. Infect. Microbiol.* **2012**, *2*, 94. [CrossRef] [PubMed]

12. Mackie, R.I.; Sghir, A.; Gaskins, H.R. Developmental microbial ecology of the neonatal gastrointestinal tract. *Am. J. Clin. Nutr.* **1999**, *69*, 1035s–1045s. [CrossRef] [PubMed]

13. Neis, E.P.J.G.; Dejong, C.H.C.; Rensen, S.S. The role of microbial amino acid metabolism in host metabolism. *Nutrients* **2015**, *7*, 2930–2946. [CrossRef]

14. Davila, A.-M.; Blachier, F.; Gotteland, M.; Andriamihaja, M.; Benetti, P.-H.; Sanz, Y.; Tomé, D. Re-print of "Intestinal luminal nitrogen metabolism: Role of the gut microbiota and consequences for the host". *Pharmacol. Res.* **2013**, *69*, 114–126. [CrossRef] [PubMed]

15. Allison, C.; Macfarlane, G.T. Influence of pH, nutrient availability, and growth rate on amine production by Bacteroides fragilis and Clostridium perfringens. *Appl. Environ. Microbiol.* **1989**, *55*, 2894–2898. [CrossRef] [PubMed]

16. Gallier, S.; Van den Abbeele, P.; Prosser, C. Comparison of the bifidogenic effects of goat and cow milk-based infant formulas to human breast milk in an in vitro gut model for 3-month-old infants. *Front. Nutr.* **2020**, *7*, 608495. [CrossRef]

17. Rossen, L.M.; Simon, A.E.; Herrick, K.A. Types of infant formulas consumed in the United States. *Clin. Pediatr.* **2016**, *55*, 278–285. [CrossRef]

18. Ventura, A.K.; San Gabriel, A.; Hirota, M.; Mennella, J.A. Free amino acid content in infant formulas. *Nutr. Food Sci.* **2012**, *42*, 271–278. [CrossRef]

19. Romaní-Pérez, M.; Bullich-Vilarrubias, C.; López-Almela, I.; Liébana-García, R.; Olivares, M.; Sanz, Y. The microbiota and the gut-brain axis in controlling food intake and energy homeostasis. *Int. J. Mol. Sci.* **2021**, *22*, 5830. [CrossRef]

20. Achour, L.; Flourié, B.; Briet, F.; Franchisseur, C.; Bornet, F.; Champ, M.; Rambaud, J.C.; Messing, B. Metabolic effects of digestible and partially indigestible cornstarch: A study in the absorptive and postabsorptive periods in healthy humans. *Am. J. Clin. Nutr.* **1997**, *66*, 1151–1159. [CrossRef]

21. Rzehak, P.; Sausenthaler, S.; Koletzko, S.; Reinhardt, D.; Von Berg, A.; Krämer, U.; Berdel, D.; Bollrath, C.; Grübl, A.; Bauer, C.P.; et al. Short- and long-term effects of feeding hydrolyzed protein infant formulas on growth at < or = 6 y of age: Results from the German Infant Nutritional Intervention Study. *Am. J. Clin. Nutr.* **2009**, *89*, 1846–1856. [CrossRef] [PubMed]

22. Mennella, J.A.; Ventura, A.K.; Beauchamp, G.K. Differential growth patterns among healthy infants fed protein hydrolysate or cow-milk formulas. *Pediatrics* **2011**, *127*, 110–118. [CrossRef] [PubMed]

23. Giovannini, M.; Agostoni, C.; Fiocchi, A.; Bellú, R.; Trojan, S.; Riva, E. Antigen-reduced infant formulas versus human milk: Growth and metabolic parameters in the first 6 months of life. *J. Am. Coll. Nutr.* **1994**, *13*, 357–363. [CrossRef] [PubMed]

24. Gappa, M.; Filipiak-Pittroff, B.; Libuda, L.; von Berg, A.; Koletzko, S.; Bauer, C.; Heinrich, J.; Schikowski, T.; Berdel, D.; Standl, M. Long-term effects of hydrolyzed formulae on atopic diseases in the GINI study. *Allergy* **2021**, *76*, 1903–1907. [CrossRef]

25. Laursen, M.F.; Andersen, L.B.B.; Michaelsen, K.F.; Mølgaard, C.; Trolle, E.; Bahl, M.I.; Licht, T.R. Infant gut microbiota development is driven by transition to family foods independent of maternal obesity. *mSphere* **2016**, *1*, e00069-15. [CrossRef]

26. Mennella, J.A.; Inamdar, L.; Pressman, N.; Schall, J.I.; Papas, M.A.; Schoeller, D.; Stallings, V.A.; Trabulsi, J.C. Type of infant formula increases early weight gain and impacts energy balance: A randomized controlled trial. *Am. J. Clin. Nutr.* **2018**, *108*, 1015–1025. [CrossRef]

27. Mennella, J.; Trabulsi, J.; Inamdar, L.; Preedy, V.R.; Watson, R.R.; Zibadi, S. The sensory world of formula-fed infants: Differences among artificial milk feedings in flavor learning and satiation. In *Handbook of Dietary and Nutritional Aspects of Bottle Feeding*; Preedy, V.R., Watson, R.R., Zibadi, S., Eds.; Wageningen Academic Publishers: Wageningen, The Netherlands, 2014; pp. 95–116.

28. Terheggen-Lagro, S.W.; Khouw, I.M.; Schaafsma, A.; Wauters, E.A. Safety of a new extensively hydrolysed formula in children with cow's milk protein allergy: A double blind crossover study. *BMC Pediatr.* **2002**, *2*, 10. [CrossRef]

29. WHO Multicentre Growth Reference Study Group. *World Health Organization (WHO) Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development*; WHO Multicentre Growth Reference Study Group: Geneva, Switzerland, 2006.

30. Roy, S.M.; Spivack, J.G.; Faith, M.S.; Chesi, A.; Mitchell, J.A.; Kelly, A.; Grant, S.F.A.; McCormack, S.E.; Zemel, B.S. Infant BMI or weight-for-length and obesity risk in early childhood. *Pediatrics* **2016**, *137*, e20153492. [CrossRef]

31. Schoeller, D.A.; Hnilicka, J.M. Reliability of the doubly labeled water method for the measurement of total daily energy expenditure in free-living subjects. *J. Nutr.* **1996**, *126*, 348S–354S.

32. Fomon, S.J.; Haschke, F.; Ziegler, E.E.; Nelson, S.E. Body composition of reference children from birth to age 10 years. *Am. J. Clin. Nutr.* **1982**, *35*, 1169–1175. [CrossRef]

33. Racette, S.B.; Schoeller, D.A.; Luke, A.H.; Shay, K.; Hnilicka, J.; Kushner, R.F. Relative dilution spaces of 2H- and 18O-labeled water in humans. *Am. J. Physiol. Metab.* **1994**, *267*, E585–E590. [CrossRef] [PubMed]

34. Lewis, J.D.; Chen, E.Z.; Baldassano, R.N.; Otley, A.R.; Griffiths, A.M.; Lee, D.; Bittinger, K.; Bailey, A.; Friedman, E.S.; Hoffmann, C.; et al. Inflammation, antibiotics, and diet as environmental stressors of the gut microbiome in pediatric Crohn's disease. *Cell Host Microbe* **2015**, *18*, 489–500. [CrossRef] [PubMed]

35. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]

36. Truong, D.T.; Franzosa, E.A.; Tickle, T.L.; Scholz, M.; Weingart, G.; Pasolli, E.; Tett, A.; Huttenhower, C.; Segata, N. Meta, Phl, An2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **2015**, *12*, 902–903. [CrossRef] [PubMed]

37. Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef] [PubMed]

38. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59–60. [CrossRef]

39. Ni, J.; Shen, T.-C.D.; Chen, E.Z.; Bittinger, K.; Bailey, A.; Roggiani, M.; Sirota-Madi, A.; Friedman, E.S.; Chau, L.; Lin, A.; et al. A role for bacterial urease in gut dysbiosis and Crohn's disease. *Sci. Transl. Med.* **2017**, *9*, eaah6888. [CrossRef]

40. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **1982**, *143*, 29–36. [CrossRef]

41. Mennella, J.A.; Reiter, A.; Brewer, B.; Pohlig, R.T.; Stallings, V.A.; Trabulsi, J.C. Early weight gain forecasts accelerated eruption of deciduous teeth and later overweight status during the first year. *J. Pediatr.* **2020**, *225*, 174–181. [CrossRef]

42. R Core Team. R: A Language and Environment for Statistical Computing. Available online: http://www.R-project.org (accessed on 22 September 2021).

43. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [CrossRef]

44. Kastl, A.J., Jr.; Terry, N.A.; Wu, G.D.; Albenberg, L.G. The structure and function of the human small intestinal microbiota: Current understanding and future directions. *Cell. Mol. Gastroenterol. Hepatol.* **2020**, *9*, 33–45. [CrossRef] [PubMed]

45. Roswall, J.; Olsson, L.M.; Kovatcheva-Datchary, P.; Nilsson, S.; Tremaroli, V.; Simon, M.-C.; Kiilerich, P.; Akrami, R.; Krämer, M.; Uhlén, M.; et al. Developmental trajectory of the healthy human gut microbiota during the first 5 years of life. *Cell Host Microbe* **2021**, *29*, 765–776.e3. [CrossRef] [PubMed]

46. Trabulsi, J.C.; Smethers, A.D.; Eosso, J.R.; Papas, M.A.; Stallings, V.A.; Mennella, J.A. Impact of early rapid weight gain on odds for overweight at one year differs between breastfed and formula-fed infants. *Pediatr. Obes.* **2020**, *15*, 12688. [CrossRef] [PubMed]

47. Johnson, C.L.; Versalovic, J. The human microbiome and its potential importance to pediatrics. *Pediatrics* **2012**, *129*, 950–960. [CrossRef] [PubMed]

48. Houghteling, P.; Walker, W.A. Why is initial bacterial colonization of the intestine important to infants' and children's health? *J. Pediatr. Gastroenterol. Nutr.* **2015**, *60*, 294–307. [CrossRef]

49. Bittinger, K.; Zhao, C.; Li, Y.; Ford, E.; Friedman, E.S.; Ni, J.; Kulkarni, C.V.; Cai, J.; Tian, Y.; Liu, Q.; et al. Bacterial colonization reprograms the neonatal gut metabolome. *Nat. Microbiol.* **2020**, *5*, 838–847. [CrossRef]

50. Atarashi, K.; Tanoue, T.; Shima, T.; Imaoka, A.; Kuwahara, T.; Momose, Y.; Cheng, G.; Yamasaki, S.; Saito, T.; Ohba, Y.; et al. Induction of colonic regulatory T cells by indigenous clostridium species. *Science* **2011**, *331*, 337–341. [CrossRef]

51. Atarashi, K.; Tanoue, T.; Oshima, K.; Suda, W.; Nagano, Y.; Nishikawa, H.; Fukuda, S.; Saito, T.; Narushima, S.; Hase, K.; et al. Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* **2013**, *500*, 232–236. [CrossRef]

52. Scaria, J.; Suzuki, H.; Ptak, C.P.; Chen, J.-W.; Zhu, Y.; Guo, X.-K.; Chang, Y.-F. Comparative genomic and phenomic analysis of Clostridium difficile and Clostridium sordellii, two related pathogens with differing host tissue preference. *BMC Genom.* **2015**, *16*, 448. [CrossRef]

53. Tracy, B.P.; Jones, S.W.; Fast, A.G.; Indurthi, D.C.; Papoutsakis, E.T. Clostridia: The importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. *Curr. Opin. Biotechnol.* **2012**, *23*, 364–381. [CrossRef]

54. Rowland, I.; Gibson, G.; Heinken, A.; Scott, K.; Swann, J.; Thiele, I.; Tuohy, K. Gut microbiota functions: Metabolism of nutrients and other food components. *Eur. J. Nutr.* **2018**, *57*, 1–24. [CrossRef] [PubMed]

55. Palmas, V.; Pisanu, S.; Madau, V.; Casula, E.; Deledda, A.; Cusano, R.; Uva, P.; Vascellari, S.; Loviselli, A.; Manzin, A.; et al. Gut microbiota markers associated with obesity and overweight in Italian adults. *Sci. Rep.* **2021**, *11*, 5532. [CrossRef] [PubMed]

56. Mihatsch, W.A.; Högel, J.; Pohlandt, F. Hydrolysed protein accelerates the gastrointestinal transport of formula in preterm infants. *Acta Paediatr.* **2001**, *90*, 196–198. [CrossRef] [PubMed]

57. Roager, H.M.; Hansen, L.B.S.; Bahl, M.I.; Frandsen, H.L.; Carvalho, V.; Gøbel, R.J.; Dalgaard, M.D.; Plichta, D.R.; Sparholt, M.H.; Vestergaard, H.; et al. Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat. Microbiol.* **2016**, *1*, 16093. [CrossRef] [PubMed]

58. Mead, G.C. The Amino Acid-fermenting Clostridia. *J. Gen. Microbiol.* **1971**, *67*, 47–56. [CrossRef] [PubMed]

59. Li, C.; Wu, A.C.; Go, R.M.; Malouf, J.; Turner, M.S.; Malde, A.K.; Mark, A.E.; Gilbert, R.G. The characterization of modified starch branching enzymes: Toward the control of starch chain-length distributions. *PLoS ONE* **2015**, *10*, e0125507. [CrossRef] [PubMed]

60. Crost, E.H.; Tailford, L.E.; Le Gall, G.; Fons, M.; Henrissat, B.; Juge, N. Utilisation of mucin glycans by the human gut symbiont Ruminococcus gnavus is strain-dependent. *PLoS ONE* **2013**, *8*, e76341. [CrossRef]

61. Png, C.W.; Lindén, S.K.; Gilshenan, K.S.; Zoetendal, E.G.; McSweeney, C.S.; Sly, L.I.; McGuckin, M.; Florin, T.H.J. Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *Am. J. Gastroenterol.* **2010**, *105*, 2420–2428. [CrossRef]

62. Desai, M.S.; Seekatz, A.M.; Koropatkin, N.M.; Kamada, N.; Hickey, C.A.; Wolter, M.; Pudlo, N.A.; Kitamoto, S.; Terrapon, N.; Muller, A.; et al. A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell* **2016**, *167*, 1339–1353.e1321. [CrossRef]

63. Walker, A.; Ince, J.; Duncan, S.H.; Webster, L.M.; Holtrop, G.; Ze, X.; Brown, D.; Stares, M.D.; Scott, P.; Bergerat, A.; et al. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J.* **2010**, *5*, 220–230. [CrossRef]

64. Wright, H.D. Direct fermentation of disaccharides and variation in sugar utilization by Streptococcus thermophilus. *J. Path Bacteriol.* **1936**, *43*, 487–501. [CrossRef]

65. Hutkins, R.; Morris, H.A.; McKay, L.L. Galactose transport in Streptococcus thermophilus. *Appl. Environ. Microbiol.* **1985**, *50*, 772–776. [CrossRef] [PubMed]

66. Kanehisa, M.; Sato, Y.; Kawashima, M. KEGG mapping tools for uncovering hidden features in biological data. *Protein Sci.* **2021**, *31*, 47–53. [CrossRef] [PubMed]

67. Monteiro, P.O.A.; Victora, C. Rapid growth in infancy and childhood and obesity in later life—A systematic review. *Obes. Rev.* **2005**, *6*, 143–154. [CrossRef] [PubMed]

68. Dos Santos Pereira Indiani, C.M.; Rizzardi, K.F.; Castelo, P.M.; Ferraz, L.F.C.; Darrieux, M.; Parisotto, T. Childhood obesity and Firmicutes/Bacteroidetes ratio in the gut microbiota: A systematic review. *Child. Obes.* **2018**, *14*, 501–509. [CrossRef]

69. Castaner, O.; Goday, A.; Park, Y.-M.; Lee, S.-H.; Magkos, F.; Shiow, S.-A.T.E.; Schröder, H. The gut microbiome profile in obesity: A systematic review. *Int. J. Endocrinol.* **2018**, *2018*, 4095789. [CrossRef]

70. Mennella, J.A.; Papas, M.A.; Reiter, A.R.; Stallings, V.A.; Trabulsi, J.C. Early rapid weight gain among formula-fed infants: Impact of formula type and maternal feeding styles. *Pediatr. Obes.* **2019**, *14*, e12503. [CrossRef]

71. Petersen, C.; Bell, R.; Klag, K.A.; Lee, S.-H.; Soto, R.; Ghazaryan, A.; Buhrke, K.; Ekiz, H.A.; Ost, K.S.; Boudina, S.; et al. T cell-mediated regulation of the microbiota protects against obesity. *Science* **2019**, *365*, eaat9351. [CrossRef]

72. Von Berg, A.; Filipiak-Pittroff, B.; Kramer, U.; Link, E.; Heinrich, J.; Koletzko, S.; Grubl, A.; Hoffmann, U.; Beckmann, C.; Reinhardt, D.; et al. The German Infant Nutritional Intervention Study (GINI) for the preventive effect of hydrolyzed infant formulas in infants at high risk for allergic diseases. Design and selected results. *Allergol. Select.* **2017**, *1*, 28–38. [CrossRef]

73. Ecklu-Mensah, G.; Gilbert, J.; Devkota, S. Dietary selection pressures and their impact on the gut microbiome. *Cell. Mol. Gastroenterol. Hepatol.* **2021**, *13*, 7–18. [CrossRef]

74. Alexander, M.; Turnbaugh, P.J. Deconstructing mechanisms of diet-microbiome-immune interactions. *Immunity* **2020**, *53*, 264–276. [CrossRef] [PubMed]

75. Ventura, A.K.; Beauchamp, G.K.; Mennella, J.A. Infant regulation of intake: The effect of free glutamate content in infant formulas. *Am. J. Clin. Nutr.* **2012**, *95*, 875–881. [CrossRef] [PubMed]

76. Koletzko, B.; von Kries, R.; Monasterolo, R.C.; Subias, J.E.; Scaglioni, S.; Giovannini, M.; Beyer, J.; Demmelmair, H.; Anton, B.; Gruszfeld, D.; et al. Infant feeding and later obesity risk. *Adv. Exp. Med. Biol.* **2009**, *646*, 15–29. [CrossRef] [PubMed]

77. Wade, K.H.; Hall, L.J. Improving causality in microbiome research: Can human genetic epidemiology help? *Wellcome Open Res.* **2020**, *4*, 199. [CrossRef]