**Supplementary Table 1. Detailed dietary information per patient**

Patient	Group	Metagenome	Day full feed 72hrs	Day first MOM	Day last MOM	Day first formula	Day last formula	Day first fortifer	Day last fortifer
367	CTRL	YES	24	2	18	9	78	-	-
369	CTRL	YES	11	2	97	76	97	26	82
379	CTRL	YES	14	1	70	33	108	21	70
381	CTRL	YES	14	1	103	52	116	21	103
394	CTRL	YES	13	2	144	-	-	17	144
403	CTRL	YES	12	1	67	-	-	17	64
405	CTRL	YES	12	1	67	-	-	17	64
407	CTRL	NO	14	0	54	-	-	34	Discharged
420	CTRL	YES	20	2	40	29	59	-	-
422	CTRL	YES	31	2	33	39	116	-	-
423	CTRL	YES	15	2	36	38	116	-	-
431	CTRL	YES	12	2	116	-	-	19	116
440	CTRL	YES	13	1	86	17	25	26	87
447	CTRL	YES	22	2	114	-	-	26	114
450	CTRL	YES	15	6	64	64	111	-	-
457	CTRL	YES	13	5	48	-	-	20	50
463	CTRL	YES	15	4	50	31	122	14	50
465	CTRL	YES	13	2	Discharged	-	-	14	94
473	CTRL	YES	13	2	54	30	54	22	26
478	CTRL	YES	12	2	38	25	84	-	-
479	CTRL	YES	15	3	70	52	70	20	70
509	CTRL	YES	15	3	22	-	-	-	-
512	CTRL	YES	10	2	31	-	-	23	31
520	CTRL	YES	23	4	58	25	106	32	58
534	CTRL	YES	21	3	82	79	100	30	98
535	CTRL	YES	13	3	45	-	-	40	45
547	CTRL	YES	17	3	Died	-	-	-	-
552	CTRL	YES	20	3	91	81	106	41	81
554	CTRL	YES	12	4	41	15	113	-	-
559	CTRL	YES	12	2	93	94	160	12	71
575	CTRL	YES	24	4	21	18	99		-
593	CTRL	YES	12	2	47	10	Discharged	-	-
594	CTRL	YES	17	3	29	18	121	-	-
610	CTRL	YES	17	4	64	65	96	18	64
652	CTRL	NO	17	4	28	27	99	-	-
676	CTRL	NO	34	1	34	-	-	-	-
682	CTRL	YES	17	4	67	57	67	27	67
364	NEC	YES	11	2	83	-	-	56	83
373	NEC	NO	11	2	44	-	-	-	-
376	NEC	NO	44	4	96	-	-	65	96
377	NEC	NO	18	1	53	54	96	26	29
384	NEC	NO	12	1	36	32	46	35	36
385	NEC	NO	30	4	Transferred	75	Transferred	70	Transferred
388	NEC	NO	15	2	36	18	115	-	-
392	NEC	NO	49	7	Discharged	-	-	76	105
393	NEC	YES	28	3	69	54	145	-	-
395	NEC	YES	41	2	69	27	154	-	-
396	NEC	NO	34	3	68	42	68	48	67
408	NEC	YES	18	3	28	24	120	-	-
410	NEC	YES	50	1	82	76	139	-	-
415	NEC	YES	25	3	33	34	93	-	-
424	NEC	YES	12	3	114	-	-	41	n/a
443	NEC	NO	15	3	32	33	156	-	-
445	NEC	NO	9	1	68	61	Discharged	43	72
451	NEC	YES	13	2	86	87	119	20	85
461	NEC	NO	20	4	Transferred	-	-	-	-
474	NEC	NO	13	2	Died	-	-	-	-
506	NEC	YES	11	2	86	82	104	65	86
529	NEC	NO	Never	3		-	-	-	-
580	NEC	YES	13	4	17	18	88	17	17
592	NEC	NO	Never	4	40	-	-	-	-
596	NEC	YES	114	4	14	26	132	-	-
629	NEC	YES	23	2	41	39	106	-	-
636	NEC	YES	12	2	58	-	-	-	-
639	NEC	NO	27	5	96	85	115	-	-
690	NEC	NO	11	2	102	-	-	62	77
691	NEC	NO	13	3	108	-	-	34	52
692	NEC	NO	26	3	140	-	-	35	46
712	NEC	YES	20	3	126	88	126	-	-
723	NEC	NO	30	2	34	35	83	-	-

Supplementary Table 2. Detailed antibiotic information per patient

Patient	Group	Metagenome	Antibiotic courses	How to interpret
367	CTRL	YES	0 (P2, G2); 25 (A2, G2, F8)	e.g., day 0 infant received 2 days of Benzyl-penicillin and 2 days of Gentamicin; day 25 infant received 2 days of Amoxicillin, 2 days of Gentamicin, 8 days of Flucloxacillin
369	CTRL	YES	0 (P2, G2); 3 (mero5)	
379	CTRL	YES	0 (P2, G2); 26 (A12); 29 (G4, F4); 87 (A1, F1, G1)	
381	CTRL	YES	0 (P2, G2); 4 (A2, F2, G2); 14 (G3, A3, F3); 26 (A7); 86 (A3, F3, G3)	
394	CTRL	YES	0 (P2, G2); 3 (V5, C5); 18 (A5, F5, G5); 41 (C5); 51 (F3, A6, G5); 65 (V2, C2)	
403	CTRL	YES	0 (P2, G2)	
405	CTRL	YES	0 (P2, G2); 47 (F2, A2, G2)	
407	CTRL	NO	0 (p2, G2); 9 (A3, F3, G3); 20 (C3, V3, M3)	
420	CTRL	YES	0 (P2, G2); 6 (C2, V2, M2); 12 (C2, V3); 15 (F12)	
422	CTRL	YES	0 (P2, G2); 8 (V2, C2); 32 (V6, C6)	
423	CTRL	YES	0 (P2, G2); 7 (V5, C5); 15 (A2, F2, G2)	
431	CTRL	YES	0 (P2, G2); 11 (A2, G2, F2); 18 (C10); 22 (G5); 27 (Tazo7); 65 (A5, G5, F5)	
440	CTRL	YES	0 (P2, G2); 6 (F4)	
447	CTRL	YES	0 (A5, G5, M5); 11 (C4, V4); 26 (C5); 33 (C5, V5); 43 (C4, V4)	
450	CTRL	YES	0 (P2, G2); 9 (V5, C5); 37 (A2, G2, F2)	
457	CTRL	YES	0 (P5, G5); 7 (V2, C2)	
463	CTRL	YES	0 (P2, G2); 14 (a3, G3, F3); 38 (A2, F2, G2); 61 (F5, A1, G5)	
465	CTRL	YES	0 (P2, G2); 10 (C6, V6); 22 (A5, F3, G3); 26 (V2)	
473	CTRL	YES	0 (P2, G2); 16 (A2, F2, G2)	
478	CTRL	YES	0 (P2, G2); 10 (F7); 66 (F7, A4)	
479	CTRL	YES	0 (P2, G2); 3 (V2, C2); 54 (A2, F2, G2)	
509	CTRL	YES	0 (P2, G2); 11 (A2, M2, G2)	
512	CTRL	YES	0(P2, G2); 21 (A2, F2, G2)	
520	CTRL	YES	0 (P5, G5); 16(A2, F2, G2); 21 (V2, C2); 37 (taz5)	
534	CTRL	YES	0 (P2, G2); 11 (V3, C3); 14 (A11)	
535	CTRL	YES	0 (P2, G2); 2 (V4, C4); 21 (V3, C3, M3); 36 (A2, F2); 41 (A2, F2)	
547	CTRL	YES	0 (P2, G2); 8 (V7, C7); 14 (F5, G5, V5, C5); 30 (F5); 56 (azith10, mero10)	
552	CTRL	YES	0 (A5, G5, M5); 5 (V5, C5); 12 (F2, A2, G2)	
554	CTRL	YES	0 (P5, G5); 6 (C5, V5); 19 (A2, G2, F2); 22 (V2, C2); 30 (A2, F2, G2); 33 (V5, C5); 44 (A7, F7, G7)	
559	CTRL	YES	0 (P2, G2); 20 (F7); 27 (co-tri5); 32 (mero 5); 39 (V5, C5, M5)	
575	CTRL	YES	0 (P2, G2); 3 (F1); 6 (V5, C5); 15 (V2, C2)	
593	CTRL	YES	0 (P2, G2)	
594	CTRL	YES	0 (P2, G2); 6 (C5, V5); 27 (A13, F2); 37 (C5, V2)	
610	CTRL	YES	0 (P2, G2); 14 (V2, C2)	
652	CTRL	NO	0 (P5, G5); 14 (V7, C2); 30 (A2, F2)	
676	CTRL	NO	0 (P2, G2); 4 (C7, V8); 13 (F3, A3, G3); 16 (V7, C7, M7); 31 (V3, C3, M3)	
682	CTRL	YES	0 (P2, G2); 3 (V2, C2)	
364	NEC	YES	0 (P2, G2); 9 (C4, V6); 38 (V9, C9); 40 (M6); 61 (V2, M2, C2)	
373	NEC	NO	0 (P2, G2); 28 (V7, C7, M7)	
376	NEC	NO	0 (P2, G2); 2 (C3, V3); 8 (V7, C7, M7); 17 (V5, M5, C5); 44 (A3, G3, F3); 63 (A3, F10, G3); 98 (V3, C3, M3)	
377	NEC	NO	34 (C7, M7, V7); 51 (A2, F2, G2); 81 (V5, C2); 87 (M6, C6)	
384	NEC	NO	10 (F5); 14 (C5, V5); 14 (A8, G8, M8); 23 (V2, C2)	
385	NEC	NO	0 (P3, G3); 9 (V10, C10, M10); 82 (A2, F2, G2)	
388	NEC	NO	0 (P2, G2); 5 (C3, V3); 31 (A6, F6, G6, M6); 57 (A3, F3, G3)	
392	NEC	NO	No data before this. 24 (C9, M9); 25 (V7); 74 (A5, F5, G5)	
393	NEC	YES	0 (P2, G2); 14 (C5, V5, M5); 53 (A2, F2, G2)	
395	NEC	YES	0 (P2, G2); 6 (V5, C5, M5); 19 (mero5); 54 (A2, F2)	
396	NEC	NO	0 (A2, F2, G2); 16 (V10, C10, M10)	
408	NEC	YES	0 (P2, G2); 11 (V7); 28 (V7, C7, M7)	
410	NEC	YES	0 (P2, G2); 7 (V17, C3); 14 (M5, C1); 15 (t4); 18 (Mero7); 25 (Ambi6, G6, Clotri8); 37 (V13); 40 (Clotri5)	
415	NEC	YES	0 (P2, G2); 5 (A7, G7, F7); 7 (M5); 17 (V2, C2)	
424	NEC	YES	0 (P2, G2); 13 (A7, F7, G5, M7); 19 (Cef5); 27 (V2, C2, M2); 54 (A2, F2, G2); 76 (G1, C3); 83 (Cef); 97 (A4, F4, G4)	
443	NEC	NO	0 (P2, G2); 5 (V5, C2); 19 (V8, C8, M8); 37 (A21); 67 (A2, G2, F2)	
445	NEC	NO	0 (P2, G2); 5 (V2, C2); 15 (V8, C8, M8)	
451	NEC	YES	0 (P1, G1); 1 (v5); 12 (A2, F2); 14 (co-amox5); 21 (co-tri 5); 38 (Cef8, M8, co-tri8)	
461	NEC	NO	0 (P2, G2); 13 (A2, F2, G2, M8); 14 (V3, C7); 17 (F7)	
474	NEC	NO	0 (P2, G2); 14 (A2, F2, G2); 19 (V7, C7, M7)	
506	NEC	YES	0 (P2, G2); 8 (Taz5); 30 (V2, Mero7); 52 (Mero5, V5); 61 (V2)	
529	NEC	NO	0 (P2, G2); 4 (V5, C5); 18 (VCM 19); 47 (VCM2); 52 (F5); 57 (VCM5); 92 (co-amox2)	
580	NEC	YES	0 (P2, G2); 24 (A5, F5, G5); 60 (a7, F7, G7, M7)	
592	NEC	NO	5 (V5, C5); 13 (co-amox7); 18 (taz7M7); 27 (F5, G5); 68 (mero10)	
596	NEC	YES	0 (P2, G2); 5 (V7, C2); 12 (V12, C12, M12); 31 (co-tri); 34 (mero26); 74 (F5)	
629	NEC	YES	0 (P2, G2); 7 (V7, C7); 10 (M5); 36 (F5)	
636	NEC	YES	0 (P5, G5); 12 (C5); 51 (A2, F2); 69 (V5, C5, M5)	
639	NEC	NO	0 (P2, G2); 8 (A2, F2, G2); 14 (V1, C1); 15 (A21); 24 (V5, C5, M5)	
690	NEC	NO	0 (P5, G5); 9(V14, C14, M14)	
691	NEC	NO	0 (P2, G2); 14 (V7, C7, M7); 42 (A2, F2, G2)	
692	NEC	NO	0 (P2, G2); 5 (V2, C2); 8 (Mero7, V7); 46 (A7, F7, G7, M7); 58 (cef21); 96 (cef5)	
712	NEC	YES	0 (P2, G2); 11 (C21, V2); 31 (A5, G5, F5); 39 (V7, C7, M5); 56 (V5, C5, M5)	
723	NEC	NO	0 (P3, G3); 5 (F2, G2, M2); 8 (VCM5); 17 (V5, C5, M5)	

A, Amoxicillin; Ambi, Ambisome; Azith, Azithromycin; C, Ceftazidime; Cef, Cefuroxime; Clotri, Cotrimoxazole; Co-amox, Co-amoxiclavulinic acid; Co-tri, Co-trimoxazole; F, Flucloxacillin; G, Gentamicin; M, Metronidazole; Mero, Meropenem; P, Benzyl-penicillin; Tazo, Tazocin; V, Vancomycin
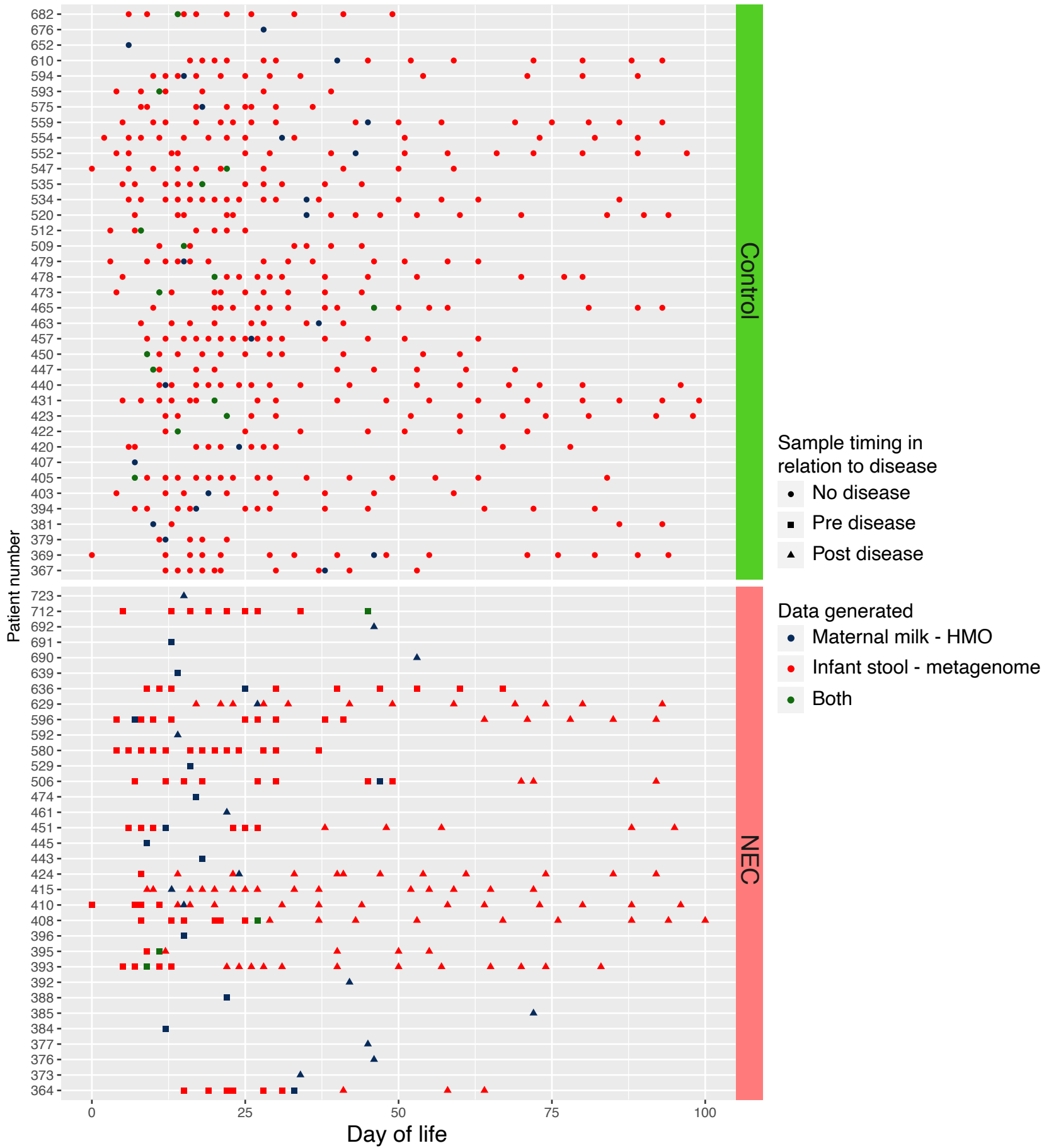
Online supplementary table 3. Sub-cohort of infants with longitudinal metagenome data from stool samples.

	Control	NEC	P value*
Number of patients	34	14	-
Number of stool samples	449	195	-
Secretors	23 (68%)	10 (71%)	0.797
Male	12 (35%)	11 (79%)	0.006
Vaginal delivery	24 (71%)	7 (50%)	0.175
Gestational age	25 [24; 26]	25 [24; 26.75]	0.908
Birthweight	645 [586.3; 747.5]	670 [562.5; 735]	0.447
Probiotics ever	34 (100%)	13 (93%)	0.871
MOM only	2 (6%)	1 (7%)	
MOM + Formula	10 (29%)	8 (57%)	
MOM + BMF	9 (27%)	2 (14%)	0.315
MOM + Formula + BMF	13 (38%)	3 (22%)	
DOL NEC onset	-	28 [13; 51]	-
NEC surgical	-	4	-

NEC, necrotising enterocolitis; MOM, mother's own breast milk; BMF, breast milk fortifier; DOL, day of life

*Differences between groups were tested applying Chi-square test and Dunn's post-hoc test where applicable.

Human milk oligosaccharide DSLNT and gut microbiome in preterm infants predicts necrotising enterocolitis

Andrea C Masi, Nicholas D Embleton, Christopher A Lamb, Gregory Young, Claire L Granger, Julia A Najera, Daniel P Smith, Kristi L. Hoffman, Joseph F. Petrosino, Lars Bode, Janet E Berrington, Christopher J Stewart

Supplementary methods

Human milk oligosaccharides analysis

MOM was collected aseptically after expression and research samples were obtained from the syringe at completion of feed. Thus, the day of the MOM samples reflects the day the infant received the milk and does not necessarily reflect the day milk was expressed. The absolute quantification for the 19 most abundant HMOs was determined by high-performance liquid chromatography (HPLC) following derivatization as per the protocol described by Bode *et al.* (*1*). Briefly, raffinose was added to every sample before analysis to work as internal standard. Lipids, proteins, lactose, salts and peptides were removed by stepwise solid-phase extraction. HMOs were labelled by adding the fluorescent tag 2-aminobenzamide to the reducing end and subsequently analysed by HPLC on an amide-80 column. The HMOs quantified account for >95% of total HMOs and included: 2'-fucosyllactose (2'FL), 3-fucosyllactose (3FL), lacto-N-neotetraose (LNnT), 3'-sialyllactose (3'SL), difucosyllactose (DFlac), 6'-sialyllactose (6'SL), lacto-N-tetraose (LNT), lacto-N-fucopentaose (LNFP) I, LNFP II, LNFP III, sialyl-LNT (LST) b, LSTc, difucosyl-LNT (DFLNT), lacto-N-hexaose (LNH), disialyllacto-N-tetraose (DSLNT), fucosyl-lacto-N-hexaose (FLNH), difucosyl-lacto-N-hexaose (DFLNH), fucosyl-disialyl-lacto-N-hexaose (FDSLNH) and disialyl-lacto-N-hexaose (DSLNH). Maternal secretor (presence of an active FUT2 gene) status was determined by presence or near-absence of 2'FL in the breast milk analysed.

Metagenomes

Infant stool was obtained directly from the nappy/diaper. DNA was extracted from ~0.1g of stool using the DNeasy PowerSoil Kit (QIAGEN) following the manufacturer's protocol. In addition to stool samples, extraction was performed on a positive (Zymo Microbial Community Standard) and negative (no sample at all) controls. A negative control was extracted in every batch of 24 samples. Library prep was performed using the Nextera DNA Flex Kit. Sequencing was performed on the HiSeq X Ten (Illumina) with a target read depth of 10M reads per sample with a read length of 150bp paired end reads.

Raw fastq files were quality trimmed and Illumina adapters removed using bbduk (BBMap version 38.69). Trimming parameters included kmer length of 19, allowing one mismatch, and a minimum Phred score of 20. Post-trimming, reads with a minimum average Phred <17 and length <50 bp were discarded. Host contamination reads were identified by mapping trimmed fastq files to a combined database containing the hg38 reference human genome and PhiX (standard Illumina spike in) using bbmap (BBMap version 37.58) with kmer length of 15, bloom filter enabled, and fast search settings. Host reads were subsequently removed, and the remaining processed fastq files were mapped against the MetaPhlan2 marker gene database (mpa_v20_m200) using bbmap with the bloom filter enabled and fast search settings (*2*). Finally, the metaphlan.py script was used to generate kingdom specific taxonomic profiles.

Statistical analysis

Statistical analysis of HMO profiles was performed using MetaboAnalyst 3.0 (*3*). Orthogonal Partial Least Squares - Discriminant Analysis (OPLS-DA) and Partial Least Squares - Discriminant Analysis (PLS-DA), for 2 or more group comparison respectively, were performed on HMO data normalised by logarithmic transformation and 2000 random permutations were used to test the significance of group separation. HMO Shannon diversity was calculated using "vegan" (version 2.5-6) package (*4*) in R. Wilcoxon rank-sum test or

Kruskal Wallis test were used for variables comparison between two or more groups, respectively, and P values were adjusted applying the Benjamini & Hochberg correction (*5*). Variables with >2 groups deemed significant with Kruskal-Wallis underwent Dunn's post-hoc test to determine P values specific to each group comparisons and resulting P values were adjusted applying Bonferroni method (*6*).

To test potential role of individual HMOs as biomarker for disease development, univariate receiver operating characteristic (ROC) curve analysis was performed, and optimal cut-off was defined by the closest point of the curve to the top-left corner. Multivariate ROC curves were also generated using linear Support Vector Machine (SVM) classification method coupled with Monte-Carlo cross validation (MCCV) to test the classification performance obtained by using 2, 3, 5, 7, 10, or 19 HMOs. In each MCCV step, two thirds of the dataset were used to determine feature importance and classification model performance was evaluated with the remaining third of the samples which was left out.

Correlation between clinical variables and individual HMOs was tested by performing a multivariate adjusted linear model in R (version 3.6.3). HMO concentrations were normalised by log-transformation prior to analysis and P values were adjusted applying the Benjamini & Hochberg correction (*5*). Clinical variables tested included delivery mode, gestational age at birth, disease status, day of life (DOL), and postmenstrual age (PMA) of sample, maternal secretor status and infant sex.

A total of 10,015,821,590 mapped reads (median 14,426,827 reads per sample) were obtained from metagenomic sequencing of the 644 preterm infant stool samples. The lowest sample contained 152,718 mapped reads. The cross-sectional cohort of stool samples collected from NEC infants before diagnosis and matched controls was analysed using MicrobiomeAnalyst (*7, 8*). Alpha diversity analysis was performed based on observed species (richness) and Shannon diversity, and beta-diversity was performed using Bray-Curtis principal coordinate

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Gut

analysis and differences between groups performed using permutational multivariate analysis of variance (PERMANOVA). MetagenomeSeq was used to assess differential abundance at the phyla and species level. This approach utilises both cumulative sum scaling normalization and zero-inflated Gaussian distribution mixture or zero-inflated Log-Normal mixture model. DMM clusters samples on the basis of microbial community structure (*9*) and was used to determine the preterm gut community types (PGCTs) from all samples, as performed previously (*10, 11*). The appropriate number of clusters was determined based on the lowest Laplace approximation score (*9*). Five PGCT was found to be optimal, and these were ordered 1-5 based on the average DOL of samples within that PGCT, where PGCT-1 contained on average the samples collected from the earliest DOL and PGCT-5 contained on average the samples collected from the oldest DOL. Analysis was performed at specific time windows, including only a single sample per infant in each time point. In cases where an infant had more than one sample within a given time window, the chosen sample reflected the PGCT which was most common among an infant's samples within the given time window. The ratios of each PGCT were compared by chi-square test.

The association of various clinical variables on the HMO and metagenome profiles was tested by applying the function "adonis" of "vegan" (version 2.5-6) package (*4*) in R. Bray-Curtis dissimilarity was used for calculating the dissimilarity matrix, and 10000 permutations were applied. Each test was performed stepwise and P values were adjusted using Benjamini & Hochberg (*5*). Clinical factors tested were delivery mode, gestational age, birthweight, sex, maternal secretor status, infant antibiotic administration (i.e., receiving antibiotics yes/no at time of sample), infant probiotic administration (i.e., receiving probiotic at time of sample, before, after, or never), diet (i.e., combinations of expressed breast milk and formula), DOL and PMA of sample, and disease status.

Random Forest algorithm was used for comparing the performance of classification models built using cross-sectional HMO profile data (nmol/ml), cross-sectional pre-NEC and matched control metagenomic data (count), and both HMO and metagenome datasets combined. The contribution given by each variable was evaluated through the Mean Decrease Accuracy (MDA) value, which indicates how much the accuracy of the model decreases when that variable is removed. The higher the MDA value, the more important that feature is. Variables associated to a negative MDA value were removed and a new model was built on the subset of variables. This step of feature filtering was performed until the model with best classification performance was obtained.

References

1. L. Bode *et al.*, Human milk oligosaccharide concentration and risk of postnatal transmission of HIV through breastfeeding. *Am J Clin Nutr* **96**, 831-839 (2012).
2. D. T. Truong *et al.*, MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods* **12**, 902-903 (2015).
3. J. Xia, I. V. Sinelnikov, B. Han, D. S. Wishart, MetaboAnalyst 3.0--making metabolomics more meaningful. *Nucleic Acids Res* **43**, W251-257 (2015).
4. F. G. B. Jari Oksanen, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs, Helene Wagner, vegan: Community Ecology Package. R package version 2.5-6. . *https://CRAN.R-project.org/package=vegan*, (2019).
5. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289-300 (1995).
6. O. J. Dunn, Multiple Comparisons Among Means. *Journal of the American Statistical Association* **56**, 52-64 (1961).
7. J. Chong, P. Liu, G. Zhou, J. Xia, Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nat Protoc* **15**, 799-821 (2020).
8. A. Dhariwal *et al.*, MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res* **45**, W180-W188 (2017).
9. I. Holmes, K. Harris, C. Quince, Dirichlet multinomial mixtures: generative models for microbial metagenomics. *PLoS One* **7**, e30126 (2012).

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Gut

10. C. J. Stewart *et al.*, Cesarean or Vaginal Birth Does Not Impact the Longitudinal Development of the Gut Microbiome in a Cohort of Exclusively Preterm Infants. *Front Microbiol* **8**, 1008 (2017).

11. C. J. Stewart *et al.*, Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583-588 (2018).

a

HMOs	AUC	CI
2	0.946	0.888-0.988
3	0.945	0.875-0.988
5	0.932	0.845-0.98
7	0.914	0.812-0.968
10	0.902	0.762-0.98
19	0.884	0.723-0.976

b

- high in NEC
- high in CTRL

c

a

P = 0.797

r2 = -0.007

b

P = 0.930

r2 = -0.009