Supplemental Material 1 - Methods

Sample collection, DNA extraction and sequencing

Stool samples were collected in DNA/RNA shield™ fecal collection tubes (Zymo Research, California) on day 1, (meconium wherever possible), at 7 ± 3 days and at 4 weeks ± 7 days. At each collection, feeding method, maternal antibiotic use, infant antibiotic use, and prebiotic use were recorded.

The samples were stored at -80°C on arrival in the laboratory (within 7 days of collection). DNA was extracted using the ZymoBIOMICS™ 96 MagBead DNA Kit (Zymo Research, California) according to the manufacturer's protocols.

Samples were randomised to five 96-well plates, along with four controls on each plate, with day 1 samples randomised to separate plates to allow maximal sequencing read coverage. Controls comprised two negative controls; water and water that had gone through the DNA extraction process, and two positive controls; *Escherichia coli* and a ZymoBIOMICS™ Microbial Community Standard (Zymo Research). One control was omitted due to well number constraints, and the sequencing facility included one negative control for each of the two sequencing runs, giving a total of 21 controls.

The Auckland Genome Sequencing facility carried out 16s rRNA amplicon sequencing on an Illumina Miseq instrument to determine microbial population structure. PCR primers used (forward 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG and reverse 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC) targeted the 16S V3 and V4 regions. Extracted DNA concentrations were measured using the Qubit™ dsDNA BR Assay Kit (Thermo Fisher Scientific, Massachusetts) for day 7 & week 4 samples and the Qubit™ dsDNA HS Assay Kit (Thermo Fisher Scientific) for day 1 samples.(1) Analysis personnel remained blinded to clinical metadata including group allocation.

Sequencing data bioinformatics

Sequencing reads were demultiplexed and adapter sequences removed by standard processing in the Illumina BaseSpace environment. Remaining primer sequences in the sequencing reads were removed using cutadapt v3.1.(2) The 16S amplicon sequences were sorted according to barcode. The DADA2 (Divisive Amplicon Denoising Algorithm) R package (3) was used for denoising the sequencing data and inferring amplicon sequence variants (ASVs).

Taxonomic assignment of ASVs was determined using the R package DECIPHER (4) and the Silva reference database SILVA_SSU_r138_2019,(5) and verified using positive controls of known microbial composition.

The R package decontam (6) was used to identify contaminant sequences that may have occurred during stool collection, sample processing or sequencing. Both frequency-based and prevalence-based methods were used, each with the default threshold P = 0.1. ASVs classified as contaminants by either method were removed.

Statistical analysis

Statistical analyses were carried out using R version 4.0.3 and phyloseq (7) and vegan (8) packages using intention-to-treat principles.

Beta-diversity (weighted UniFrac distance) was compared between groups using Permutational Multivariate Analysis of Variance using Distance Matrices (PERMANOVA) with the adonis2 function using 9999 permutations and assessing the marginal effects, adjusted for prespecified potential confounders of hospital, feeding (exclusively breastmilk, formula with or without breastmilk), prioritised ethnicity (Pacific was included in “Other” due to low numbers),(9) mode of birth and primary reason for risk of hypoglycaemia. The permutation-based p values < 0.05 were considered statistically significant. Principal Coordinates Analysis ordination was generated using weighted UniFrac as the distance measure for the week 4 samples.

DNA concentrations were compared between groups using mixed linear models, with timepoint as a fixed effect and subject ID as a random effect.

Microbial community stability was compared between groups using individual dissimilarity (1 – stability) by weighted UniFrac distance between day 7 and week 4 samples from subjects who provided samples at both times and Kruskal-Wallis tests. Within-subject and between-subject stability were compared using Student's t-test.

Within-sample diversity was compared using Shannon diversity index calculated for each day 7 and week 4 sample and mixed linear models with adjustment for the same confounders as the primary outcome. Samples with incomplete covariate data were excluded.

Microbial taxa abundance was compared between groups at the genus level for the day 7 and week 4 samples using mixed models implemented in MaAsLin2.(10) Group, timepoint, hospital, feeding, ethnicity, delivery mode and reason for risk of hypoglycaemia were included as fixed effects and subject ID as a random effect.

In the pre-specified exploratory analyses, differences in beta-diversity with respect to covariates of interest (delivery mode, type of feeding, hospital, ethnicity, reason for risk of hypoglycaemia) were tested using PERMANOVA at each time point and for each covariate separately. Alpha-diversity (Shannon diversity index), DNA concentrations and genus abundance (MaAsLin2) were assessed using mixed linear models with timepoint and covariate as fixed effects, and subject ID as a random effect. Ethnicity groups were compared using Wilcoxon tests. Associations with false discovery rate (FDR) corrected p-value (q-value) <0.05 were considered significant.

To assess the potentially confounding effect of treatment dextrose gel, we undertook a sensitivity analysis excluding infants who received treatment gel. In an additional exploratory analysis we also compared infants who received any dextrose gel (prophylaxis or treatment) with those who received neither.

Sample size

Based on the Dirichlet-multinomial distribution,(11) we estimated that 60 infants per group would have >90% probability of detecting 20% differences in the frequency of the five most common taxa between groups, allowing for 10% diversity within groups.

References

1. Stinson LF, Keelan JA, Payne MS. Comparison of Meconium DNA Extraction Methods for Use in Microbiome Studies. *Front Microbiol* 2018 Feb 20;9:270.

2. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011;17(1).

3. Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016 Jul;13(7):581-3.

4. Wright ES. Using DECIPHER v2.0 to analyze big biological sequence data in R. *R Journal* 2016;8:352-9.

5. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013 Jan;41(Database issue):590.

6. Davis N, Proctor D, Holmes S, et al. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 2018;6:226.

7. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013 Apr 22;8(4):e61217.

8. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan: Community Ecology Package. R package version 2.5-7. 2020 Nov 28.

9. Ministry of Health, N Z. HISO 10001:2017 Ethnicity Data Protocols. 2017.

10. Mallick H, Rahnavard A, McIver L, Ma S, Zhang Y, Nguyen L, et al. Multivariable Association Discovery in Population-scale Meta-omics Studies. bioRxiv 2021 [Preprint]. January 20, 2021 [cited 2021 Feb 4] https://doi.org/2021.01.20.427420.

11. Mattiello F, Verbist B, Faust K, Raes J, Shannon WD, Bijnens L, et al. A web application for sample size and power calculation in case-control microbiome studies. *Bioinformatics* 2016 Jul 1;32(13):2038-40.