



Maternal milk microbiota and oligosaccharides contribute to the infant gut microbiota assembly

Martin Frederik Laursen¹, Ceyda T. Pekmez², Melanie Wange Larsson^{2,3}, Mads Vendelbo Lind², Chloe Yonemitsu⁴, Anni Larnkjær², Christian Mølgaard², Lars Bode⁴, Lars Ove Dragsted², Kim F. Michaelsen^{1,2}, Tine Rask Licht¹ and Martin Iain Bahl¹✉

© The Author(s) 2021

Breastfeeding protects against diseases, with potential mechanisms driving this being human milk oligosaccharides (HMOs) and the seeding of milk-associated bacteria in the infant gut. In a cohort of 34 mother–infant dyads we analyzed the microbiota and HMO profiles in breast milk samples and infant’s feces. The microbiota in foremilk and hindmilk samples of breast milk was compositionally similar, however hindmilk had higher bacterial load and absolute abundance of oral-associated bacteria, but a lower absolute abundance of skin-associated *Staphylococcus* spp. The microbial communities within both milk and infant’s feces changed significantly over the lactation period. On average 33% and 23% of the bacterial taxa detected in infant’s feces were shared with the corresponding mother’s milk at 5 and 9 months of age, respectively, with *Streptococcus*, *Veillonella* and *Bifidobacterium* spp. among the most frequently shared. The predominant HMOs in feces associated with the infant’s fecal microbiota, and the dominating infant species *B. longum* ssp. *infantis* and *B. bifidum* correlated inversely with HMOs. Our results show that breast milk microbiota changes over time and within a feeding session, likely due to transfer of infant oral bacteria during breastfeeding and suggest that milk-associated bacteria and HMOs direct the assembly of the infant gut microbiota.

ISME Communications (2021)1:21; <https://doi.org/10.1038/s43705-021-00021-3>

INTRODUCTION

Breastfeeding reduces the risk of infectious diseases and may additionally reduce the risk of obesity, metabolic, and immune-related diseases later in life.^{1,2} Breast milk contains all the macro- and micronutrients necessary for infant growth, as well as bioactive compounds such as antimicrobial compounds, immunoglobulins, cytokines, peptide hormones and growth factors.^{3,4} Additionally, research within the last decade has revealed that breast milk contains a diverse microbiota,^{5–8} which may act as a reservoir of microbes for colonization of the infant gut and subsequently impact current and long-term health.^{9–11} The origins of bacteria in human breast milk are not fully elucidated but thought to be maternal skin,¹² infant oral cavity¹³ and environmental,⁷ as well as the maternal gut and oral cavity through the proposed entero-mammary¹⁴ and oro-mammary pathways,¹⁵ respectively. Breast milk from a complete feeding session varies in macronutrient composition and energy content during the feed.¹⁶ Practically, different fractions from a feeding session can be obtained with the terms *foremilk*, the very first milk when emptying the breast and *hindmilk*, the remainder or last milk when emptying the breast, commonly being used.^{16–18} By sampling and analyzing both fore- and hindmilk from the same feeding session prior to and after nursing the infant at the breast, it is possible to assess the extent to which retrograde inoculation of infant oral bacteria occurs. However, to our knowledge, no previous studies have assessed how the bacterial load and

composition changes along a feeding session. It has been estimated that exclusively breastfed infants, by consuming ~800 ml breast milk per day, ingest somewhere between 10⁵ and 10⁷ bacteria daily through breast milk.¹⁰ However, knowledge on which milk-associated bacterial taxa that establish and colonize the infant gut is scarce, and only a limited number of studies have estimated the extent to which breast milk microbes are transferred to the infant gut longitudinally.^{13,19,20} Especially, information about potential continued milk-to-gut transmission of bacteria spanning the complementary feeding period is lacking. Breast milk also contains a range of human milk oligosaccharides (HMOs), which may impact infant health both directly and indirectly (through selective microbial utilization by infant gut microbes).²¹ The HMOs, which constitute the third largest fraction of breast milk and comprises 5–20 g/l in mature milk, largely reach the intestinal tract undigested and thus become available as substrates for growth of specific gut microbes.²² Some individuals harboring a non-functional fucosyltransferase-2 (FUT2) gene, termed non-secretors, have markedly reduced breast milk levels of α1-2-fucosylated HMOs (e.g., 2FL) and previous studies suggest that this impacts the gut microbiota of the breastfed infant.^{23–25} While previous in vitro studies have shown that some groups of gut microbes are capable of degrading HMOs,^{26,27} there is a lack of human studies confirming this. *Bifidobacterium* spp. are archetypical HMO-degraders²⁸ and given the dominance of *Bifidobacterium* in the gut of breastfed infants,^{29–31} it is of

¹National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark. ²Department of Nutrition, Exercise, and Sports, University of Copenhagen, Frederiksberg, Denmark. ³Department of Nursing and Nutrition, University College Copenhagen, Copenhagen, Denmark. ⁴Department of Pediatrics and Larsson-Rosenquist Foundation Mother-Milk-Infant Center of Research Excellence, University of California San Diego, La Jolla, CA, USA. ✉email: mbah@food.dtu.dk

Received: 7 January 2021 Revised: 10 May 2021 Accepted: 20 May 2021
Published online: 07 June 2021

particular interest to investigate whether seeding of these from breast milk may occur, and whether HMOs contribute to their dominance. Therefore, across the lactation period, we characterized the microbiota in maternal foremilk and hindmilk samples and in infant's fecal samples and compared these with the HMO profiles in both sample types, in order to assess the impact of maternal milk microbes and HMOs on infant gut microbiota composition.

MATERIALS AND METHODS

Cohort subjects

The present study includes maternal breast milk and infant fecal samples from mother–infant pairs enrolled in the SKOT III cohort.³² The study protocol was approved by the Regional Ethical Committee of the Capital Region of Denmark in accordance with the Helsinki declaration (H-15008948) and registered at the Danish Data Protection Agency (2015-57-0117 and 2015-57-0116). Written informed consent was obtained from the parents. This cohort consist of breastfed infants that exhibit either excessive weight gain (EWG; $n = 13$, defined as increment of ≥ 1 SDs in weight-for-age Z-score during the first five months postpartum and a weight-for-age Z-score > 2.00 at recruitment time) or normal weight gain (NWG; $n = 17$, defined as weight-for-age Z-score between -1.00 and $+1.00$ SDs at first examination) during exclusive breastfeeding. In the present study, we included four additional participants that were previously excluded due to growth patterns, which did not adhere to the definition of growth pattern for the two groups,³² as our aim was to assess influence of breast milk microbiome and HMOs on infant gut microbiome, independent of infant growth patterns (Table S1). Recruitment of participants and criteria of inclusion and exclusion have been described previously.³² All infants were required to be within 5–6.5 months of age at first visit and at that time have breastfeeding as the primary energy source, allowing maximum two meals per day of solid foods. Infants were seen for a second visit at age 9 months ± 2 weeks. Maternal breast milk samples and infant fecal samples were obtained at both visits (Table S2). Additionally, we obtained information on age, gender, birth mode, gestational age at birth, presence of siblings, breastfeeding practices and use of antibiotics from parental questionnaires as described previously.³²

Fecal and breast milk samples and DNA extraction

On the morning of the visits parents collected infant's fecal samples from diapers and transported them on ice to the Department Nutrition, Exercise and Sports, University of Copenhagen, where they were stored at -80°C until DNA extraction. Before the first and second visit, the mothers were instructed to collect 10 ml breast milk from the start of the breastfeeding session (foremilk), then breastfeed the infant and subsequently collect 10 ml at the end of the same breastfeeding session (hindmilk). Foremilk and hindmilk samples were collected from the same breast, without cleaning the breast before or in between samplings. However, for three individuals at both visits, foremilk and hindmilk were collected on 2 separate days, 1–5 days apart (Table S2). Milk samples were collected into disposable tubes using a manual breast pump (Type HarmonyTM, Medela AG, Baar, Switzerland) and stored at -20°C in participants homes until transportation (on ice) to the University of Copenhagen, where the samples were stored at -80°C until DNA extraction. In order to avoid fecal microbial contaminants in the breast milk samples, extraction of DNA from infant's feces and maternal breast milk samples was performed separately on different occasions. Further, a total of 10 blank DNA extraction controls (one for each batch of DNA extraction from milk samples) were included through all the same steps as in the procedure for the milk samples. DNA was extracted from ~ 200 mg infant feces in random order using the DNeasy PowerLyzer PowerSoil kit (Qiagen, 12855-100) with a few modifications: Bead beating was performed at 30 cycles/s for 10 min (Retsch MM 300 mixer mill) and the initial centrifugation steps were performed at $10,000 \times g$ for 3 min, as recommended for clay matter. Prior to DNA extraction of milk samples, ~ 2 ml was centrifuged at $20,000 \times g$ for 2 min in micro-centrifuge tubes, after which the upper fat layer was removed using a pipette tip and the supernatant discarded by inverting the tube. The pellets were then re-suspended in 750 μl bead solution (from the DNeasy PowerLyzer PowerSoil kit) and transferred to a new tube. Total DNA was extracted in random order from these suspensions using the same protocol as for fecal samples. Concentrations of DNA in both feces and breast milk samples were measured by Qubit[®] dsDNA HS assay (InvitrogenTM, Q32851).

Analysis of human milk oligosaccharides in breast milk and infant feces

Human milk oligosaccharides in breast milk. Mothers collected the entire content of both breasts using a manual breast pump at infant age 5 and 9 months, as described above. The HMOs were analyzed in well-mixed samples of right and left breast at the University of California, San Diego. Human milk was spiked with raffinose (a non-HMO carbohydrate) as an internal standard at the beginning of sample preparation to correct for sample losses during sample processing and allow for absolute oligosaccharide quantification. Oligosaccharides were extracted by high-throughput solid phase extraction over C18 (Hypercarb-96, 25 mg bed weight, Thermo Scientific) and Carbohydrate microcolumns (Hypersep-96 C18, 25 mg bed weight, Thermo Scientific) using a controlled vacuum manifold. Use of high-throughput microcolumns was validated in multiple different ways: (i) establishing parallelism in serial dilutions, (ii) spiking milk with individual HMO standards to determine recovery, and (iii) comparison with direct in-sample derivatization as used by others.³³ Oligosaccharides were fluorescently labeled with 2-aminobenzamide (2AB, Sigma) in a 96-well thermocycler at 65°C for exactly 2 h. The reaction was stopped abruptly by reducing the thermocycler temperature to 4°C . The amount of 2AB was titrated to be in excess to account for the high and variable amount of lactose and other glycans in milk samples. Unreacted 2AB was removed by high-throughput solid phase extraction over silica microcolumns (Hypersep silica, 25 mg bed weight, Thermo Scientific). Labeled oligosaccharides were analyzed by HPLC (Dionex Ultimate 3000, Thermo Scientific) on an amide-80 column (15 cm length, 2 mm inner diameter, 3 μm particle size; Tosoh Bioscience) with a 50-mmol/l ammonium formate–acetonitrile buffer system. Separation was performed at 25°C and monitored with a fluorescence detector at 360 nm excitation and 425 nm emission. Peak annotation was based on standard retention times of commercially available HMO standards (Sigma, Dextra, Elicityl) and a synthetic HMO library³⁴ and offline mass spectrometric analysis on a Thermo LCQ Duo Ion trap mass spectrometer equipped with a Nano-ESI-source. Absolute concentrations were calculated based on HMO standard response curves for each of the annotated HMO (Oligosaccharide detection limit: ~ 20 pmol, dynamic range between 20 and 5000 pmol; milk samples were diluted accordingly). The total concentration of HMOs was calculated as the sum of the annotated oligosaccharides. HMO-bound fucose was calculated on a molar basis. One mole HMO with one fucose residue counted as one mole HMO-bound fucose while one mole HMO with two or more fucose residues counted as two or more moles of HMO-bound fucose. The same was calculated for HMO-bound sialic acid. Maternal Secretor status was determined by the high abundance (Secretor) or near absence (Non-Secretor) of the HMO 2'-fucosyllactose in the respective milk samples with a cutoff of 100 nmol 2FL per ml. HMO diversity was calculated using the Simpson's diversity index.

Human milk oligosaccharides in feces. A fecal slurry (1:1 v/v feces:water) from 50 ± 2 mg feces was mixed with 1362.5 μl of 96% ethanol and spiked with 60 μl of internal standard mixture of seven compounds³⁵ for quality control. The samples were mixed at 60°C for 2 min in a Thermo mixer (Eppendorf, East Lime, USA) at 1400 rpm, centrifuged for 2 min at 14,000 rpm in an eppendorf centrifuge and the supernatants were filtered through a 0.2 μm filter into the wells of a 2 ml 96-well plate. A blank sample (Solvent A) treated similar to the fecal supernatants, an external standard containing 44 biologically relevant metabolites (metabolomics standard)³⁵ and a pooled sample containing equal amounts of each sample were added to unused wells and analyzed repeatedly (every fifteen samples) throughout the sequence for quality control. Quality control samples were used to evaluate possible contamination, and monitoring drifts in mass accuracy, retention time and instrumental sensitivity. The samples were analyzed by ultra performance liquid chromatography (UPLC) coupled with a quadrupole-time of flight mass spectrometer (q-TOF-MS) equipped with electrospray ionization (ESI) (SYNAPT, Waters, Manchester, UK) in positive ionization mode. A reverse phase HSS T3 C18 column (2.1 \times 100 mm, 1.8 μm) coupled with a pre-column (VanGuard HSS T3 C18 column (2.1 \times 5 mm, 1.8 μm)) was used for chromatographic separation. Experimental conditions are given in Supplementary File 1. The data was preprocessed using MZmine version 2.28.³⁶ A subset of samples was used to optimize the preprocessing parameters by estimating, e.g., minimum and maximum peak widths and m/z precision across peak scans. Optimized preprocessing parameters are given in Supplementary File 2. The data matrix was imported into MATLAB R2015b (The Mathworks, Inc., MA, USA). Features that are present in the blanks, early and late eluting features ($RT < 0.3$ and $RT > 6.4$), potential isotope peaks as well as features with implausible masses based on elemental composition were removed from the dataset with an in-house algorithm similar to the one recently

published.³⁷ Features within 0.01 min RT tolerance for the sequence and having correlation coefficients >0.7 were grouped as belonging to the same compound. The compounds of interest were confirmed using the authentic standards of 2'FL, 3'FL, 3'SL, 6'SL, LNT and LNT (Glycom A/S, Hørsholm, Denmark). The authentic standards were run in MS mode with the same experimental conditions but in a separate sequence, hence a 0.02 min tolerance between sequences. Further, MS/MS experiments were performed with the same experimental conditions as for MS scan, except for the collision-induced dissociation energies of 10, 20, and 30 eV. Parent ion and the major fragment masses were searched in the cleaned dataset with 0.02 Da *m/z* and 0.02 min retention time tolerance. Parent ions and the corresponding fragments were grouped together within 0.01 retention time tolerance and at least 0.7 correlation coefficient. Additionally, similarity of peak shapes was inspected for the corresponding fragments of each parent ion. Pooled samples were used for linear correction of signal drift throughout the analysis.³⁸ The major fragment peak intensities observed for each compound 2'FL/3'FL (major fragment *m/z*: 163.059 for [M + H]⁺: 489.182 with retention time (rt): 0.56 min), LNT/LNT (major fragment *m/z*: 366.135 for [M + H]⁺: 708,256 with rt: 0.56 min) and 3'SL/6'SL (major fragment *m/z*: 292.103 for [M + H]⁺: 634.220 with rt: 0.57 min), when running the authentic standards (Supplementary File 3) were used for infant feces samples to estimate the relative abundances of fucosylactoses (2'FL + 3'FL), lacto-N-(neo)tetraoses (LNT + LNT) and sialyllactoses (3'SL + 6'SL) between individuals (Supplementary File 4 and 5).

Quantitative PCR

Total bacterial load in breast milk samples was estimated by quantitative PCR (qPCR) on DNA extracted from these, using universal primers (341F: 5'-CCTACGGGAGGAGCAG-3', 518R: 5'-ATTACCGCGGCTGCTGG-3', final concentration 0.5 μM each, annealing temperature 60 °C) targeting the V3 region of the 16S rRNA gene. Using species/subspecies-specific primers, absolute abundances of *B. longum* subsp. *longum*³⁹ (lon_0274_F: 5'-GAGGCGATGGTC TGAAGTT-3', lon_0274_R: 5'-CCACATCGCCGAGAAGATT-3', final concentration 0.75 μM each, annealing temperature 50 °C), *B. longum* subsp. *infantis*⁴⁰ (Blon0915F: 5'-CGTATTGGCTTTGTACCGCAITTT-3', Blon0915R: 5'-ATCGTGCCGGT-GAGATTAC-3', final concentration 0.75 μM each, annealing temperature 50 °C) and *B. bifidum*⁴¹ (BiBIF-1: 5'-CCACATGATCGCATGTGATTG-3', BiBIF-2: 5'-CCG AAGCTTGCTCCAAA-3', final concentration 0.5 μM each, annealing temperature 60 °C) were estimated by qPCR of DNA extracted from infant fecal samples obtained at age 5 months. Each reaction was performed (in triplicates) with 2 μl template DNA, the specified primer concentrations and 2x SYBR Green I Master Mix solution (LightCycler[®] 480 SYBR Green I Master, Roche). A total of 10 blank DNA extractions controls, extracted in parallel with breast milk samples, were included as negative controls on plates containing DNA extracted from breast milk. Standard curves were generated from triplicate 10-fold serial dilutions of linearized plasmid (containing 10⁰–10⁸ gene copies/μl), constructed by cloning a PCR amplified 199 bp fragment of the 16S rRNA gene (V3-region) of *E. coli* ATCC 25922⁴² for estimation of total bacterial load in breast milk or by triplicate 10-fold serial dilutions of DNA extracted from pure cultures of *B. longum* subsp. *longum* DSM 20219, *B. longum* subsp. *infantis* DSM 20088 or *B. bifidum* DSM 20456 for quantification of these in infant's feces. Plates were run on the LightCycler[®] 480 Instrument II (Roche) with the program including 5 min pre-incubation at 95 °C, followed by 45 cycles with 15 s at 95 °C, 15 s annealing at 50–60 °C and 15 s at 72 °C and a subsequent melting curve analysis including 5 min at 95 °C, 1 min at 65 °C and continuous temperature increase (ramp rate 0.11 °C/s) until 98 °C. Data were analyzed with the LightCycler[®] 480 Software (v1.5) (Roche).

Library preparation and 16S rRNA gene amplicon sequencing

PCR amplification for library preparation and subsequent sequencing on the Ion Torrent platform was performed with DNA extracted from infant's fecal material (*n* = 77), maternal breast milk (*n* = 87) and blank DNA extraction controls (*n* = 6) essentially as previously reported.⁴³ Briefly, amplification of the V3-region of the 16S rRNA gene was performed as follows: For fecal samples the PCR mixture consisted of 0.2 μl Phusion High-Fidelity DNA polymerase (ThermoFisher Scientific, F-553L), 4 μl HF-buffer, 0.4 μl dNTP (10 mM of each base), 1 μM forward primer (PBU; 5'-A-adapter-TCAG-barcode-CCTACGGGAGGAGCAG-3') and 1 μM reverse primer (PBR; 5'-trP1-adapter-ATTACCGGCTGCTGG-3') and 1–5 ng community fecal DNA in 20 μl total reaction volume. Both primers (TAG Copenhagen A/S) were linked to sequencing adaptors and the forward primer additionally contained a unique 10 bp barcode (Ion Xpress™ Barcode Adapters) for each sample. The PCR program consisted of initial

denaturation for 30 s at 98 °C, followed by 24 cycles of 98 °C for 15 s and 72 °C for 30 s, and lastly 72 °C for 5 min to allow final extension before cooling to 4 °C. For milk and blank samples the PCR was performed with 0.08 μl Accuprime Taq polymerase (ThermoFisher Scientific, 12346086), 2 μl 10x Accuprime PCR buffer II, 0.5 μM of each of the abovementioned primers and 1 μl of DNA extracted from milk or blank controls in a total volume of 20 μl. The PCR program consisted of initial denaturation for 120 s at 94 °C, followed by 40 cycles of 94 °C for 20 s, 58 °C for 20 s and 68 °C for 30 s, and lastly 68 °C for 5 min to allow final extension before cooling to 4 °C. No-template controls were included for each PCR run, all resulting in less 0.05 ng/μl. The PCR products were purified by use of HighPrep™ PCR Magnetic Beads (MAGBIO[®], AC-60005) with a 96-well magnet stand (MAGBIO[®], MyMag 96), according to the manufacturers recommendations. DNA quantity was measured using Qubit[®] dsDNA HS assay (Invitrogen™, Q32851) and samples were pooled to obtain equimolar libraries and sequenced on three separate chips using the Ion OneTouch™ and Ion PGM systems with a 318-Chip v2 incorporating the Hi-Q chemistry in 200 bp runs. Sequence data has been deposited at NCBI's Sequence Read Archive under BioProject number PRJNA553872.

Bioinformatics, data analysis, and presentation

Raw-sequencing reads were de-multiplexed according to barcode and trimmed to remove barcodes and primers, maintaining only reads containing both forward and reverse primers and discarding reads below 125 bp or above 180 bp using the CLC Genomic Workbench (v8.5) software (CLCbio, Qiagen, Aarhus, DK). An amplicon sequence variant (ASV) table including all samples was generated in R (R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>) using the DADA2 pipeline v1.12.⁴⁴ Briefly, the reads were quality filtered (maxEE = 2, maxN = 0, truncQ = 2), denoised using pooled data and increased homopolymer gap penalty and band size as recommended for Ion Torrent reads. Chimeric sequences were removed and the resulting ASVs were assigned taxonomy using the RDP 16S rRNA database.⁴⁵ In addition, for selected ASVs, nBLAST⁴⁶ analysis was performed in order to confirm taxonomy assigned by the RDP database, and in some cases to explore species level classifications not obtained with RDP. Species level classification/grouping was only adopted/confirmed if the top BLAST hit showed 100% homology to the ASV sequence, and no other species showed 100% homology. Potential reagent contaminating ASVs in the breast milk samples were identified based on the frequency method, using qPCR-estimated bacterial load, within the *decontam* R package software (v1.6.0).⁴⁷ ASVs identified as contaminants (72 out of 575) were then removed from the complete ASV table (with the exception of two ASVs; ASV_7 and ASV_24, which were kept due to their abundance in infant feces, and thus non-likelihood to represent reagent contaminants) prior to downstream community analysis with the QIIME2 software.⁴⁸ Within the QIIME2 environment ASVs assigned to Cyanobacteria/chloroplast or with a frequency <100 reads across all samples were removed and the *core diversity metrics* function was applied with a rarefaction depth of 18,000 reads per sample to generate phylogenetic (weighted and unweighted UniFrac) and non-phylogenetic (Bray–Curtis and Jaccard) distance matrices and principal coordinates, as well as alpha diversity measures (Shannon index, Observed ASVs and Evenness index). The Jaccard similarity index was calculated as (1-Jaccard distance). Using the *rarefy table* function, the ASV table was rarefied to 18,000 reads per sample and collapsed to higher taxonomical levels using the *taxa collapse* function. Relative abundances at all levels were calculated by total sum scaling. Absolute abundances in the milk microbiota dataset was calculated by multiplying qPCR-estimated total bacterial load with the relative abundances for each taxa. A heatmap was generated using the *heatmap.2* function within the R package *gplots* (v3.1.1), using euclidean distances and UPGMA clustering on genus level relative abundance data with a cutoff of 1% average relative abundance within infant feces or maternal hindmilk at either 5 or 9 months of age. Color gradient PCoA plots were generated using the *ggplot* function within the R package *ggplot2* (v3.3.3). All other figures were generated using the GraphPad Prism Software v 8.3 (GraphPad Software Inc., La Jolla, CA).

Statistics

Statistics were performed with QIIME2, R or GraphPad Prism Software. For non-paired data Mann–Whitney *U* or unpaired *t*-tests were used (depending on normality distribution), whereas Wilcoxon-signed rank test or paired *t*-test were used for paired data. For comparisons of more than

two groups one-way analysis of variance (ANOVA) with Fisher's least significant difference (LSD) post hoc test were used. Correlation between variables was assessed by Spearman's rank correlation analyses. When appropriate, *p*-values were corrected for multiple testing by the false discovery rate (FDR) method.⁴⁹ ANOSIM test (*beta diversity group significance* function, within QIIME2) were used to assess differences between groups in beta diversity distance measures, whereas ADONIS tests (*adonis* function, within R package *vegan* v2.5-7) were used to evaluate the impact of continuous variables on beta diversity distance measures, in both cases using 999 permutations.

RESULTS

Study participant characteristics

Mother–infant dyads included in this study were enrolled as participants in the SKOT III cohort, including two groups of infants experiencing either normal weight gain (NWG) or excessive weight gain (EWG) during exclusive breastfeeding and additionally four participants that were previously excluded due to non-conform growth patterns.³² The 34 breastfed infants were all born at term (79.3% vaginally), there was an equal gender distribution and 48.5% the infants had older siblings. Thirteen (38.2%) of the infants were included in the group with excessive weight gain. Antibiotics exposure at time of sampling was limited to two mother–infant dyads, of which one mother received oral antibiotics at the first sampling (infant age 5–6 months) and one infant received oral antibiotics at second sampling (age 9 months). In terms of maternal secretor status, 14.6% of the mothers had highly reduced levels of α 1-2-fucosylated HMOs in the breast milk, and were thus classified as non-secretors. Infant feces and maternal breast milk samples (both foremilk and hindmilk samples, collected before and after nursing the infant, respectively) were obtained at infant age 5.8 ± 0.4 and 9.0 ± 0.3 months (Tables S1–2).

Breast milk samples contain a microbiota distinguishable from negative controls

Since human breast milk is low in microbial biomass,^{6,50} care must be taken during DNA extraction and PCR-based amplification when profiling the bacterial communities in this sample type, due to the risk of reagent or cross-contamination.^{51,52} Therefore, we initially measured bacterial load in both foremilk and hindmilk collected at infant age 5 and 9 months, as well as blank DNA extraction controls (Fig. 1a). The median bacterial loads were at least one order of magnitude higher than those found in DNA extraction negative controls (Fig. 1a). Interestingly, hindmilk samples, taken after feeding the infant at the breast, had a significantly higher bacterial load compared to foremilk samples (taken before the infant was latched at the breast) at both infant age 5 and 9 months, suggesting that a significant transfer of infant oral bacteria occurs during breastfeeding. Bacterial load increased significantly over time in hindmilk, but not in foremilk. We observed clear correlations between the bacterial load in foremilk and hindmilk, obtained from the same individuals, at both 5 months ($\rho = 0.70$, $p < 0.0001$) and 9 months ($\rho = 0.55$, $p = 0.010$) of age (Fig. 1b, c), supporting that the data represents biological variation between individual milk samples, rather than stochastic reagent contamination. We also found significant correlations across time (Fig. S1a, b), within paired samples of hindmilk ($\rho = 0.54$, $p = 0.01$) but not foremilk ($\rho = 0.18$, $p = 0.43$). Next, we profiled the microbial communities residing in the milk samples using 16S rRNA amplicon sequencing. In order to remove any reads potentially originating from reagent contamination we used the *decontam* pipeline,⁴⁷ which filters out amplicon sequence variants (ASVs) with abundances across samples inversely related to the measured bacterial load, since these are highly likely to represent contaminants, as also recently demonstrated for breast milk samples.⁵¹ This procedure removed 57.5% of the reads in the DNA extraction negative control

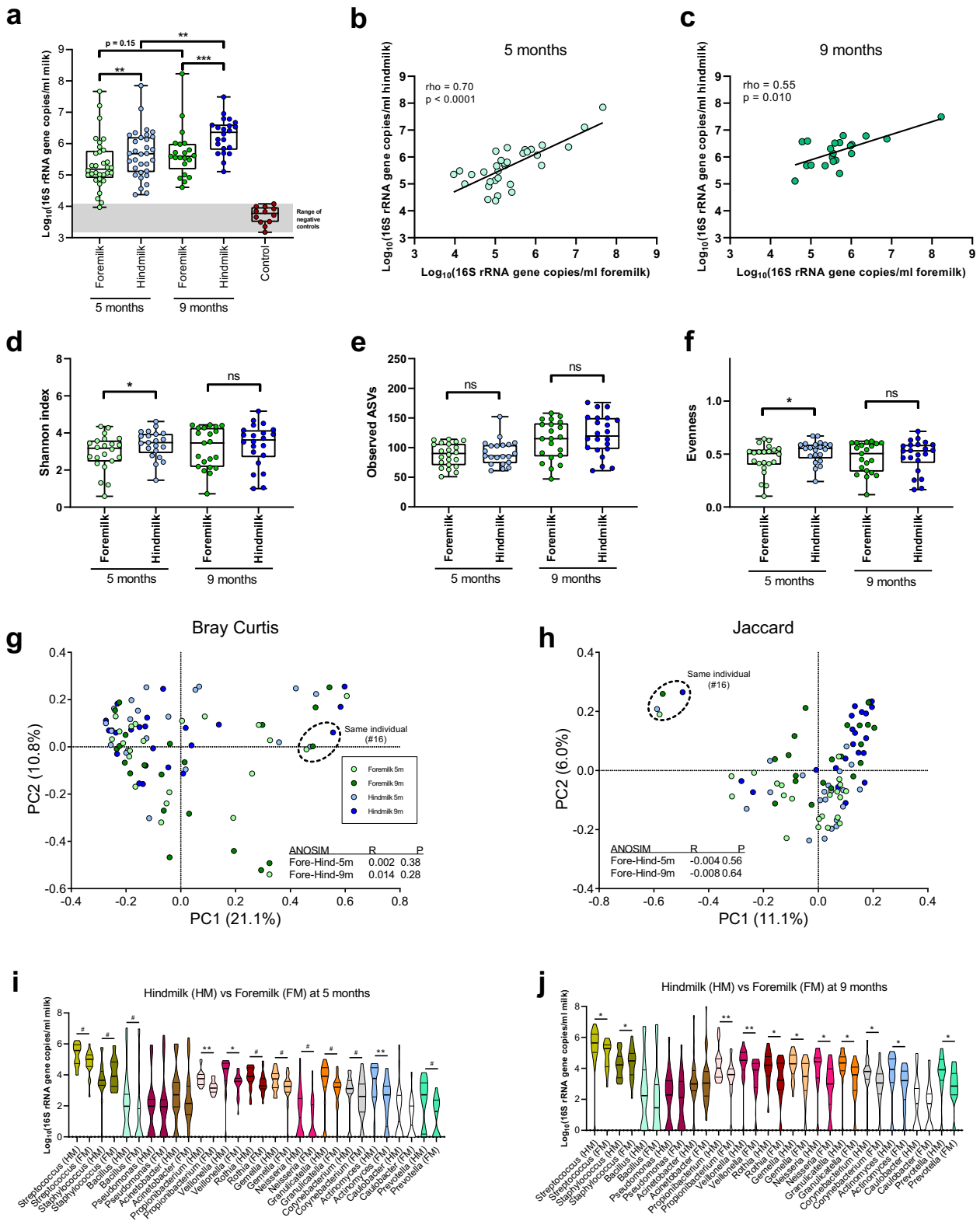
samples, but only 2.8–11.2% of the reads in the milk samples and <0.5% of the reads in corresponding infant fecal samples (Fig. S2a). The community composition in the remaining reads of the negative control samples was different from those in both maternal milk and infant feces samples (Fig. S2b) and also clustered separately in principal coordinate analysis (PCoA) plots (Fig. S2c–f).

Foremilk and hindmilk samples contain a similar compositional microbiota differing mainly in absolute abundances of skin-associated versus infant oral-associated bacteria

We found no significant differences in maternal foremilk or hindmilk between infants experiencing EWG and NWG in terms of bacterial load or composition (Fig. S3). We did, however, find a slight increase in alpha diversity measures at 5 months in the EWG group compared to NWG in foremilk, yet this was not consistent as the reverse was observed for hindmilk (Fig. S3). We then compared the microbial communities in the fore- and hindmilk samples irrespective of weight-gain group and found slightly, but significantly higher Shannon index and evenness in the hindmilk compared to foremilk at infant age 5 months but not 9 months (Fig. 1d–f). The hindmilk and foremilk microbial communities were not distinguishable based on Bray–Curtis and Jaccard distance matrices, neither at infant age 5 nor 9 months (Fig. 1g, h). Interestingly, both foremilk and hindmilk samples from the only mother receiving oral antibiotics (#16) at the time of sampling at infant age 5 months displayed a distinct microbiota compared to all other milk samples and this difference was sustained at infant age 9 months (Fig. 1h). Comparison of genus level relative abundances between the two milk sample types (Table S3) revealed that although generally similar, foremilk consistently (at both infant age 5 and 9 months) contained a pronounced and significantly higher relative abundance of *Staphylococcus* (FDR $p < 0.00001$), with the differentially abundant ASVs classified as *S. epidermidis* or *S. hominis* (Table S4), both species recognized as common skin-commensals.⁵³ No other bacterial taxa were differentially relative abundant at infant age 5 and 9 months after FDR correction (Tables S3–4). However, when calculating absolute abundances (by multiplying relative abundances for each taxa by qPCR-estimated total bacterial load) we found that actually most bacterial genera, notably many of which are often found as abundant members of the oral microbiota, e.g., *Streptococcus*, *Veillonella*, *Rothia*, *Neisseria*, *Prevotella*, *Gemella*, *Actinomyces* and *Granulicatella*,⁵⁴ were significantly or borderline significantly more abundant in hindmilk (Fig. 1i, j). In contrast, non-oral-associated bacteria such as *Acinetobacter*, *Pseudomonas* and *Caulobacter*⁵⁴ did not differ significantly. The only genera with higher absolute abundance in foremilk was *Staphylococcus*. Collectively, these observations suggest that a significant amount of infant oral-associated bacteria are transferred to the breast milk when the infant is breastfed and that foremilk harbors a higher proportion of staphylococci likely originating from the breast skin environment. Since hindmilk samples contained a higher bacterial load (thus reduced risk of contaminating reads) and since it seems to be a better representative of the bulk milk-microbiota that the infants receive,¹⁸ we chose to focus the remainder of our analyses on the hindmilk samples.

The microbiota of maternal breast milk and infant feces undergo temporal development and are distinct but overlapping in community composition

The bacterial alpha diversity in infant fecal samples increased over time in terms of Shannon Index, number of observed ASVs and evenness, whereas an increase in hindmilk bacterial alpha diversity was only found in terms of the number of observed ASVs (Fig. 2a–c). Temporal changes in the microbial composition of both infant feces and maternal hindmilk were also observed



with a clear difference between these two compartments (Fig. 2d, e). We found a higher degree of microbial community similarity at 5 versus 9 months within individuals than between individuals in both compartments (Fig. 2f), highlighting the individuality of both these microbial ecosystems. Cluster analysis of all samples, based on relative abundances of dominant genera, supported both

compartmental and temporal differences and further indicated three main clusters of bacterial genera within the samples (Fig. 3). Cluster 1 included genera often found in both sample types, e.g., *Bifidobacterium*, *Streptococcus* and *Veillonella*, suggesting that milk-to-gut transmission of bacteria occurs primarily within these genera. Cluster 2 contained genera found mostly or only

Fig. 1 Bacterial load, diversity, and composition in foremilk versus hindmilk. **a** Boxplot (median; IQR; range) of bacterial load in foremilk and hindmilk at infant age 5 and 9 months, as well as DNA blank extraction controls. Statistical significance was evaluated by paired *t*-tests. Asterisks indicate $**p < 0.01$ and $***p < 0.001$. **b, c** Scatter dot plots of bacterial load in foremilk versus hindmilk at infant age 5 (**a**) and 9 (**b**) months. A linear regression line was fitted to the log-transformed data and the relationship was evaluated by Spearman's rank correlation analyses. **d–f** Boxplots (Median; IQR; Range) of bacterial alpha diversity measures, Shannon index (**d**), Observed ASVs (**e**) and Evenness index (**f**), comparing foremilk and hindmilk at both infant age 5 and 9 months. Statistical significance was evaluated by paired *t*-tests and asterisks indicate $*p < 0.05$. **g, h** PCoA plots based on Bray–Curtis (**g**) and Jaccard (**h**) dissimilarities, including foremilk and hindmilk samples at both infant age 5 and 9 months, with statistically differential clustering evaluated by ANOSIM tests. **i, j** Violin plots of the absolute abundances of the major genera (average relative abundance cutoff across all milk samples of 1.0%) comparing hindmilk to foremilk at infant age 5 months (**i**) and 9 months (**j**). Thick black lines indicate median and thin black lines 25 and 75 percentiles. Statistical difference was evaluated by Wilcoxon-signed rank tests and symbols indicate $^{\#}FDRp = 0.05–0.1$, $*FDRp < 0.05$, and $**FDRp < 0.01$.

in milk (e.g., *Staphylococcus*, *Rothia*, *Gemella*, *Corynebacterium*, *Granulicatella*, *Actinomyces*, *Propionibacterium*, *Prevotella*, *Neisseria*, *Caulobacter*, *Pseudomonas*, *Acinetobacter* and *Bacillus*), although both *Actinomyces* and *Granulicatella* were also prevalent, but not abundant, in infant feces. Of note, four genera (*Bacillus*, *Acinetobacter*, *Pseudomonas* and *Caulobacter*) showed highly varying abundance patterns in hindmilk and dominated in a few mothers, but were almost never detected in infant feces. In mother–infant pair #16, where maternal oral antibiotics were consumed at the time of sampling at 5 months, we observed a lack of milk streptococci and instead a strong dominance by *Acinetobacter* and *Bacillus* or *Pseudomonas* at both the 5 and 9 month samplings. The implicated infant had a pronounced depletion of fecal *Bifidobacterium* and was instead dominated by *Klebsiella* at the 5 month sampling. Cluster 3 comprises genera found mainly or only in infant feces (e.g., *Lachnospiraceae*, *Clostridiaceae*, *Roseburia*, *Fecalibacterium*, *Blautia*, *Anaerostipes*, *Collinsella*, *Klebsiella*, *Bacteroides* and *Parabacteroides*) although a few genera in this cluster (e.g., *Escherichia*, *Enterococcus* and *Lactobacillus*) were also occasionally found in hindmilk samples. In infant feces, the relative abundance of *Bifidobacterium* and *Streptococcus* decreased with age (Fig. S4), whereas genera/families from the *Clostridiales* order increased with age (*Ruminococcus*, *Roseburia*, *Faecalibacterium*, *Clostridiaceae*, *Anaerostipes*, *Blautia* and *Lachnospiraceae*). The temporal changes in maternal hindmilk were more subtle and mainly included significant increases in relative abundance of *Prevotella* and *Neisseria* (Fig. S4).

Maternal breast milk taxa are shared with the infant gut

In order to investigate the potential transmission of bacterial taxa from maternal breast milk to the infant gut, we compared microbial community similarities between related and unrelated mother/infant pairs. At infant age 5 and 9 months, the microbiota similarity was higher within mother–infant pairs than between mothers and unrelated infants for both hindmilk (Fig. 4a) and foremilk (Fig. S5a), suggesting that maternal breast milk microbes are transmitted to the infant gut. On average, 18.8% of the ASVs detected in hindmilk were shared with paired infant feces at 5 months of age, which dropped to 15.3% at 9 months of age (Fig. 4b). Very similar results were obtained using the foremilk microbiota (Fig. S5b), suggesting that there is a slight decrease over time in the fraction of the milk microbiota that is transmitted to the infant gut. We then calculated the percentages of total ASVs detected in infant feces that were shared with hindmilk and found this to be 33.2% at 5 months, but this number significantly decreased to 22.6% at 9 months of age (Fig. 4b). Again, this was very consistent when using the foremilk samples (Fig. S5b). Thus, the contribution of milk microbes to the gut microbial community decreases when the infants start eating other types of food that may also contribute to the gut microbiota. In fact, the actual number of shared ASVs did not significantly change over time (Fig. S6), thus the temporal decrease in percent of shared ASVs reflect the increased richness observed in both milk and infant feces microbiota over time (Fig. 2b), rather than a decrease in

actual number of shared ASVs. The ASVs that were most frequently shared ($n = 40$, Table S5) between breast milk and infant feces belonged to the genus *Streptococcus*, followed by *Veillonella*, *Actinomyces*, *Haemophilus*, *Bifidobacterium*, *Gemella* and *Granulicatella* (Fig. 4c and Fig. S5c). At infant age 5 months (Fig. 4d and Fig. S5d), we found that sharing of various *Streptococcus* ASVs (e.g., *S. mitis* group, *S. salivarius*, *S. peroris* and *S. lactarius*) was highly prevalent (>80% of mother–infant pairs). Interestingly, we found two *Bifidobacterium longum* ASVs (ASV_1 and ASV_7) to be shared in >50% of mother–infant pairs. Both had a low relative abundance in hindmilk, but were highly abundant in infant feces (Fig. 4d). We also found evidence of sharing of typical skin-microbes such as *P. acne* (ASV_5) and *S. epidermidis* (ASV_3), although both were not abundant in infant feces. At infant age 9 months, ASVs within the genus *Veillonella* (classified as, e.g., *V. dispar*, *V. atypica* and *V. rogosae*) were more prevalently shared (Fig. 4e and Fig. S5e). Strikingly, while most of the shared ASVs were found in higher proportions in hindmilk compared to feces (e.g., ASVs within *Streptococcus*, *Staphylococcus*, *Gemella*, *Granulicatella* and *Actinomyces*), a few had a much higher relative abundance in the feces than in the hindmilk (e.g., *Bifidobacterium*, *Enterococcus*, *Lactobacillus*, *Escherichia*, and some *Veillonella* ASVs), indicating that the latter were selected for by the gut ecosystem. Given the risk of low abundant hindmilk ASVs to represent reagent contamination, we checked the presence and absolute abundance of the 40 most frequently shared ASVs in sequenced DNA extraction controls, and found that all, except two (ASV_7 and ASV_13), were either not detected in controls ($n = 19/40$) or markedly less abundant ($n = 19/40$) in controls compared to milk (Fig. S7). Owing to our special interest in the two shared *Bifidobacterium longum* ASVs we further investigated these and found that ASV_1, but not ASV_7, was significantly more abundant ($p = 0.009$) in milk samples compared to controls and correlated significantly between paired hindmilk and foremilk samples ($\rho = 0.31$, $p = 0.048$), suggesting genuine presence of this ASV in breast milk. Even though ASVs belonging to the genera *Bacillus*, *Acinetobacter*, and *Pseudomonas* were frequently found in breast milk and were even highly dominating in some samples (Fig. 3), they were never shared with the infant gut, suggesting that they are not able to survive and establish in the infant gut ecosystem. Although typical gut colonizers such as *Bacteroides*, *Parabacteroides* and several genera within *Ruminococcaceae* and *Lachnospiraceae* were sporadically detected in maternal breast milk samples (Fig. 3), specific ASVs within these genera were never shared with infant feces, pointing towards other sources than breast milk for their establishment in the infant gut. Furthermore, we consistently noted a very low amount of shared ASVs in the mother–infant pair where the mother was in treatment with antibiotics at first sampling, even at infant age 9 months (Fig. S6). Altogether, these data suggest that some (incl. *Streptococcus*, *Veillonella*, and *Bifidobacterium* spp.), but not all, maternal breast milk microbes are transmitted to the infant gut and that breast milk is an unlikely source of fecal *Bacteroides*, *Lachnospiraceae*, and *Ruminococcaceae* species.

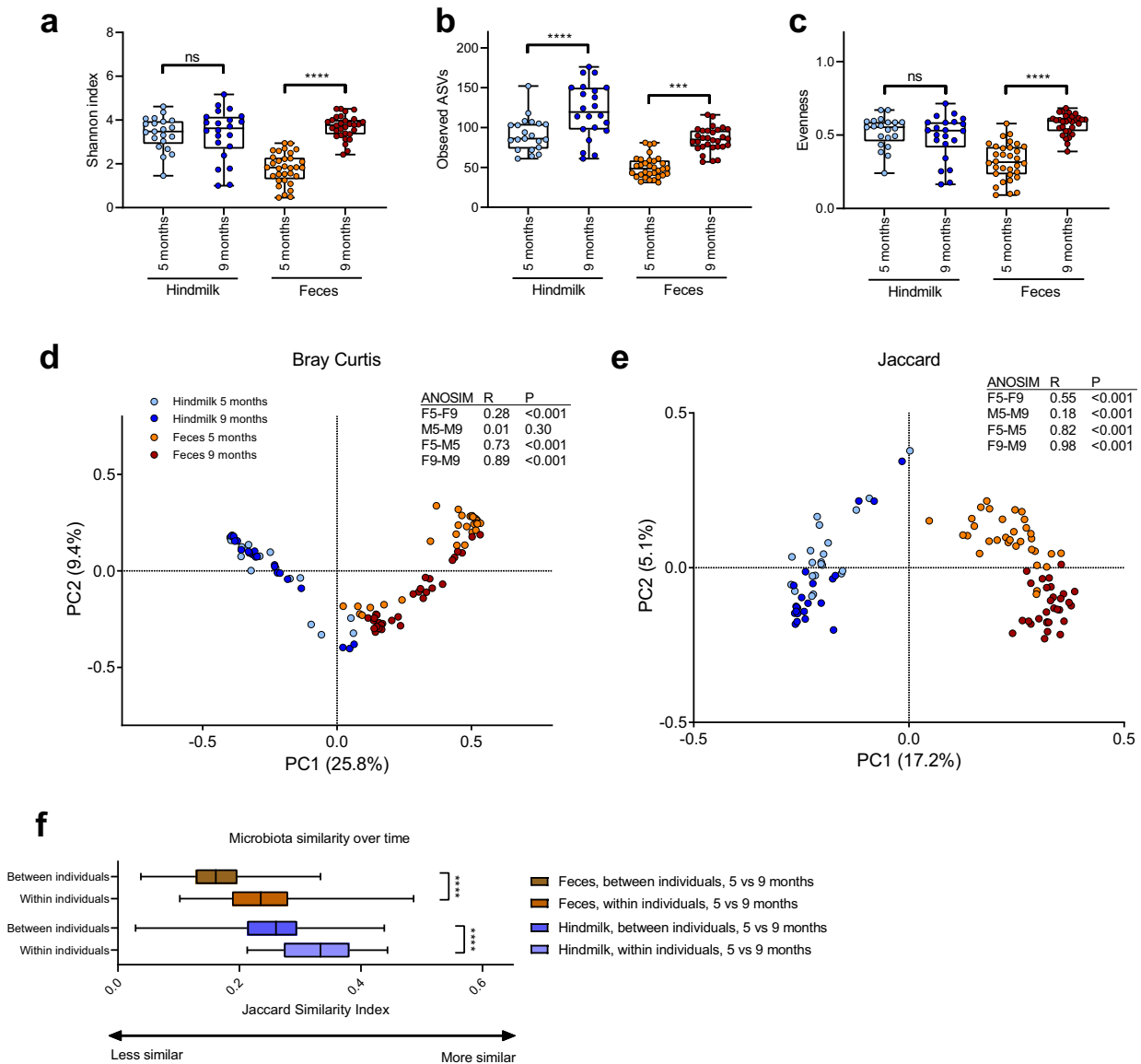


Fig. 2 Temporal development of the microbiota in hindmilk and infant's feces. **a–c** Boxplots (Median; IQR; Range) of bacterial alpha diversity measures, Shannon index (**a**), Observed ASVs (**b**), and Evenness index (**c**) in hindmilk and infant's feces at infant age 5 and 9 months. Statistical significance was evaluated by paired *t*-tests, with asterisks indicating ****p* < 0.001 and *****p* < 0.0001. **d, e** PCoA plots based on Bray–Curtis (**d**) and Jaccard (**e**) dissimilarities, comparing across both sample type and age, with statistical differentially clustering evaluated by ANOSIM tests. **f** Boxplots (median; IQR; range) of microbiota similarity over time in hindmilk and infant's feces, comparing between to within individuals. Statistical significance evaluated by Mann–Whitney test, with asterisks indicating *****p* < 0.0001.

Secretor status and human milk oligosaccharide composition in breast milk show limited correlation with gut microbiota composition

Maternal secretor status did not appear to affect the overall infant gut microbial community (Fig. S8a, b). Since *in vitro* studies have demonstrated that specific *Bifidobacterium* spp. (e.g., *B. longum* subsp. *infantis* and *B. bifidum*) are major consumers of HMOs,¹¹ including 2FL,⁵⁵ and an impact of secretor status on the abundance of this genus has previously been reported,²³ we expected that specific species within this genus would be affected by maternal secretor status. However, we found no significant differences in relative abundances of any of the *Bifidobacterium* ASVs or other ASVs in feces of infants with secretor mothers versus non-secretor mothers (Tables S6 and 7). Neither HMO diversity nor specific HMO

concentrations (measured in mixtures of combined content of breast milk from both breasts), including cumulative fucosylated, sialylated and total HMOs, associated with the overall infant fecal microbiota at age 5 months (Table S8) nor specific genera or ASVs (Tables S9 and 10) after correction for multiple testing, with the exception of *Escherichia*/ASV_13 (*Escherichia* spp.) correlating positively with 6'SL (FDR_p < 0.02) and a tendency for negative correlation with fucosylated HMOs (FDR_p < 0.1). These results imply that secretor status and the specific HMO composition of breast milk may have limited influence on the gut microbiota of the infant, at least at this particular age of 5 months. This could possibly be due to excess HMO amounts in the infant gut and that specific individual HMOs may be redundant for broad-range HMO-utilizers such as *B. bifidum* and *B. longum* subsp. *infantis*.

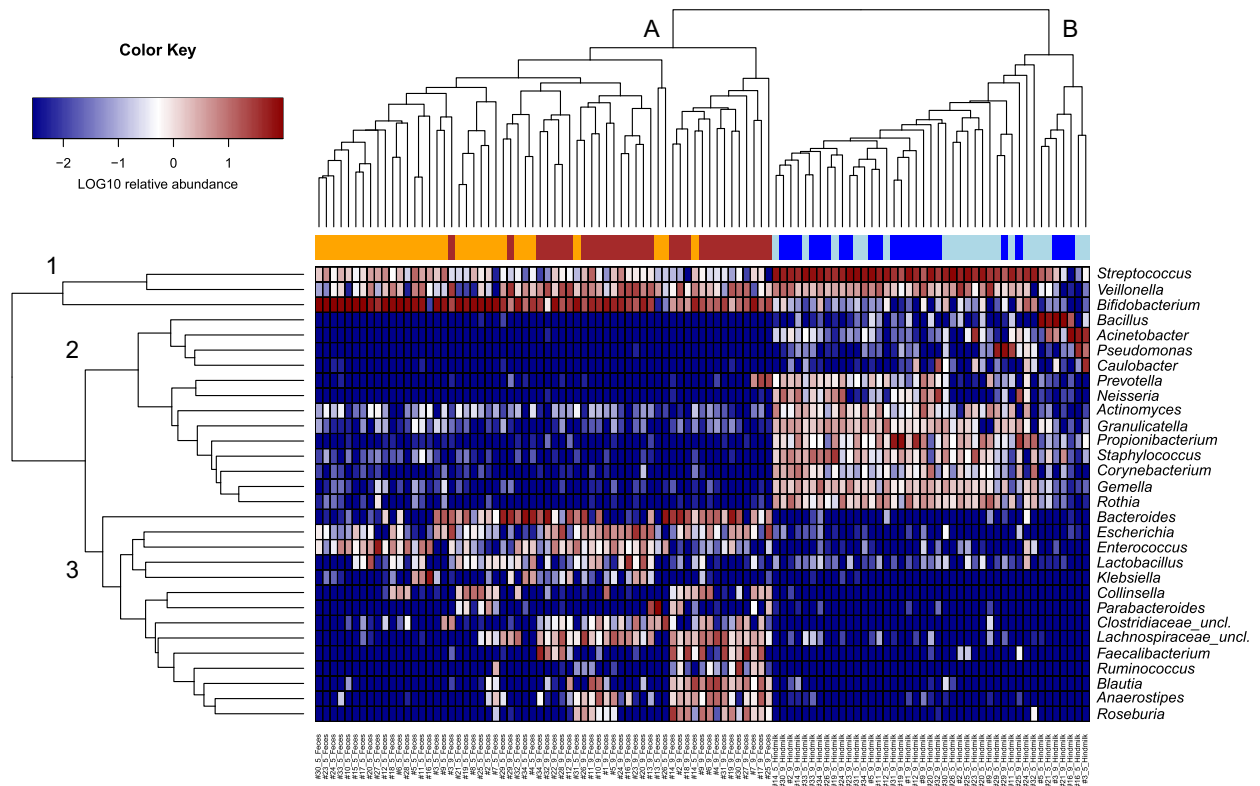


Fig. 3 Clustering of bacterial composition in hindmilk and infant's feces. Heatmap of relative abundances of major genera (>1% average relative abundance in either milk or feces) detected in hindmilk and infant's feces. Samples cluster according to sample type, **A** infant's feces and **B** hindmilk. Genera cluster into (1) bacterial genera common to both sample types, (2) bacterial genera most common to hindmilk samples, and (3) bacterial genera most common to infant's fecal samples. Orange; infant's feces at infant age 5 months, dark-red; infant's feces at infant age 9 months, light-blue; maternal hindmilk at infant age 5 months, dark-blue; maternal hindmilk at infant age 9 months.

Fecal residuals of human milk oligosaccharides associate with the infant gut microbiota composition and explains the abundance of HMO-degrading *Bifidobacterium* species

In order to investigate in vivo microbial utilization of HMOs, we compared the fecal relative abundance of fucosyllactoses (2FL+3FL), sialyllactoses (3SL+6SL) and lacto-N-(neo)tetraose (LNT+LNnT) with the infant fecal microbiota at age 5 months. Relative abundances of all three categories of HMOs were significantly associated with microbiota composition (Fig. 5a–c and Table S11). Further analyses revealed that several of the most abundant/prevalent genera and ASVs detected at age 5 months correlated with the relative abundances of HMOs (Table S12 and 13), including *Bifidobacterium*, *Bacteroides*, *Lactobacillus*, and *Collinsella* (negative associations), as well as *Streptococcus* and *Klebsiella* (positive associations). We focused hereafter on *Bifidobacterium* ASVs, due to their dominance in this cohort, representing on average 69% of the fecal communities and being the dominant genera (>50% of the total community) in 78% of the infants at age 5 months. Among ASVs belonging to *Bifidobacterium*, only ASV_1 (*B. longum*, average relative abundance 54%) and ASV_18 (*B. bifidum*, average relative abundance 2%) correlated negatively with HMO abundances, while the others, including ASV_7 (also assigned to *B. longum*, average relative abundance 5%) did not (Table S13). To differentiate between *B. longum* subsp. *longum* and *B. longum* subsp. *infantis* we performed qPCR using subspecies-specific primers^{39,40} and compared this data to the ASV relative abundances. We found significant positive correlations between ASV_1 and *B. longum* subsp. *infantis* ($\rho = 0.57$, $p = 0.0006$) and between ASV_7 and *B. longum* subsp. *longum* ($\rho = 0.45$, $p = 0.0095$). We also performed qPCR using *B. bifidum*-specific primers and were able to validate the taxonomy of ASV_18, which correlated

significantly with *B. bifidum* ($\rho = 0.82$, $p < 0.00001$). Indeed, qPCR-estimated absolute abundances of *B. longum* subsp. *infantis* and *B. bifidum* show significant negative associations with all three HMO categories (Fig. 5d–i), also when excluding infant's from non-secretor mothers in case of 2FL+3FL. Thus, largely in agreement with the literature on genomic potential to degrade HMOs and in vitro evidence of HMOs utilization,²⁸ we find that specific *Bifidobacterium* species associate with HMOs in feces, suggesting in vivo HMO degradation and providing a rationale for their dominance in the infant gut.

DISCUSSION

Our study overall revealed a high degree of similarity in microbial composition of foremilk and hindmilk samples obtained, respectively, before and after breastfeeding the infant. However, interestingly, the hindmilk contained a significantly higher bacterial load at both infant age 5 and 9 months, suggesting that transfer of infant oral bacteria to the milk occurs during breastfeeding. Indeed many oral-associated bacteria such as *Veillonella*, *Streptococcus*, and *Rothia*^{54,56} showed higher absolute abundance in hindmilk than in foremilk at both infant age 5 and 9 months. Conversely, we consistently found a higher relative and absolute abundance of *Staphylococcus* in the foremilk, likely originating from the areolar skin microbiota, which is dominated by this genus.¹⁹ These findings are supported by a recent study comparing the milk microbiota of breastfeeding mothers who either nursed their infant by latching or provided breast milk in a bottle (never-latched).¹² Here, the authors found a strong dominance of *Staphylococcus*

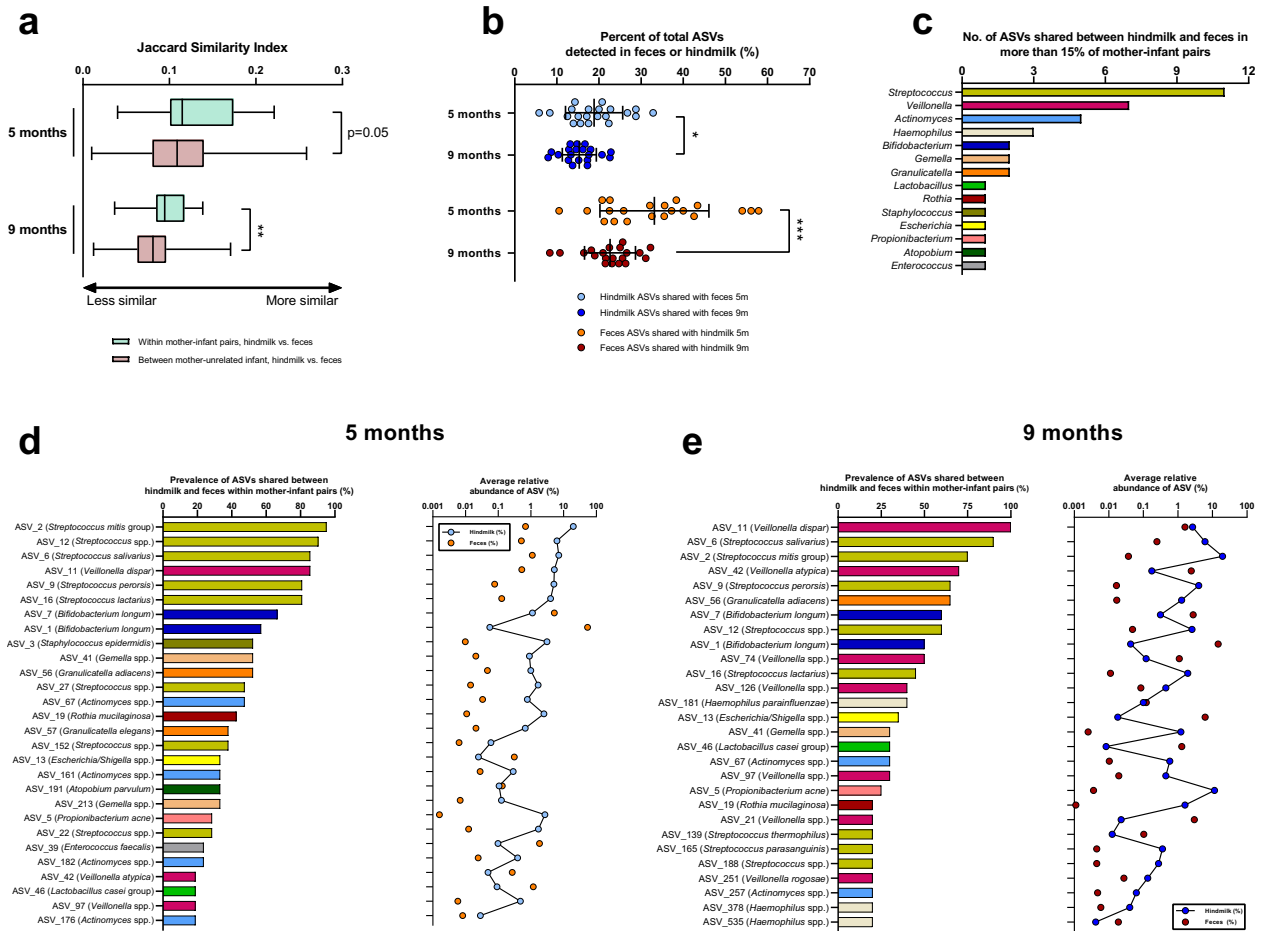


Fig. 4 Sharing of bacterial taxa between hindmilk and infant's feces. **a** Boxplots (median; IQR; range) of microbiota similarity between hindmilk and infant's feces, comparing within mother–infant pairs to between mothers and unrelated infants at both 5 and 9 months of infant age. Statistical significance was evaluated by unpaired *t*-tests, with asterisks indicating $**p < 0.01$. **b** Scatter dot plots of percentage of ASVs shared between hindmilk and infant's feces at 5 and 9 months of infant age. Line and error bars indicate average \pm sd. Statistical significance was evaluated by paired *t*-tests with asterisks indicating $*p < 0.05$ and $***p < 0.001$. **c** Barplots of bacterial genera with most ASVs shared between hindmilk and infant's feces. Only ASVs shared in at least 15% of all mother–infant pairs at either 5 or 9 months of infant age were included. **d, e** Bar and dot plots of prevalence (within mother–infant pairs) and average relative abundance of ASVs shared between hindmilk and infant feces, in at least 15% of mother–infant pairs at **d** 5 months and **e** 9 months of infant age.

in breast milk when nursing did not include physical latching, whereas *Streptococcus* and *Veillonella* only appeared with latching.¹² Thus, it is likely that *Staphylococcus* spp. spread to breast milk primarily from the areolar/nipple skin, whereas breast milk-associated *Streptococcus* and *Veillonella* originate from infant oral cavity.

We observed a clear overlap between hindmilk and infant feces with regards to ASVs classified as *Streptococcus*, *Veillonella*, and *Bifidobacterium*, and found that ~33 and 23 percent of the infant fecal ASVs were shared with the maternal milk at infant ages 5 and 9 months, respectively. Importantly, this does not seem to imply that fewer bacteria are transmitted from milk to gut over time, since we found no differences in the absolute number of ASVs shared within dyads between 5 and 9 months of infant age. However, it seems reasonable that the relative contribution of breast milk microbes to the developing infant gut microbiota decreases gradually over time as the infants is exposed to and colonized with more microbes from diet and the environment (as reflected by the increase in richness from 5 to 9 months of age) and as the diet-related-selective forces (nutrient availability) driving microbial establishment in the gut changes with progression of complementary feeding. Within

the ASVs most frequently shared between breast milk and infant feces, *S. mitis* group (e.g., *S. mitis*, *S. oralis* and *S. infantis*), and *S. salivarius* are common oral cavity inhabitants,⁵⁷ *S. peroris* was first isolated from the infant pharynx⁵⁸ and *S. lactarius* from human breast milk.⁵⁹ *Veillonella* spp. shared such as *V. atypica*, *V. dispar*, and *V. rogosae* are common and abundant members of the oral cavity.⁶⁰ Further, we found sharing of ASVs belonging to *Granulicatella* spp. (e.g., *G. elegans* and *G. adiacens*), *Rothia mucilaginoso*, *Actinomyces* spp., *Gemella* spp. and *Haemophilus* spp., which are also all common oral microbes in infants and toddlers.^{54,61} We cannot exclude the possibility that the presence of these typical oral species in the infant gut results mainly from direct oral-to-gut transmission in the infant,⁶² rather than indirect (through maternal milk). However, it seems plausible that the large amount of bacteria ingested daily contained in the milk, constitutes at least a contributing route for bacteria of suspected oral origin to seed the infant gut. Also, we did find sharing of ASVs belonging to typical skin microbes such as *Staphylococcus epidermidis* and *Propionibacterium acne*. Although typical strict anaerobic gut bacteria, such as *Bacteroides*, *Parabacteroides*, *Faecalibacterium*, *Blautia*, *Roseburia*, and *Anaerostipes* were sporadically detected in breast milk samples,

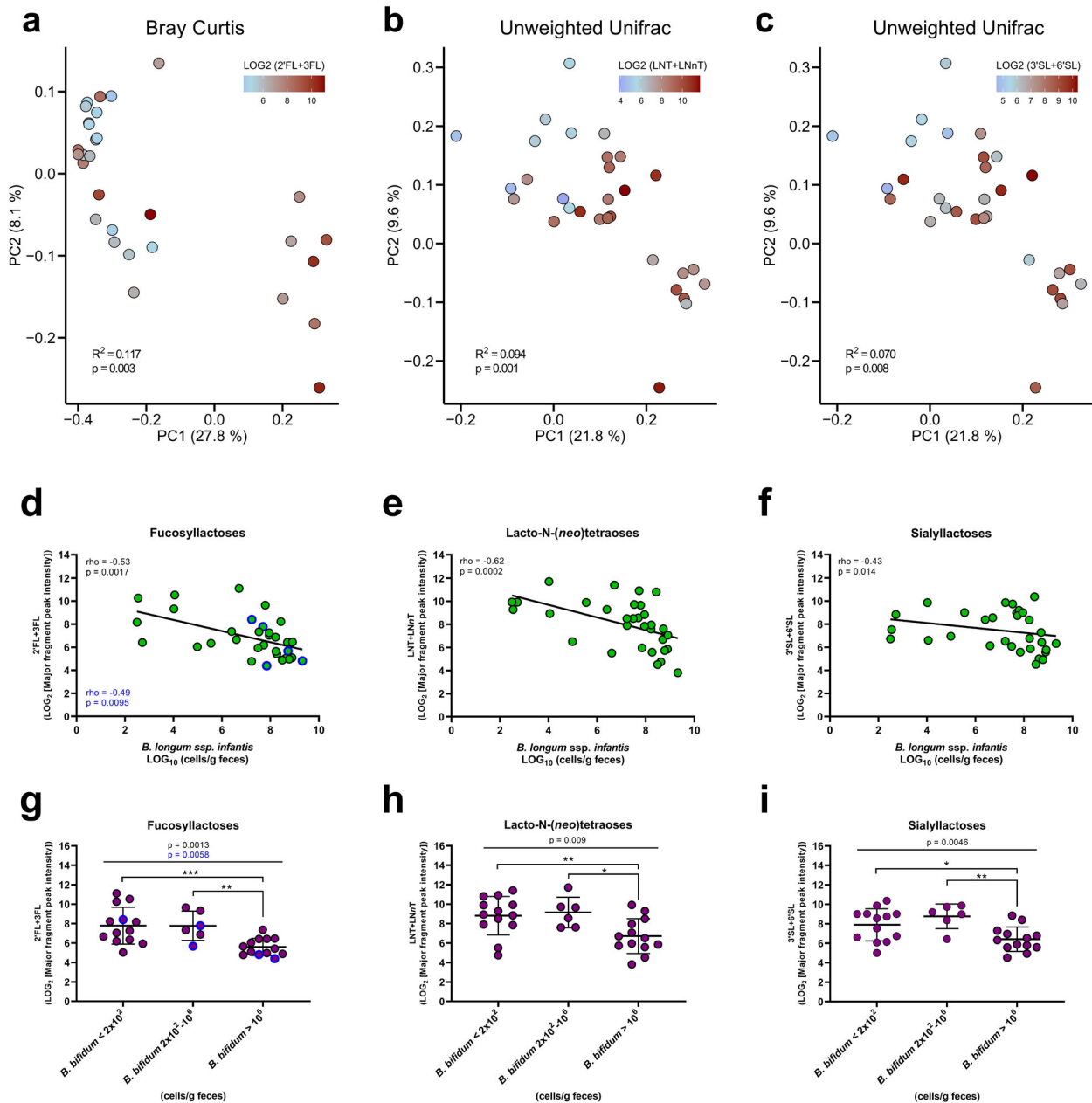


Fig. 5 Human milk oligosaccharides associate with infant fecal microbiota composition and explains abundance of specific *Bifidobacterium* species. **a–c** PCoA plots based on **a** Bray–Curtis dissimilarities or **b, c** unweighted Unifrac distances, colored according to fecal relative abundance (LOG_2 [major fragment peak intensity]) of **a** 2FL+3FL, **b** LNT+LNnT, and **c** 3'SL+6'SL at infant age 5 months. Statistical significance evaluated by ADONIS tests. **d–f** Scatter dot plots of qPCR-estimated absolute abundance of *B. longum* subsp. *infantis* versus relative abundance (LOG_2 [major fragment peak intensity]) of the three HMO classes in feces at infant age 5 months. A linear regression line was fitted to the data points and associations between variables were assessed by Spearman's rank correlation analyses. **g–i** Dot plots of relative abundance (LOG_2 [major fragment peak intensity]) of the three HMO classes stratified according to qPCR-estimated absolute abundance of *B. bifidum* in feces at infant age 5 months ($B. bifidum < 2 \times 10^2$ cells/g feces indicates values below limit of detection). Statistical significance was evaluated by one-way ANOVA (exact p -value indicated) with Fisher's LSD post hoc tests with asterisks indicating * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. For 2FL+3FL (fucosyllactoses) statistical analyses (marked in blue text) were repeated excluding infants from non-secretor mothers (dots marked with blue border).

they were never shared with corresponding infant feces, suggesting that breast milk is not a common source of these. However, typical infant gut colonizers *Escherichia*, *Lactobacillus*, *Enterococcus*, and especially *Bifidobacterium* were indeed shared (prevalence between 20 and 70%). These genera are not common inhabitants of the oral cavity⁶¹ nor human skin,⁵³ thus suggesting that not all milk-to-gut transmitted bacteria originate from infant saliva or maternal skin.

We detected ASVs classified as *Bifidobacterium* in ~80% of the breast milk samples. Two *B. longum* ASVs (ASV_1 and ASV_7, indicated by specific qPCR to represent *B. longum* subsp. *longum* and *B. longum* subsp. *infantis*, respectively) were among the top 10 most frequently shared (50–70%), and were both prevalent and highly abundant in the infant fecal communities. Although both these ASVs were also detected in negative sequencing controls, the observed higher absolute abundance in milk samples of

ASV_1 compared to sequencing controls and a positive correlation in paired foremilk versus hindmilk samples support the genuine presence of ASV_1 in the milk samples. In contrast, no difference in abundance was found in milk samples versus negative sequencing controls for ASV_7. This specific ASV was identified as a potential contaminant during decontamination filtering but was retained due to the very high abundance in fecal samples. Therefore, it is possible that the observation of ASV_7 in milk samples may be due to contamination. The *B. breve*, *B. bifidum*, and *B. catenulatum* group ASVs were only shared in one or two mother–infant pairs, but also only detected in one to three milk samples at each sampling time (and never in negative sequencing controls). Both *B. longum* and *B. breve* have previously been detected in breast milk in around 80% of individuals, whereas *B. bifidum* and *B. pseudocatenulatum* were previously found in 25% and 15% of individuals, respectively.⁶³ Isolation frequency of *Bifidobacterium* species from breast milk typically range from 10 to 30%,^{6,64–66} with the same strains of *Bifidobacterium* species, notably often belonging to *B. longum* (including both subspecies) and *B. breve*, recovered from maternal breast milk and corresponding infant's feces.^{67–72} Thus, although breast milk-seeding probably does not fully account for the diversity and prevalence of *Bifidobacterium* species in the infant gut, collectively, data support that it is a contributing factor.

Our results suggest that maternal secretor status does not have a major impact on the infant fecal microbiota, although the limited size of our cohort, and especially low a number of non-secretors, prevents us from making statistically solid conclusions. Our results are in line with recent studies showing limited effect of maternal secretor status on infant gut microbiota composition during the first 3–6 months of life,^{25,73,74} although other studies suggest a more marked impact.^{23,24} Further, in our study the concentrations of individual HMOs in breast milk did not correlate strongly with infant gut microbiota composition. Thus, it is possible that efficient HMO consumers will thrive in the intestinal environments of both infants from secretor and non-secretor mothers and somewhat independent of the specific HMO composition in breast milk, since some gut microbes, especially *B. longum* subsp. *infantis* and *B. bifidum*, harbor the gene sets required for degradation of a range of HMO structures from breast milk of non-secretors as well as secretors.²⁸ Indeed, these taxa have been shown to be the major in vitro HMO-utilizing *Bifidobacterium* species,⁵⁵ and represent two of the most abundant and prevalent gut species detected in this and other⁷⁵ cohorts of breastfed infants. In accordance with this, we observe significant negative correlations between absolute abundances of *B. longum* subsp. *infantis* and *B. bifidum* and levels of all three major classes of HMOs (fucosylactoses, sialylactoses and lacto-N-(neo)tetraoses) in infant's feces. This supports broad-range HMO utilization within these species in vivo and suggests that HMOs are contributing to the dominance of these species in the breastfed infant's gut.

Despite the relatively low number of mother–infant dyads included in this study, which limits the power of analysis, we were thus able to uncover several interesting findings. Collectively, our results show that breast milk microbiota composition changes during a feeding session and over the lactation period and that it probably contributes to infant gut seeding of milk-associated bacteria, especially *Streptococcus*, *Veillonella*, and *Bifidobacterium* spp., also during complementary feeding. Along with HMOs, milk-associated bacteria likely direct a specific assembly of the infant gut microbiota, notably dominated by HMO-degrading *Bifidobacterium* species.

REFERENCES

- Victora, C. G. et al. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *Lancet* **387**, 475–490 (2016).
- Vieira Borba, V., Sharif, K. & Shoenfeld, Y. Breastfeeding and autoimmunity: programming health from the beginning. *Am. J. Reprod. Immunol.* **79**, e12778 (2018).
- Andreas, N. J., Kampmann, B. & Mehning Le-Doare, K. Human breast milk: a review on its composition and bioactivity. *Early Hum. Dev.* **91**, 629–635 (2015).
- Eriksen, K. G., Christensen, S. H., Lind, M. V. & Michaelsen, K. F. Human milk composition and infant growth. *Curr. Opin. Clin. Nutr. Metab. Care* **21**, 200–206 (2018).
- Hunt, K. M. et al. Characterization of the diversity and temporal stability of bacterial communities in human milk. *PLoS ONE* **6**, 1–8 (2011).
- Jost, T., Lacroix, C., Braegger, C. & Chassard, C. Assessment of bacterial diversity in breast milk using culture-dependent and culture-independent approaches. *Br. J. Nutr.* **110**, 1253–1262 (2013).
- Moossavi, S. et al. Composition and variation of the human milk microbiota are influenced by maternal and early-life factors. *Cell Host Microbe* **25**, 324–335 (2019).
- Lundgren, S. N. et al. Microbial communities in human milk relate to measures of maternal weight. *Front. Microbiol.* **10**, 1–15 (2019).
- Oikonomou, G. et al. Milk microbiota: what are we exactly talking about? *Front. Microbiol.* **11**, 1–15 (2020).
- Gomez-Gallego, C., Garcia-Mantrana, I., Salminen, S. & Collado, M. C. The human milk microbiome and factors influencing its composition and activity. *Semin. Fetal Neonatal Med.* **21**, 400–405 (2016).
- Jost, T., Lacroix, C., Braegger, C. & Chassard, C. Impact of human milk bacteria and oligosaccharides on neonatal gut microbiota establishment and gut health. *Nutr. Rev.* **73**, 426–437 (2015).
- Kordy, K. et al. Contributions to human breast milk microbiome and enteromammary transfer of *Bifidobacterium breve*. *PLoS ONE* **15**, e0219633 (2020).
- Williams, J. E. et al. Strong multivariate relations exist among milk, oral, and fecal microbiomes in mother–infant dyads during the first six months postpartum. *J. Nutr.* **149**, 902–914 (2019).
- Rodriguez, J. M. The origin of human milk bacteria: is there a bacterial enteromammary pathway during late pregnancy and lactation? *Adv. Nutr. An. Int. Rev. J.* **5**, 779–784 (2014).
- Moossavi, S. & Azad, M. B. Origins of human milk microbiota: new evidence and arising questions. *Gut Microbes* **12**, e1667722 (2020).
- Saarela, T., Kokkonen, J. & Koivisto, M. Macronutrient and energy contents of human milk fractions during the first six months of lactation. *Acta Paediatr. Int. J. Paediatr.* **94**, 1176–1181 (2005).
- van Sadelhoff, J. H. J. et al. Short communication: differences in levels of free amino acids and total protein in human foremilk and hindmilk. *Nutrients* **10**, 1828 (2018).
- Bishara, R., Dunn, M. S., Merko, S. E. & Darling, P. Volume of foremilk, hindmilk, and total milk produced by mothers of very preterm infants born at less than 28 weeks of gestation. *J. Hum. Lact.* **25**, 272–279 (2009).
- Pannaraj, P. S. et al. Association between breast milk bacterial communities and establishment and development of the infant gut microbiome. *JAMA Pediatr.* **171**, 647 (2017).
- Fehr, K. et al. Breastmilk feeding practices are associated with the co-occurrence of bacteria in mothers' milk and the infant gut: the CHILD cohort study. *Cell Host Microbe* **28**, 285–297 (2020).
- Bode, L. The functional biology of human milk oligosaccharides. *Early Hum. Dev.* **91**, 619–622 (2015).
- Bode, L. Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology* **22**, 1147–1162 (2012).
- Lewis, Z. T. et al. Maternal fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants. *Microbiome* **3**, 13 (2015).
- Korpela, K. et al. Fucosylated oligosaccharides in mother's milk alleviate the effects of caesarean birth on infant gut microbiota. *Sci. Rep.* **8**, 13757 (2018).
- Bai, Y. et al. Fucosylated human milk oligosaccharides and N-glycans in the milk of Chinese mothers regulate the gut microbiome of their breast-fed infants during different lactation stages. *mSystems* **3**, e00206–e00218 (2018).
- Yu, Z. T., Chen, C. & Newburg, D. S. Utilization of major fucosylated and sialylated human milk oligosaccharides by isolated human gut microbes. *Glycobiology* **23**, 1281–1292 (2013).
- Marcobal, A. et al. Consumption of human milk oligosaccharides by gut-related microbes. *J. Agric. Food Chem.* **58**, 5334–5340 (2010).
- Sakanaka, M. et al. Varied pathways of infant gut-associated *Bifidobacterium* to assimilate human milk oligosaccharides: Prevalence of the gene set and its correlation with bifidobacteria-rich microbiota formation. *Nutrients* **12**, 71 (2019).
- Bezirtzoglou, E. et al. Microbiota profile in feces of breast- and formula-fed newborns by using fluorescence in situ hybridization (FISH). *Anaerobe* **17**, 478–482 (2011).

30. Bäckhed, F. et al. Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe* **17**, 690–703 (2015).
31. Jost, T., Lacroix, C., Braegger, C. P. & Chassard, C. New insights in gut microbiota establishment in healthy breast fed neonates. *PLoS ONE* **7**, e44595 (2012).
32. Larsson, M. W. et al. Excessive weight gain followed by catch-down in exclusively breastfed infants: an exploratory study. *Nutrients* **10**, 1290 (2018).
33. Austin, S. & Bénét, T. Quantitative determination of non-lactose milk oligosaccharides. *Anal. Chim. Acta* **1010**, 86–96 (2018).
34. Prudden, A. R. et al. Synthesis of asymmetrical multiantennary human milk oligosaccharides. *Proc. Natl. Acad. Sci. USA* **114**, 6954–6959 (2017).
35. Barri, T., Holmer-Jensen, J., Hermansen, K. & Dragsted, L. O. Metabolic fingerprinting of high-fat plasma samples processed by centrifugation- and filtration-based protein precipitation delineates significant differences in metabolite information coverage. *Anal. Chim. Acta* **718**, 47–57 (2012).
36. Pluskal, T., Castillo, S., Villar-Briones, A. & Oresic, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395 (2010).
37. Stanstrup, J. et al. The metaRbionomics toolbox in bioconductor and beyond. *Metabolites* **23**, 200 (2019).
38. van der Kloet, F. M., Bobeldijk, I., Verheij, E. R. & Jellema, R. H. Analytical error reduction using single point calibration for accurate and precise metabolomic phenotyping. *J. Proteome Res.* **8**, 5132–5141 (2009).
39. Lawley, B. et al. Differentiation of *Bifidobacterium longum* subspecies *longum* and *infantis* by quantitative PCR using functional gene targets. *PeerJ.* **5**, e3375 (2017).
40. Frese, S. A. et al. Persistence of supplemented *Bifidobacterium longum* subsp. *infantis* EVC001 in breastfed infants. *mSphere* **2**, e00501–e00517 (2017).
41. Matsuki, T., Watanabe, K., Tanaka, R. & Oyaizu, H. Rapid identification of human intestinal bifidobacteria by 16S rRNA-targeted species- and group-specific primers. *FEMS Microbiol. Lett.* **167**, 113–121 (1998).
42. Tulstrup, M. V. L. et al. Antibiotic treatment affects intestinal permeability and gut microbial composition in Wistar rats dependent on antibiotic class. *PLoS ONE* **10**, e0144854 (2015).
43. Laursen, M. F. et al. Intestinal *Enterococcus* abundance correlates inversely with excessive weight gain and increased plasma leptin in breastfed infants. *FEMS Microbiol. Ecol.* **96**, fiae066 (2020).
44. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
45. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
46. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
47. Davis, N. M., Proctor, D. 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* **6**, 226 (2018).
48. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
49. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).
50. Heikkilä, M. P. & Saris, P. E. J. Inhibition of *Staphylococcus aureus* by the commensal bacteria of human milk. *J. Appl. Microbiol.* **95**, 471–478 (2003).
51. Douglas, C. A. et al. DNA extraction approaches substantially influence the assessment of the human breast milk microbiome. *Sci. Rep.* **10**, 123 (2020).
52. de Goffau, M. C. et al. Recognizing the reagent microbiome. *Nat. Microbiol.* **3**, 851–853 (2018).
53. Byrd, A. L., Belkaid, Y. & Segre, J. A. The human skin microbiome. *Nat. Rev. Microbiol.* **16**, 143–155 (2018).
54. Dzidic, M. et al. Oral microbiome development during childhood: an ecological succession influenced by postnatal factors and associated with tooth decay. *ISME J.* **12**, 2292–2306 (2018).
55. Garrido, D. et al. Comparative transcriptomics reveals key differences in the response to milk oligosaccharides of infant gut-associated bifidobacteria. *Sci. Rep.* **5**, 1–18 (2015).
56. Könönen, E., Kanervo, A., Takala, A., Asikainen, S. & Jousimies-Somer, H. Establishment of oral anaerobes during the first year of life. *J. Dent. Res.* **78**, 1634–1639 (1999).
57. Abranches, J. et al. Biology of oral Streptococci. *Microbiol. Spectr.* **6**, 1128 (2018).
58. Kawamura, Y. et al. *Streptococcus peroris* sp. nov. and *Streptococcus infantis* sp. nov., new members of the *Streptococcus mitis* group, isolated from human clinical specimens. *Int. J. Syst. Bacteriol.* **48**, 921–927 (1998).
59. Martín, V., Mañes-Lázaro, R., Rodríguez, J. M. & Maldonado-Barragán, A. *Streptococcus lactarius* sp. nov., isolated from breast milk of healthy women. *Int. J. Syst. Evol. Microbiol.* **61**, 1048–1052 (2011).
60. Djais, A. A. et al. Identification and phylogenetic analysis of oral *Veillonella* species isolated from the saliva of Japanese children. *F1000 Res.* **8**, 616 (2019).
61. Sulyanto, R. M., Thompson, Z. A., Beall, C. J., Leys, E. J. & Griffen, A. L. The predominant oral microbiota is acquired early in an organized pattern. *Sci. Rep.* **9**, 10550 (2019).
62. Schmidt, T. S. B. et al. Extensive transmission of microbes along the gastrointestinal tract. *Elife* **8**, e42693 (2019).
63. Lugli, G. A. et al. Investigating bifidobacteria and human milk oligosaccharide composition of lactating mothers. *FEMS Microbiol. Ecol.* **96**, fiae049 (2020).
64. Martín, R. et al. Isolation of bifidobacteria from breast milk and assessment of the bifidobacterial population by PCR-denaturing gradient gel electrophoresis and quantitative real-time PCR. *Appl. Environ. Microbiol.* **75**, 965–969 (2009).
65. Treven, P. et al. Evaluation of human milk microbiota by 16S rRNA gene next-generation sequencing (NGS) and cultivation/MALDI-TOF mass spectrometry identification. *Front. Microbiol.* **10**, 2612 (2019).
66. Soto, A. et al. Lactobacilli and bifidobacteria in human breast milk: Influence of antibiotherapy and other host and clinical factors. *J. Pediatr. Gastroenterol. Nutr.* **59**, 78–88 (2014).
67. Jost, T., Lacroix, C., Braegger, C. P., Rochat, F. & Chassard, C. Vertical mother-neonate transfer of maternal gut bacteria via breastfeeding. *Environ. Microbiol.* **16**, 2891–2904 (2013).
68. Martín, V. et al. Sharing of bacterial strains between breast milk and infant feces. *J. Hum. Lact.* **28**, 36–44 (2012).
69. Makino, H. et al. Transmission of intestinal *Bifidobacterium longum* subsp. *longum* strains from mother to infant, determined by multilocus sequencing typing and amplified fragment length polymorphism. *Appl. Environ. Microbiol.* **77**, 6788–6793 (2011).
70. Makino, H. et al. Multilocus sequence typing of bifidobacterial strains from infant's faeces and human milk: Are bifidobacteria being sustainably shared during breastfeeding?. *Benef. Microbes* **6**, 563–572 (2015).
71. Duranti, S. et al. Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission. *Microbiome* **5**, 282 (2017).
72. Peirotnén, A., Arqués, J. L., Medina, M. & Rodríguez-Minguez, E. Bifidobacterial strains shared by mother and child as source of probiotics. *Benef. Microbes* **9**, 231–238 (2018).
73. Borewicz, K. et al. The association between breastmilk oligosaccharides and faecal microbiota in healthy breastfed infants at two, six, and twelve weeks of age. *Sci. Rep.* **10**, 4270 (2020).
74. Lewis, Z. T. et al. The fecal microbial community of breast-fed infants from Armenia and Georgia. *Sci. Rep.* **7**, 40932 (2017).
75. Stewart, C. J. et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583–588 (2018).

ACKNOWLEDGEMENTS

We thank the families participating the SKOT III cohort study. We thank Birgitte Hermansen and Vivian Anker for sample collection and technical assistance, Katrine Pedersen, Barbara Thorslund, and Erik Kiehl for their contribution to DNA extraction, PCR, qPCR, and 16S rRNA amplicon library preparation of material from breast milk and feces samples. We thank Mikiyasu Sakanaka for critical review and feedback on the manuscript. We thank Glycom A/S for kindly donating HMOs. Ion Torrent sequencing was performed by the DTU in-house facility DTU Multi-Assay Core (DMAC), Technical University of Denmark. The SKOT III study was funded by University of Copenhagen, University College Copenhagen and the Family Larsson-Rosenquist Foundation. Work on HMO quantification in feces was supported by a Carlsberg Foundation grant (Semper Ardens) on food intake biomarkers to L.O.D.

AUTHOR CONTRIBUTIONS

K.F.M., M.W.L., M.V.K., C.M., and A.L. designed the SKOT III study and were in charge of sample and initial data collection. M.F.L., M.I.B., and T.R.L. conceived and designed the studies related to analysis of breast milk and infant fecal microbiota. M.F.L. and M.I.B. were in charge of DNA extraction, PCR, qPCR, and 16S rRNA amplicon library preparation. M.F.L. and M.I.B. analyzed the 16S rRNA amplicon sequence data and interpreted it in the context of the study. L.B. and C.Y. analyzed HMOs in maternal breast milk and L.O.D. and C.T.P. analyzed selected HMOs in infant feces. M.F.L. wrote the manuscript with substantial contributions from M.I.B. and T.R.L. All authors read and approved the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43705-021-00021-3>.

Correspondence and requests for materials should be addressed to M.I.B.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021