Diurnal rhythmicity of fecal microbiota and metabolite profiles in the first year of life: a randomized controlled interventional trial with infant formula

Nina Heppner[1,2], Sandra Reitmeier[1,2], Marjolein Heddes[1,2], Michael Vig Merino[1,2], Leon Schwartz[3], Alexander Dietrich[3], Markus List[3], Michael Gigl[4], Chen Meng[4], Hélène Omer[1,2], Karin Kleigrewe[4], Melanie Schirmer[2], Silke Kiessling[1,2,5*], Dirk Haller [1,2*]

[1] Chair of Nutrition and Immunology, Technical University of Munich, Gregor-Mendel-Str. 2, 85354 Freising, Germany

[2] ZIEL - Institute for Food & Health, Technical University of Munich, 85354 Freising, Germany

[3] Chair of Experimental Bioinformatics, Technical University of Munich, Maximus-von-Imhof Forum 3, 85354 Freising, Germany

[4] Bavarian Center for Biomolecular Mass Spectrometry, Technical University of Munich, Gregor-Mendel-Str. 4, 85354 Freising, Germany

[5] Faculty of Health and Biomedical Science, University of Surrey, 388 Stag Hill Campus, 17 GU27XH, Guildford, United Kingdom

* Authors contributed equally

Correspondence:
Prof. Dr. Dirk Haller
Technical University of Munich
Chair of Nutrition and Immunology
Gregor-Mendel-Str. 2
85354 Freising, Germany
Email: dirk.haller@tum.de
Phone: + 49 8161-71 2026

Keywords

Circadian rhythmicity, infant microbiome, untargeted metabolomics, Bifidobacteria, Galacto-oligosaccharides, intervention trial, infant formula; gut chemostat

Summary (150 words)

Microbiota assembly in the infant gut is influenced by time and duration of dietary exposure to breast-milk, infant formula and solid foods. In this randomized controlled intervention study, longitudinal sampling of infant stools (n=998) showed similar development of fecal bacterial communities between formula- and breast-fed infants during the first year of life (N=210). Infant formula supplemented with galacto-oligosaccharides (GOS) was most efficient to sustain high levels of bifidobacteria compared to formula containing *B. longum* and *B. breve* or placebo. Metabolite (untargeted) and bacterial profiling (16S *rRNA*/shallow metagenomics sequencing) revealed 24-hour oscillations and integrated data analysis identified circadian networks. Rhythmicity in bacterial diversity, specific taxa and functional pathways increased with age and was most pronounced following breast-feeding and GOS-supplementation. Circadian rhythms in dominant taxa were discovered *ex-vivo* in a chemostat model. Hence microbiota rhythmicity develops early in life, likely due to bacterial intrinsic clock mechanism and is affected by diet.

Main

Characteristic of the developing infant gut microbiota is the high degree of individual variation during the first years of life[1]. The deterministic or stochastic succession of microbial communities in the healthy infant gut is largely linked to environmental and dietary exposures[2-4]. Colonization of the sterile infant gut starts at birth and depends on fetal exposure to microbial metabolites from the mother (immune imprinting)[5], birth mode, and the presence of microbial pioneers from the fecal, vaginal or skin environment[3]. Timing and duration of breast- and formula feeding are subsequent drivers of early life microbial succession[6,7]. Breast-fed infants are typically characterized by low bacterial diversity and dominance of bifidobacteria[8-10]. Human breastmilk oligosaccharides (HMOs) selectively shape the early life colonization[11-13] and the supplementation of infant formula with complex oligosaccharides such as galacto-oligosaccharides (GOS) partially mimics the effect of breastmilk[14,15]. Considering the well-documented effects of the early life microbiota on health outcomes later in life[8,16,17], the attempts to optimize infant formula are numerous[18-21] but often lack well-controlled intervention strategies to directly compare supplements in infant formula.

Circadian clocks have developed in almost all organisms to anticipate recurring environmental changes. In mammals, a central clock in the brain orchestrates subordinated peripheral clocks of various tissues and single cells[22]. Thereby the circadian system drives 24-hour oscillations to control body and organ functions, however, the contribution of circadian rhythms in the developing gut including the process of microbial colonization is completely unknown[23]. So far, circadian clock mechanisms were believed to be restricted to photosynthetic bacteria, such as cyanobacteria, but, in single bacterial cultures circadian oscillations were found in gene expression of non-photosynthetic *Bacillus subtilis*[24] and swarming behavior of the gut bacterium *Klebsiella aerogenes*[25], suggesting the presence of endogenous circadian clocks in specific gut bacteria. In recent years, time-of-day dependent fluctuations in the abundance of certain gut bacterial species were shown at the individual and population level[26,27]. Growing evidence provides a link between loss of rhythmicity and disease pathology[27]. Animal studies determined the origin of these bacterial oscillations to reside in the circadian system (lat. *circa* = approximately, *dies* = day)[28-30], particularly the intestinal circadian clock[31]. The mammalian

circadian system develops *in utero*, which continues during aging and plays an important role during development[32]. Gut microbiota rhythms can influence clock gene expression in the host [33,34] and disturbance of the circadian crosstalk between bacteria and the host was shown to affect gastrointestinal health and metabolism[30,31,34]. Whether bacteria endogenous clock mechanisms can drive rhythmicity of complex microbiota and thereby impact host physiology remains unknown.

Here we established a randomized, controlled intervention trial including 210 infants to address the role of cow's-milk based infant formula supplemented with bifidobacteria and GOS on early life development of the gut microbiota. Longitudinal sampling of infant stool (n=998) allowed us to temporally resolve microbiota and metabolite profiles and to identify diurnal oscillations of the microbiome in relation to dietary exposure. Circadian rhythmicity of dominant bacteria taxa was confirmed *ex-vivo* in a gut chemostat model, supporting the hypothesis that microbiota rhythmicity develops early in life and is affected by diet. Moreover, these results suggest that circadian rhythms for specific taxa established independent of the host by endogenous clock mechanisms of the bacteria itself.

Results

Eligibility assessment of 265 pregnant women led to the inclusion of 223 infants in the intervention study. Of those infants, 210 actively participated and were randomly assigned to one of the four different formula groups (**Fig. 1**). 63 infants were breast-fed for the duration of the intervention (1 year). 35 infants were randomized to receive the placebo containing no additional supplements (Formula A), 39 infants were randomized to Formula B containing *B. longum* and *B. breve*, 36 infants received formula containing prebiotics (GOS) (Formula C) and 37 infants consumed formula containing both bifidobacteria and GOS (Formula D) (**Fig. 1**). Fecal samples were collected at five time points during the first year of life (**Fig. 2A**). 87% of participants contributed to all sampling time points. The characteristics of participants in the five study arms were comparable (**Table 1**) and the weight gain of infants was similar across groups (**Suppl. Fig. 1A**). Daily formula amounts consumed were similar across interventional groups (**Suppl. Fig. 1B-C**). Rates for study discontinuation were low (**Table 1**) with suspected cow's milk protein allergy (6 out of 9 cases) being the main reason for study withdrawal.

Temporal development of infant microbiota

Longitudinal infant and parental fecal sampling (n_i=998; n_p=152) demonstrated that the temporal assembly of infant microbiota developed towards the parental microbial composition (**Fig. 2B-D, Suppl. Fig. 2A-B**). Supervised beta-diversity analysis illustrated significant shifts of microbiota profiles at month 3-24, yet showing clearly distinct patterns compared to parental samples, whereas the early sampling time points (0.5, 1 and 3 months) are characterized by indistinguishable clusters of microbiota profiles (**Fig. 2B, Suppl. Fig. 2C**). Unsupervised analysis based on generalized UniFrac distances identified remarkable heterogeneity in microbiota profiles largely independent of the feeding type or birth groups (C-section vs vaginal birth) (**Fig. 2C**). Parental samples appeared in two distinct phylogenetic clusters and were flanked by samples from the older sampling time points with Bacteroidota (synonym Bacteroidetes) and Bacillota (syn. Firmicutes) becoming more dominant. Relative abundance of Actinomycetota (syn. Actinobacteria) was highly variable at early sampling time points and decreased with age (**Fig. 2C**). Gradual development of the infant microbiota is highlighted by increasing richness and alpha-diversity reaching significance at 3 months of age. Significant differences in average individual richness and Shannon effective number of species between

infants and parents dissolved after 24 months (**Fig. 2D, Suppl. Fig. 2A-B**). Unsupervised hierarchical clustering identified three distinct clusters across all infant samples during the first year of life (**Fig. 2F**). The clusters were characterized by a disproportionally high relative abundance (%) of the orders Bifidobacteriales (Cluster 1), Enterobacterales (Cluster 2) and Bacteroidales (Cluster 3) (**Fig. 2G**). Minor differences in bacterial abundance between the clusters included Eubacteriales (syn. Lachnospirales), Lactobacillales and Veillonellales-Selenomonadales (**Suppl. Fig. 2D**). The Bifidobacteriales-dominated cluster was largely composed of fecal samples from early time points (0.5, 1 and 3 months). Formula consumption peaked at 3 months of age and breast-feeding simultaneously decreased with increasing formula milk consumption (**Fig. 2E**). Concordant with decreasing formula intake after 3 months (**Fig. 2E**), 97.5% of infants received solid foods by month 7 (**Table 1**). The two other clusters also included samples from older infants (7 and 12 months) (**Fig. 2H**). Samples from interventions and breast-fed infants (**Fig. 2H**) were equally distributed in all three clusters. In addition, fecal samples from babies born via caesarian section were distributed homogenously (**Suppl. Fig. 2E**).

Impact of bifidobacterial and GOS intervention on infant microbiota assembly

Generalized UniFrac distance showed no significant differences between community structures of the four differentially supplemented infant formulas (A-D) and the exclusively breast-fed group (**Fig. 3A**), suggesting that inter-individual variations largely trumped the stratification of microbiota profiles based on feeding regiments. Species richness increased equally across feeding groups as infants aged (**Fig. 3B**). Beta-diversity was not significantly different between breast-feeders and the formula groups (**Suppl. Fig. 3A),** but differed between the placebo (Formula A) and other supplemented groups at age 3 months (**Suppl. Fig. 3B)**. Progression in relative abundance of the top ten taxa over time are summarized in the horizon blot (**Fig. 3C**). *Bifidobacterium* was the most abundant genus (41.2%) across all infant fecal samples reaching peak abundance at the age of 3 months (**Fig. 3C**). Average abundance of *Bifidobacterium* was lowest in the placebo group A (32.3%) followed by the bifidobacteria-supplemented group B (39.7%). *Bacteroides, Escherichia-Shigella, Veillonella, Klebsiella, Clostridium, Streptococcus* and *Enterococcus* contribute to the very early life infant communities (months 0.5, 1, 3) at substantially lower abundance levels (<13%) (**Fig. 3C**). Relative to the taxon median, the genera *Bacteroides* and *Blautia* clearly increased in average abundance at month 7 and 12 (**Fig. 3C**). Despite some variation over time, formula-fed infants showed on average a 29% higher abundance of *Veillonella* compared to breast-fed infants (**Fig. 3C**). GOS supplementation in the interventional groups C and group D (*Bifidobacterium* and GOS) increased the average bifidobacterial abundance above values in breast-fed and reached significance at month 3 and 7 of age (**Fig. 3D**). No additional differences were detected between interventions (data not shown). Shallow metagenomic analysis identified *B. longum*, *B. breve*, *V. parvula*, *E. coli* and *B. bifidum* as the 5 most abundant and prevalent bacterial species at month 3 and 7 independent of the feeding groups (**Fig. 3E**). Some breast-fed infants lacked the most common *Bifidobacterium* strains and instead harbored different *Veillonella* species and *Enterococcus faecalis* (**Fig. 3E**). Correlation analysis based on shallow metagenomics identified pathways related to amino acid biosynthesis (yellow boxes), nucleoside and nucleotide biosynthesis (blue), cell structure biosynthesis (green) and cofactor, carrier and vitamin biosynthesis (orange) in decreasing abundance (**Fig. 3F**). In accordance with the low fiber availability at these early life stages, fermentation was the least abundant pathway. No clear clustering occurred in relation to age and feeding (**Fig. 3F**).

Metabolites clearly differ between feeding groups and display diurnal rhythmicity

We next characterized the metabolite milieu at three months of age since solid food hasn't been introduced yet and infants consumed the highest volume of formula at this time point. Supervised and unsupervised statistical analysis was performed using untargeted metabolomics data. Contrary to the supervised analysis (**Fig. 4A**, left panel), the unsupervised approach identified three distinct metabolite clusters (**Fig. 4A**, right panel). Cluster 1 and 2 were clearly dominated by fecal samples from breast-fed infants, while the majority of formula-fed infants resided in cluster 3 (**Fig. 4A**, right panel). Cluster analysis of metabolite abundances was used to stratify patterns according to the feeding groups (**Fig. 4B**). As illustrated in the heatmap for the 49 most abundant features, breast- and formula-fed infants differentially clustered with only occasional appearance of samples from formula-fed infants in the breast-fed cluster (**Fig. 4B**). A more detailed analysis on the respective metabolites, phosphocholines as well as some specialized sugar derivatives like sugar acid, N-acetylated quadruple sugar or N-acetylated trisaccharide are increased in the breast-fed group, whereas carnitines is increased in all formula-fed groups. Indole derivatives such as bacterial-derived indolelactic acid is also increased in formula-fed groups B, C and D (**Fig. 4C**). Network correlational analysis was employed to identify the association of bacterial species and fecal metabolites (**Fig. 4D**). The top ten significantly bacterial species correlated positively or negatively with metabolites, suggesting functional redundancy between taxa (**Fig. 4E**). Variation in abundance of metabolites and bacteria according to the sampling daytime was previously identified in adult populations[27,28]. To evaluate whether this also occurs in infants in the first year of life, we analyzed daytime-dependent fluctuations in metabolite abundance at age 3 months. Peak abundance of fecal metabolites across all fecal samples is illustrated in the heatmap, supporting diurnal rhythms (**Fig. 4F**). Network correlational analysis further correlated bacteria (zOTUs), such as *B. longum* and *Veillonella* and rhythmic metabolites (**Fig. 4G**). These findings support the hypothesis that circadian regulation starts at early stages of life with possible functional contribution to the infant microbiome.

The development of diurnal rhythmicity of the infant microbiota is affected by age and formula supplementation

In accordance to daytime-dependent differences in metabolite abundance and their link to distinct bacterial taxa (**Fig. 4G**), significant diurnal rhythms were found in bacterial richness and alpha-diversity of fecal samples collected from infants throughout day and night (**Fig. 5A, Suppl. Fig. 4A, B**). Diurnal rhythmicity was also present at the level of individual zOTUs as illustrated by their normalized peak abundance across the 24 h day (**Fig. 5B**). Interestingly, significant rhythmicity was detected for the highly abundant genera *Bifidobacterium* (**Fig. 5C**). The amount of diurnal zOTUs increased as infants mature in parallel to the increasing diversification of infant gut microbiota, with highest numbers of rhythmic zOTUs at age 12 months (**Fig. 5D**). Of note, rhythmic zOTUs were most numerous in cluster 3 containing samples of older infants (**Fig. 2H, Suppl. Fig. 4C**). Interestingly, the number of rhythmic bacteria was highest in formula groups supplemented with bifidobacteria and GOS, which even exceeded the number of rhythmic zOTUs in breast-fed infants (**Fig. 5E**). Independent of the groups, the highest numbers of rhythmic zOTUs belonged to the genera *Veillonella* (7 zOTUs), *Bacteroides* (6 zOTUs), *Bifidobacterium* (6 zOTUs), *Streptococcus* (6 zOTUs) and *Clostridium* (5 zOTUs). PICRUSt2 analysis revealed that the total amount of significantly rhythmic assigned pathways dominated in samples from formula group D at age 7 months (604), followed by the breast-fed group showing rhythmicity in 162 pathways (**Fig. 5F**). Furthermore, LEfSe analysis[35] of these rhythmic metabolic pathways identified that the placebo group (formula A)

differed the most from the other feeding groups and particularly 29 rhythmic pathways differed between placebo and breast-fed infants (**Suppl. Fig. 4D**). Interestingly, in all group-to-placebo comparisons, fatty acid beta-degradation was enriched in placebo-fed infants (**Suppl. Fig. 4D**). In contrast to the placebo group, results from the groups supplemented with formula containing bifidobacteria and GOS most closely resembled the breast-fed group.

Circadian rhythms in dominant infant microbial taxa persist *in vitro*

In order to investigate whether microbial rhythms observed during early life are host-driven or rely on microbial intrinsic clock mechanisms, fecal samples from selected exclusive breast-fed and formula-fed (group A) infants at age 3 months were cultured in a chemostat model. To establish stable growth conditions, an initial adaptation phase (batch culture) was set up followed by continuous culturing for 7 days in the chemostat to ensure host-independent conditions (**Fig. 5G**). Beta diversity was similar between chemostats containing stool samples from the breast-fed donor, but differed significantly between feeding conditions (**Suppl. Fig. 5A**). In contrast to the inoculum, which was highly enriched in *Bifidobacterium* (>90% abundance) in the breast-fed sample, *Bifidobacterium* reached notably lower average abundance in the chemostat (12.79 < 89.73 in breast-fed; 4.74 < 75.49 in group A) (**Suppl. Fig. 5B**). In the seconded paired chemostat, the inoculum came from placebo formula group A, and the genera *Clostridium* and *Escherichia−Shigella* were similarly enriched, whereas *Bifidobacterium* was low in abundance (**Suppl. Fig. 5B)**. To investigate potential rhythmic alterations in abundance, beginning at day seven of continuous culture, samples were taken every 4 hours over the course of 48 hours (**Fig. 5G)**. Circadian analysis on microbiota composition revealed truly endogenous circadian rhythms in all chemostats illustrated at the level of dominant phyla, families and genera (**Fig 5H, Suppl. Fig. 5B, C**). Rhythmic changes in abundance are highlighted for highly abundant zOTUs (**Fig. 5I, Suppl. Fig. 5D**). In accordance with results demonstrated in infant stool samples (**Fig. 5C, E**), significant rhythms were identified for zOTUs of the genera *Bifidobacterium* and *Enterobacteriaceae,* which were overrepresented in the breast-fed samples, and *Clostridium sensu stricto* 1, which were overrepresented in chemostat samples from formula A (**Fig. 5I, Suppl. Fig. 5D**). These results are the first demonstration that specific taxa from a complex microbiota undergo circadian rhythmicity after seven days of continuous culture, pointing to a host-independent, intrinsic establishment of rhythmicity in the infant bacteria.

Discussion

In this randomized, controlled intervention trial, we demonstrated that formula- and breast-fed infants share characteristic features of microbiota assembly in the first year of life including age-related increase in bacterial richness and diversity as well as dominant abundance of specific bifidobacterial species. Interestingly 24-hour bacterial oscillations were present in all groups and could be integrated into rhythmic metabolite networks. Of note, bacterial profiles highly varied between individuals, even though the study population was homogenous related to geographical area, formula consumption, daily routines (sleep, eat, diaper change) as well as birth, family and health characteristics of the cohort. Interestingly, formula- and breast-fed infants substantially differed in their metabolite profiles, despite considerably overlapping bacterial community structures, emphasizing the unique composition of breast milk.[36] Nevertheless, and consistent with previous findings,[15,37,38] GOS supplemented infant formula was most efficient to sustain high levels of the genus *Bifidobacterium* compared to formula containing bifidobacteria or placebo. Remarkably, the presence of GOS in infant formula

resulted in the high numbers of rhythmic zOTUs across a range of different bacterial taxa, even exceeding the number of rhythmic zOTUs in breastfed infants. Similar but less pronounced was the effect of formula supplementation with *B. longum* and *B. breve*, despite representing the most abundant species in the infant gut, however placebo showed the fewest number of rhythmic bacterial taxa. Host independent circadian regulation of bacterial rhythmicity was confirmed in breast-fed and infant formula derived stool microbiota using an *ex vivo* gut chemostat model. These findings demonstrate for the first time that bacterial intrinsic clock mechanisms contribute to the development of circadian rhythmicity in the microbiota at early life stages.

The infant microbiome gradually develops towards a more diverse adult-like microbiome over time.[7] Interestingly, shallow metagenomic sequencing of samples from 3- and 7-months old infants demonstrated that *Bifidobacterium* (*B. breve*, *B. longum* and *B. bifidum*), *Veillonella* and *Enterobacteriaceae* dominate this early life stage, irrespective of the diet received. These results are consistent with previously published baby cohort where fecal samples of 40 days, 3-months and 6-months old infants were compared, showing that *Bifidobacterium* and *Enterobacteriaceae* were the most predominant taxa, regardless of diet and age.[39] Alpha-diversity significant increased from age 3 months onwards, gradually approaching parental diversity levels at age 24 months. This is in line with previous reports indicating significant changes in the dominant phyla and diversity appearing from age 3 months onwards[16] or at the end of the first year of life[40] until the parent's diversity and composition is reached by month 14[16] or on the second birthday.[41] None of the three types of supplementation investigated in this study were able to fully recreate the breast milk-related microbial environment. Accordingly, the metabolic profile of 3-months old, breast-fed babies was clearly distinguishable from formula-consuming infants. For example, breast-fed babies showed a high number of organic oxygen metabolites, whereas formula-fed babies were associated with a lipid-enriched metabolome composition. Similarly, metabolic profiles of 12 month old infants were reported to be distinguishable between diets.[42]

Previous work in adult populations, done by others as well as our group, documented that microbiota composition and function undergoes diurnal rhythmicity.[26,43] However, whether these rhythms appear at early age remained unknown. We provide the first evidence that during the first year of life in all feeding groups, the abundance of the most dominant taxa, including *Bifidobacterium*, varied depending on the time of defecation and circadian rhythm analysis revealed that these fluctuations follow the 24-hour day. These diurnal oscillations could indeed explain the opposing results on abundance fluctuations reported, e.g. for bifidobacteria[44] if samples were taken at different day times in previous studies. Here we discovered significant daytime-dependent fluctuations in bacterial diversity, at the level of genera and zOTUs in infant stool samples. Although rhythmicity was present in all feeding groups and at all ages examined, including the genera *Veillonella*, *Bacteroides*, *Bifidobacterium*, *Streptococcus* and *Clostridium*, dramatic differences were detected between rhythmic zOTUs based on infant, age and diet. The overall highest number of rhythmic zOTUs was found in infants fed with formula supplemented with the combination of bifidobacteria and GOS, which exceeded numbers seen in breast-fed infants. Moreover, circadian analysis identified the presence of diurnal rhythms in microbiota composition and function already in 2-weeks old infants demonstrating that microbial oscillations are an early event in the infant gut development. The number of rhythmic bacteria gradually increased as the infant gut matured and regular diet was introduced to the infant nutrition.

Circadian regulation of microbiota composition and function was recently shown to depend on the host circadian clock, particularly in intestinal epithelial cells.[45] Conversely, microbiota have been shown to influence host circadian rhythms,[46,47] indicating a cross-talk between the host interface and bacteria. Microbiota transfer experiments further demonstrated the physiological relevance of microbial oscillations for the host's metabolic and gastrointestinal health.[26,45,48] Since environmental factors, such as the light-dark-cycle, or time of food intake, are potent signals impacting microbial rhythms,[26,45,48,49] abnormal light or food exposure can alter microbial oscillations and thus might lead to the development of pathologies and disease. To gain further insights into the possible functionality of microbiota oscillations in infants, rhythmic zOTU-associated pathways were examined. Similar to results obtained from the taxonomic analysis, most rhythmic assigned pathways were found in 3- and 7-months old infants fed with formula containing GOS or both the prebiotic supplement and bifidobacteria respectively, followed by breast-fed infants. Importantly, results from the placebo group differed the most from breast-fed infants, whereas results from infants fed with formula containing bifidobacteria and GOS most closely resembled the breast-fed group. The association between the bifidogenic effect of GOS supplementation and the augmentation of the number of rhythmic pathways needs to be further investigated in future studies.

Recent work on single bacterial cultures discovered rhythmic gene expression in *Bacillus subtilis*[24] and recorded circadian rhythms in the swarming behavior of the gut bacterium *Klebsiella aerogenes*[25]. To identify whether the diurnal rhythms observed in infants gut microbiota are generated by the host circadian clock or bacterial intrinsic mechanism, we cultured complex microbiota communities in a chemostat model. For this purpose, samples from either exclusively breastfed or formula A-fed infants were cultured and supplemented with media containing infant formula. Of note, the relative abundance of taxa as well as bacterial diversity changed significantly when compared to the original inoculum. Despite both inocula being dominated by bifidobacteria, their relative abundance decreased after the continuous flow with culture media begun until a fluctuating but steady abundance was achieved. Importantly, after 7 days *in vitro* when the communities were established, circadian rhythm analysis of microbiota composition provides the first evidence of circadian rhythms in the abundance of dominant bacteria, such as *Bifidobacterium*, *Veillonella*, *Clostridium* and *Bacteroides*, in complex human microbiota communities. Since oscillations from 3-months old infant stool samples are documented outside and independent of the host, these are truly endogenous circadian rhythms and therefore must be generated by bacterial intrinsic clock mechanisms. Importantly, we observed that rhythmic bacterial taxa differed depending on whether the original sample was obtained from an exclusively breastfed infant or exclusively fed with formula A. For example, the composition of breast-fed infants was dominated by rhythmic oscillations of zOTUs belonging to *Enterobacteriacea* and *Bifidobacterium*, whereas formula A-fed samples were enriched with rhythmic zOTUs belonging to *Clostridium* and *Bacteroides*. However, independent of which taxa dominated in both extreme infant feeding exposures (exclusive BF and exclusive formula A), all of them had the potential to become rhythmic *in vitro*, suggesting the importance of bacterial intrinsic clock mechanism for the establishment of circadian microbial communities in the developing infant gut.

In summary, this controlled intervention study provided compelling evidence of circadian rhythmicity in the developing infant gut microbiota affected by diet. Our findings warrant the need for further analysis of circadian fluctuations of both bacteria and metabolites and their functional role in contributing to the benefits of infant nutrition.

References

1. Martino, C., Dilmore, A.H., Burcham, Z.M., Metcalf, J.L., Jeste, D., and Knight, R. (2022). Microbiota succession throughout life from the cradle to the grave. Nature Reviews Microbiology *20*, 707-720.
2. Penders, J., Thijs, C., Vink, C., Stelma, F.F., Snijders, B., Kummeling, I., van den Brandt, P.A., and Stobberingh, E.E. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. Pediatrics *118*, 511-521.
3. Ferretti, P., Pasolli, E., Tett, A., Asnicar, F., Gorfer, V., Fedi, S., Armanini, F., Truong, D.T., Manara, S., and Zolfo, M. (2018). Mother-to-infant microbial transmission from different body sites shapes the developing infant gut microbiome. Cell host & microbe *24*, 133-145. e135.
4. Yang, I., Corwin, E.J., Brennan, P.A., Jordan, S., Murphy, J.R., and Dunlop, A. (2016). The infant microbiome: implications for infant health and neurocognitive development. Nursing research *65*, 76.
5. Al Nabhani, Z., and Eberl, G. (2020). Imprinting of the immune system by the microbiota early in life. Mucosal immunology *13*, 183-189.
6. Mueller, N.T., Bakacs, E., Combellick, J., Grigoryan, Z., and Dominguez-Bello, M.G. (2015). The infant microbiome development: mom matters. Trends in molecular medicine *21*, 109-117.
7. Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y., Xie, H., and Zhong, H. (2015). Dynamics and stabilization of the human gut microbiome during the first year of life. Cell host & microbe *17*, 690-703.
8. Milani, C., Duranti, S., Bottacini, F., Casey, E., Turroni, F., Mahony, J., Belzer, C., Palacio, S.D., Montes, S.A., and Mancabelli, L. (2017). The first microbial colonizers of the human gut: composition, activities, and health implications of the infant gut microbiota. Microbiology and Molecular Biology Reviews *81*.
9. Lawson, M.A., O'Neill, I.J., Kujawska, M., Gowrinadh Javvadi, S., Wijeyesekera, A., Flegg, Z., Chalklen, L., and Hall, L.J. (2020). Breast milk-derived human milk oligosaccharides promote Bifidobacterium interactions within a single ecosystem. The ISME journal *14*, 635-648.
10. Bazanella, M., Maier, T.V., Clavel, T., Lagkouvardos, I., Lucio, M., Maldonado-Gòmez, M.X., Autran, C., Walter, J., Bode, L., and Schmitt-Kopplin, P. (2017). Randomized controlled trial on the impact of early-life intervention with bifidobacteria on the healthy infant fecal microbiota and metabolome. The American journal of clinical nutrition *106*, 1274-1286.
11. Bode, L. (2012). Human milk oligosaccharides: every baby needs a sugar mama. Glycobiology *22*, 1147-1162.
12. Ballard, O., and Morrow, A.L. (2013). Human milk composition: nutrients and bioactive factors. Pediatric Clinics of North America *60*, 49-74.
13. Bode, L., and Jantscher-Krenn, E. (2012). Structure-function relationships of human milk oligosaccharides. Advances in Nutrition *3*, 383S-391S.
14. Veereman-Wauters, G., Staelens, S., Van de Broek, H., Plaskie, K., Wesling, F., Roger, L., McCartney, A., and Assam, P. (2011). Physiological and bifidogenic effects of prebiotic supplements in infant formulae. Journal of pediatric gastroenterology and nutrition *52*, 763-771.
15. Sierra, C., Bernal, M.-J., Blasco, J., Martínez, R., Dalmau, J., Ortuno, I., Espín, B., Vasallo, M.-I., Gil, D., and Vidal, M.-L. (2015). Prebiotic effect during the first year of life in healthy infants fed formula containing GOS as the only prebiotic: a multicentre, randomised, double-blind and placebo-controlled trial. European journal of nutrition *54*, 89-99.
16. Stewart, C.J., Ajami, N.J., O'Brien, J.L., Hutchinson, D.S., Smith, D.P., Wong, M.C., Ross, M.C., Lloyd, R.E., Doddapaneni, H., and Metcalf, G.A. (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature *562*, 583-588.
17. Cox, L.M., Yamanishi, S., Sohn, J., Alekseyenko, A.V., Leung, J.M., Cho, I., Kim, S.G., Li, H., Gao, Z., and Mahana, D. (2014). Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. Cell *158*, 705-721.
18. Vandenplas, Y., Greef, E.D., and Veereman, G. (2014). Prebiotics in infant formula. Gut microbes *5*, 681-687.

19. Mugambi, M.N., Musekiwa, A., Lombard, M., Young, T., and Blaauw, R. (2012). Synbiotics, probiotics or prebiotics in infant formula for full term infants: a systematic review. Nutrition journal *11*, 1-32.
20. Indrio, F., Gutierrez Castrellon, P., Vandenplas, Y., Cagri Dinleyici, E., Francavilla, R., Mantovani, M.P., Grillo, A., Beghetti, I., Corvaglia, L., and Aceti, A. (2022). Health Effects of Infant Formula Supplemented with Probiotics or Synbiotics in Infants and Toddlers: Systematic Review with Network Meta-Analysis. Nutrients *14*, 5175.
21. Braegger, C., Chmielewska, A., Decsi, T., Kolacek, S., Mihatsch, W., Moreno, L., Piescik, M., Puntis, J., Shamir, R., and Szajewska, H. (2011). Supplementation of infant formula with probiotics and/or prebiotics: a systematic review and comment by the ESPGHAN committee on nutrition. Journal of pediatric gastroenterology and nutrition *52*, 238-250.
22. Mohawk, J.A., Green, C.B., and Takahashi, J.S. (2012). Central and peripheral circadian clocks in mammals. Annu Rev Neurosci *35*, 445-462. 10.1146/annurev-neuro-060909-153128.
23. Roenneberg, T., and Merrow, M. (2016). The circadian clock and human health. Current biology *26*, R432-R443.
24. Eelderink-Chen, Z., Bosman, J., Sartor, F., Dodd, A.N., Kovacs, A.T., and Merrow, M. (2021). A circadian clock in a nonphotosynthetic prokaryote. Sci Adv *7*. 10.1126/sciadv.abe2086.
25. Paulose, J.K., Cassone, C.V., Graniczkowska, K.B., and Cassone, V.M. (2019). Entrainment of the Circadian Clock of the Enteric Bacterium Klebsiella aerogenes by Temperature Cycles. iScience *19*, 1202-1213. 10.1016/j.isci.2019.09.007.
26. Thaiss, C.A., Zeevi, D., Levy, M., Zilberman-Schapira, G., Suez, J., Tengeler, A.C., Abramson, L., Katz, M.N., Korem, T., and Zmora, N. (2014). Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. Cell *159*, 514-529.
27. Reitmeier, S., Kiessling, S., Clavel, T., List, M., Almeida, E.L., Ghosh, T.S., Neuhaus, K., Grallert, H., Linseisen, J., Skurk, T., et al. (2020). Arrhythmic Gut Microbiome Signatures Predict Risk of Type 2 Diabetes. Cell Host Microbe *28*, 258-272 e256. 10.1016/j.chom.2020.06.004.
28. Thaiss, C.A., Zeevi, D., Levy, M., Zilberman-Schapira, G., Suez, J., Tengeler, A.C., Abramson, L., Katz, M.N., Korem, T., Zmora, N., et al. (2014). Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. Cell *159*, 514-529. 10.1016/j.cell.2014.09.048.
29. Liang, X., Bushman, F.D., and FitzGerald, G.A. (2015). Rhythmicity of the intestinal microbiota is regulated by gender and the host circadian clock. Proc Natl Acad Sci U S A *112*, 10479-10484. 10.1073/pnas.1501305112.
30. Altaha, B., Heddes, M., Pilorz, V., Niu, Y., Gorbunova, E., Gigl, M., Kleigrewe, K., Oster, H., Haller, D., and Kiessling, S. (2022). Genetic and environmental circadian disruption induce weight gain through changes in the gut microbiome. Mol Metab *66*, 101628. 10.1016/j.molmet.2022.101628.
31. Heddes, M., Altaha, B., Niu, Y., Reitmeier, S., Kleigrewe, K., Haller, D., and Kiessling, S. (2022). The intestinal clock drives the microbiome to maintain gastrointestinal homeostasis. Nat Commun *13*, 6068. 10.1038/s41467-022-33609-x.
32. Escobar, C., Rojas-Granados, A., and Angeles-Castellanos, M. (2021). Development of the circadian system and relevance of periodic signals for neonatal development. Handb Clin Neurol *179*, 249-258. 10.1016/B978-0-12-819975-6.00015-7.
33. Mukherji, A., Kobiita, A., Ye, T., and Chambon, P. (2013). Homeostasis in intestinal epithelium is orchestrated by the circadian clock and microbiota cues transduced by TLRs. Cell *153*, 812-827. 10.1016/j.cell.2013.04.020.
34. Leone, V., Gibbons, S.M., Martinez, K., Hutchison, A.L., Huang, E.Y., Cham, C.M., Pierre, J.F., Heneghan, A.F., Nadimpalli, A., Hubert, N., et al. (2015). Effects of diurnal variation of gut microbes and high-fat feeding on host circadian clock function and metabolism. Cell Host Microbe *17*, 681-689. 10.1016/j.chom.2015.03.006.
35. Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. Genome biology *12*, 1-18.
36. Andreas, N.J., Kampmann, B., and Le-Doare, K.M. (2015). Human breast milk: A review on its composition and bioactivity. Early human development *91*, 629-635.

37. Ben, X.-M., Li, J., Feng, Z.-T., Shi, S.-Y., Lu, Y.-D., Chen, R., and Zhou, X.-Y. (2008). Low level of galacto-oligosaccharide in infant formula stimulates growth of intestinal Bifidobacteria and Lactobacilli. World journal of gastroenterology: WJG *14*, 6564.
38. Matsuki, T., Tajima, S., Hara, T., Yahagi, K., Ogawa, E., and Kodama, H. (2016). Infant formula with galacto-oligosaccharides (OM55N) stimulates the growth of indigenous bifidobacteria in healthy term infants. Beneficial microbes *7*, 453-461.
39. Ma, J., Li, Z., Zhang, W., Zhang, C., Zhang, Y., Mei, H., Zhuo, N., Wang, H., Wang, L., and Wu, D. (2020). Comparison of gut microbiota in exclusively breast-fed and formula-fed babies: A study of 91 term infants. Scientific Reports *10*, 15792.
40. Clemente, J.C., Ursell, L.K., Parfrey, L.W., and Knight, R. (2012). The impact of the gut microbiota on human health: an integrative view. Cell *148*, 1258-1270.
41. Wernroth, M.-L., Peura, S., Hedman, A.M., Hetty, S., Vicenzi, S., Kennedy, B., Fall, K., Svennblad, B., Andolf, E., and Pershagen, G. (2022). Development of gut microbiota during the first 2 years of life. Scientific Reports *12*, 9080.
42. Sillner, N., Walker, A., Lucio, M., Maier, T.V., Bazanella, M., Rychlik, M., Haller, D., and Schmitt-Kopplin, P. (2021). Longitudinal profiles of dietary and microbial metabolites in formula-and breastfed infants. Frontiers in Molecular Biosciences *8*, 660456.
43. Reitmeier, S., Kiessling, S., Clavel, T., List, M., Almeida, E.L., Ghosh, T.S., Neuhaus, K., Grallert, H., Linseisen, J., and Skurk, T. (2020). Arrhythmic gut microbiome signatures predict risk of type 2 diabetes. Cell host & microbe *28*, 258-272. e256.
44. Taft, D.H., Liu, J., Maldonado-Gomez, M.X., Akre, S., Huda, M.N., Ahmad, S., Stephensen, C.B., and Mills, D.A. (2018). Bifidobacterial dominance of the gut in early life and acquisition of antimicrobial resistance. Msphere *3*, e00441-00418.
45. Heddes, M., Altaha, B., Niu, Y., Reitmeier, S., Kleigrewe, K., Haller, D., and Kiessling, S. (2022). The intestinal clock drives the microbiome to maintain gastrointestinal homeostasis. Nature Communications *13*, 6068.
46. Brooks, J.F., Behrendt, C.L., Ruhn, K.A., Lee, S., Raj, P., Takahashi, J.S., and Hooper, L.V. (2021). The microbiota coordinates diurnal rhythms in innate immunity with the circadian clock. Cell *184*, 4154-4167. e4112.
47. Mukherji, A., Kobiita, A., Ye, T., and Chambon, P. (2013). Homeostasis in intestinal epithelium is orchestrated by the circadian clock and microbiota cues transduced by TLRs. Cell *153*, 812-827.
48. Altaha, B., Heddes, M., Pilorz, V., Niu, Y., Gorbunova, E., Gigl, M., Kleigrewe, K., Oster, H., Haller, D., and Kiessling, S. (2022). Genetic and environmental circadian disruption induce weight gain through changes in the gut microbiome. Molecular Metabolism *66*, 101628.
49. Voigt, R.M., Forsyth, C.B., Green, S.J., Mutlu, E., Engen, P., Vitaterna, M.H., Turek, F.W., and Keshavarzian, A. (2014). Circadian disorganization alters intestinal microbiota. PloS one *9*, e97500.

Additional information

Corresponding author

Correspondence to Prof. Dr. Dirk Haller, Chair of Nutrition and Immunology, Technical University of Munich, Gregor-Mendel-Str. 2, 85354 Freising, Germany; dirk.haller@tum.de

Acknowledgements

The Infantibio-II study was initiated and financed by Töpfer GmbH (Dietmannsried, Germany). Besides manufacturing the study formula, they had no role in the conduct and management of the study, analysis and interpretation of data nor the creation of this manuscript. D.H. and S.K. received funding from the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) SFB 1371 (no. 395357507). D.H. also received funding in the frame of the Joint Programming Initiative of the European Union (project name EcoBiotic) and the German Ministry of Education and Research (BMBF; FKZ 01EA2207). The Technical University of Munich provided technical support through the Core Facility Microbiome of the ZIEL Institute for Food & Health (Klaus Neuhaus, Lukas Mix, Angela Sachsenhauser and Caroline Ziegler) for 16S rRNA gene amplicon sequencing. We thank Nikolai Köhler for giving us advice on integrative data analysis as well as Katharina Sontheimer, Claudia Seegerer and Franziska Kummert for support with sample management and handling. We are most grateful to the participating families and infants of the Infantibio-II study for their time, effort and enthusiasm in our study.

Author contributions

D.H. conceived and coordinated the project and secured funding. S.K. performed and supervised the circadian analysis. N.H. and D.H. were responsible for creation of trial documents and ethics approval, N.H. was responsible for participant recruitment, enrolment, data collection, management and coordination of collection of all trial samples. N.H., S.R., H.O., S.K. and M.VM. performed 16S rRNA sequencing and subsequent analysis with support from L.S. and M.L., and M.G., S.R. C.M. and K.K. were responsible for metabolomics analysis. S.R. and M.S. performed shotgun metagenomics analysis, H.O., M.H. and M.VM. performed the chemostat experiments and analysis, N.H., S.R., and M.L. performed bioinformatics analysis. D.H., S.K. and H.O. supervised the work and data analysis. N.H., S.K. and D.H. wrote the manuscript. All authors reviewed the manuscript.

Competing interest declaration

The authors declare no competing interests.

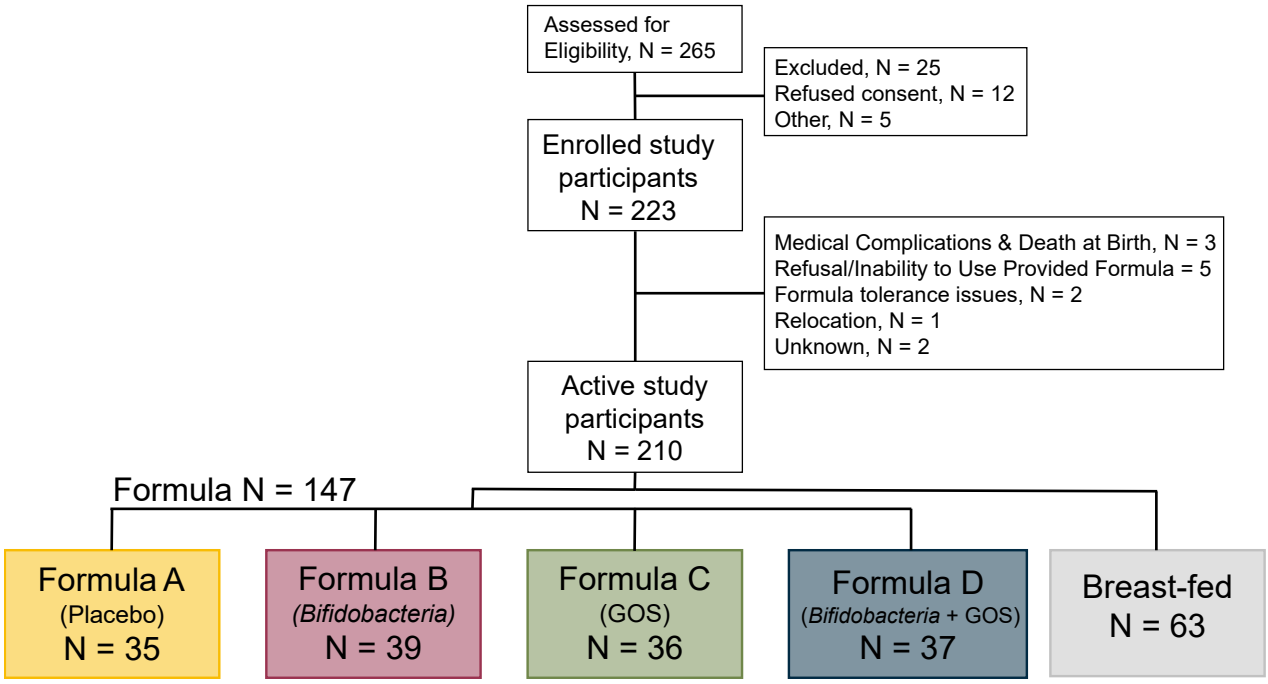

Assessed for Eligibility, N = 265
Excluded, N = 25
Refused consent, N = 12
Other, N = 5
Enrolled study participants N = 223
Medical Complications & Death at Birth, N = 3
Refusal/Inability to Use Provided Formula = 5
Formula tolerance issues, N = 2
Relocation, N = 1
Unknown, N = 2
Active study participants N = 210
Formula N = 147
Formula A (Placebo) N = 35
Formula B (Bifidobacteria) N = 39
Formula C (GOS) N = 36
Formula D (Bifidobacteria + GOS) N = 37
Breast-fed N = 63

Figure 1 Study schematic of participant enrollment and randomization.
Schematic of participant enrolment and reasons for drop-out before start of study. Active study participants (N =210), provided at least one fecal sample. Newborns were randomly assigned to one of four formula groups: Placebo, *Bifidobacterium*-supplemented (Formula B), galacto-oligosaccharide (GOS)-supplemented (Formula C) or formula containing GOS and *Bifidobacteria* (Formula D). Breast-fed infants that never consumed any of the provided formula served as reference group.

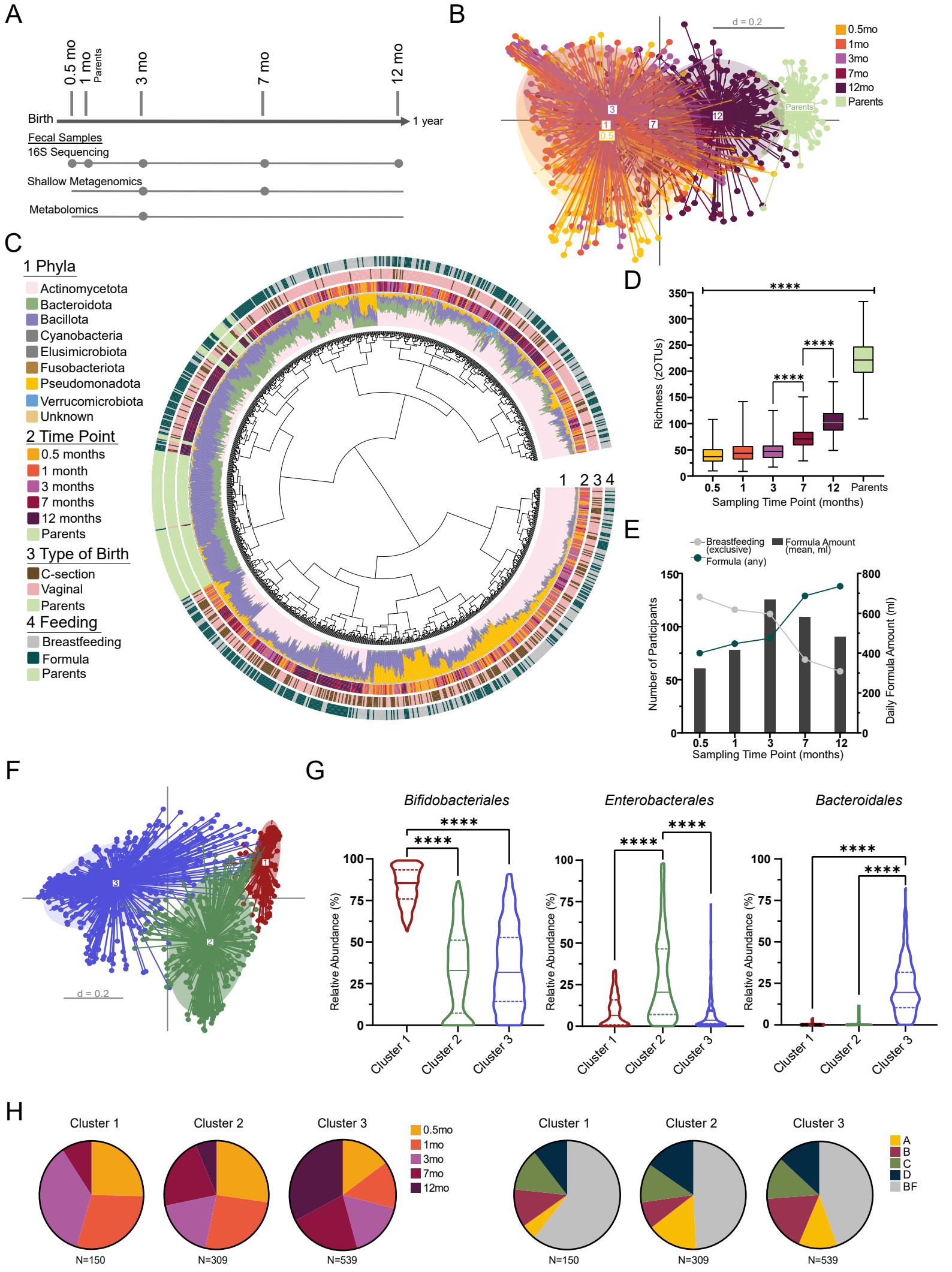

A
0.5 mo
1 mo
Parents
3 mo
7 mo
12 mo
Birth
1 year
Fecal Samples
16S Sequencing
Shallow Metagenomics
Metabolomics
B
d = 0.2
0.5mo
1mo
3mo
7mo
12mo
Parents
C
1 Phyla
Actinomycetota
Bacteroidota
Bacillota
Cyanobacteria
Elusimicrobiota
Fusobacteriota
Pseudomonadota
Verrucomicrobiota
Unknown
2 Time Point
0.5 months
1 month
3 months
7 months
12 months
Parents
3 Type of Birth
C-section
Vaginal
Parents
4 Feeding
Breastfeeding
Formula
Parents
1 2 3 4
D
Richness (zOTUs)
Sampling Time Point (months)
****
E
Breastfeeding (exclusive)
Formula Amount (mean, ml)
Formula (any)
Number of Participants
Daily Formula Amount (ml)
Sampling Time Point (months)
F
d = 0.2
G
Bifidobacteriales
Enterobacterales
Bacteroidales
Relative Abundance (%)
Cluster 1
Cluster 2
Cluster 3
****
H
Cluster 1
Cluster 2
Cluster 3
N=150
N=309
N=539
0.5mo
1mo
3mo
7mo
12mo
A
B
C
D
BF

Figure 2 Temporal development of infant microbiota

(A) Schematic of study fecal sample collection and time points of analysis of fecal samples.

(B) Nonmetric multidimensional scaling (NMDS) plot of fecal samples grouped according to infant age (months) or parent sample (separation of groups as whole is significant: $p \leq 0.001$, permutational multivariate analysis based on distance matrices).

(C) *De novo* phylogenetic clustering of all study fecal samples (1150 fecal samples from 210 infants (multiple time points) and 152 parents (single time point)) displayed as a dendrogram, phyla per sample as stacked bar plots (second innermost ring), metadata for each sample in the outermost three rings (feed type, type of birth and sampling time point).

(D) Alpha-diversity (richness, zOTUs) at different infant sampling time points and for parental samples. (**** = $p \leq 0.0001$, 25th to the 75th percentile box plot, median: solid line, whiskers: min to max).

(E) Mean daily formula intake (ml, bar plot) and exclusive breastfeeding rates (grey line) versus formula consumers (turquoise line).

(F) Metric multidimensional scaling (MDS) plot of beta-diversity (unsupervised clustering) of all infant samples (n= 998) producing three distinct clusters (separation of groups as whole is significant: $p \leq 0.001$, permutational multivariate analysis based on distance matrices).

(G) Top 3 relative abundance (%) cluster dominating bacterial order violin plot (**** = $p \leq 0.0001$, median: solid line, 25th and 75th percentile: dashed lines), de novo generated clusters.

(H) Descriptive cluster composition regarding sampling time point and feed type (breastfeeding: grey, formula A (placebo): yellow, formula B (bifidobacteria): red, formula C (GOS): green, formula D (bifidobacteria + GOS): dark blue).

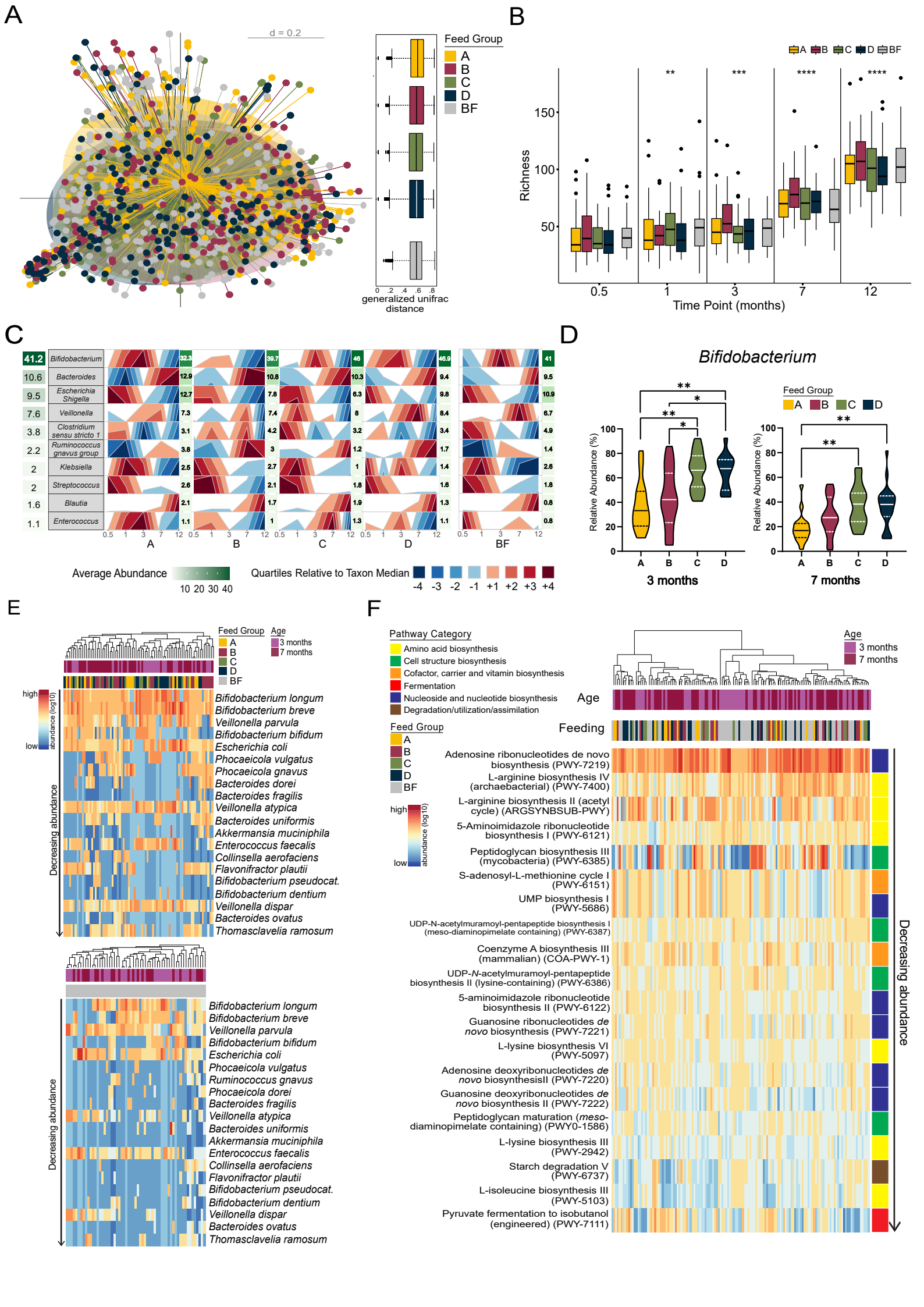

A
d = 0.2
Feed Group
A
B
C
D
BF
0 .2 .4 .6 .8
generalized unifrac distance
B
A B C D BF
Richness
150
100
50
0.5
1
3
7
12
Time Point (months)
**
***
****
****
C
41.2 Bifidobacterium
10.6 Bacteroides
9.5 Escherichia Shigella
7.6 Veillonella
3.8 Clostridium sensu stricto 1
2.2 Ruminococcus gnavus group
2 Klebsiella
2 Streptococcus
1.6 Blautia
1.1 Enterococcus
A
B
C
D
BF
Average Abundance
10 20 30 40
Quartiles Relative to Taxon Median
-4 -3 -2 -1 +1 +2 +3 +4
D
Bifidobacterium
Feed Group
A B C D
Relative Abundance (%)
3 months
7 months
E
Feed Group
A
B
C
D
BF
Age
3 months
7 months
high
low
abundance (log10)
Decreasing abundance
Bifidobacterium longum
Bifidobacterium breve
Veillonella parvula
Bifidobacterium bifidum
Escherichia coli
Phocaeicola vulgatus
Phocaeicola gnavus
Bacteroides dorei
Bacteroides fragilis
Veillonella atypica
Bacteroides uniformis
Akkermansia muciniphila
Enterococcus faecalis
Collinsella aerofaciens
Flavonifractor plautii
Bifidobacterium pseudocat.
Bifidobacterium dentium
Veillonella dispar
Bacteroides ovatus
Thomasclavelia ramosum
Ruminococcus gnavus
F
Pathway Category
Amino acid biosynthesis
Cell structure biosynthesis
Cofactor, carrier and vitamin biosynthesis
Fermentation
Nucleoside and nucleotide biosynthesis
Degradation/utilization/assimilation
Age
Feeding
Feed Group
A
B
C
D
BF
Adenosine ribonucleotides de novo biosynthesis (PWY-7219)
L-arginine biosynthesis IV (archaebacterial) (PWY-7400)
L-arginine biosynthesis II (acetyl cycle) (ARGSYNBSUB-PWY)
5-Aminoimidazole ribonucleotide biosynthesis I (PWY-6121)
Peptidoglycan biosynthesis III (mycobacteria) (PWY-6385)
S-adenosyl-L-methionine cycle I (PWY-6151)
UMP biosynthesis I (PWY-5686)
UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing) (PWY-6387)
Coenzyme A biosynthesis III (mammalian) (COA-PWY-1)
UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing) (PWY-6386)
5-aminoimidazole ribonucleotide biosynthesis II (PWY-6122)
Guanosine ribonucleotides de novo biosynthesis (PWY-7221)
L-lysine biosynthesis VI (PWY-5097)
Adenosine deoxyribonucleotides de novo biosynthesisII (PWY-7220)
Guanosine deoxyribonucleotides de novo biosynthesis II (PWY-7222)
Peptidoglycan maturation (meso-diaminopimelate containing) (PWY0-1586)
L-lysine biosynthesis III (PWY-2942)
Starch degradation V (PWY-6737)
L-isoleucine biosynthesis III (PWY-5103)
Pyruvate fermentation to isobutanol (engineered) (PWY-7111)
Decreasing abundance

Figure 3 Impact of bifidobacterial and GOS intervention on infant microbiota assembly
(A) Infant beta-diversity (MDS plot) based on feed group. Boxplots show generalized UniFrac distance for each of the five feed groups.

(B) Richness over time in the different feed groups (0.5:1 mo: ** = $p \leq 0.01$, 0.5:3 mo: *** = $p \leq 0.001$, 0.5:7 mo and 0.5:12 mo **** = $p \leq 0.0001$, Anova $p < 2.2e{-}16$).

(C) Top 10 relative abundance (%) taxa (genus) found in >80% of samples (average abundance >= 0.5). Color scale corresponding to increase (red) or decrease (blue) abundance relative to the taxon median, average abundance per feed group in green boxes.

(D) Frequency distribution of the relative abundance of the genus *Bifidobacterium* (median: solid line, upper and lower quartile: dotted line, ** = $p \leq 0.01$, * = $p \leq 0.05$). Minimal formula consumers were omitted (lowest quartile based on daily formula volume consumption).

(E) Relative abundance (log values) of the top 20 species (shallow metagenomics) of individual samples. Metadata as bars above samples (feeding, age) with hierarchal clustering at the very top.

(F) Top 20 identified pathways (column normalized) associated with shallow metagenomic sequencing samples sorted by decreasing abundance. Metadata as bars above samples (feeding, age) with hierarchal clustering at the very top; pathway (BioCyc identifier), and pathway categories on the left and right of the heatmap respectively.

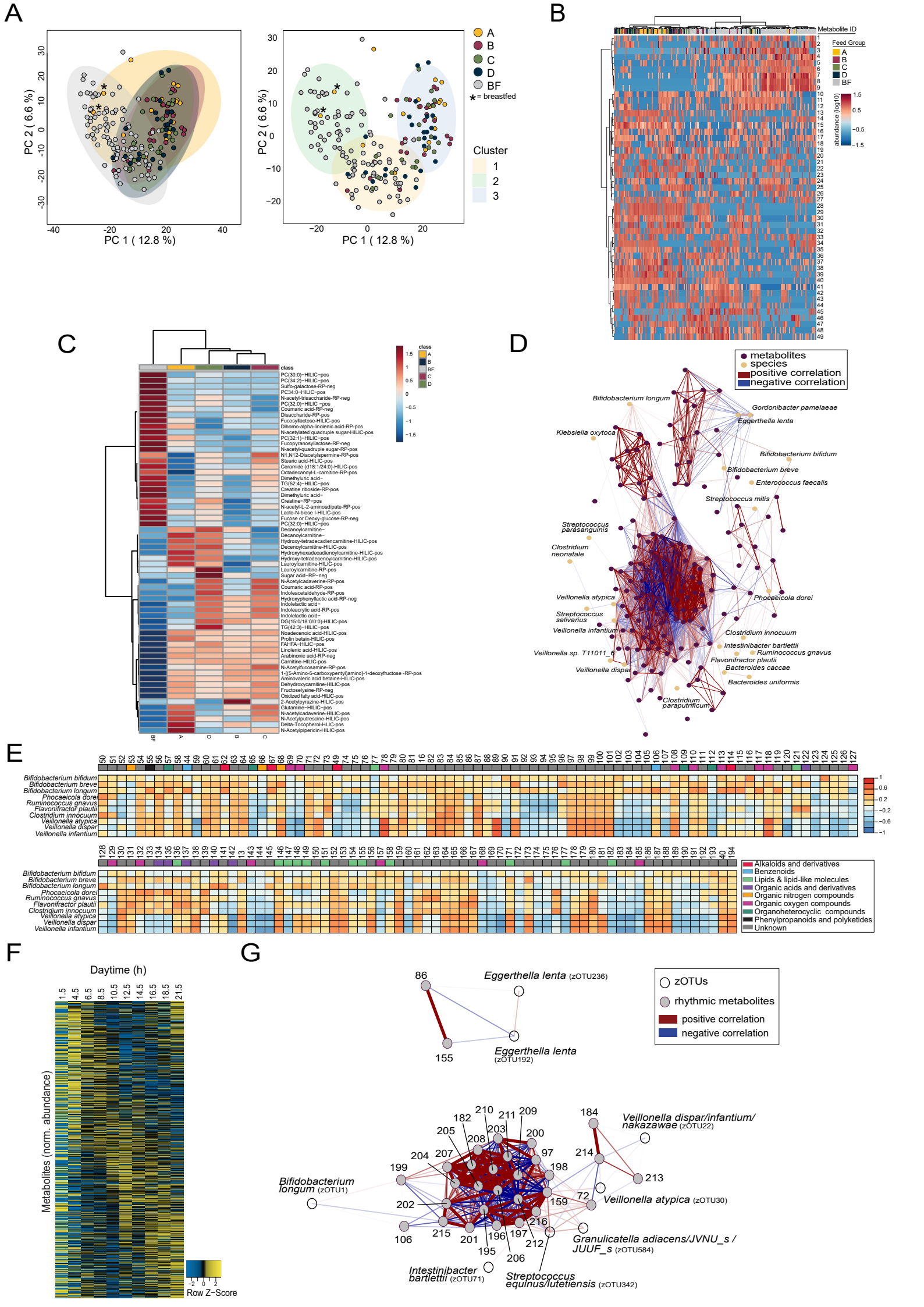

A
PC 2 (6.6 %)
PC 1 (12.8 %)
A
B
C
D
BF
* = breastfed
Cluster
1
2
3
B
Metabolite ID
Feed Group
A
B
C
D
BF
abundance (log10)
C
class
D
metabolites
species
positive correlation
negative correlation
Bifidobacterium longum
Gordonibacter pamelaeae
Eggerthella lenta
Klebsiella oxytoca
Bifidobacterium bifidum
Bifidobacterium breve
Enterococcus faecalis
Streptococcus mitis
Streptococcus parasanguinis
Clostridium neonatale
Veillonella atypica
Phocaeicola dorei
Streptococcus salivarius
Veillonella infantium
Clostridium innocuum
Intestinibacter bartlettii
Ruminococcus gnavus
Veillonella sp. T11011_6
Flavonifractor plautii
Bacteroides caccae
Veillonella dispar
Bacteroides uniformis
Clostridium paraputrificum
E
Bifidobacterium bifidum
Bifidobacterium breve
Bifidobacterium longum
Phocaeicola dorei
Ruminococcus gnavus
Flavonifractor plautii
Clostridium innocuum
Veillonella atypica
Veillonella dispar
Veillonella infantium
Alkaloids and derivatives
Benzenoids
Lipids & lipid-like molecules
Organic acids and derivatives
Organic nitrogen compounds
Organic oxygen compounds
Organoheterocyclic compounds
Phenylpropanoids and polyketides
Unknown
F
Daytime (h)
Metabolites (norm. abundance)
Row Z-Score
G
Eggerthella lenta (zOTU236)
Eggerthella lenta (zOTU192)
zOTUs
rhythmic metabolites
positive correlation
negative correlation
Veillonella dispar/infantium/ nakazawae (zOTU22)
Bifidobacterium longum (zOTU1)
Veillonella atypica (zOTU30)
Granulicatella adiacens/JVNU_s / JUUF_s (zOTU584)
Intestinibacter bartlettii (zOTU71)
Streptococcus equinus/lutetiensis (zOTU342)

Figure 4 Metabolites clearly differentiate feeding groups and display circadian changes

(A) Principal component analysis of infant metabolite profiles (3 months), stratified according to feed type (left panel) and unsupervised clustering (right panel). * = formula infants but breast-fed at 3-month time point.

(B) Intra-sample-normalized heatmap depicting the relative abundance of the top 49 metabolites found in infant fecal samples (3 months).

(C) Mean abundance of annotated fecal metabolites (3 months) displayed per feed group.

(D) Correlational network of fecal metabolites and species identified using shallow metagenomics sequencing.

(E) Heatmap of significantly correlated ($r > 0.5$) species (shallow metagenomic sequencing) and fecal metabolites, each metabolite is color-coded according to category affiliation to MetaCyc pathways.

(F) Heatmap of normalized abundance of fecal metabolites at time-of-day intervals. Metabolites are sorted by their peak abundance determined with a cosine-wave regression.

(G) Network correlational analysis of zOTUs and rhythmic metabolites. Thickness of connecting lines indicates strength of correlation.

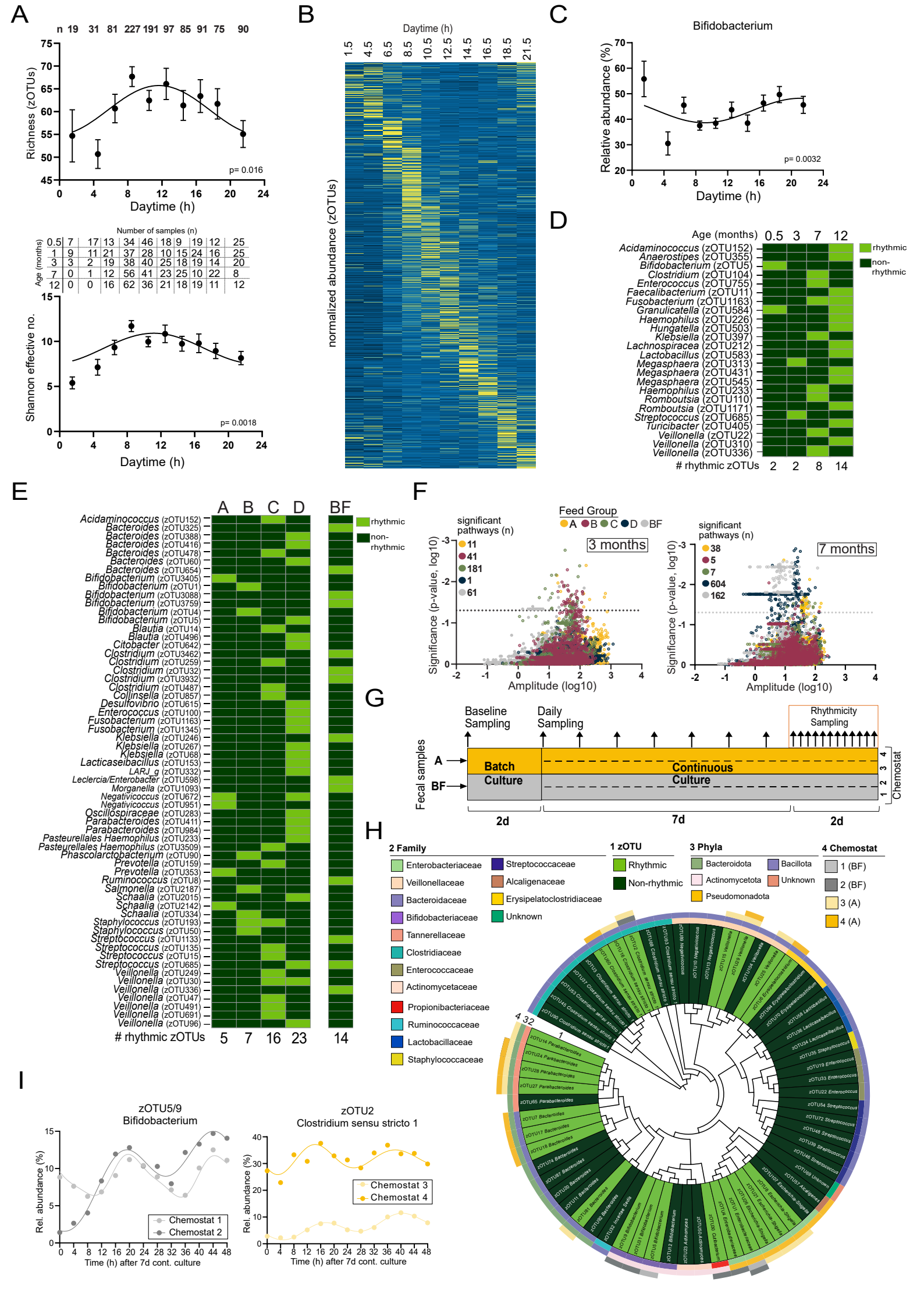

A
n 19 31 81 227 191 97 85 91 75 90
Richness (zOTUs)
p= 0.016
Daytime (h)
Number of samples (n)
Age (months)
0.5 7 17 13 34 46 18 9 19 12 25
1 9 11 21 37 28 10 15 24 16 25
3 3 2 19 38 40 25 18 19 14 20
7 0 1 12 56 41 23 25 10 22 8
12 0 0 16 62 36 21 18 19 11 12
Shannon effective no.
p= 0.0018
Daytime (h)
B
Daytime (h)
1.5 4.5 6.5 8.5 10.5 12.5 14.5 16.5 18.5 21.5
normalized abundance (zOTUs)
C
Bifidobacterium
Relative abundance (%)
p= 0.0032
Daytime (h)
D
Age (months) 0.5 3 7 12
Acidaminococcus (zOTU152)
Anaerostipes (zOTU355)
Bifidobacterium (zOTU5)
Clostridium (zOTU104)
Enterococcus (zOTU755)
Faecalibacterium (zOTU11)
Fusobacterium (zOTU1163)
Granulicatella (zOTU584)
Haemophilus (zOTU226)
Hungatella (zOTU503)
Klebsiella (zOTU397)
Lachnospiracea (zOTU212)
Lactobacillus (zOTU583)
Megasphaera (zOTU313)
Megasphaera (zOTU431)
Megasphaera (zOTU545)
Haemophilus (zOTU233)
Romboutsia (zOTU110)
Romboutsia (zOTU1171)
Streptococcus (zOTU685)
Turicibacter (zOTU405)
Veillonella (zOTU22)
Veillonella (zOTU310)
Veillonella (zOTU336)
rhythmic zOTUs 2 2 8 14
rhythmic
non-rhythmic
E
A B C D BF
Acidaminococcus (zOTU152)
Bacteroides (zOTU325)
Bacteroides (zOTU388)
Bacteroides (zOTU416)
Bacteroides (zOTU478)
Bacteroides (zOTU60)
Bacteroides (zOTU654)
Bifidobacterium (zOTU3405)
Bifidobacterium (zOTU1)
Bifidobacterium (zOTU3088)
Bifidobacterium (zOTU3759)
Bifidobacterium (zOTU4)
Bifidobacterium (zOTU5)
Blautia (zOTU14)
Blautia (zOTU496)
Citobacter (zOTU642)
Clostridium (zOTU3462)
Clostridium (zOTU259)
Clostridium (zOTU32)
Clostridium (zOTU3932)
Clostridium (zOTU487)
Collinsella (zOTU857)
Desulfovibrio (zOTU615)
Enterococcus (zOTU100)
Fusobacterium (zOTU1163)
Fusobacterium (zOTU1345)
Klebsiella (zOTU246)
Klebsiella (zOTU267)
Klebsiella (zOTU68)
Lacticaseibacillus (zOTU153)
LARJ_g (zOTU332)
Leclercia/Enterobacter (zOTU598)
Morganella (zOTU1093)
Negativicoccus (zOTU672)
Negativicoccus (zOTU951)
Oscillospiraceae (zOTU283)
Parabacteroides (zOTU411)
Parabacteroides (zOTU984)
Pasteurellales Haemophilus (zOTU233)
Pasteurellales Haemophilus (zOTU3509)
Phascolarctobacterium (zOTU90)
Prevotella (zOTU159)
Prevotella (zOTU353)
Ruminococcus (zOTU8)
Salmonella (zOTU2187)
Schaalia (zOTU2015)
Schaalia (zOTU2142)
Schaalia (zOTU334)
Staphylococcus (zOTU193)
Staphylococcus (zOTU50)
Streptococcus (zOTU1133)
Streptococcus (zOTU135)
Streptococcus (zOTU15)
Streptococcus (zOTU685)
Veillonella (zOTU249)
Veillonella (zOTU30)
Veillonella (zOTU336)
Veillonella (zOTU47)
Veillonella (zOTU491)
Veillonella (zOTU691)
Veillonella (zOTU96)
rhythmic zOTUs 5 7 16 23 14
rhythmic
non-rhythmic
F
Feed Group
A B C D BF
3 months
significant pathways (n)
11
41
181
1
61
Significance (p-value, log10)
Amplitude (log10)
7 months
significant pathways (n)
38
5
7
604
162
Significance (p-value, log10)
Amplitude (log10)
G
Baseline Sampling
Daily Sampling
Rhythmicity Sampling
Fecal samples
A
BF
Batch Culture
Continuous Culture
Chemostat
1 2 3 4
2d
7d
2d
H
2 Family
Enterobacteriaceae
Veillonellaceae
Bacteroidaceae
Bifidobacteriaceae
Tannerellaceae
Clostridiaceae
Enterococcaceae
Actinomycetaceae
Propionibacteriaceae
Ruminococcaceae
Lactobacillaceae
Staphylococcaceae
Streptococcaceae
Alcaligenaceae
Erysipelatoclostridiaceae
Unknown
1 zOTU
Rhythmic
Non-rhythmic
3 Phyla
Bacteroidota
Actinomycetota
Pseudomonadota
Bacillota
Unknown
4 Chemostat
1 (BF)
2 (BF)
3 (A)
4 (A)
I
zOTU5/9
Bifidobacterium
Rel. abundance (%)
Chemostat 1
Chemostat 2
Time (h) after 7d cont. culture
zOTU2
Clostridium sensu stricto 1
Rel. abundance (%)
Chemostat 3
Chemostat 4
Time (h) after 7d cont. culture

Figure 5 Developing daytime-related rhythmicity of the infant microbiota is affected by formula supplementation

(A) Diurnal profiles of alpha-diversity of samples with known time of defecation (n = 987), displayed as mean ± SEM. Number of total samples at each time point are shown above the graph, number of samples per time of day per age group are displayed in the table in graph.

(B) Heatmap of relative abundance of 777 zOTUs detected in infant samples (mean relative abundance > 1%, normalized to peak per zOTU) ordered by time of cosine-wave amplitude peak of the zOTUs.

(C) Diurnal profile of the genus *Bifidobacterium* across all infant samples.

(D) Rhythmic zOTUs at age 0.5, 3, 7 and 12 months with the total number of rhythmic zOTUs displayed at the bottom of the graph.

(E) Rhythmic zOTUs in the five feeding groups with the total number of rhythmic zOTUs displayed at the bottom of the graph.

(F) Quantification and significance of rhythmic and arrhythmic pathways in infant samples (3 and 7 months), separated by feed group. Significantly rhythmic pathways are located above the dashed line ($p < 0.05$).

(G) Schematic illustration of study design and sampling time points of chemostats 1-4 containing stool cultures obtained from breastfed (BF) and placebo formula (A) fed infants.

(H) Phylogenetic clustering of zOTUs identified in chemostat samples, displayed as a cladogram. Rhythmicity is indicated for the individuals zOTUS (ring 1), family (ring 2), phyla (ring 3) and in which chemostat the rhythmic zOTUs were found (ring 4).

(I) Circadian profiles of zOTUs from chemostat cultures after 7 days of continuous culture. Significant rhythms are illustrated with fitted cosine-wave curves (cosine-wave regression, $p \leq 0.05$). Data are represented as mean ± SEM.

Tables

Table 1 Cohort characteristics of enrolled infants

Values are given in median with interquartile range in parenthesis unless otherwise indicated.

	A		B		C		D		BF	
N	35	-	39	-	36	-	37	-	63	-
Samples (n)	119	-	136	-	125	-	135	-	483	-
Discontinue study (N)	4	-	2	-	1	-	2	-	2	-
Birth characteristics										
Gestational age (days)	279.0	(272.0-284.0)	277.0	(271.0-284.0)	279.5	(273.5-284.0)	278.0	(273.0-282.5)	278.0	(270.0-285.0)
Weight (g)	3450	(3100-3725)	3250	(2960-3640)	3410	(2903-3893)	3370	(3060-3585)	3360	(3015-3690)
Size (cm)	52.0	(50.0-53.0)	51.0	(49.0-53.0)	51.5	(50.0-53.0)	51.0	(50.0-52.5)	52.0	(50.0-53.0)
Cesaerean delivery, %	37.1	-	25.6	-	38.9	-	40.5	-	27.0	-
Female, %	51.4	-	61.5	-	30.6	-	58.3	-	50.8	-
Male, %	48.6	-	38.5	-	69.4	-	43.2	-	49.2	-
Twin, %	0	-	10.3	-	11.1	-	10.8	-	6.3	-
Maternal characteristics										
BMI pre pregnancy, kg/m2, mean	23.2		22.7		22.7		23.9		22.0	
Age at inclusion (years), mean	34.2		33.1		33.3		32.7		33.4	
Nuclear family characteristics										
Atopic disease, %	57.1	-	59.0	-	50.0	-	51.4	-	60.3	-
Autoimmune disease, %	8.6	-	2.6	-	5.6	-	8.1	-	11.1	-
Environment										
Animal contact, %	31.4	-	38.5	-	27.8	-	35.1	-	30.2	-
Distance to study center (km),mean	36.60		36.2		41.20		39.9		41.7	
Siblings in household, %	22.9	-	30.8	-	25	-	37.8	-	44.4	-
Sample characteristics										
Breastfeeding during first year, % (N)	88.6	-	87.2	-	91.7	-	83.8	-	100	-
Exclusive formula-feeding, % (N)	11.4	-	12.8	-	8.3	-	16.2	-	0	-
Daily formula intake (ml)	480.0	(200.0-720.0)	500	(242.5-700.0)	400.0	(220.0-750.0)	505	(217.5-657.5)	0	0
Age start solid food (months)	4.9	(4.4-5.8)	4.9	(4.2-5.4)	5	(4.6-5.8)	4.6	(4.0-6.1)	5.5	(5.0-6.1)

STAR Methods

Ethics

The study was carried out in accordance with the Declaration of Helsinki and was approved by the Ethics committee of the Technical University of Munich under the study number 254/17S and included the written informed consent of the parents or legal guardians of the participating infants. The study is registered at the German Clinical Trials Registry (DRKS00012313).

Study participants

Healthy women in their third trimester of pregnancy were recruited in the greater Munich, Germany area and eligibility for participation in the study was assessed. Exclusion criteria were premature birth (gestational age being less than 37 weeks), physical deformities that affect feeding (such as cleft lip and palate) and serious medical conditions that required an intensive care stay exceeding 3 days. Based on current healthcare guidelines in Germany, parents were explicitly encouraged to breastfeed. For participation in the study, parents agreed to use the provided study infant formula in case formula feeding became necessary.

Controlled intervention trial design

The Infantibio-II study was a double-blind, randomized and placebo-controlled interventional trial. Infant fecal samples were collected at age 0.5, 1, 3, 7 and 12 months with an additional, voluntary follow-up sample at age 2 years. Infants were randomized to one of four formula groups. The letters A-D were used for blinding of the different formula groups. Formula group A (placebo) contained no additional supplements. Formula group B contained two bacterial isolates from native fecal samples of infants[10] belonging to the species *Bifidobacterium longum* subsp. *infanti*s and *B. breve*. Lactosan International GmbH & Co. used freeze-dried bacterial powder to supplement infant formula with 10^7 cfu/g of each strain. Formula group C contained a syrup-based galacto-oligosaccharide (GOS) mixture (5g/L, FrieslandCampina). Formula group D contained both *Bifidobacterium* strains and GOS. Parents were free to decide if, when and how much formula was fed to the infants in addition to or after cessation of breastfeeding. The reference group were infants that were fully breastfed for the entirety of the study period with the exception of solid food in addition to breastmilk.

Infant formula

Infant formula was manufactured by Töpfer GmbH (Dietmannsried, Germany) and provided blinded to the study team for distribution to participants. The four formulas could not be distinguished from one another by looks, smell or other characteristics. Unblinding of formula groups took place after study completion. The study formula was a powdered organic cows-milk based infant formula containing all the necessary nutrients based on laws and regulations at the time of manufacturing (2018-2020). Parents received detailed instructions on how to prepare drink-ready formula including a digital thermometer for correct preparation.

Fecal sample collection

Parents collected fecal samples from the diapers of participating infants. Stool sampling kits were provided including collection tubes with DNA stabilizing solution (Invitek Moleculuar, 1038111300) and empty fecal collection tubes (Süsse Labortechnik, H8555T). Two spoons of fecal matter were added to the tubes and immediately refrigerated until pick-up by study team, latest 24 hours after defecation. Samples were aliquoted and stored at -80°C at the study center at the Technical University Munich in Freising, Germany (ZIEL- Institute for Food & Health). Case report forms detailing birth characteristics, familial health status, environmental factors

and the current feeding habits of the infant were completed at enrollment of study and with every stool sample collection.

16S rRNA gene amplicon sequencing and analysis

Metagenomic DNA was isolated from 600 µl fecal matter in DNA stabilizing solution using a previously described DNA-isolation protocol.[50] Negative controls in form of DNA extracted from water and DNA extracted from stool stabilizer (without fecal matter) were included every 45 samples. DNA extraction was followed by two-step polymerase chain reaction (PCR) amplification and sequencing on Illumina HiSeq in paired-end mode (2x 250 bp) using the HiSeq Rapid Kit v2 from Illumina.[50]

The UPARSE-based[51] platform "Integrated Microbial Next Generation Sequencing" (IMNGS)[52] was adapted as an in-house offline version (NGSToolkit, v6.0.2-beta.5) for processing the sequencing output. Reads were trimmed to the first base. USEARCH 8.0[53] was used for paring of samples with 1 allowed mismatch per barcode. Remaining reads were trimmed (5 nucelotides each) followed by denoising and clustering by 100% identity, creating zero-radius operational taxonomic units (zOTUS) for optimal resolution of microbial strains in 16S rRNA sequencing. Chimera sequences were removed using UCHIME.[54] Following best practices, zOTUs with a relative abundance below 0.25% were removed.[55] Taxonomic labels were assigned by SINA 1.6.1 with the SILVA database (release 128).[56] The sequences of a subset of zOTUs of interest were identified using the EzBioCloud database.[57] Maximum-likelihood phylogeny calculations were done with FastTree[58] and visualized with EvolView.[59] Longitudinal changes in microbial abundances were visualized using horizon plots generated with the BiomeHorizon R package (v1.0.0).[60]

RHEA was used for microbial profile analysis.[61] Statistical analysis was performed in GraphPad Prism, version 9.4.1 (GraphPad Software, LLC) or in R (2022.02.3+492). Normalization (no random subsampling, no rounding; cutoff was set to 10000 reads for 1st year samples, 20000 reads in case of the fermenter samples and infant samples including the 24-month samples. The different thresholds were chosen in relation to the number of reads of each dataset.) and analysis such as alpha- and beta-diversity as well as taxonomic binning were performed with Rhea. Correction for multiple testing was done with Benjamini-Hochberg procedure. For comparison between more than two groups, a one-way ANOVA was used for normally distributed data (Shapiro Wilk test, alpha = 0.05). Otherwise, we used a Kruskal-Wallis test with Dunn's procedure as post-hoc multiple comparison test. Stars indicate significance level: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. The "NBClust (v3.0)" package in R was used for hierarchical clustering of samples. The optimal number of clusters was determined based on the Calinski and Harabasz index[62] and Ward's hierarchical clustering criterion.[63] Differences in microbial composition were calculated using a PERMANOVA test on generalized UniFrac distances, where the length of the bar in the graph provides context for dissimilarities of samples.

PICRUSt2 (v2.4.2)[64] was used for functional analysis on all zOTUs identified in infant samples. zOTU sequences were used for predicting Metacyc pathway abundances (Caspi, Billington et al. 2018) without considering super-classes. The estimated Metacyc pathway abundances were used as a basis for linear discriminant analysis (LDA) effect size (LEfSe) calculations (alpha: 0.01) comparing formula group A to the other feed groups.[65]

Diurnal and circadian analysis of infant microbiota and *in vitro* samples

To identify diurnal rhythms, a cosine-wave equation with a fixed 24-h period:

y = baseline + (amplitude*cos(2*π*((x-[phase shift)/24))) was fitted on alpha-diversity, taxa level and relative abundance of individual zOTUs. Significance was determined with an F-Test and goodness of fit was corrected for multiple testing. Significant rhythmicity was assumed to occur when $p \leq 0.05$. An adjusted version of the "CompareRhythms" script[66] created for circadian rhythm comparison between groups, which performs JTK_cycle[67] followed by DODR.[68] Heatmaps were created using the heatmap script in R. The amplitude shown in the Manhattan plot is based on JTK_cycle and the phase is calculated using cosine-wave regression. Circadian cosine-fitting for *in vitro* sample was performed using an in-house code by Isaiah Ting based on.[69]

Shallow metagenomics sequencing

To further enhance the identification and the understanding of the role of individual bacterial strains, shallow shotgun metagenomic sequencing was performed using isolated gDNA on a NovaSeq machine (1G raw data with paired-end 150 bp) on a subset of paired infant samples at time points month 3 and month 7. The resulting metagenomic data was demultiplexed and from the raw sequencing reads adapters were trimmed using "Trim Galore" (developed by F. Krueger, The Babraham Institute: https://github.com/FelixKrueger/TrimGalore, v.0.6.7). Quality control of sequencing reads (removal of low-quality reads and contaminants) was done using the pipeline "Knead data" (https://github.com/biobakery/kneaddata, v.0.7.7-alpha). Species-level taxonomic annotation was done using MetaPhlan 3.0.13 and HUMAnN 3.0.[70] Gene families that were detected via HUMAnN 3.0 were mapped to the KEGG pathway orthologs.[71] Centered log ratio scaling was applied.

Untargeted metabolomics

Sample preparation: 100 mg native infant fecal matter was added to a 15 ml tube (CKMix 50, Bertin Technologies) containing ceramic beads (mixture of 2.8 mm and 5.0 mm beads). Then, 5 ml of methanol-based dehydrocholic acid extraction solvent (c = 1.3 µmol/L) was added and the sample was homogenized using a bead beater system (Precellys Evolution, Bertin Technologies) with a cooling adapter containing liquid nitrogen (Bertin Technologies). The bead beating procedure was performed three times per sample for 20 seconds at 10 000 rpm with a 15 second break between each round of bead beating. The sample was then centrifuged at 8 000 rpm for 10 minutes at 10°C. 500 ml of clear supernatant was transferred to a LC-MS/MS vial for untargeted metabolomics analysis. Multiple smaller volumes of samples were pooled to prepare a QC-sample.

Untargeted mass spectrometric measurement**:** analysis was performed using a Nexera UHPLC system (Shimadzu, Duisburg, Germany) coupled to a Q-TOF mass spectrometer (TripleTOF 6600 AB Sciex, Darmstadt, Germany). Separation of the stool samples was performed either using a UPLC BEH Amide 2.1 × 100 mm, 1.7 µm analytic column (Waters, Eschborn, Germany) with a 400 µL/min flow rate or with a Kinetex XB18 2.1 x 100 mm, 1.7 µm (Phenomenex, Aschaffenburg, Germany) with a 300 µL/min flow rate. For the HILIC-separation the settings were as follows. The mobile phase was 5 mM ammonium acetate in water (eluent A) and 5 mM ammonium acetate in acetonitrile/water (95/5, v/v) (eluent B). The gradient profile was 100% B from 0 to 1.5 min, 60% B at 8 min and 20% B at 10 min to 11.5 min and 100% B at 12 to 15 min. For the reversed-phase separation eluent A was 0.1% formic acid and eluent B was 0.1% formic acid in acetonitrile. The gradient profile started with 0.2% B which was held for 0.5 min. Afterwards the concentration of eluent B was increased to 100% until 10 min which was held for 3.25 min. Afterward the column was equilibrated at starting conditions. A volume of 5 µL per sample was injected. The autosampler was cooled to 10 °C

and the column oven heated to 40 °C. Every tenth run a quality control (QC) sample which was pooled from all samples was injected. The samples were measured in a randomized order and in the Information Dependent Acquisition (IDA) mode. MS settings in the positive mode were as follows: Gas 1 55, Gas 2 65, Curtain gas 35, Temperature 500 °C, Ion Spray Voltage 5500, declustering potential 80. The mass range of the TOF MS and MS/MS scans were 50–2000 m/z and the collision energy was ramped from 15–55 V. MS settings in the negative mode were as follows: Gas 1 55, Gas 2 65, Cur 35, Temperature 500 °C, Ion Spray Voltage –4500, declustering potential –80. The mass range of the TOF MS and MS/MS scans were 50–2000 m/z and the collision energy was ramped from –15–55 V.

Data processing: The "msconvert" from ProteoWizard72 was used to convert raw files to mzXML (de-noised by centroid peaks). The bioconductor/R package xcms73 was used for data processing and feature identification. More specifically, the matched filter algorithm was used to identify peaks (full width at half maximum set to 7.5 seconds). Then the peaks were grouped into features using the "peak density" method.73 The area under the peak was integrated to represent the abundance of features. The retention time was adjusted based on the peak groups present in most samples. To annotate features with names of metabolites, the exact mass and MS2 fragmentation pattern of the measured features were compared to the records in HMBD74 and the public MS/MS spectra in MSDIAL,75 referred to as MS1 and MS2 annotation, respectively. Missing values were imputed with half of the limit of detection method, i.e., for every feature, the missing values were replaced with half of the minimal measured value of that feature in all measurements. To confirm a MS2 spectra is well annotated, we manually reviewed our MS2 fragmentation pattern and compared it with records in the public database, using SIRIUS76 or previously measured reference standards to evaluate the correctness of the annotation. The mass spectrometric data is made publically available using the massive repository (https://massive.ucsd.edu) under the accession number MassIVE MSV000093140.

.

Metabolite analysis

To obtain information about potential associations between metabolites and species (data from shallow metagenomic sequencing) or zOTUs (data from 16S sequencing), network analysis was performed. For this, a correlation matrix was calculated. A spearman correlation was calculated and only features with a significant correlation and an absolute correlation value above 0.5 were selected. For plotting of the network, the R qgraph version 1.9 was used. Therefore payout spring was used, with a minimum absolut edge weight of 0.5.

Data from positive, negative and reverse phase were merged, i.e. some of the metabolites might appear more than once (due to being detected in different phases). To reduce the number of metabolites according to the post prevalent metabolites, the input tables were ordered by decreasing sum over all individuals. Afterwards a spearman correlation was calculated and used for further analysis. For 4F a subset of significantly different species was selected. Based on this subset a correlation matrix was calculated. In the heatmaps, only significant and high correlated (absolute correlation coefficient above 0.5) species - metabolites correlations are shown.

***Ex vivo* Infant Colonic Continuous Fermentation**

Infant Stool Samples Selection, Preparation, and Inoculation**:** Fecal samples of two 3-month old infants, one exclusively breast-fed and one exclusively fed with formula A, were selected to be representative of their feeding group based on Principal Component Analysis (PCA) and

Bifidobacterium abundance. Fecal slurries were prepared under anaerobic conditions by diluting each sample (previously thawed on ice and weighted: BF = 1.03 g, A = 1.98 g) in reduced PBS (3% L-Cys HCl). Each slurry was filtered (70 µm cell strainer) and an equal amount was loaded in two syringes to be injected in duplicate chemostat vessels (BF = 2x6.5mL, A = 2x8mL). The whole procedure was finished ~1.5 hours after the samples were thawed.

Preparation of the Culture Medium: A modified culture medium was prepared to mimic the physiological conditions of the infant colon. The media composition was based on the composition and calculations described in literature[77] adapted to the composition of the infant formula A. A fat content of 10% was chosen for this experiment,[78] and a full nutrient ratio of carbohydrate:nitrogen:fats after digestion with 13 g/L of carbohydrates was adjusted to 44:46:10 (corresponding to 11.88 g/L of fat). Therefore, the total concentration of undigested compounds in the medium were 2.96 g/L of total fat and 13.72 g/L of undigested N-compounds. The full composition of the culture medium, in g/L, was as follows (all the medium components were provided by Sigma-Aldrich, unless specified otherwise): *base media* (NaCl: 4.47; KCl: 4.43; $MgSO_4$: 1.24; $CaCl_2$: 0.1; $NaHCO_3$: 1.5; KH_2PO_4: 0.5; $FeSO_4$: 0.005; Bile salts: 0.05; Tryptone: 0.5; Peptone: 0.5; Yeast Extract: 2.5; Tween 80: 1; L-carnitine: 0.0015); *mucin media* (porcine gastric mucin type II: 4); and *milk media* (whey protein hydrolysate (Myprotein, UK): 11.21; casein: 0.47; pre infant formula A: 11.88). Each solution was sterilized by autoclaving at 121ºC for 15 min (110ºC for *milk media* after adjusting pH to 8.0[79] and then mixed thoroughly. Two solutions were added after sterile filtration (0.2 µm): the first containing 0.8 g/L of L-cysteine, 5.86 g/L of lactose, and 0.004 g/L of inositol and the second was a vitamin solution as prepared by[80] which also included hemin (0.01 g/L).

Ex vivo gut chemostat: The two infant fecal samples were cultured in duplicate chemostats (Bioreactor Model Multifors 2, Infors HT, Switzerland) according to the illustrated experimental scheme (**Fig. 5G**). In the first stage of the *ex vivo* gut chemostat model, a batch culture (total 48 h) was employed to allow for sample adaptation after inoculation. After 24 h without adjustment of pH or flow of fresh culture medium, the culture volume in the bioreactors were doubled and maintained at 800 mL until the end of the batch culture. In the second stage of the gut chemostat, a continuous culture was established for 7 days applying a retention time of 12.5 h.[77] Samples were taken once each day at the same time (1.5 mL for DNA extraction and 16S rRNA Sequencing). Finally, the third stage were 48 hours of rhythmicity sampling during which samples were taken every 4 h. The following conditions were applied during the 9 days of continuous culturing: 400 mL of volume in the bioreactors, feed rate of 32 ml/h/chemostat, pH was set to 6.5 and then maintained (pH electrodes, Hamilton, Germany). Temperature was kept at 37ºC, anaerobic conditions were controlled by supplying a continuous flow of N_2 and redox values were measured (Redox electrodes, 1 per chemostat, Hamilton, Germany).

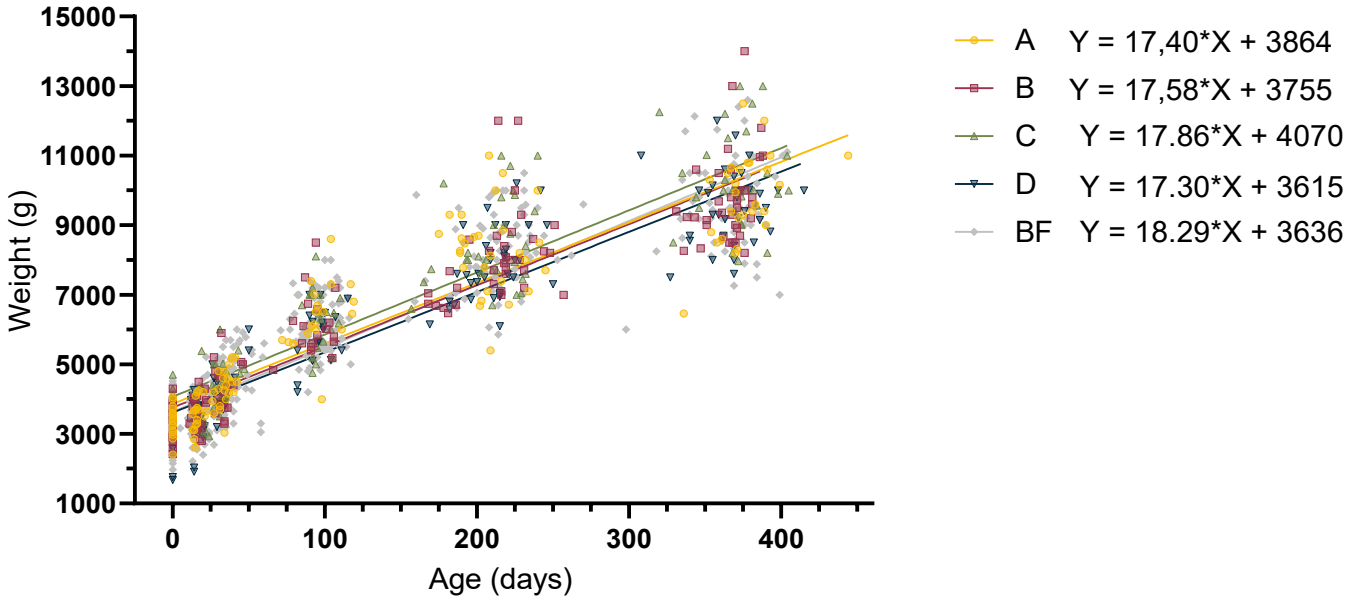
References STAR Methods

10. Bazanella, M., Maier, T.V., Clavel, T., Lagkouvardos, I., Lucio, M., Maldonado-Gòmez, M.X., Autran, C., Walter, J., Bode, L., and Schmitt-Kopplin, P. (2017). Randomized controlled trial on the impact of early-life intervention with bifidobacteria on the healthy infant fecal microbiota and metabolome. The American journal of clinical nutrition *106*, 1274-1286.
50. Reitmeier, S., Kiessling, S., Neuhaus, K., and Haller, D. (2020). Comparing circadian rhythmicity in the human gut microbiome. STAR protocols *1*, 100148.
51. Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature methods *10*, 996-998.
52. Lagkouvardos, I., Joseph, D., Kapfhammer, M., Giritli, S., Horn, M., Haller, D., and Clavel, T. (2016). IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. Scientific reports *6*, 33721.
53. Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics *26*, 2460-2461. 10.1093/bioinformatics/btq461.
54. Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics *27*, 2194-2200.
55. Reitmeier, S., Hitch, T.C., Fikas, N., Hausmann, B., Ramer-Tait, A.E., Neuhaus, K., Berry, D., Haller, D., Lagkouvardos, I., and Clavel, T. (2020). Handling of spurious sequences affects the outcome of high-throughput 16S rRNA gene amplicon profiling.
56. Pruesse, E., Peplies, J., and Glöckner, F.O. (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics *28*, 1823-1829.
57. Yoon, S.-H., Ha, S.-M., Kwon, S., Lim, J., Kim, Y., Seo, H., and Chun, J. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. International journal of systematic and evolutionary microbiology *67*, 1613.
58. Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. PloS one *5*, e9490.
59. Subramanian, B., Gao, S., Lercher, M.J., Hu, S., and Chen, W.-H. (2019). Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. Nucleic acids research *47*, W270-W275.
60. Fink, I., Abdill, R.J., Blekhman, R., and Grieneisen, L. (2022). BiomeHorizon: visualizing microbiome time series data in R. Msystems *7*, e01380-01321.
61. Lagkouvardos, I., Fischer, S., Kumar, N., and Clavel, T. (2017). Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. PeerJ *5*, e2836.
62. Caliński, T., and Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics-theory and Methods *3*, 1-27.
63. Murtagh, F., and Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? Journal of classification *31*, 274-295.
64. Douglas, G.M., Maffei, V.J., Zaneveld, J.R., Yurgel, S.N., Brown, J.R., Taylor, C.M., Huttenhower, C., and Langille, M.G. (2020). PICRUSt2 for prediction of metagenome functions. Nature biotechnology *38*, 685-688.
65. Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. Genome biology *12*, 1-18.
66. Pelikan, A., Herzel, H., Kramer, A., and Ananthasubramaniam, B. (2022). Venn diagram analysis overestimates the extent of circadian rhythm reprogramming. The FEBS journal *289*, 6605-6621.
67. Hughes, M.E., Hogenesch, J.B., and Kornacker, K. (2010). JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. Journal of biological rhythms *25*, 372-380.
68. Thaben, P.F., and Westermark, P.O. (2016). Differential rhythmicity: detecting altered rhythmicity in biological data. Bioinformatics *32*, 2800-2808.
69. van der Veen, D.R., and Gerkema, M.P. (2017). Unmasking ultradian rhythms in gene expression. FASEB J *31*, 743-750. 10.1096/fj.201600872R.
70. Beghini, F., McIver, L.J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., and Thomas, A.M. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. elife *10*, e65088.

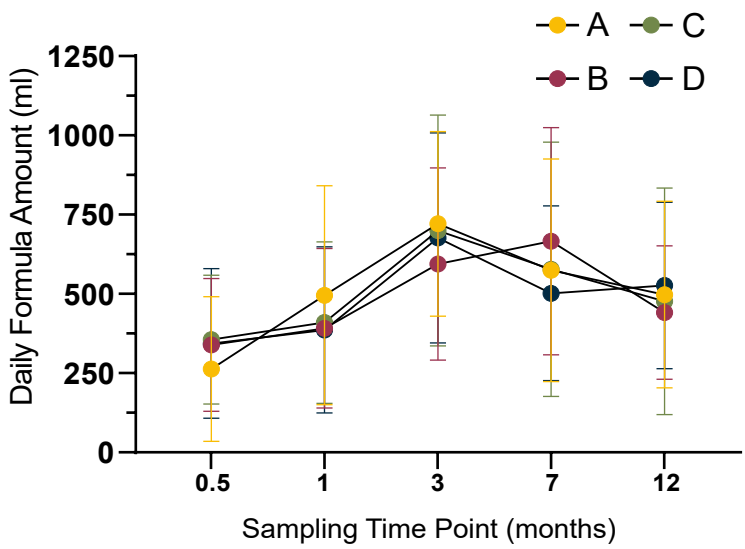
71. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research *28*, 27-30.
72. Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008). ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics *24*, 2534-2536.
73. Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical chemistry *78*, 779-787.
74. Wishart, D.S., Feunang, Y.D., Marcu, A., Guo, A.C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., and Karu, N. (2018). HMDB 4.0: the human metabolome database for 2018. Nucleic acids research *46*, D608-D617.
75. Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., and Arita, M. (2015). MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nature methods *12*, 523-526.
76. Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A.A., Melnik, A.V., Meusel, M., Dorrestein, P.C., Rousu, J., and Böcker, S. (2019). SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. Nature methods *16*, 299-302.
77. Cinquin, C., Le Blay, G., Fliss, I., and Lacroix, C. (2004). Immobilization of infant fecal microbiota and utilization in an in vitro colonic fermentation model. Microbial ecology *48*, 128-138.
78. Wieczorek-Filipiak, M., Drzymala-Czyz, S., Szczepanik, M., Szydlowski, J., Walkowiak, D., Nowak, J.K., Moczko, J.A., and Walkowiak, J. (2019). Fecal fat concentration and excretion in the first 2 years of life: a cross-sectional study. Journal of Pediatric Gastroenterology and Nutrition *68*, 285-289.
79. Pelegrine, D.H.G., and Gomes, M.T.M.S. (2012). Analysis of Whey Proteins Solubility at High Temperatures. International Journal of Food Engineering *8*. doi:10.1515/1556-3758.1265.
80. Gibson, G.R., and Wang, X. (1994). Enrichment of bifidobacteria from human gut contents by oligofructose using continuous culture. FEMS microbiology letters *118*, 121-127.

Supplemental Figure 1

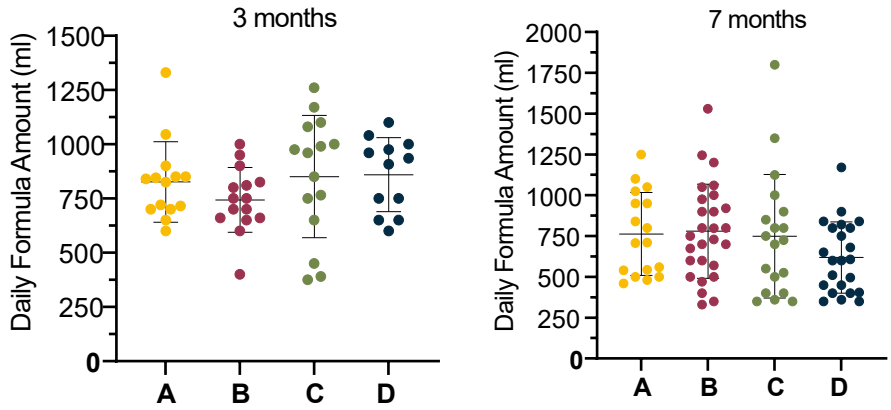
A



B



C



Supplemental Figure 1: Intervention groups and breastfeeders are comparable with regards to weight gain and formula consumption

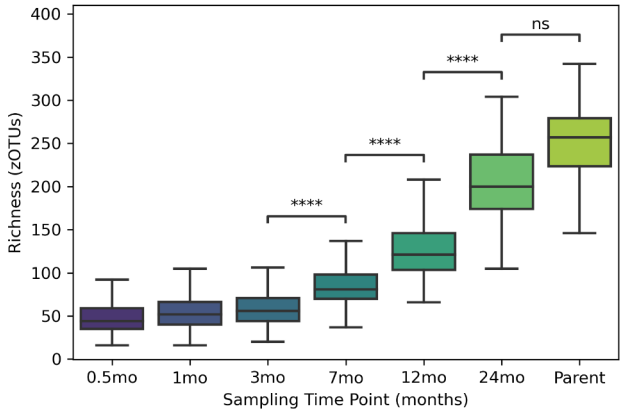
(A) Longitudinal weight gain of study participants in the different feeding groups (linear regression, slopes are not significantly different).

(B) Daily amount of formula (ml) consumed (mean with SD) per interventional group at different sampling time points.

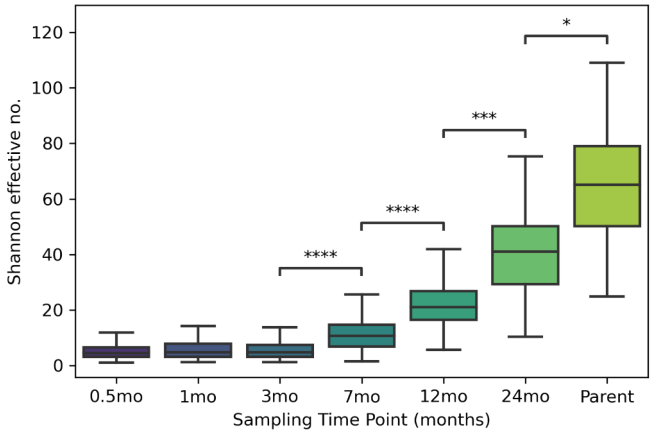
(C) Formula amount consumption (ml) of infants where minimal formula consumers were omitted (lowest quartile based on daily formula volume consumption) at age 3 and 7 months (mean, SD).

Supplemental Figure 2

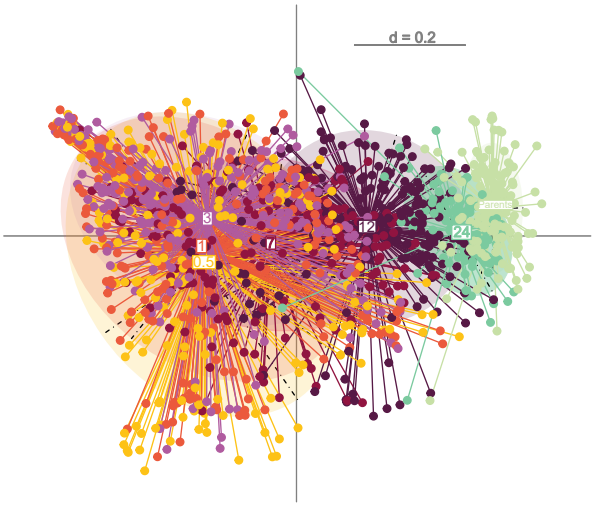
A



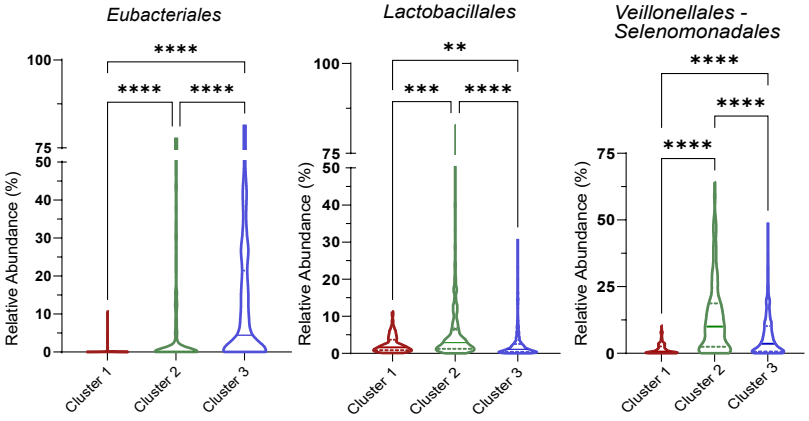
B



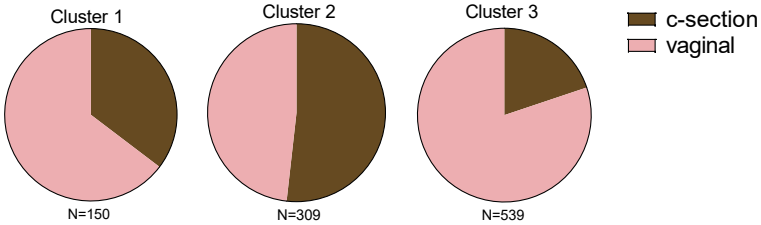
C



D



E



Supplemental Figure 2: Infant fecal microbiota has further moved toward an adult microbiota at follow-up sampling age 24 months

(A) Richness (zOTUs) at all infant sampling time points (including 24 months) and for parental samples (**** = $p \leq 0.0001$, 25th to the 75th percentile box plot, median: solid line, whiskers: min to max).

(B) Alpha-diversity (Shannon effective number) at all infant sampling time points (including 24 months) and for parental samples (**** = $p \leq 0.0001$, 25th to the 75th percentile box plot, median: solid line, whiskers: min to max).

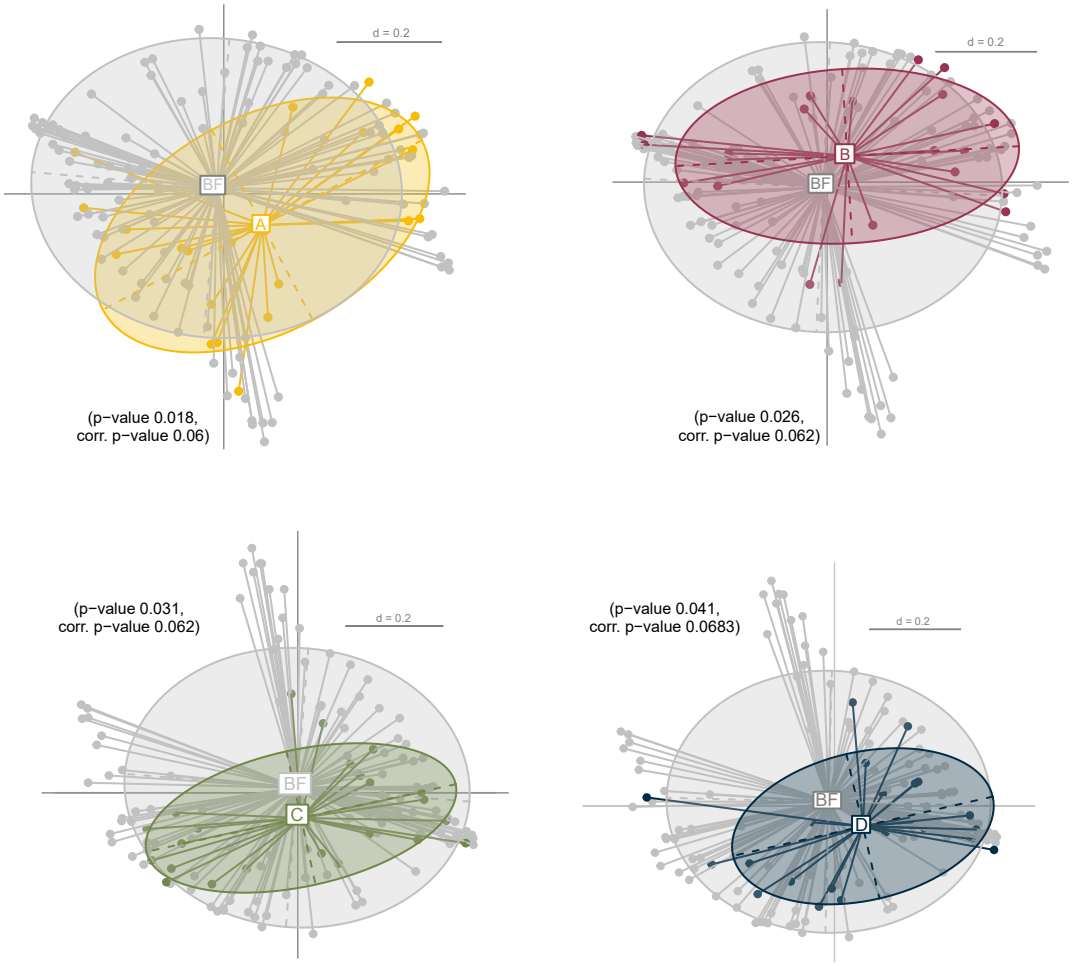
(C) Nonmetric multidimensional scaling (NMDS) plot of fecal samples grouped according to infant age (including 24-month samples) or parent sample. 0.5 months-1 month: corr.p-value 0.057, 1 month-3 months: corr. p-value 0.0011, 3 months-7 months: corr. p-value 0.0011, 7 months-12 months: corr. p-value 0.0011, 12 months-24 months: corr. p-value 0.0011, 24 months-Parent: corr. p-value 0.0011.

(D) Top 4-6 relative abundance (%) cluster dominating bacterial order violin plot (**** = $p \leq 0.0001$, median: solid line, 25th and 75th percentile: dashed lines), de novo generated clusters.

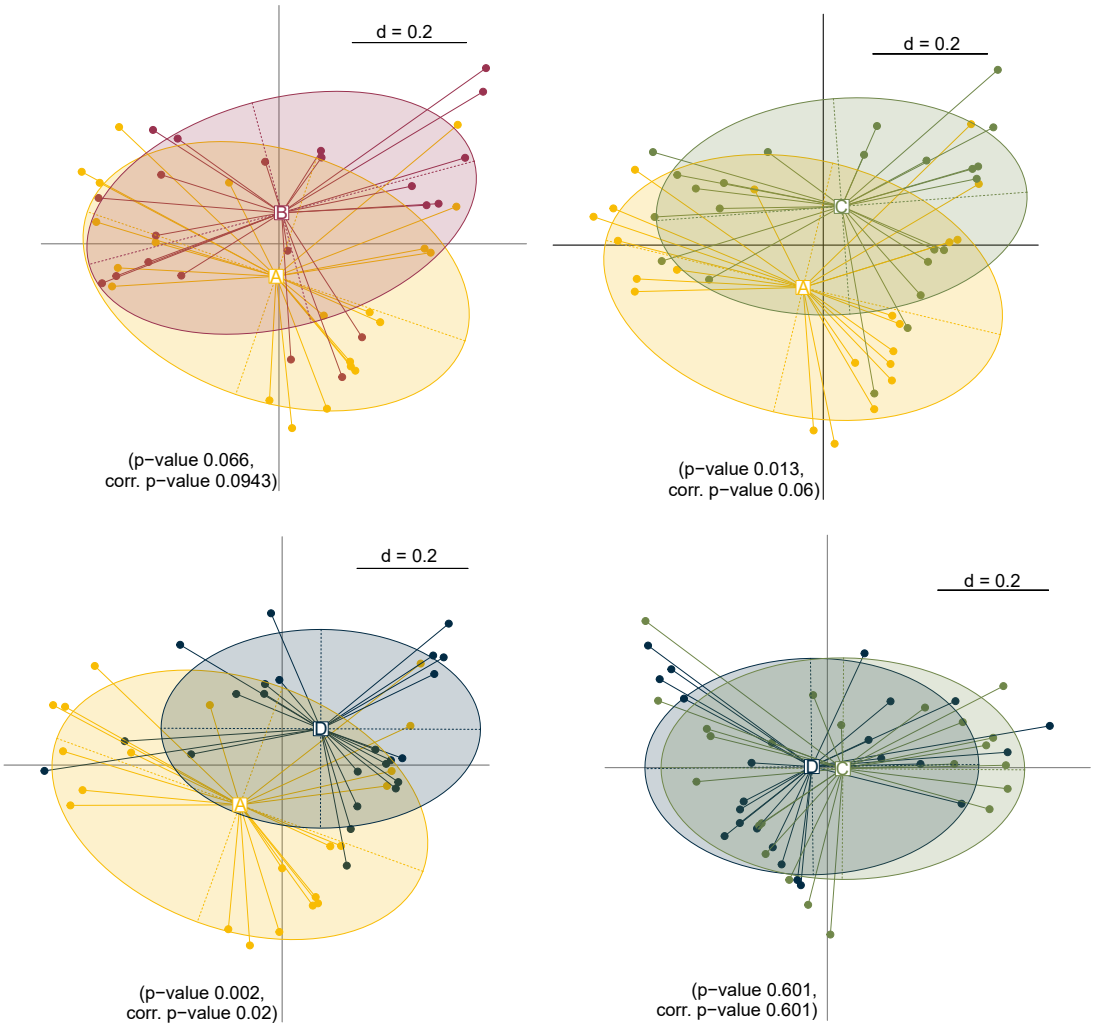
(E) Birth mode (c-section vs. vaginal) distribution in the three clusters.

Supplemental Figure 3

A



B

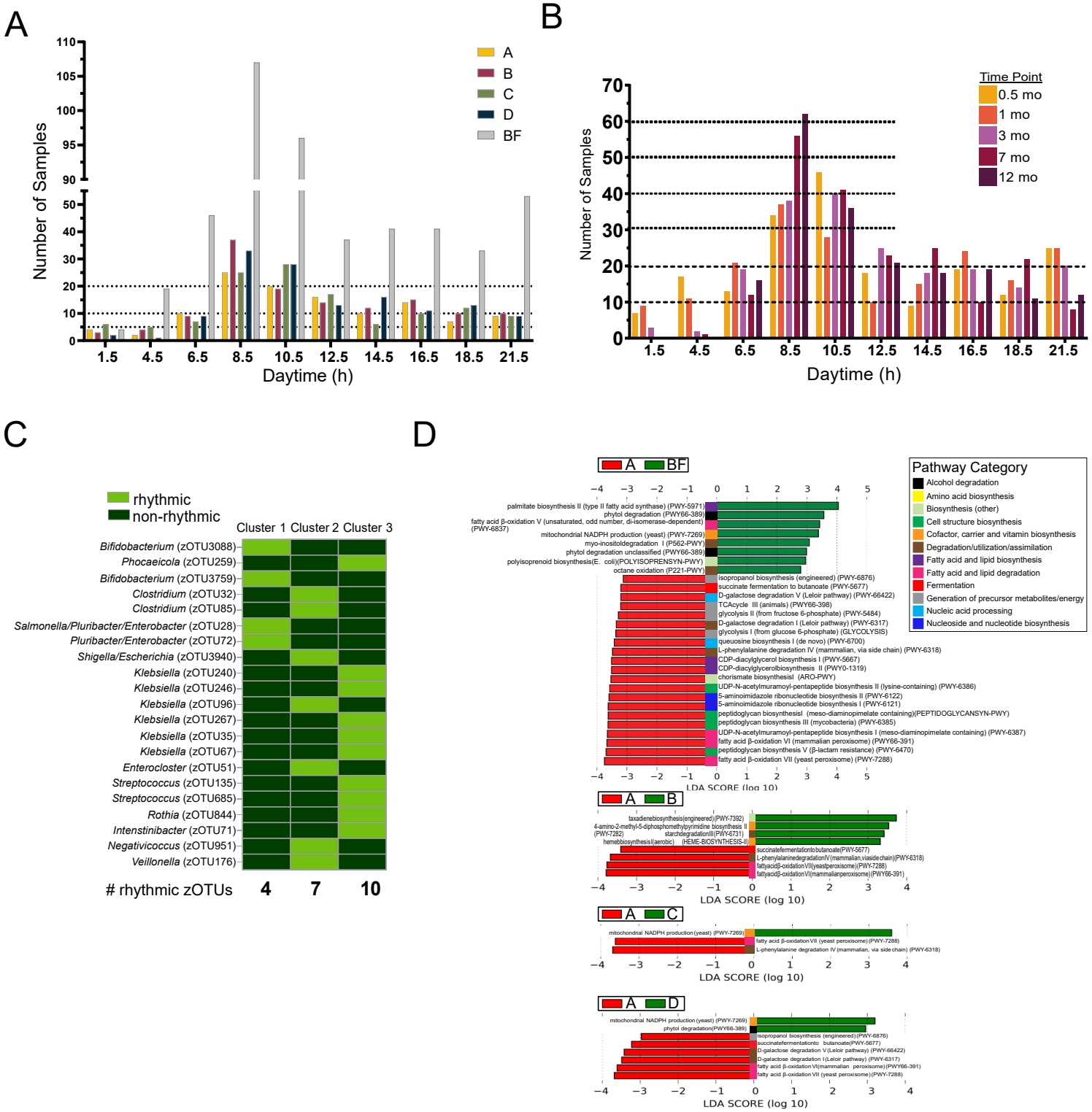


Supplemental Figure 3: Beta diversity at 3 months of age significantly differs between placebo formula and formula containing Bifidobacteria and GOS but does not reach significance in comparing all other groups

(A) MDS plots of beta-diversity comparison between breastfeeders and formula groups at 3 months (BF-A: corr. p-value 0.6, BF-B: corr. p-value 0.062, BF-C: corr. p-value 0.062, corr. p-value 0.0683).

(B) MDS plots of beta-diversity comparison between placebo formula A and other groups at 3 months (A-B: corr. p-value 0.0943, A-C: corr. p-value 0.06, A-D: corr. p-value 0.02) and comparison between A and D (corr. p-value: 0.601).

Supplemental Figure 4



Supplemental Figure 4: Differences in daytime sampling and rhythmicity in unsupervised clusters and between feed groups

(A) Number of fecal samples at different times of day (hours) per feed group.

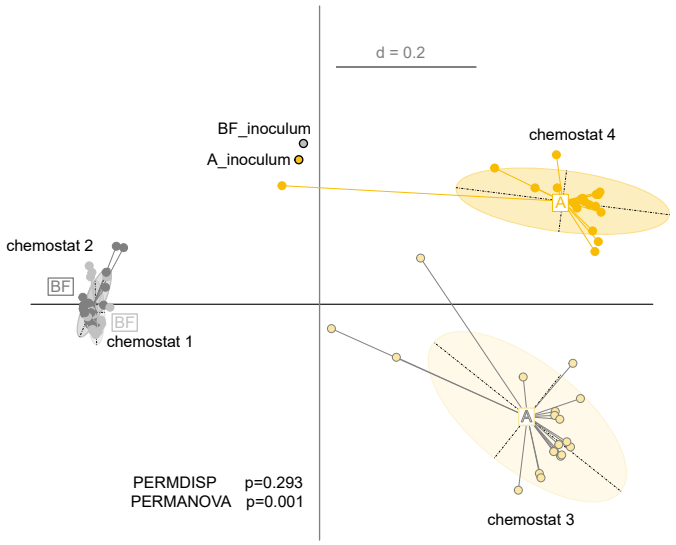
(B) Number of fecal samples at different times of day (hours) per sampling time point.

(C) Rhythmic zOTUs in the three de novo clusters with the total number of rhythmic zOTUs displayed at the bottom of the graph.

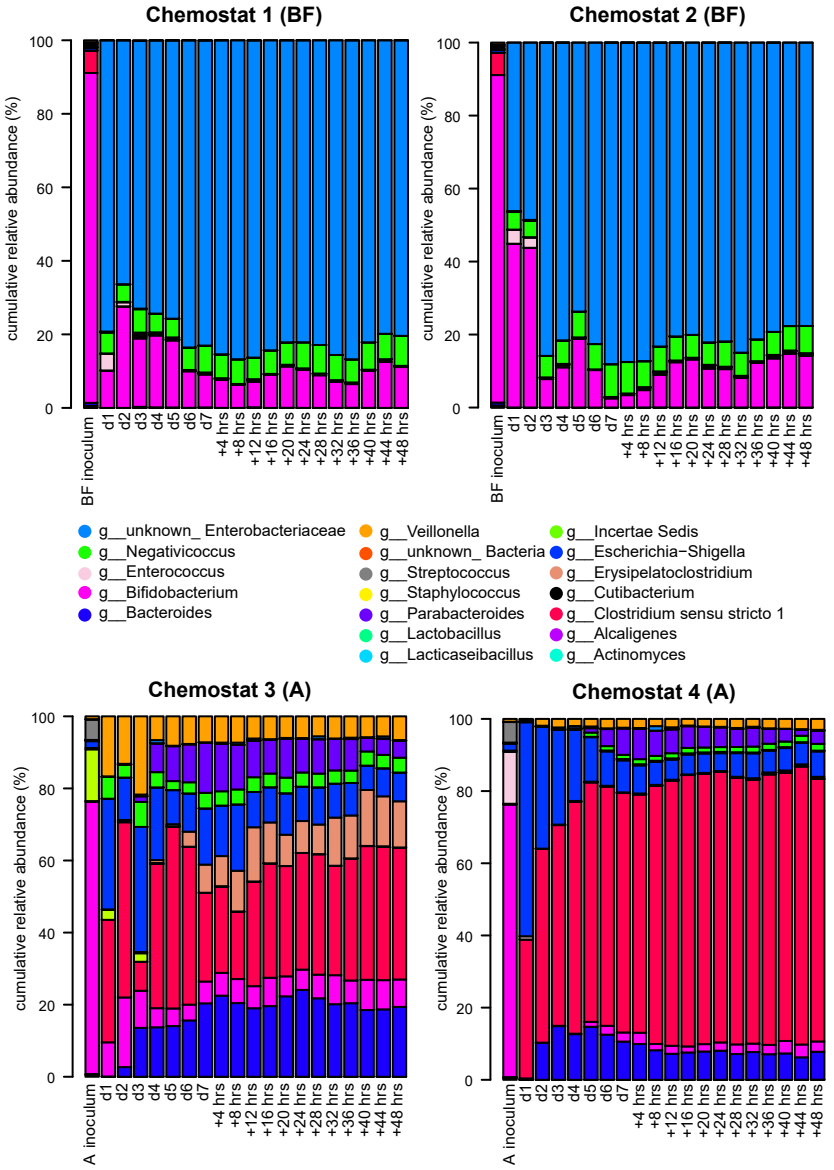
(D) Linear discriminant analysis (LDA) effect size (LEfSe) of pathways at age 7 months between placebo formula group (A) and all other feeding groups (formula B, C, D and breastfeeding (BF)).

Supplemental Figure 5

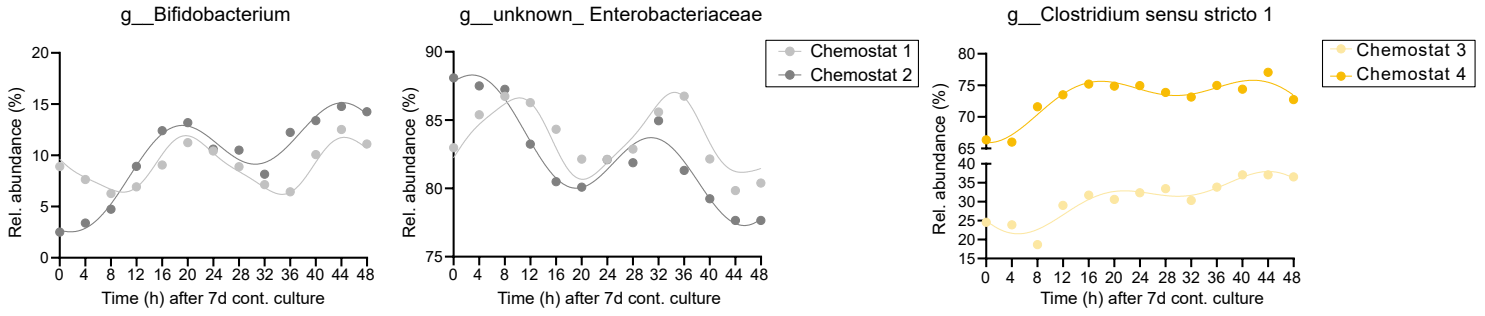
A



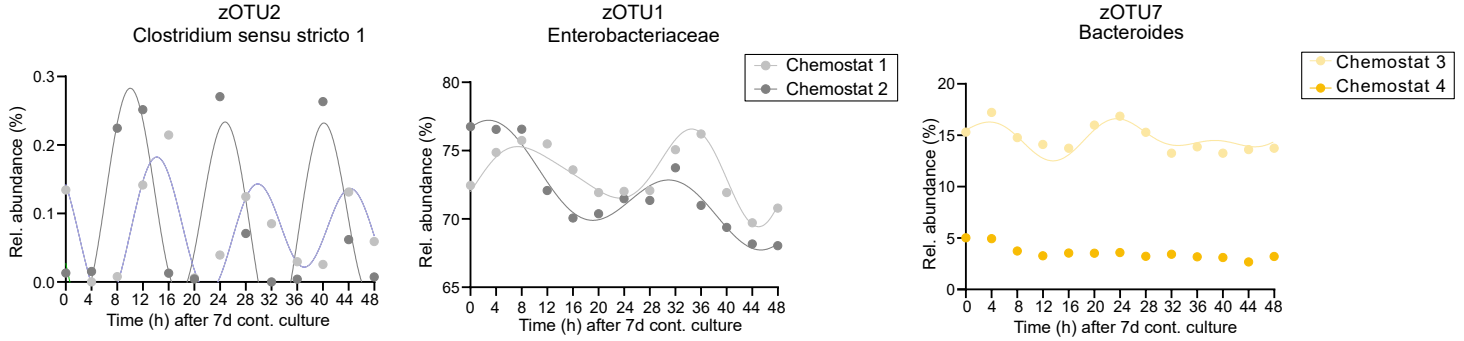
B



C



D



Supplemental Figure 5: Chemostats differ in taxonomic composition and diversity but consistently show rhythmicity in all vessels

(A) MDS plot of beta-diversity comparison between placebo formula A and BF of the *ex vivo* chemostat samples and the inoculum samples (separation of groups as whole is significant: $p \leq 0.001$, permutational multivariate analysis based on distance matrices). BF_chemostat_1-BF_chemostat_2: corr. p-value 0.0627, BF_chemostat_1-A_chemostat_3: corr. p-value 0.003, BF_chemostat_1-A_chemostat_4: corr. p-value 0.003, BF_chemostat_2-A_chemostat_3: corr. p-value 0.003, BF_chemostat_2-A_chemostat_4: corr. p-value 0.003, A_chemostat_3-A_chemostat_4: corr. p-value 0.003

(B) Taxonomic binning barplots showing the differences in taxonomic composition of each chemostat throughout the continuous culture and rhythmicity sampling period.

(C, D) Circadian profiles of genera (C) and zOTUs (D) from chemostat cultures after 7 days of continuous culture. Significant rhythms are illustrated with fitted cosine-wave curves (cosine-wave regression, $p \leq 0.05$). Fpr zOTU7, chemostat 4 is not significantly rhythmic. Data are represented as mean ± SEM.

Supplemental Table 1: untargeted metabolomics

Nr in Figure	Fig	Original ID	RT [sec]	m/z	Compound Name
1	4C/4D	ID1878_RP_pos	445.2965	584.2579	unknown
2	4C/4D	ID2147_HILIC_pos	41.3172	872.7654	TG(52:4)
3	4C/4D	ID823_RP_neg	47.9137	855.3091	unknown
4	4C/4D	ID1240_RP_pos	45.7357	381.0789	Disaccharide
5	4C/4D	ID1843_HILIC_pos	46.7741	732.5525	PC(32:1)
6	4C/4D	ID1295_HILIC_pos	325.4208	527.3134	unknown
7	4C/4D	ID1979_RP_pos	47.5239	708.2544	N-acetyl-quadruple sugar
8	4C/4D	ID2155_HILIC_pos	530.2239	878.2997	unknown
9	4C/4D	ID2205_HILIC_pos	569.4872	1387.4841	unknown
10	4C/4D	ID666_HILIC_pos	265.5851	337.2509	unknown
11	4C/4D	ID666_RP_neg	520.0732	539.2167	unknown
12	4C/4D	ID784_RP_neg	519.3508	805.5629	unknown
13	4C/4D/4F	ID338_HILIC_pos	45.5292	213.1824	unknown
14	4C/4D	ID947_RP_pos	645.0067	325.2108	unknown
15	4C/4D	ID554_HILIC_pos	41.3418	297.2767	Noadecenoic acid
16	4C/4D/4F	ID1067_RP_pos	449.214	345.2824	unknown
17	4C/4D	ID194_RP_neg	632.1141	295.189	unknown
18	4C/4D	ID1495_HILIC_pos	466.705	604.2497	unknown
19	4C/4D	ID1524_RP_pos	671.0764	430.3485	unknown
20	4C/4D	ID1462_HILIC_pos	345.6149	591.3413	unknown
21	4C/4D	ID203_HILIC_pos	44.6892	165.0937	unknown
22	4C/4D	ID951_HILIC_pos	43.9798	424.3549	unknown
23	4C/4D	ID1501_HILIC_pos	276.4828	604.5999	unknown
24	4C/4D	ID1075_RP_pos	361.9261	351.1558	unknown
25	4C/4D	ID1989_RP_pos	708.0827	763.6034	unknown
26	4C/4D	ID276_RP_neg	544.2773	375.33	unknown
27	4C/4D	ID1872_HILIC_pos	45.5417	747.5897	unknown
28	4C/4D	ID279_RP_pos	587.4899	182.161	unknown
29	4C/4D	ID528_RP_pos	587.51	249.2476	unknown
30	4C/4D	ID1448_HILIC_pos	44.9654	583.5245	DG(15:0/18:0/0:0)
31	4C/4D	ID932_RP_pos	560.2321	322.2425	unknown
32	4C/4D	ID328_HILIC_pos	45.7918	209.187	unknown
33	4C/4D	ID425_HILIC_pos	70.027	255.0003	unknown
34	4C/4D	ID52_RP_pos	882.4914	97.9682	unknown
35	4C/4D	ID1543_RP_pos	298.375	433.2828	unknown
36	4C/4D	ID579_RP_neg	36.9573	502.914	unknown
37	4C/4D	ID957_HILIC_pos	44.3644	425.3386	Delta-Tocopherol
38	4C/4D	ID131_HILIC_pos	409.5396	144.0998	Prolin betain
39	4C/4D	ID962_HILIC_pos	78.185	426.1919	unknown
40	4C/4D/4F	ID2104_HILIC_pos	44.2302	832.669	unknown
41	4C/4D	ID1884_HILIC_pos	107.927	755.6154	TG(42:3)
42	4C/4D	ID1586_HILIC_pos	130.8504	635.6108	unknown
43	4C/4D	ID2017_HILIC_pos	129.2633	801.6593	unknown

44	4C/4D/4F	ID205_RP_pos	276.8442	165.0539	Coumaric acid
45	4C/4D	ID668_RP_pos	254.763	280.1055	unknown
46	4C/4D	ID1138_HILIC_pos	396.2653	472.1415	unknown
47	4C/4D	ID519_RP_neg	439.0721	474.2547	unknown
48	4C/4D	ID136_RP_pos	68.6502	145.1328	N-Acetylcadaverine
49	4C/4D/4F	ID719_RP_pos	47.173	287.241	N1,N12-Diacetylspermine
50	4F	ID13_RP_pos	45.6086	58.0649	unknown
51	4F	ID32_RP_pos	48.5084	72.0808	unknown
52	4F	ID61_RP_pos	43.7121	104.1066	unknown
53	4F	ID98_RP_pos	47.0312	132.0761	Creatine
54	4F	ID125_RP_pos	587.6851	143.1058	unknown
55	4F	ID143_RP_pos	276.9252	147.0429	Coumarin (low hit rate 48.72%)
56	4F	ID170_RP_pos	587.6044	155.178	unknown
57	4F	ID184_RP_pos	337.5045	160.0748	Indoleacetaldehyde
58	4F	ID200_RP_pos	587.5189	164.1504	unknown
59	4F	ID250_RP_pos	507.8812	174.1303	unknown
60	4F	ID262_RP_pos	587.4744	178.1659	unknown
61	4F	ID300_RP_pos	587.5718	186.1558	unknown
62	4F	ID305_RP_pos	337.5045	188.07	Indoleacrylic acid
63	4F	ID318_RP_pos	587.5566	192.1818	unknown
64	4F	ID336_RP_pos	587.5244	196.1768	unknown
65	4F	ID340_RP_pos	268.5427	197.0664	Dimethyluric acid
66	4F	ID365_RP_pos	47.7627	204.0862	N-acetyl-L-2-aminoadipate
67	4F	ID375_RP_pos	338.6819	206.0805	Indolelactic acid
68	4F	ID377_RP_pos	43.3248	206.1021	N-Acetylfucosamine
69	4F	ID515_RP_pos	47.3348	247.0927	unknown
70	4F	ID586_RP_pos	47.2125	264.1182	Creatine riboside
71	4F	ID591_RP_pos	591.3177	265.2523	unknown
72	4F/4H	ID637_RP_pos	43.2368	273.1433	unknown
73	4F	ID699_RP_pos	38.3207	284.9259	unknown
74	4F	ID782_RP_pos	709.7177	299.1386	unknown
75	4F	ID797_RP_pos	709.6879	300.1412	unknown
76	4F	ID814_RP_pos	374.105	302.2304	unknown
77	4F	ID843_RP_pos	655.1938	307.2623	Dihomo-alpha-linolenic acid
78	4F	ID853_RP_pos	43.1228	309.1652	1-[(5-Amino-5-carboxypentyl)amino] -1-deoxyfructose
79	4F	ID899_RP_pos	408.5222	316.2476	Decanoylcarnitine
80	4F	ID918_RP_pos	587.8349	318.2994	unknown
81	4F	ID1062_RP_pos	449.149	344.2793	Lauroylcarnitine
82	4B/4F	ID1148_RP_pos	269.0495	367.1495	unknown
83	4B/4F	ID1182_RP_pos	279.9274	371.2175	unknown
84	4F	ID1195_RP_pos	279.7921	372.2206	unknown
85	4F	ID1249_RP_pos	47.249	383.1374	unknown
86	4F/4H	ID1276_RP_pos	424.3017	386.2411	unknown
87	4F	ID1299_RP_pos	279.4737	389.2283	unknown

88	4F	ID1385_RP_pos	385.0854	402.283	unknown
89	4F	ID1706_RP_pos	44.7933	471.2175	unknown
90	4F	ID1843_RP_pos	437.6435	552.3348	unknown
91	4F	ID1858_RP_pos	275.0814	569.2906	unknown
92	4F	ID1867_RP_pos	445.3817	583.2545	unknown
93	4F	ID1882_RP_pos	709.7493	585.2672	unknown
94	4F	ID1907_RP_pos	709.232	604.2229	unknown
95	4F	ID1914_RP_pos	709.8671	607.2511	unknown
96	4F	ID1923_RP_pos	548.5101	617.3865	unknown
97	4F/4H	ID1932_RP_pos	587.3156	639.485	unknown
98	4F	ID1948_RP_pos	587.6053	657.459	unknown
99	4F	ID1953_RP_pos	587.5339	658.462	unknown
100	4F	ID1955_RP_pos	587.5718	659.4563	unknown
101	4F	ID1959_RP_pos	587.5208	660.4575	unknown
102	4F	ID4_RP_neg	44.3665	88.0397	unknown
103	4F	ID54_RP_neg	45.4521	162.0766	unknown
104	4F	ID56_RP_neg	49.2802	163.0616	Fucose or Deoxy-glucose
105	4F	ID57_RP_neg	71.257	163.0615	unknown
106	4F/4H	ID58_RP_neg	280.2972	163.0388	Coumaric acid
107	4F	ID62_RP_neg	49.3001	164.0641	unknown
108	4F	ID67_RP_neg	47.5997	165.0402	Arabinonic acid (98.25%)
109	4F	ID72_RP_neg	88.6861	167.021	unknown
110	4F	ID82_RP_neg	280.4396	181.0507	Hydroxyphenyllactic acid
111	4F	ID85_RP_neg	280.1192	182.0535	unknown
112	4F	ID97_RP_neg	271.8839	195.0522	Dimethyluric acid
113	4F	ID98_RP_neg	47.1844	195.0506	Sugar acid
114	4F	ID104_RP_neg	346.4323	204.0671	Indolelactic acid
115	4F	ID107_RP_neg	346.0053	205.0695	unknown
116	4F	ID136_RP_neg	53.2571	259.0136	3-O-sulfo-β-D-galactose
117	4F	ID147_RP_neg	347.5474	272.054	unknown
118	4F	ID215_RP_neg	44.3295	307.1504	Fructoselysine
119	4F	ID269_RP_neg	570.8473	365.2306	unknown
120	4F	ID280_RP_neg	46.2214	380.1617	unknown
121	4F	ID297_RP_neg	492.7667	389.2698	unknown
122	4F	ID398_RP_neg	492.7135	435.2752	unknown
123	4F	ID416_RP_neg	39.8382	440.8856	unknown
124	4F	ID481_RP_neg	58.5597	462.0931	unknown
125	4F	ID573_RP_neg	418.5118	499.3269	unknown
126	4F	ID599_RP_neg	429.2894	513.3428	unknown
127	4F	ID630_RP_neg	53.4444	523.1444	fucopyranosyllactose
128	4F	ID656_RP_neg	420.7276	536.3211	unknown
129	4F	ID677_RP_neg	47.9845	546.2038	N-acetyl-trisaccharide
130	4B/4F	ID788_RP_neg	282.616	811.4574	unknown
131	4F	ID8_HILIC_pos	424.4444	60.0803	unknown
132	4F	ID80_HILIC_pos	385.1537	123.0541	2-Acetylpyrazine
133	4F	ID85_HILIC_pos	351.9125	128.1056	N-Acetylpiperidin

134	4F	ID97_HILIC_pos	369.6011	131.1162	N-Acetylputrescine
135	4F	ID137_HILIC_pos	352.4391	145.1325	N-acetylcadaverine
136	4F	ID141_HILIC_pos 100%	424.348	146.1171	3-Dehydroxycarnitine
137	4F	ID145_HILIC_pos	424.3643	147.1195	Glutamine
138	4F	ID161_HILIC_pos	97.413	153.0638	unknown
139	4F	ID170_HILIC_pos	332.2044	159.1481	unknown
140	4F	ID189_HILIC_pos	405.4134	162.1117	Carnitine
141	4F	ID193_HILIC_pos	405.2537	163.1126	unknown
142	4F	ID245_HILIC_pos	312.4109	175.0703	unknown
143	4F	ID449_HILIC_pos	439.5683	264.1192	unknown
144	4F	ID454_HILIC_pos	439.7345	265.1215	unknown
145	4F	ID518_HILIC_pos	357.7949	284.062	unknown
146	4B/4F	ID550_HILIC_pos	45.7972	295.2258	oxidized fatty acid
147	4F	ID582_HILIC_pos	44.1117	307.261	stearic acid
148	4F	ID607_HILIC_pos	291.0532	314.231	Decenoylcarnitine
149	4F	ID614_HILIC_pos	289.283	316.247	Decanoylcarnitine
150	4F	ID680_HILIC_pos	284.8135	342.2629	unknown
151	4F	ID688_HILIC_pos	283.692	344.2791	Lauroylcarnitine
152	4B/4F	ID768_HILIC_pos	307.4853	371.218	unknown
153	4F	ID784_HILIC_pos	459.039	381.1279	unknown
154	4F	ID795_HILIC_pos	292.5264	384.2727	Hydroxy-tetradecadiencarnitine
155	4F/4H	ID807_HILIC_pos	289.7417	386.2895	Hydroxy-tetradecenoylcarnitine
156	4F	ID839_HILIC_pos	50.7572	397.378	unknown
157	4F	ID876_HILIC_pos	416.8625	406.1316	Lacto-N-biose I
158	4F	ID899_HILIC_pos	287.1428	412.3056	Hydroxyhexadecadienoylcarnitine
159	4F/4H	ID948_HILIC_pos	43.7895	423.3233	unknown
160	4F	ID983_HILIC_pos	274.8781	429.3742	unknown
161	4F	ID1001_HILIC_pos	43.2323	431.3727	unknown
162	4F	ID1015_HILIC_pos	395.6925	434.1883	unknown
163	4F	ID1021_HILIC_pos	395.6508	435.1907	unknown
164	4F	ID1062_HILIC_pos	42.4714	446.397	unknown
165	4F	ID1071_HILIC_pos	42.6837	447.4	unknown
166	4F	ID1074_HILIC_pos	42.8372	448.3988	unknown
167	4F	ID1105_HILIC_pos	52.8675	459.3401	unknown
168	4F	ID1231_HILIC_pos	461.7946	511.1628	3-Fucosyllactose
169	4F	ID1345_HILIC_pos	43.8975	548.5346	unknown
170	4F	ID1346_HILIC_pos	97.4613	548.5369	unknown
171	4F	ID1380_HILIC_pos	44.23	561.4887	FAHFA
172	4F	ID1399_HILIC_pos	80.4379	567.4422	unknown
173	4F	ID1425_HILIC_pos	46.2204	573.4763	unknown
174	4F	ID1431_HILIC_pos	43.6126	576.566	unknown
175	4F	ID1565_HILIC_pos	92.2118	630.6151	unknown
176	4F	ID1604_HILIC_pos	227.4001	642.5123	unknown
177	4F	ID1621_HILIC_pos	44.7479	650.6385	Ceramide (d18:1/24:0)
178	4F	ID1661_HILIC_pos	40.811	663.4519	unknown

179	4F	ID1663_HILIC_pos	41.092	664.4597	unknown
180	4F	ID1707_HILIC_pos	45.8138	678.5186	unknown
181	4F	ID1726_HILIC_pos	512.4292	689.2044	unknown
182	4F/4H	ID1780_HILIC_pos	256.9355	706.5366	PC(30:0)
183	4F	ID1790_HILIC_pos	98.0462	710.5901	unknown
184	4F/4H	ID1792_HILIC_pos	97.5934	711.5925	unknown
185	4F	ID1794_HILIC_pos	494.3389	714.242	N-acetylated quadruple sugar
186	4F	ID1812_HILIC_pos	46.2407	719.57	unknown
187	4F	ID1822_HILIC_pos	46.1982	721.5722	unknown
188	4F	ID1876_HILIC_pos	44.335	749.601	unknown
189	4F	ID1989_HILIC_pos	84.2318	792.6685	unknown
190	4F	ID1992_HILIC_pos	86.5068	793.6716	unknown
191	4F	ID1994_HILIC_pos	92.0972	794.6834	unknown
192	4F	ID1998_HILIC_pos	76.3574	795.6861	unknown
193	4F	ID2074_HILIC_pos	200.7976	819.6667	unknown
194	4F	ID2163_HILIC_pos	42.3567	880.7645	unknown
195	4B/4H	ID990_HILIC_pos	42.9535	430.3752	unknown
196	4H	ID989_HILIC_pos	43.0046	429.3716	unknown
197	4H	ID971_HILIC_pos	43.8123	427.3523	unknown
198	4H	ID924_RP_pos	40.5003	320.8688	unknown
199	4H	ID580_RP_pos	560.4512	263.2366	unknown
200	4H	ID503_HILIC_pos	46.0612	279.2291	Linolenic acid
201	4H	ID2187_HILIC_pos	548.1752	1022.3515	unknown
202	4H	ID2152_HILIC_pos	530.8375	877.2995	unknown
203	4H	ID196_HILIC_pos	44.9749	163.1463	unknown
204	4H	ID1912_HILIC_pos	254.2377	763.599	unknown
205	4H	ID1906_HILIC_pos	253.9488	762.5947	PC34:0
206	4B/4H	ID1904_HILIC_pos	253.4206	761.5871	unknown
207	4H	ID1889_HILIC_pos	253.2523	758.5674	PC(34:2)
208	4H	ID1858_HILIC_pos	256.0458	736.572	unknown
209	4H	ID1855_HILIC_pos	255.7154	735.5715	unknown
210	4H	ID1852_HILIC_pos	47.5158	734.5681	PC(32:0)
211	4H	ID1850_HILIC_pos	253.782	734.5695	PC(32:0)
212	4B/4H	ID176_HILIC_pos	427.8128	160.1324	Aminovaleric acid betaine (could be 4- or 5-)
213	4H	ID1581_HILIC_pos	44.1712	634.6276	unknown
214	4H	ID1501_RP_pos	583.4205	428.3723	Octadecanoyl-L-carnitine
215	4H	ID1182_HILIC_pos	471.4868	495.1841	unknown
216	4B/4H	ID1058_HILIC_pos	42.9742	444.3705	unknown
Fig4B	4B	ID397_HILIC_pos	357.6501	245.0794	unknown
Fig4B	4B	ID162_HILIC_pos	45.5128	153.1254	unknown
Fig4B	4B	ID978_HILIC_pos	43.2362	428.3824	unknown
Fig4B	4B	ID920_HILIC_pos	60.6639	418.1883	unknown
Fig4B	4B	ID322_RP_neg	516.766	399.1485	unknown
Fig4B	4B	ID2150_HILIC_pos	529.651	876.2944	unknown