Supplemental material: Biological samples collection and measurement

All biological samples are temporarily stored at -80° freezer at Tianjin Women and Children's Health Center (TJWCH), and regularly shipped on dry ice to different labs for analysis.

- Fecal samples:
    - Fecal samples are collected a maximum of 72 hours before the planned visit, stored at 2-8°C and bring to the hospital using a cooling bag. For subjects aged 6 months and under (<= 6m), 1 flat spoon of stool sample is collected. For subjects over 6 months of age (> 6m), 2 swabs of stool samples is collected. The collected samples are put in a pre-prepared tube and the stabilizer should merge the sample. A sampling guideline is included in the stool sample kits.
    - The methods of library construction and sequencing of microbiota in fecal samples are adopted from the protocol described previously[1]. Genomic DNA is firstly isolated from stool samples using Magpure Stool DNA KF Kit B (Magen, China) following the manufacturer's instruction. The isolated DNA is subjected to random fragmentation with Covaris E220 (Covaris, Brighton, UK). After ligation, a MGIEasyTM DNA Library Prep Kit (MGI, Shenzhen, China) is applied and the resulting ssDNA circles are used to generate DNA nanoballs (DNBs) by rolling circle amplification (RCA)[2 3]. After RCA and the formation of DNBs, the final products are measured by Qubit using the ssDNA HS Assay kit (Invitrogen), and loaded on a DIPSEQ platform (MGI, Shenzhen, China) for sequencing using paired-end 100 bp mode following the manufacturer's instructions[4]. Metagenome analysis by DNA shotgun sequencing using DIPSEQ platforms has been previously established to study human gut microbiome[5 6].
    - For bioinformatics analysis, the high-quality cleaned reads are used for the annotation and profile acquisition of taxons using MetaPhlan3 (version 3.0.14, code: metaphlan.py input.fastq -input_type fastq -nproc 10 > profiled_metagenome.txt)[7]. For functional abundance calculation, the HumanN3 pipeline (version 3.0.1, code:humann --input input.fastq --threads 10 --output output)[7] is used to map the sequences against the UniRef90 database with default parameters to obtain functional profiles. Taxonomic profiles include the stratified relative abundances from phylum to species levels. The stratified relative abundances are extracted according to the taxonomic levels of interest. The Chao1 and Shannon indexes are calculated to estimate microbial alpha-diversity.
- Breast milk samples:
    - Breast milk samples are collected in group 1 at 1-month and 4-month timepoints from mothers who provide exclusive breastfeeding to their infants. The volume of a breast milk sample is 30ml. A breast milk pump is available at the site and milk is collected in the morning from 9 to 11am. For mothers who cannot complete such collection during the visit, the study team will arrange another time (within 48h). A delay of additional 48h is accepted if the mother is sick.
    - Macronutrient analyses, including protein, lactose, energy, fat, and calcium, are performed on HLIFE MR-1011 automatic breast milk analyzer in the lab in TJWCH.
    - Quantitative analysis of Vitamin D in human milk. Vitamin D_3 and 25-OH D_3 are target analytes for quantitative analysis. Analysis is performed on Agilent 1290

Infinity liquid chromatography coupled to Agilent 6495 tandem mass spectrometry (Agilent). The analysis is conducted by SMQ Group Medical Laboratory (Shen Zhen, China). Phenomenex Kinetex PFP (2.1×150 mm, 1.7 μm) has been chosen as the analytical column. The separation gradient is shown in Supplementary Table S1. Mass spectrometric detection is carried out on Atmospheric Pressure Chemical Ionization operating in positive mode at unit resolution ($APCI^+$). Compound parameters are shown in Supplementary Table S2. The detection limit and quantitation limit of Vitamin D_3 are 0.02 ng/mL and 0.07 ng/mL, respectively. For 25-OH D_3, the detection limit is 0.05 ng/ml and the quantitation limit is 0.16 ng/ml.

- Quantification of HMOs in human milk by UHPLC-FLD. The method has been developed for the quantification of the major human milk oligosaccharides (HMOs) in human milk using UHPLC with fluorometric detection (Thermo U3000). External standard used for the quantification of identified HMOs is listed in Supplemental Table S3.

- Child urine sample

  - Urine samples are collected using urine collection sterile bags and then transferred into the two 5ml cryotubes. The samples are collected a maximum of 48hr before the planned visit, stored at 2-8°C. If samples are not collected prior the visit, the study team will propose to collect the sample during the visit or to collect the samples at home in the following 48 hours, stored at 2-8°C.

  - Analysis for urine creatinine and pyridinoline (PYD) is performed using commercial ELISA kits (MicroVue Creatinine 8009 and MicroVue PYD 8010, respectively), according to the manufacturer's instructions. The analysis is conducted by SMQ Group Medical Laboratory (Shen Zhen, China).

Reference

1. Fang C, Zhong H, Lin Y, et al. Assessment of the cPAS-based BGISEQ-500 platform for metagenomic sequencing. *Gigascience* 2018;7(3):1-8. doi: 10.1093/gigascience/gix133
2. Xu Y, Lin Z, Tang C, et al. A new massively parallel nanoball sequencing platform for whole exome research. *BMC Bioinformatics* 2019;20(1):153. doi: 10.1186/s12859-019-2751-3
3. Drmanac R, Sparks AB, Callow MJ, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* 2010;327(5961):78-81. doi: 10.1126/science.1181498
4. Huang J, Liang X, Xuan Y, et al. A reference human genome dataset of the BGISEQ-500 sequencer. *Gigascience* 2017;6(5):1-9. doi: 10.1093/gigascience/gix024
5. Zhang Y, Gu Y, Ren H, et al. Gut microbiome-related effects of berberine and probiotics on type 2 diabetes (the PREMOTE study). *Nat Commun* 2020;11(1):5015. doi: 10.1038/s41467-020-18414-8
6. Jie Z, Yu X, Liu Y, et al. The Baseline Gut Microbiota Directs Dieting-Induced Weight Loss Trajectories. *Gastroenterology* 2021;160(6):2029-42 e16. doi: 10.1053/j.gastro.2021.01.029
7. Beghini F, McIver LJ, Blanco-Miguez A, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife* 2021;10 doi: 10.7554/eLife.65088

Supplemental Table S1 Separation gradient of chromatographic separation

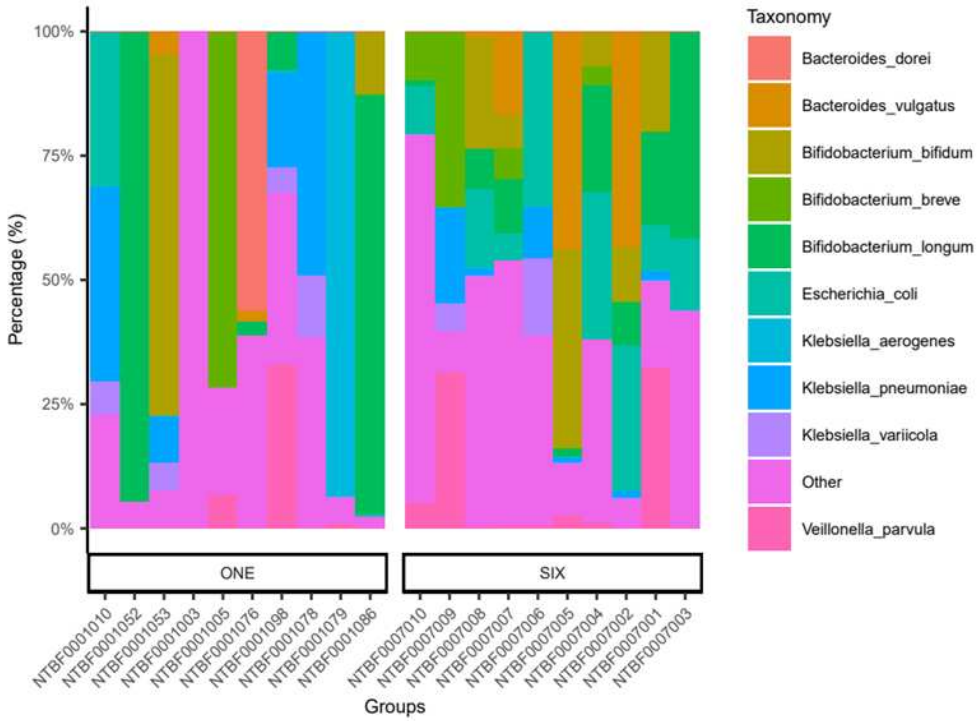
Time (min)	Flow rate (mL/min)	Mobile phase A (%)	Mobile phase B (%)
0	0.300	74.0	26.0
0.50	0.300	74.0	26.0
5.00	0.300	76.0	24.0
6.00	0.300	80.0	20.0
13.00	0.300	82.0	18.0
13.01	0.300	74.0	26.0

Supplemental Table S2. Compound parameters of mass spectrometric detection for vitamin D

Analyte	Transition Reactions (m/z) used for Quantification
D3-PTAD	560.31 →298.1
25-OHD3-PTAD	558.31 →298.1
D3- d3-PTAD	563.31 →301.1
25-OHD3-d6-PTAD	564.41 →298.0

Supplemental Table S3 External standard used for the quantification of identified HMOs

Analyte	Unit	Concentration range	
		Min	Max
2′FL	mg/L	19.4	5324.2
3FL	mg/L	17.8	4890.3
LDFT	mg/L	16.9	1011.3
3′SL	mg/L	16.0	799.3
LNT	mg/L	18.3	1643.8
LNnT	mg/L	18.1	723.7
6′SL	mg/L	17.7	1237.1
LNFP-I	mg/L	29.2	2923.5
Maltotriose	mg/L	30.1	6867.2



Supplemental Figure S1. Relative abundances of top 10 Species in the two groups.