**Supplementary Material for**

**Influence of early life factors on the breast milk and fecal microbiota of mother-newborn dyads**

Emmanuel Cervantes-Monroy, Imelda C. Zarzoza-Mendoza, Samuel Canizales-Quinteros, Sofia Moran-Ramos, Judith Villa-Morales, Blanca López-Contreras, Fairt V. Carmona-Sierra 5,and Maricela Rodríguez-Cruz

Corresponding author: Email:maricela.rodriguez.cruz@gmail.com

**The Supplementary Material includes.**

1. **Procedures**
	1. *DNA extraction from fecal and BM samples*

*1.2 PCR Amplification and 16S rRNA Sequencing*

*1.3 Bioinformatic analyses*

1. **Results**
	1. *Flowchart of the Participants and Sample Analysis*
2. **Supplementary Tables**
	1. *Supplementary Table 1*
	2. *Supplementary Table 2*
	3. *Supplementary Table 3*
	4. *Supplementary Table 4*
	5. *Supplementary Table 5*
	6. *Supplementary Table 6*
3. **Procedures**

During the home visit, the mother's body weight and body composition were assessed by bioelectrical impedance using a scale (BC-585F FitScan, TANITA, U.S.A.) as described elsewhere [1]. Pre-gestational BMI was calculated using the self-reported body weight (kg) divided by the height (m) square. The BMI at 20-28DPP was obtained using the body weight measured during the visit. The weight gain during pregnancy was calculated based on the self-reported last gestational weight. Newborn body weight and length measurements were performed supine, without clothes and diapers, using an electronic baby scale and an infantometer (Seca 334, Mobile digital baby scale, Hamburg, Germany) on a flat, solid, and even surface, as recommended by WHO guidelines. The body weight gain from birth (WGFB) and length gain from birth (LGFB) were calculated as the difference between the day of birth and 20-28 postnatal days (20-28PND) data. Finally, the head circumference was assessed using an ergonomic circumference measuring tape (Seca 201) placed around the most prominent area of the forehead, above the eyebrows and ears, and behind the occipital protuberance. All measurements were performed in triplicate.

* 1. *DNA extraction from fecal and BM samples*

For DNA isolation from fecal samples (0.2 g), the commercial PowerFecal kit (QIAamp® Power Fecal® DNA, QIAGEN, U.S.A.) was used following the manufacturer’s instructions, including a mechanical disruption by FastPrep (MP Biomedicals; Solon, Ohio, U.S.A.). The DNA was eluted in a final volume of 100 µL and stored at -70 ºC for further analysis. For BM DNA extraction, the samples (3 mL) were thawed to 5 ºC and centrifuged at 20,000 g for 20 min at 4ºC (Allegra X-22, Beckman Coulter Co., U.S.A.). After that, the fat layer was removed, and the supernatant was discarded. Finally, according to the manufacturer’s protocol, the genomic DNA was isolated from the pellet using the aforesaid commercial kit. The purified DNA was eluted in a final volume of 50 µL and stored at -70ºC for further analysis. The extracted genomic DNA concentration from fecal and BM samples was measured by spectrophotometry (Nanodrop 2000c, Thermo Scientific, Wilmington, DE, U.S.A.). The A260/280 ratio was used as a parameter of DNA purity.

* 1. *PCR Amplification and 16S rRNA Sequencing*

The metataxonomic profiling of the fecal and BM microbiota was carried out by the 16S ribosomal RNA (16S rRNA) gene sequencing technique described by López-Contreras *et al.*, 2018 [2]. The hypervariable region 4 (V4) was amplified using the oligonucleotides recommended by the Earth Microbiome Project (515F-806R). Briefly, the first PCR amplification was performed with 100 ng of DNA using Platinum TM Taq and Polymerase (Invitrogen TM, Thermo Scientific, U.S.A.) under the following PCR conditions: 94°C for 3 min, 30 cycles of 94°C for 30 s, 50°C for 30 s and 68°C for 1.5 min and finally 68°C for 5 min. The DNA integrity was verified by agarose gel electrophoresis (2%). The amplicons were purified through the Agencourt AMPure XP beads (Beckman Coulter, Fullerton, CA, U.S.A.) using the Agilent Bravo Automated Liquid Handling Platform (Agilent Technologies, CA, U.S.A.). A second PCR was performed to attach the Illumina sequencing adapters to the amplicons with the following conditions: 94°C for 2 min, five cycles of 94°C for 30 s, 60°C for 30 s, and 68°C for 1.5 min, and finally 68°C for 5 min. The amplicons were purified, and the libraries' quantity was assessed using a Qubit® dsDNA HS assay kit on the Qubit 2.0 fluorometer (Life Technologies, ThermoFisher Scientific; USA). The final library size and concentration were determined by an Agilent D1000 ScreenTape for 4200 TapeStation System (Agilent Technologies, Waldbronn, Germany) and a Qubit 2.0 fluorometer, respectively. Finally, the amplicons were pooled in equimolar ratios and sequenced using the Illumina Miseq 2x250-bp platform at the core unit of *Instituto Nacional de Medicina Genómica*. Negative controls included sterile DNA-free water for fecal and BM DNA extraction and sequencing library preparation. A mock community was used as a positive control (Zymo Research Corp, CA, U.S.A.).

* 1. *Bioinformatic analyses*

The raw sequences underwent quality control analysis using the FASTQC (version 0.11.9 [3]) tool. Then, the reads were processed and analyzed through the QIIME2 pipeline (version 2020.6 [4]). The sequences were demultiplexed, and DADA2 [5] was used for trimming and filtering to remove sequencing errors in chimeric and or singleton reads, obtaining the amplicon sequence variants (ASVs). The ASVs were referenced against the SILVA 138 database (version 123 [6]) for the 99% sequence similarity taxonomic assignment. Afterward, the metadata, ASV, and taxonomic tables were imported into R software (version 4.1.2 [7]) through the Qiime2R package (version 0.99.6 [8]). The diversity, structure, and taxonomic composition analysis were conducted in Rstudio (version 1.4.1717 [9]) using the phyloseq (version 1.38.0 [10]) and microbiome (version 1.16.0 [11]) packages. Other packages, such as Tidyverse (version 1.3.1 [12]) and Vegan (version 2.6-2 [13]), were also used. A phylogenetic tree was created using the Ape package (version 5.6-2 [14]). The *Decontam* package(version 1.14.0 [15]) used the prevalence method to detect and remove environment or reagent contaminant sequences. Moreover, quality filtering was applied, excluding all samples with less than 5,000 reads, ASVs without reads, and those belonging to domains other than Bacteria and Archaea. Core microbiota, which refers to microbial taxa occurring above a particular occupancy frequency threshold, is typically defined using a prevalence threshold ranging from 30% to 95% [16]. This study used a prevalence threshold of 70% to identify the core microbiota from maternal and newborn fecal and BM samples. The alpha diversity (within-sample diversity) was estimated with the Chao (richness) and Shannon (diversity) metrics, rarefying the sequences to 5,000 reads. The unweighted and weighted UniFrac distances were calculated, and a Principal Coordinate Analysis (PCoA) was performed to assess the dissimilarity in microbiota community structure (β diversity) between samples. A differential abundance test was conducted employing the linear discriminant analysis (LDA) effect size (LefSe [17]) tool to identify genus-level differences between groups based on maternal nutrition status and age, delivery mode, and newborn sex.

1. **Results**
	1. *Flowchart of the Participants and Sample Analysis*

Nine hundred eighty-six pregnant women participated in the Program of Educational Strategies for Health during Pregnancy and Lactation between November 2018 and March 2020, of which 40.87% decided to participate (Supplementary Figure1). During the telephone follow-up and home visit at 20-28DPP, 86.85% of the candidates were excluded due to failing to meet the inclusion criteria or refusing to continue their participation. Finally, fifty-three mother-newborn dyads were enrolled in the study. However, 13 dyads were excluded due to technical issues with the maternal fecal, BM, or newborn fecal samples. These samples comprised 20 mother fecal-BM-newborn fecal triads, 15 mother fecal-newborn fecal pairs, two mother fecal-BM pairs, one BM-newborn fecal pair, one mother fecal sample, and one BM sample for data analysis.

**References**

1. Badillo-Suárez, P.A.; Rodríguez-Cruz, M.; Bernabe-García, M.; Villa-Morales, J.; Iglesias-Rodríguez, R.; Canizales-Quinteros, S.; et al. Influence of maternal body fat on insulin levels, insulin-like growth factor-1, and obestatin. *J. Hum. Lact.* **2022**, 38(4), 619–632.
2. López-Contreras, B.E.; Morán-Ramos, S.; Villarruel-Vázquez, R.; Macías-Kauffer, L.; Villamil-Ramírez, H.; León-Mimila, P.; et al. Composition of gut microbiota in obese and normal-weight Mexican school-age children and its association with metabolic traits. *Pediatr. Obes.* **2017**, 13(6), 381–388.
3. Andrews, S. **2015**. FastQC. URL <https://qubeshub.org/resources/fastqc>
4. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, X.X.; Al-Ghalith, G.A.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME2. *Nat. Biotechnol*. **2019**, 37(8), 852–857.
5. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, 13(7), 581–583.
6. Yilmaz, P.; Parfrey, L.W.; Yarza, P.; Gerken, J.; Pruesse, E.; Quast, C.; et al. The SILVA and “all-species Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Res*. **2014**, 42(D1), D643–D648.
7. R Core Team. **2021**. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
8. Bisanz, J.E. **2018**. qiime2R: Importing QIIME2 artifacts and associated data into R sessions.
9. RStudio Team. **2022**. RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA.
10. McMurdie, P.J.; Holmes, S. Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **2013**, 8(4), e61217.
11. Lahti, L.; Shetty, S. **2012-2019**. Microbiome R package.
12. Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L.; François, R.; et al. Welcome to the tidyverse. *J. Open Source Softw*. **2019**, 21;4(43), 1686.
13. Oksanen, J.; Simpson, G.L.; Blanchet, F.G.; Kindt, R.; Legendre, P.; Minchin, P.R.; et al. **2020**. vegan: *Community Ecology Package*. R package version 2.5–7.
14. Paradis, E.; Schliep, K. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **2019**, 35(3), 526–528.
15. Davis, N.M.; Proctor, Di.M.; Holmes, S.P.; Relman, D.A.; Callahan, B.J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **2018**, 6:226, 1–13.
16. Risely, A. Applying the core microbiome to understand host–microbe systems. *J. Anim. Ecol*. **2020**, 89, 1549– 1558.
17. Segata, N.; Izard, J.; Waldron, L.; Gevers, D.; Miropolsky, L.; Garrett, W.S.; et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* **2011,** 12, R60, 1–18.
18. **Supplementary Tables**
	1. *Supplementary Table 1*

**Supplementary Table 1.** Maternal gut microbiota composition at phylum level.

|  |  |  |
| --- | --- | --- |
| **Phylum** | **Number of genera** | **Mean relative abundance (%)** |
| Firmicutes | 181 | 60.2665 |
| Bacteroidota | 26 | 31.3971 |
| Actinobacteriota | 31 | 3.9583 |
| Proteobacteria | 44 | 2.6712 |
| Verrucomicrobiota | 7 | 1.0345 |
| Desulfobacterota | 4 | 0.2800 |
| Euryarchaeota | 1 | 0.2095 |
| Cyanobacteria | 1 | 0.0989 |
| Elusimicrobiota | 1 | 0.0432 |
| Fusobacteriota | 3 | 0.0230 |
| Unclassified | 1 | 0.0122 |
| Campilobacterota | 1 | 0.0039 |
| Thermoplasmatota | 1 | 0.0011 |
| Synergistota | 1 | 0.0005 |
| Bdellovibrionota | 1 | 0.0002 |
| **Total** | **304** | **100** |

*Supplementary Table 2*

**Supplementary Table 2.** Maternal gut microbiota composition at genus level stratified by phylum.

|  |  |
| --- | --- |
| **Genus** | **Mean relative abundance (%)** |
| Firmicutes | **60.27** |
| Other Firmicutes | 41.88 |
| *Blautia* | 6.05 |
| *Faecalibacterium* | 5.41 |
| *Oscillospiraceae UCG-002* | 2.63 |
| *Christensenellaceae R-7* | 2.35 |
| *Subdoligranulum* | 1.95 |
| Bacteroidota | **31.40** |
| *Bacteroides* | 17.12 |
| *Prevotella* | 9.27 |
| *Alistipes* | 1.69 |
| Other Bacteroidota | 1.32 |
| *Parabacteroides* | 1.15 |
| *Paraprevotella* | 0.85 |
| Actinobacteriota | **3.96** |
| *Bifidobacterium* | 3.28 |
| *Collinsella* | 0.48 |
| Other Actinobacteriota | 0.09 |
| *Eggerthella* | 0.05 |
| *Gordonibacter* | 0.03 |
| *Adlercreutzia* | 0.02 |
| Proteobacteria | **2.67** |
| Other Proteobacteria | 1.18 |
| *Sutterella* | 0.70 |
| *Escherichia-Shigella* | 0.58 |
| *Parasutterella* | 0.16 |
| *Pseudomonas* | 0.05 |
| *Succinivibrio* | 0.01 |
| OthersOther taxa | **1.71**1.71 |
| **Total** | **100** |

*Supplementary Table 3*

**Supplementary Table 3.** Breast milk microbiota composition at phylum level.

|  |  |  |
| --- | --- | --- |
| **Phylum** | **Number of genera** | **Mean relative abundance (%)** |
| Firmicutes | 157 | 59.1724 |
| Proteobacteria | 108 | 27.1315 |
| Actinobacteriota | 42 | 7.0116 |
| Bacteroidota | 45 | 5.7910 |
| Unclassified | 1 | 0.2938 |
| Fusobacteriota | 4 | 0.1665 |
| Verrucomicrobiota | 7 | 0.1371 |
| Planctomycetota | 6 | 0.1171 |
| Acidobacteriota | 8 | 0.0704 |
| Desulfobacterota | 4 | 0.0324 |
| Myxococcota | 4 | 0.0197 |
| Spirochaetota | 1 | 0.0146 |
| Cyanobacteria | 2 | 0.0110 |
| Bdellovibrionota | 1 | 0.0075 |
| Euryarchaeota | 1 | 0.0060 |
| Synergistota | 1 | 0.0051 |
| Patescibacteria | 3 | 0.0034 |
| Elusimicrobiota | 1 | 0.0028 |
| WPS-2 | 1 | 0.0026 |
| Campilobacterota | 2 | 0.0023 |
| Chloroflexi | 1 | 0.0006 |
| Gemmatimonadota | 1 | 0.0003 |
| Fibrobacterota | 1 | 0.0003 |
| **Total** | **402** | **100** |

*Supplementary Table 4*

**Supplementary Table 4.** Breast milk microbiota composition at genus level stratified by phylum.

|  |  |
| --- | --- |
| **Genus** | **Mean relative abundance (%)** |
| Firmicutes | **59.17** |
| *Streptococcus* | 31.52 |
| *Staphylococcus* | 17.46 |
| Other Firmicutes | 6.77 |
| *Gemella* | 1.91 |
| *Faecalibacterium* | 0.87 |
| *Blautia* | 0.64 |
| Proteobacteria  | **27.13** |
| Other Proteobacteria | 14.95 |
| *Pseudomonas* | 4.90 |
| *Stenotrophomonas* | 2.55 |
| *Herbaspirillum* | 2.09 |
| *Escherichia-Shigella* | 1.40 |
| *Acinetobacter* | 1.25 |
| Actinobacteriota | **7.01** |
| *Bifidobacterium* | 4.05 |
| *Corynebacterium* | 2.22 |
| *Rothia* | 0.48 |
| Other Actinobacteriota | 0.12 |
| *Actinomyces* | 0.11 |
| *Collinsella* | 0.02 |
| Bacteroidota | **5.79** |
| *Bacteroides* | 2.07 |
| *Prevotella* | 1.83 |
| Other Bacteroidota | 1.04 |
| Alistipes | 0.49 |
| *Nubsella* | 0.20 |
| *Sphingobacterium* | 0.17 |
| OthersOther taxa | **0.89**0.89 |
| **Total** | **100** |

*Supplementary Table 5*

**Supplementary Table 5.** Newborn gut microbiota composition at phylum level.

|  |  |  |
| --- | --- | --- |
| **Phylum** | **Number of genera** | **Mean relative abundance (%)** |
| Proteobacteria | 41 | 35.5038 |
| Actinobacteriota | 23 | 30.4233 |
| Firmicutes | 113 | 24.0218 |
| Bacteroidota | 17 | 10.0156 |
| Verrucomicrobiota | 1 | 0.0222 |
| Desulfobacterota | 4 | 0.0051 |
| Unclassified | 1 | 0.0050 |
| Campilobacterota | 1 | 0.0008 |
| Elusimicrobiota | 1 | 0.0006 |
| Fusobacteriota | 1 | 0.0005 |
| Synergistota | 1 | 0.0005 |
| Euryarchaeota | 1 | 0.0004 |
| Cyanobacteria | 1 | 0.0001 |
| Bdellovibrionota | 1 | 0.0001 |
| **Total** | **402** | **100** |

*Supplementary Table 6*

**Supplementary Table 6.** Newborn gut microbiota composition at genus level stratified by phylum.

|  |  |
| --- | --- |
| **Genus** | **Mean relative abundance (%)** |
| Proteobacteria  | **35.50** |
| Other Proteobacteria | 18.26 |
| *Escherichia-Shigella* | 13.42 |
| *Pseudomonas* | 2.80 |
| *Sutterella* | 1.00 |
| *Herbaspirillum* | 0.01 |
| *Stenotrophomonas* | 0.01 |
| Actinobacteriota | **30.42** |
| *Bifidobacterium* | 29.16 |
| *Actinomyces* | 0.74 |
| *Collinsella* | 0.38 |
| *Rothia* | 0.08 |
| Other Actinobacteriota | 0.04 |
| *Eggerthella* | 0.02 |
| Firmicutes | **24.02** |
| *Clostridium sensu stricto 1* | 9.52 |
| Other Firmicutes | 5.26 |
| *Streptococcus* | 4.70 |
| *Erysipelatoclostridium* | 2.10 |
| *Hungatella* | 1.30 |
| *Enterococcus* | 1.14 |
| Bacteroidota | **10.02** |
| *Bacteroides* | 8.97 |
| *Parabacteroides* | 0.67 |
| *Prevotella* | 0.30 |
| Other Bacteroidota | 0.04 |
| Grupo *dgA-11* | 0.02 |
| *Alistipes* | 0.02 |
| OthersOther taxa | **0.04**0.04 |
| **Total** | **100** |