Temporal dynamics of the gut microbiome and metabolome in preterm and term infants from birth through the first year of life

Polly Soo Xi Yap
University of Malaya Faculty of Medicine https://orcid.org/0000-0002-1615-3138
Chun Wie Chong
Monash University - Malaysia Campus
Azanna Ahmad Kamar
Universiti Malaya
Ivan Kok Seng Yap
Sarawak Research and Development Council
Yao Mun Choo
Universiti Malaya
Nai Ming Lai
Taylor's University
Cindy Shuan Ju Teh (cindysjteh@um.edu.my)
https://orcid.org/0000-0002-9062-3839

Research

Keywords: 16S rRNA sequencing, nuclear magnetic resonance, human gut microbiome, metabolomics, preterm infant, neonatal intensive care unit

Posted Date: January 31st, 2020

DOI: https://doi.org/10.21203/rs.2.22382/v1

Abstract

Background: Emerging evidence has shown a link between the perturbations and development of the gut microbiota in infants to their immediate and long-term health. In comparison to the healthy full-term neonate, preterm neonates experience disparate gut bacterial establishment (e.g. duration in the womb), colonisation (e.g. mode of delivery), and development (e.g. frequent use of antibiotics). To better understand the assembly of the gut microbiota in preterm infants, faecal samples were longitudinally collected from preterm (n = 19) and term (n = 20) infants, up to 12 months after birth. We characterised bacterial compositions by 16S rRNA gene sequencing (n = 141) and metabolomics profiling (n = 141) using nuclear magnetic resonance (NMR) spectroscopy.

Results: Significant differences in faecal bacterial composition between term and preterm infants were detected in sample collected in week 2, month 6 and month 12. Interestingly, separation of the bacterial composition between term and preterm infants was more pronounced at month 12 as compared to the earlier time-points, suggesting distinct level of microbial maturation in gut between the two groups. Intestinal microbiota of preterm neonates was consistently characterised by dominance of pathogenic bacteria from the Enterobacteriaceae family and a paucity of strictly anaerobic taxa including Veillonella and Bacteroides relative to infants born at term. Consistent result was observed in the stool NMR spectroscopy in which clear separation in stool metabolomics profiles was observed between the term and preterm neonates.

Conclusion: Overall, we identified a panel of amino acid metabolites and central metabolism intermediates in the preterm infants' stool, possibly indicating incomplete fermentation of complex polysaccharides in the guts of these infants. In contrast, the term infant stool had significantly higher levels of metabolites which are commonly found in milk such as fucose and β-hydroxybutyrate (BHBA). Birth weight was selected as the best explanatory variable for the metabolomics profiles, pointing to the strong relationship between protein synthesis, as well as fucose and BHBA in physical development. By following both term and preterm infants for 12 months, our study reported the dynamic of gut microbial composition and their contribution to metabolism and potential impact to growth in neonates.

Background

Preterm birth is one of the major determinants of neonatal mortality and morbidity. Prematurity is also well-recognised to carry long term consequences on health well into the infants' adult life. Children born prematurely have higher rates of developing diabetes (1, 2), hypertensions (3, 4) and obesity (5) as compared to those born at term. However, there is still a lack of evidence of the association of prematurity with later adult metabolic syndromes. There have been various studies of patterns in gut microbiome of preterm infants in the last two decades, most reporting correlations between the gut microbiota and various factors during early life (6–8). Comparison of the stool microbiota of full-term and preterm neonates revealed significant differences. A paucity of Bacteroides, Bifidobacterium, Streptococcus and Lactobacillus among the preterm infants has been described previously (7, 9). When

compared to term infants, preterm infants tend to be dominated by facultative anaerobes including Enterobacteriaceae and Enterococcaceae (10). Also, a number of studies have characterised the gut microbiome composition in infants subjected to different modes of delivery (11, 12), modes of feeding (7), geographical locations (13–15) and antibiotic exposures (16, 17). However, limited studies have advanced interpretations in terms of longitudinal associated patterns, especially gut microbiota maturation after weaning, despite increasing evidence of the population dynamics of the human gut microbiome (18). Understanding how host factors influence diversity is an important step towards clarifying the functioning of human gut towards the host health conditions. Additionally, the integration of metagenomes and metabolome at which these ecosystems operate provides the basis for the assessment of the functioning of human system biology (19). Given that host metabolism is strongly influenced by gut microbes through their collective activities and cooperative responses, the association of health status and gut microbial dynamics should therefore be viewed holistically.

In Malaysia, the rate of preterm birth was reported to be 12.4% based on the latest publicly available National Obstetrics Registry report in 2015 (20). Specifically, preterm birth rates were highest among the indigenous population in Peninsular Malaysia, mothers at extreme ages, parity > 6, and mothers with preeclampsia, eclampsia and renal disease (20). The complications of preterm birth and morbidity arise from immature organ systems that are incapable of supporting life in the extrauterine environment (21). Despite the significant health burden of preterm birth, the consequences of impaired maturation of gut microbiota to the development of preterm infants to adult life remained poorly studied, especially in South East Asia. Further, there is a lack of information regarding the contribution of gut microbial composition to host metabolism, and the interplay between both in the growth of preterm infants. There is therefore a need to characterise the temporal dynamics of the human gut microbiome and metabolome to understand their roles in the preterm infants' health and development. In this study, we compared the variations in the faecal metagenome and metabolome between preterm and term infants at varying time points from birth to month 12. The pattern of gut bacterial assemblages and metabolic profiles were modelled using selected clinical characteristics while the relationship of microbial taxa and metabolites that differed significantly across different time points was assessed through network analysis.

Results

Characteristics of the preterm and term infants

Of the 19 preterm infants studied, two were lost-to-follow-up due to early death. The remaining preterm infants included three sets of twins and one set of triplets. For comparison, a total of 20 healthy full-term infants born at a gestational age of more than 37 weeks were included in this study. Characteristics of the preterm and term infants are reported in Table 1. Term infants were mainly born by spontaneous vaginal delivery and no exposure to antibiotics while preterm infants were mainly born by caesarean delivery (94.7%, $p < 0.0001$) and exposed to antibiotics (57.9%, $p < 0.0001$). Clinical characteristics of the preterm infants during admission are reported in Table 2. Eleven preterm infants had respiratory distress syndrome (RDS) and five patients received invasive mechanical ventilation. Eleven patients received

parental nutrition (PN) and three of them had peripherally inserted central catheter (PICC). Eleven patients were treated with antibiotics for three days to two week with various combinations of penicillin (8 patients), gentamicin (10 patients), vancomycin (4 patients) and meropenem (4 patients).

Preterm infants' stool samples collected during NICU admission were subjected for isolation for multidrug resistant Enterobacteriaceae, as published in Yap et al., (22). Antibiotic susceptibility profile of the isolates was correlated with patient clinical data for further statistical analyses.

Stool samples were collected at: day 1 (meconium), week 1, week 2, month 6 and month 12 of life from all infants. However, some time points were missed for selected infants due to the lack of adherence to the sampling schedule. In total, we collected 141 samples, for an average of 3.05 samples per preterm infants and 4.15 samples per term infants.

Microbial composition comparison between term and preterm infants

The alpha diversity of the faecal microbial community between term and preterm infants and time points were evaluated using Simpson, Shannon and Pielou's evenness indices (Supplementary Figure 1). Overall, no significant difference in bacterial richness and evenness were observed between term and preterm samples. However, among the five sampling time points, statistically significant differences ($p < 0.01$) were achieved for all three indices between samples from week 2 with month 12. Separately, beta diversity was assessed using PERMANOVA and PLS-DA. Overall, significant differences in faecal bacterial composition between the term and preterm infants were detected in PERMANOVA (pseudo-F = 2.4834, P(perm) = 0.001). From the first PLS-DA plot (Figure 1), it is apparent that the faecal bacterial compositions from both groups are highly variable. When time points were added as factor, it was observed that samples from week 2 were overlapped with meconium and week 1. For month 6 and month 12, however, both groups formed two tight clusters with little overlapped. Consistent observation was also obtained using pairwise PERMANOVA analysis (Supplementary Table 2).

The overall distributions of the phyla and genera were provided in Supplementary Figure 2. Differentially abundance OTUs were identified based on negative binomial model. A total of 55 OTUs from 3 phyla (Proteobacteria, Firmicutes and Bacteroidetes) whose abundance differed between two groups across the five time points were identified. The species which matched highest sequence homology with the input were included in the Supplementary Table 3. When the comparison was made at phylum level, the dominant phyla for term infants were Firmicutes and Bacteroidetes; while enriched levels of Proteobacteria were observed in preterm infants' stool at the first two weeks of life. At 12 months of life, significant abundance of Bacteroidetes (OTU0012, OTU0029, OTU0051 and OTU0110) were observed among the preterm group. Conversely, term infants were observed with abundant species of Proteobacteria at month 12. The preterm infants' stool samples had significantly enriched *Klebsiella* OTUs (mainly *K. pneumoniae*) during the first two weeks of life while the level of *Klebsiella* OTUs started to elevate in term infant stools only after 6 months of life. It was observed that OTU0029 (*Bacteroides fragilis*) was elevated in week 1 and week 2 stools of term infants while the significant elevation only observed at month 12 for preterm.

Temporal differences in metabolomics profiles between term and preterm infants

PLS-DA plot inferred using metabolomics profiles showed less apparent separation in comparison to the microbial composition, although distinct clusters specific to term and preterm are still discernible (Figure 1). When time points were added as factor, samples from meconium, week 1 and week 2 were loosely clustered especially among the preterm infants (Figure 2). Consistent with bacterial composition, samples from month 6 were significantly different (*P*(perm) = 0.005, *P*(MC) = 0.005). Nonetheless, both sample groups (i.e. term and preterm) from month 12 did not show significant difference with most of the data points overlapped with month 6. This observation was also reflected in the significant metabolites detected consistently in month 6 and month 12 in the term infants. Significantly expressed metabolites according to time points were further identified using permutation test (number of permutations = 1000). Only metabolites with a *P*-value of 0.01 and below were selected and summarised in Table 3. The corresponding covariance plots of preterm vs term derived from stools samples obtained from meconium, week 1, week 2, month 6 and month 12 of life were included in Supplementary Figure 3. Meconium collected from preterm infants showed elevated glycerol. At week 1, preterm group showed elevated faecal valine, leucine, isoleucine, tyrosine and phenylalanine, whereas α-glucose and methylmalonic acid (MMA) were elevated in the term group. However, no significant metabolic changes in preterm and term infants were observed at week 2. Changes in stool metabolites were more pronounced at month 6 with the preterm group showing higher levels of faecal succinate, citrate and trimethylamine-N-oxide (TMAO), whereas the term infants showed higher levels of faecal β-hydroxybutyric acid (BHBA), fucose and pyruvatoxime. Additionally, the latter consistently showed elevated levels of BHBA and fucose up to month 12, whereas the preterm infants exhibited higher levels of faecal tyrosine and phenylalanine at month 12 as well as week 1.

Association between differentially expressed OTUs and significant metabolites

Network analysis was performed based on sparse least square model (sPLS-DA) (Figure 3). Three *Erysipelatoclostridium*-related OTUs (OTU0035, OTU0090 and OTU0112) were positively correlated to the branched-chain amino acids (BCAAs): valine, leucine and isoleucine. In addition, OTU0035 (*Erysipelatoclostridium ramosum*) was negatively correlated to BHBA and pyruvatoxime while OTU0090 (*Clostridium cocleatum*) and OTU0112 (*Clostridium spiroforme*) were negatively correlated to pyruvatoxime and α-glucose. *Bacteroides fragilis* (OTU0029) which was differentially expressed in term group at week 1 and week 2 and subsequently in preterm group at month 12, showed moderate to strong negative correlations to BHBA, fucose, MMA and pyruvatoxime. *Lactobacillus mucosae* (OTU0048) showed moderate positive correlations with the three BCAAs and strong negative correlations with pyruvatoxime. *Veillonella seminalis* (OTU0017) also showed strong correlations with the three BCAAs. OTU0050 which belongs to the family *Comamonadaceae*, showed distinct clustering with negative correlations with BHBA and fucose. OTU0050 was consistently significantly expressed in preterm group at week 1, month 6 and month 12, while term group, conversely, was enriched with the same OTU at week 2.

Correlations between omics data with demographic and clinical parameters

Distance based linear modelling was carried out to identify the demographic and clinical predictors for the differentially expressed 16S metagenomic and NMR metabolomic profiles. We modelled the faecal microbial and metabolic composition by splitting the data into each respective time point. Birth weight was consistently selected as the best explanatory variable for the total variation in the metabolic profiles all infants at birth and at month 12 (Table 4). On the other hand, gestational age was selected to explain the elevated 16S metagenomics profile for all infants at month 12. When the data was analysed by considering clinical parameters, the step-wise selection algorithm selected “PICC line insertion” and “isolation of bacteria resistant to 3rd generation cephalosporins” as the best explanatory parameters for the faecal metabolic composition of the preterm group at month 6 and month 12 respectively.

Discussion

The development of infant gut microbiome is generally believed to begin from amniotic fluid (23, 24), although the microbial biomass in the fluid is very low (25). In comparison to preterm infants, full term infants have completed in utero microbial and organ development. The relatively more matured organs provided distinct ecological landscape for the colonisation and establishment of the gut microbiota (26). We showed however that no categorical difference in alpha diversity was detected between term and preterm infants. The lack of differences is surprising given that preterm infants born via caesarean section were previously found to have less diverse microbiome as they are not exposed to maternal vaginal, faecal and epithelial microbes (27). Nonetheless, significant differences in faecal microbiota diversity were obtained across the time points especially week 2 with month 6 and week 2 with month 12. Among the measured time points, week 2 exhibited the lowest diversity. When the comparison was made by considering beta-diversity, significant difference in the taxonomic compositions between chronological age and sampling time points became apparent (Figs. 1 & 2, Supplementary Fig. 1).

In this study, we observed a higher Firmicutes:Bacteriodetes ratio in term than preterm neonates in week 1, week 2 and month 6. High Firmicutes:Bacteriodetes ratio of the gut microbes has been previously associated with a “healthier” gut. However, recent publications suggested that such link is over-simplified (18). Owing to the diverse functional capability of each phylum, variation in “beneficial gut microbes” might not induce uniform phylum-wide metabolic responses. A change of associated external environment such as diet, infancy infections, living conditions are known to impact the composition of infant gut microbiome (28). As compared to meconium, the bacterial membership in the stool collected in week 1 and week 2 displayed relatively higher consistency across individuals. The homogenisation effect observed might be a natural response of the gut microbiome towards the changes from womb to birth in which a smaller subset of available microbial populations was preferentially selected. It is notable that the higher Tenericutes observed in preterms’ meconium compared to the terms is consistent with the composition of placental microbiome in preterm neonates (29, 30). It is interesting to note that the divergence in preterm and term microbiomes was strongest at month 12. The dichotomous development of the microbiome may be due to the available of the different species pools during birth (for e.g. natural

birth vs caesarean). The variation in physical development further resulted in the different taxa being selected. These observations were consistent with the ecological processes such as “historical contingency” and “niche specialisation” (31, 32).

In comparison to gut microbial composition, the metabolic profiles were relatively less sensitive to the infant development in terms of sampling time points. Temporal stability of the gut metabolome can be related to the concept of high redundancy of metabolic pathways across different microbial species (33, 34). Our finding is also in line with previous literature suggesting that human metabolic profiles are more conserved than gut microbiome (34). Elevated serum MMA and stool glucose have been associated with vitamin B12 (35) and carbohydrate malabsorption (36) in children respectively. In this study, MMA and α-glucose were found to be significantly increased in week 1 stools of the term group. Our findings suggest that negative feedback loop is likely playing a role in primary microbial community succession during early life. If the abundance of a particular microbe exceeded a certain threshold, a negative feedback loop may occur with a strong interplay between host and microbial metabolism, thus resulting in a change in the gut environment to decrease that microbe’s growth in relation to other species (33). The negative feedback loops may involve the accumulation of a specific toxic metabolite (in this case the MMA) while paradoxically promoting high microbiome diversity and resilience in the host (37). Our network analysis shows that the elevation of stool MMA is negatively correlated with two Parabacteroides OTUs (OTU0051 and OTU0110) and two Bacteroides OTUs (OTU0012 and OTU0029). These four OTUs were differentially expressed in the stool samples of term group at week 1 and week 2, and preterm group at month 12. We postulate that these microbes may play a role in stabilising the microbiota equilibrium which consistent with the level of physical development in the term and preterm infants.

The functional output of the gut microbiota, including amino acids and short-chain fatty acids (SCFAs) are thought to be important modulators underlying the pathogenesis of metabolic diseases such as obesity (38), insulin resistance and type 2 diabetes mellitus (39, 40). Metabolites identified in the preterm group at week 1 include BCAAs such as valine, leucine and isoleucine, and aromatic amino acids including tyrosine and phenylalanine. Protein catabolism and absorption generally take place in the small intestine. During this process, amino acids that are not re-absorbed will be transferred to colon. We detected a positive correlation between BCAAs with three OTUs of Erysipelatoclostridium (former Clostridium spp.) and Lactobacillus (OTU0048). The former has been reported to be prevalent among the type 2 diabetes mellitus patients (41) while Clostridium ramosum was detected in high-fat fed mice (42). It was postulated that amino acids can serve as precursors for the synthesis of SCFAs by intestinal microbes to deliver additional energy to the host (38, 43). Thus, the elevated Erysipelatoclostridium (Clostridium) OTUs in preterm infants may play an important role in the energy metabolism that fuel rapid physical development. Succinate is a key intermediate in the tricarboxylic acid (TCA) cycle in host metabolism. The major producers of succinate in the mammalian gut are Prevotella spp. (44). Although significantly higher abundance of Prevotella spp. in preterm as compared to term infants was detected, high succinate concentration was only observed in the stool samples collected at month 6. The lack of correlation might be due to the fact that succinate is only transiently available as it is rapidly converted as an intermediate in the production of propionate by the succinate-utilising bacteria (45, 46). Further, the

high expression of succinate may also due to the extended administration of antibiotic among the preterm infants (47, 48). Another TCA cycle intermediate identified in preterm at month 6 is citrate, possibly indicating incomplete fermentation of complex polysaccharides in the gut of preterm. Additionally, the elevated TMAO identified in the preterm group at month 6 might be a response of weaning as the metabolite has been associated with animal proteins such as eggs, red meat and dairy (49, 50).

At 6 months of life, increased milk metabolites such as fucose and β-hydroxybutyrate (BHBA), as well as energy–related intermediate, pyruvatoxime, were observed in the term infants. Emerging evidence suggests the presence of fucose in mammalian gut could improve host health by supporting the competitive growth of beneficial members of the gut community (51), suppressing virulence genes of some gut pathogens (52), and interact directly with host immune cells (53). Besides compositional shift in the gut, some human intestinal symbionts for example Bacteroides fragilis, has the ability to incorporate fucose into their own glycans, which is crucial for its fitness in the gut (54, 55). In this study, B. fragilis was strongly and negatively associated with fucose. Besides being identified abundantly in milk, BHBA is long viewed as a ketone body that serves as a circulating energy source by the brain when blood glucose is low in times of fasting or prolonged exercise (56). An increase in BHBA was observed in preterm infants with postnatal growth failure (57). In comparison, pyruvatoxime is an oxime derivative from pyruvate, which is the key output of the anaerobic metabolism of glucose that facilitates the growth of pyruvatoxime-oxidising bacteria such Alcaligenes spp. (58) and Pseudomonas aeruginosa (59). Conversely, the metabolite also negatively associated with L. mucosae, E. ramosum, B. fragilis, C. cocleatum and C. spiroforme. One of the caveats of this study is the lack of detailed dietary data as the duration of breastfeeding, introduction to solid foods and cessation of milk feeding may influence the gut microbiota and subsequently the metabolites levels in the stool (60, 61).

Based on the DISTLM analysis of the overall elevated metabolites profile, birth weight was found to be the best explanatory variable at birth and at month 12 of life. In comparison, when the model was constructed by selecting the bacterial taxa that differed significantly in all infants, gestational age was associated with the gut microbial profiles at month 12. When the same analysis was repeated on the differentially expressed metabolites of preterm infants, “PICC line insertion” and “isolation of bacteria resistant to 3rd generation cephalosporins” were selected as the best explanatory variable at month 6 and month 12 respectively. PICCs are routinely used in the NICU to provide prolonged therapy and parenteral nutrition. However, PICC-related complications such as bloodstream-infections and thrombosis have been frequently reported (62, 63). It should also be highlighted that host faecal metabolite is a manifestation of the host homeostasis-gut microbes interplay with the gut environment. Hence, faecal metabolic composition is more regulated by the host as compared to faecal microbial composition which interacts directly with the diet and the environment.

Conclusion

Based on the metabolomic profiles, it was apparent that the preterm infants harboured significantly more carbohydrate metabolism intermediates (citrate, succinate), BCAAs (leucine, isoleucine, valine) and aromatic amino acids (tyrosine, phenylalanine) in the stools through the first year of life as opposed to the term infants. Such differences might imply impaired sugar metabolism among the preterm infants. The metabolism of sugar is an important pathway for energy harvesting and essential of physical and mental growth in children (64). Despite our modest sample size, the current study is one of the largest longitudinal studies that applied system biology approach to compare the variation of host metagenomics and metabolic responses between preterm and term infants in Asia. In order to generate greater understanding in disease pathogenesis, there is a pressing need to extend the temporal coverage of infant gut microbial research. As the first targeted report on Malaysian preterm infant gut microbiome, our study contributed to the understanding of the least assessed South East Asian gut microbiota. In addition, the baseline data provided may facilitate the understanding of the dynamic link between gut microbial composition and systemic metabolomics in the infants' growth and health.

Methods

Sample collection

Preterm infants were recruited among patients admitted to the neonatal intensive care unit (NICU) at University Malaya Medical Centre (Kuala Lumpur, Malaysia) from June 2014 to December 2014. Enrolled preterm infants were born at a gestational age (GA) of < 37 weeks. Clinical data and stool samples from diapers of preterm neonates were collected. Term neonates at a GA > 37 weeks and with no congenital malformations and metabolic diseases were recruited in this study as a comparative group. Stool samples were collected up to 12 months of life. Samples were immediately stored at -20ºC and then transferred to -80ºC no more than 3 days post-collection.

DNA extraction, 16S rRNA sequencing and data processing

Nucleic acid extraction of stool was carried out on 220 mg of sample using the Qiagen DNA Isolation Kit in accordance with the manufacturer's instructions. PCR amplification of the 16S V3 and V4 regions was performed using primers V3-V4F (5'- TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCT ACG GGN GGC WGC AG -3') and V3-V4R (5'- GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGA CTA CHV GGG TAT CTA ATC C -3'). The reaction was performed with 1 µM each primer, 2x KAPA HiFi HotStart Ready Mix, 5 ng/µL of gDNA in 10 mM Tris pH 8.5 made up to 25 µL of reaction mixture. The thermoprofile used was as followed: initial denaturation for 3 min at 95 ºC followed by 25 cycles of 95 ºC for 30 s, 55 ºC for 30 s and 72 ºC for 30 s with a final extension at 72 ºC for 5 mins and hold at 4 ºC. Total DNA concentration of the purified PCR amplicons was measured using a Qubit® 3.0 Fluorometer (Life Technologies, Carlsbad, CA, United States) and normalised to 5ng/µL for 16S rRNA gene amplification using Nextera XT Index Kit. Amplicons were checked with a Bioanalyzer DNA 1000 chip and libraries were sequenced using a 2 x 300 bp paired-end run (MiSeq Reagent kit v3) on a HiSeq-Illumina platform. Controls during DNA extraction and PCR amplification were also included and sequenced. The

16S rRNA gene amplicon sequencing targeting variable regions V3 – V4 was performed using Hiseq rapid SBS kit v2 (Illumina) based on the standard operation procedure of the Hiseq 2500 platform.

Raw reads generated from Illumina paired-end sequencing were imported into Mothur software (version 1.39) for quality filtering and processing (65). Briefly, 99,122,189 paired-end sequences were joined into contigs. Contigs with ambiguous bases, homopolymer >6bp and sequence length below/ exceed the cut-off (< 460 bp and > 466 bp) were trimmed. Approximately 40% of the sequences was removed based on the filter and a total of 59,308,069 sequences were used for subsequent procedures. Chimeric sequences detected using Uchime command were removed before clustering the sequences into operational taxonomic units (OTUs). The taxonomic assignment was carried out in reference to SILVA database. Sequences affiliated with chloroplast, mitochondria, unknown, archaea, and eukaryote linages were further removed. Sequence for each OTU was blasted using NCBI database to identify the microorganisms at species level. The data was rarefied to equal depth for ease of comparison. Alpha diversity metrics including Shannon Diversity Index, Simpson Diversity Index and Pielou's Evenness were calculated and projected in bar graphs using microbiomeSeq R package (https://github.com/umerijaz/microbiomeSeq). In additional, bar charts were constructed using phyloseq package in R (66) to display the proportional differences in genus and phylum across groups. The beta-diversity was evaluated using statistical ordinations: Partial least squares discriminant analysis (PLS-DA) and Permutational Multivariate Analysis of Variance (PERMANOVA). PLS-DA models were built by using the MixOmics R package (67). PERMANOVA testing the inter-subject significance with time points was performed by using Primer + PERMANOVA software (version 7, Primer-E Ltd., Ivybridge, UK). Further, taxa showing significant differences in abundance between term and preterm infants at the respective time points were identified using negative log binomial model implemented in DESeq2 R package (68). Correction of multiple corrections was conducted using Benjamini-Hochberg (BH) procedure implemented in DESeq2. Differentially expressed phylum genus and OTUs were selected by using BH adjusted P-value cut off of 0.01.

Metabolomic analysis through Nuclear Magnetic Resonance (NMR)

The faecal metabolite extraction was adapted from fecal metabolite extraction protocol (69). The faecal water was homogenised with 1 mL of phosphate buffer (90% D2O,1 mM 3-trimethylsilyl-1-[2,2,3,3-2H4] propionate (TSP) and 3 mM sodium azide;pH 7.4). The homogenates were sonicated for 30 min in room temperature using a water bath sonicator followed by centrifugation at 13,000 rpm for 10 min. A total of 600 µL of supernatant was transferred to a 5 mm (outer diameter) NMR tube (Norell, USA) for NMR analysis.

A standard 1-dimensional (1D) 1H-NMR spectrum was acquired for each sample with a pulse [recycle delay (RD)–90º–t1–90º–tm–90º–acquire free induction decay (FID)] on a Bruker 600 MHz spectrometer (Bruker Biospin, Fallenden, Switzerland) with a 5 mm BB(F)O broadband probe operating at 600.13 MHz (ambient probe temperature 27 °C). The spectra were acquired according to parameters used previously (70). The field frequency was locked on D2O solvent and water peak suppression was performed by

gradient water pre-saturation during RD of 4 s and a mixing time (tm) of 0.01 s. The 90° pulse length was adjusted to ~10 µs and an acquisition time of 2.65 s was used.

Phasing and baseline correction of the 1H-NMR spectra were performed manually using Bruker TopSpin (version 4.0.5, Bruker Biospin, Fallenden, Switzerland). All spectra were referenced to the TSP resonance at δ 0.00. The spectra were digitised into 10 k data points using an in-house developed MATLAB (version R2014a, Natwick, USA) script (69). The regions containing water resonances (δ 4.5 – 6.5) were removed from each spectrum to eliminate distortion effects of water peak on the baseline. Additionally, the regions, δ 0.0 – 0.5 and δ 9.2 – 10.0 in the stool water spectra which contain only noise were removed. Normalisation to the total sum of residual spectrum and data scaling to unit variance were also carried out prior to pattern recognition analysis.

Beta-diversity analyses: PLS-DA and PERMANOVA were performed as mentioned in the 16S bacterial composition analyses. The covariance plot was used to aid interpretation of the significance of each metabolite from the permutation tests. The colours projected onto the spectrum indicate the significance of the metabolites with the blue indicating no significance difference at $p > 0.05$ confidence levels and red indicating high significance difference at $p < 0.05$. In this study, only metabolites with $p < 0.01$ was considered significant.

Integrated analysis of microbiome and metabolomics datasets

The relative abundance of the dominant bacterial taxa from 16S rRNA gene sequencing and the concentration of selected NMR metabolites was integrated using sparse partial least squares regression (sPLS) implemented under MixOmics R package (67). Network graph was constructed to observe the association of metabolomics and 16S metagenomics features.

Distance-based linear modelling (DISTLM)

To identify the association between the clinical and demographic parameters with microbial assemblage pattern and changes in metabolic profiles, we conducted DISTLM as described in Yap et al. (70). In brief, the clinical and demographic parameters were selected and fitted to the overall changes in gut microbial composition and metabolic profiles using stepwise regression under the second-order bias-corrected Akaike Information Criterion (AIC).

Abbreviations

AIC: Akaike Information Criterion; BCAA: branched-chain amino acid; BHBA: β-hydroxybutyrate; DISTLM: Distance-based linear modelling; GA: gestational age; MMA: methylmalonic acid; NICU: neonatal intensive care unit; NMR: nuclear magnetic resonance; OTU: operational taxonomic units; PERMAOVA: permutational multivariate analysis of variance; PICC: peripherally inserted central catheter; PLS-DA: partial lease squares discriminant analysis; PN: parental nutrition; RD: recycle delay; RDS: respiratory

distress syndrome; SCFA: short-chain fatty acid; sPLS: sparse partial least squares regression; TCA: tricarboxylic acid; TMAO: trimethylamine-N-oxide; TSP: trimethylsilyl-1-[2,2,3,3-2H4] propionate

Declarations

Ethics approval and consent to participate

Written informed consent was obtained from the infants' parents and investigations were conducted according to the principles approved by the University of Malaya Research Ethics Committee (UMREC) with ethical approval number 201310-0267.

Consent for publication

Not applicable.

Availability of data and materials

The dataset generated and analysed during the current study is available in the NCBI Sequence Read Archive (PRJNA578822; available at https://www.ncbi.nlm.nih.gov/sra/PRJNA578822) and additional data is available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was funded by University of Malaya Research Grant (UMRG: RG353-13HTM) and Postgraduate Research Funding (PPP: PG179-2015A).

Authors' contributions

CSJT, CWC and IKSY conceived and designed the experiments and contributed analytic tools. PSXY collected the samples and clinical data, did laboratory work, bioinformatics, and statistical analysis, and wrote the manuscript. CWC assisted in bioinformatics and statistical analysis and helped write the manuscript. AAK helped obtain clinical data and data interpretation. YMC and NML contributed to the administrative, ethics and technical duties. CSJT obtained the funding and critical review of manuscript. All authors have read and approved the final version to be published.

Acknowledgement

We gratefully acknowledge the help and support of subjects' parents who have participated our studies. We further acknowledge University of Malaya for providing research facilities.

References

1. Morrison KM, Ramsingh L, Gunn E, Streiner D, Van Lieshout R, Boyle M, et al. Cardiometabolic Health in Adults Born Premature With Extremely Low Birth Weight. Pediatrics. 2016;138(4).
2. Crump C, Winkleby MA, Sundquist K, Sundquist J. Risk of diabetes among young adults born preterm in Sweden. Diabetes care. 2011;34(5):1109-13.
3. Naumburg E, Soderstrom L. Increased risk of pulmonary hypertension following premature birth. BMC pediatrics. 2019;19(1):288.
4. Crump C, Winkleby MA, Sundquist K, Sundquist J. Risk of hypertension among young adults who were born preterm: a Swedish national study of 636,000 births. American journal of epidemiology. 2011;173(7):797-803.
5. Mathai S, Derraik JG, Cutfield WS, Dalziel SR, Harding JE, Biggs J, et al. Increased adiposity in adults born preterm and their children. PloS one. 2013;8(11):e81840.
6. Arboleya S, Binetti A, Salazar N, Fernandez N, Solis G, Hernandez-Barranco A, et al. Establishment and development of intestinal microbiota in preterm neonates. FEMS microbiology ecology. 2012;79(3):763-72.
7. Chernikova DA, Madan JC, Housman ML, Zain-Ul-Abideen M, Lundgren SN, Morrison HG, et al. The premature infant gut microbiome during the first 6 weeks of life differs based on gestational maturity at birth. Pediatric research. 2018;84(1):71-9.
8. Korpela K, Blakstad EW, Moltu SJ, Strommen K, Nakstad B, Ronnestad AE, et al. Intestinal microbiota development and gestational age in preterm neonates. Scientific reports. 2018;8(1):2453.
9. Itani T, Ayoub Moubareck C, Melki I, Rousseau C, Mangin I, Butel MJ, et al. Establishment and development of the intestinal microbiota of preterm infants in a Lebanese tertiary hospital. Anaerobe. 2017;43:4-14.
10. Jacquot A, Neveu D, Aujoulat F, Mercier G, Marchandin H, Jumas-Bilak E, et al. Dynamics and clinical evolution of bacterial gut microflora in extremely premature patients. The Journal of pediatrics. 2011;158(3):390-6.
11. Rutayisire E, Huang K, Liu Y, Tao F. The mode of delivery affects the diversity and colonization pattern of the gut microbiota during the first year of infants' life: a systematic review. BMC gastroenterology. 2016;16(1):86.
12. Shi YC, Guo H, Chen J, Sun G, Ren RR, Guo MZ, et al. Initial meconium microbiome in Chinese neonates delivered naturally or by cesarean section. Scientific reports. 2018;8(1):3255.
13. Kuang YS, Li SH, Guo Y, Lu JH, He JR, Luo BJ, et al. Composition of gut microbiota in infants in China and global comparison. Scientific reports. 2016;6:36666.
14. Subramanian S, Huq S, Yatsunenko T, Haque R, Mahfuz M, Alam MA, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature. 2014;510(7505):417-21.
15. Taddei CR, Oliveira FF, Duarte RT, Talarico ST, Takagi EH, Ramos C, II, et al. High abundance of Escherichia during the establishment of fecal microbiota in Brazilian children. Microbial ecology.

2014;67(3):624-34.
16. Gibson MK, Wang B, Ahmadi S, Burnham CA, Tarr PI, Warner BB, et al. Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. Nature microbiology. 2016;1:16024.
17. Aloisio I, Quagliariello A, De Fanti S, Luiselli D, De Filippo C, Albanese D, et al. Evaluation of the effects of intrapartum antibiotic prophylaxis on newborn intestinal microbiota using a sequencing approach targeted to multi hypervariable 16S rDNA regions. Applied microbiology and biotechnology. 2016;100(12):5537-46.
18. Priya S, Blekhman R. Population dynamics of the human gut microbiome: change is the only constant. Genome biology. 2019;20(1):150.
19. Turnbaugh PJ, Gordon JI. An invitation to the marriage of metagenomics and metabolomics. Cell. 2008;134(5):708-13.
20. Preliminary Report of National Obstetrics Registry, Jan 2013 – Dec 2015. Malaysia National Obstetrics Registry. 2017.
21. Lawn JE, Gravett MG, Nunes TM, Rubens CE, Stanton C, Group GR. Global report on preterm birth and stillbirth (1 of 7): definitions, description of the burden and opportunities to improve data. BMC pregnancy and childbirth. 2010;10 Suppl 1:S1.
22. Yap PS, Ahmad Kamar A, Chong CW, Yap IK, Thong KL, Choo YM, et al. Intestinal carriage of multidrug-resistant gram-negative bacteria in preterm-infants during hospitalization in neonatal intensive care unit (NICU). Pathogens and global health. 2016;110(6):238-46.
23. Collado MC, Rautava S, Aakko J, Isolauri E, Salminen S. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. Scientific reports. 2016;6:23129.
24. Wassenaar TM, Panigrahi P. Is a foetus developing in a sterile environment? Letters in applied microbiology. 2014;59(6):572-9.
25. Lim ES, Rodriguez C, Holtz LR. Amniotic fluid from healthy term pregnancies does not harbor a detectable microbial community. Microbiome. 2018;6(1):87.
26. DiBartolomeo ME, Claud EC. The Developing Microbiome of the Preterm Infant. Clinical therapeutics. 2016;38(4):733-9.
27. Makino H, Kushiro A, Ishikawa E, Kubota H, Gawad A, Sakai T, et al. Mother-to-infant transmission of intestinal bifidobacterial strains has an impact on the early development of vaginally delivered infant's microbiota. PloS one. 2013;8(11):e78331.
28. Gschwendtner S, Kang H, Thiering E, Kublik S, Fosel B, Schulz H, et al. Early life determinants induce sustainable changes in the gut microbiome of six-year-old children. Scientific reports. 2019;9(1):12675.
29. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. The placenta harbors a unique microbiome. Science translational medicine. 2014;6(237):237ra65.

30. Tirone C, Pezza L, Paladini A, Tana M, Aurilia C, Lio A, et al. Gut and Lung Microbiota in Preterm Infants: Immunological Modulation and Implication in Neonatal Outcomes. Frontiers in immunology. 2019;10:2910.
31. Costello EK, Stagaman K, Dethlefsen L, Bohannan BJ, Relman DA. The application of ecological theory toward an understanding of the human microbiome. Science. 2012;336(6086):1255-62.
32. Vieira-Silva S, Falony G, Darzi Y, Lima-Mendez G, Garcia Yunta R, Okuda S, et al. Species-function relationships shape ecological properties of the human gut microbiome. Nature microbiology. 2016;1(8):16088.
33. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. Nature. 2012;489(7415):220-30.
34. Visconti A, Le Roy CI, Rosa F, Rossi N, Martin TC, Mohney RP, et al. Interplay between the human gut microbiome and host metabolism. Nature communications. 2019;10(1):4505.
35. Sentongo TA, Azzam R, Charrow J. Vitamin B12 status, methylmalonic acidemia, and bacterial overgrowth in short bowel syndrome. Journal of pediatric gastroenterology and nutrition. 2009;48(4):495-7.
36. Hammer HF, Fine KD, Santa Ana CA, Porter JL, Schiller LR, Fordtran JS. Carbohydrate malabsorption. Its measurement and its contribution to diarrhea. The Journal of clinical investigation. 1990;86(6):1936-44.
37. Bever JD, Westover KM, Antonovics J. Incorporating the Soil Community into Plant Population Dynamics: The Utility of the Feedback Approach. Journal of Ecology. 1997;85(5):561-73.
38. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006;444(7122):1027-31.
39. Chen S, Akter S, Kuwahara K, Matsushita Y, Nakagawa T, Konishi M, et al. Serum amino acid profiles and risk of type 2 diabetes among Japanese adults in the Hitachi Health Study. Scientific reports. 2019;9(1):7010.
40. Palomino-Schatzlein M, Mayneris-Perxachs J, Caballano-Infantes E, Rodriguez MA, Palomo-Buitrago ME, Xiao X, et al. Combining metabolic profiling of plasma and faeces as a fingerprint of insulin resistance in obesity. Clinical nutrition. 2019.
41. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature. 2012;490(7418):55-60.
42. Woting A, Pfeiffer N, Loh G, Klaus S, Blaut M. Clostridium ramosum promotes high-fat diet-induced obesity in gnotobiotic mouse models. mBio. 2014;5(5):e01530-14.
43. Serena C, Ceperuelo-Mallafre V, Keiran N, Queipo-Ortuno MI, Bernal R, Gomez-Huelgas R, et al. Elevated circulating levels of succinate in human obesity are linked to specific gut microbiota. The ISME journal. 2018;12(7):1642-57.
44. Reichardt N, Duncan SH, Young P, Belenguer A, McWilliam Leitch C, Scott KP, et al. Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. The ISME journal. 2014;8(6):1323-35.

45. Watanabe Y, Nagai F, Morotomi M. Characterization of Phascolarctobacterium succinatutens sp. nov., an asaccharolytic, succinate-utilizing bacterium isolated from human feces. Applied and environmental microbiology. 2012;78(2):511-8.
46. Louis P, Flint HJ. Formation of propionate and butyrate by the human colonic microbiota. Environmental microbiology. 2017;19(1):29-41.
47. Tulstrup MV, Christensen EG, Carvalho V, Linninge C, Ahrne S, Hojberg O, et al. Antibiotic Treatment Affects Intestinal Permeability and Gut Microbial Composition in Wistar Rats Dependent on Antibiotic Class. PloS one. 2015;10(12):e0144854.
48. Woodmansey EJ, McMurdo ME, Macfarlane GT, Macfarlane S. Comparison of compositions and metabolic activities of fecal microbiotas in young adults and in antibiotic-treated and non-antibiotic-treated elderly subjects. Applied and environmental microbiology. 2004;70(10):6113-22.
49. Janeiro MH, Ramirez MJ, Milagro FI, Martinez JA, Solas M. Implication of Trimethylamine N-Oxide (TMAO) in Disease: Potential Biomarker or New Therapeutic Target. Nutrients. 2018;10(10).
50. Razavi AC, Potts KS, Kelly TN, Bazzano LA. Sex, gut microbiome, and cardiovascular disease risk. Biology of sex differences. 2019;10(1):29.
51. Pham TA, Clare S, Goulding D, Arasteh JM, Stares MD, Browne HP, et al. Epithelial IL-22RA1-mediated fucosylation promotes intestinal colonization resistance to an opportunistic pathogen. Cell host & microbe. 2014;16(4):504-16.
52. Pacheco AR, Curtis MM, Ritchie JM, Munera D, Waldor MK, Moreira CG, et al. Fucose sensing regulates bacterial intestinal colonization. Nature. 2012;492(7427):113-7.
53. Pickard JM, Chervonsky AV. Intestinal fucose as a mediator of host-microbe symbiosis. Journal of immunology. 2015;194(12):5588-93.
54. Coyne MJ, Reinap B, Lee MM, Comstock LE. Human symbionts use a host-like pathway for surface fucosylation. Science. 2005;307(5716):1778-81.
55. Fletcher CM, Coyne MJ, Villa OF, Chatzidaki-Livanis M, Comstock LE. A general O-glycosylation system important to the physiology of a major human intestinal symbiont. Cell. 2009;137(2):321-31.
56. Newman JC, Verdin E. Ketone bodies as signaling metabolites. Trends in endocrinology and metabolism: TEM. 2014;25(1):42-52.
57. Younge NE, Newgard CB, Cotten CM, Goldberg RN, Muehlbauer MJ, Bain JR, et al. Disrupted Maturation of the Microbiota and Metabolome among Extremely Preterm Infants with Postnatal Growth Failure. Scientific reports. 2019;9(1):8167.
58. Castignetti D, Petithory JR, Hollocher TC. Pathway of oxidation of pyruvic oxime by a heterotrophic nitrifier of the genus Alcaligenes: evidence against hydrolysis to pyruvate and hydroxylamine. Archives of biochemistry and biophysics. 1983;224(2):587-93.
59. Amarger N, Alexander M. Nitrite formation from hydroxylamine and oximes by Pseudomonas aeruginosa. Journal of bacteriology. 1968;95(5):1651-7.

60. Thompson AL, Monteagudo-Mera A, Cadenas MB, Lampl ML, Azcarate-Peril MA. Milk- and solid-feeding practices and daycare attendance are associated with differences in bacterial diversity, predominant communities, and metabolic and immune function of the infant gut microbiome. Frontiers in cellular and infection microbiology. 2015;5:3.
61. Laursen MF, Bahl MI, Michaelsen KF, Licht TR. First Foods and Gut Microbes. Frontiers in microbiology. 2017;8:356.
62. Grau D, Clarivet B, Lotthe A, Bommart S, Parer S. Complications with peripherally inserted central catheters (PICCs) used in hospitalized patients and outpatients: a prospective cohort study. Antimicrobial resistance and infection control. 2017;6:18.
63. Yu X, Yue S, Wang M, Cao C, Liao Z, Ding Y, et al. Risk Factors Related to Peripherally Inserted Central Venous Catheter Nonselective Removal in Neonates. BioMed research international. 2018;2018:3769376.
64. Henderickx JGE, Zwittink RD, van Lingen RA, Knol J, Belzer C. The Preterm Gut Microbiota: An Inconspicuous Challenge in Nutritional Neonatal Care. Frontiers in cellular and infection microbiology. 2019;9:85.
65. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and environmental microbiology. 2009;75(23):7537-41.
66. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PloS one. 2013;8(4):e61217.
67. Rohart F, Gautier B, Singh A, Le Cao KA. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS computational biology. 2017;13(11):e1005752.
68. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology. 2014;15(12):550.
69. Yap IK, Li JV, Saric J, Martin FP, Davies H, Wang Y, et al. Metabonomic and microbiological analysis of the dynamic effect of vancomycin-induced gut microbiota modification in the mouse. J Proteome Res. 2008;7(9):3718-28.
70. Yap IK, Kho MT, Lim SH, Ismail NH, Yam WK, Chong CW. Acclimatisation-induced stress influenced host metabolic and gut microbial composition change. Molecular bioSystems. 2015;11(1):297-306.

Tables

Table 1. Demographics and characteristics of all infants studied.

	Preterm ($n = 19$)	Term ($n = 20$)	P value
Number of stool samples for omics studies	58	83	-
Mean gestational age, weeks (SD)	31.79 (2.97)	39.85 (0.81)	< 0.0001
Mean birth weight, g (SD)	1349.21 (515.98)	3036.5 (214.44)	< 0.0001
Delivery: caesarean/ vaginal	18/1	5/15	< 0.0001
Gender: male/ female	8/11	10/10	0.6211
Race: Malay/ Chinese/ Indian	11/7/1	5/11/4	0.0865
Number of twins/ triplets	3/1	0/0	-
Antibiotic treatment: yes/no	11/8	0/20	< 0.0001

P values < 0.05 were considered as significant.

t test for continuous data; Chi-square test for categorical data.

Table 2. Clinical characteristics of the preterm infants studied.

Infant	Gestational age (weeks)	Birth weight (g)	Hospital stay (days)	Parenteral nutrition (days)	Invasive ventilation (days)	Respiratory Distress Syndrome (RDS)	Antibiotherapy (days)	Death	Sample collection*	
									Culture[a]	16s rRNA
B1	33	1570	21	0	0	No	0	No	m, w1, w2	m, m6, m12
B2	33	1725	13	0	0	No	0	No	m, w1	w1, m6, m12
B4	30	1510	44	20	0	Yes	12	No	m, w1, w2	m, w1, m6, m12
B6	35	1390	36	12	3	Yes	7	No	m, w1, w2	m, w1, m6, m12
B7	33	490	34	34	14	Yes	14	Yes	m, w1, w2	m, w1, w2
B8	30	1060	5	13	5	Yes	3	No	m	m, w1, w2
B9	36	1480	20	0	0	No	0	No	m, w1, w2	w1, w2, m6
B11	36	1555	29	0	0	No	7	No	m, w1, w2	m, w1, w2
B12	34	2845	6	0	0	No	0	No	m	m, m6
B13	32	1640	25	11	0	Yes	0	No	m, w1, w2	m, w1, w2, m6, m12
B14	32	785	57	10	3	Yes	4	No	m, w1, w2	m, w1, w2, m6, m12
B15	32	1540	25	0	0	Yes	3	No	m	m, w1, w2
B16	30	925	56	17	0	Yes	3	No	m, w2	m, w2
B17	30	1240	39	4	0	Yes	6	No	m, w2	w2
B18	30	1220	30	0	0	Yes	0	No	m, w2	m, w2
B22	23	550	18	8	18	Yes	10	Yes	m, w1	m, w1
B25	29	1125	42	14	0	No	9	No	m, w1, w2	m, w1, w2, m6

B26	33	1690	14	0	0	No	0	No	m, w1, w2	m, w1, m6, m12
B27	33	1295	33	6	0	No	0	No	m, w1, w2	m, w1
Median (95% CI)	32(31, 33)	1390 (1141, 1639)	25 (18, 32)	6 (2, 10)	0 (0, 2)		3 (1, 5)			

*m, meconium; w1, week 1; w2, week 2; m6, month 6; m12, month 12.

[a]culture for samples collected during hospitalisation.

Table 3. List of 1H NMR-derived metabolites that differ significantly between preterm and term infants at different time-points.

	Metabolites		
Time-point	**Relatively higher in Preterm**	**Relatively higher in Term**	**Chemical shifts, ppm (multiplicity)**
Meconium Preterm (n =15) vs Term(n = 19); R2X= 0.621; Q2Y= 0.179	Glycerol		3.5698(m), 3.6625(m), 3.7978(tt)
Week 1 Preterm (n =14) vs Term(n = 15); R2X= 0.39; Q2Y= 0.487		a-glucose	5.24(d), 3.56(dd), 3.70(t), 3.40(t), 3.83(m), 3.72(dd), 3.85(m)
		Methylmalonic acid (MMA)	1.2362(d), 3.3074(q)
	Valine		0.9936 (d), 1.0128 (d)
	Leucine		0.9543 (m)
	Isoleucine		0.9945(t), 1.012(d)
	Tyrosine		7.2004 (d), 6.9097 (d)
	Phenylalanine		7.335 (m), 7.3863 (m) ,7.4347 (m)
Week 2 Preterm (n =11) vs Term(n = 15); R2X= 0.228; Q2Y= 0.143	Nil	Nil	
Month 6 Preterm (n =10) vs Term(n = 17); R2X= 0.7; Q2Y= 0.66	Succinate		2.4103 (s)
	Trimethylamine-N-oxide (TMAO)		3.3645 (s)
	Citrate		2.5488 (d), 2.6988 (d)
		BHBA	1.2117 (d)
		Fucose	1.2525 (d)
		Pyruvatoxime	2.0642 (s)
Month 12 Preterm (n =8) vs Term(n = 17); R2X= 0.33; Q2Y= 0.216		Taurine	3.2572 (t), 3.4584 (t)
		BHBA	1.2117 (d)
		Fucose	1.2525 (d)
	Tyrosine		7.2004 (d), 6.9097 (d)
	Phenylalanine		7.335 (m), 7.3863 (m) ,7.4347 (m)

Table 4. Cross-modelling of the stool metabolites profiles using a stepwise selection procedure.

Collection time	Variable	AICc	SS (trace)	Pseudo-F	*P*	Prop.	Cumul.	Res.df
(A) Modelling of the differentially expressed gut microbial pattern of all infants using infant demographics as the predictor								
Month 12	(+)Gestational age	172.85	5188.1	4.2861	0.001	0.16305	0.16305	22
(B) Modelling of the elevated stool metabolites pattern of all infants using infant demographics as the predictor								
Meconium	(+)Birth weight	-186.74	0.041532	9.2098	0.003	0.21819	0.21819	33
Month 12	(+)Birth weight	-146.81	0.0090759	4.565	0.019	0.17184	0.17184	22
(C) Modelling of the elevated stool metabolites pattern of preterm infants using patient clinical information as the predictor								
Month 6	(+)PICC line insertion (Yes/ No)	-51.058	0.011965	3.7559	0.015	0.25031	0.53349	7
Month 12	(+)Isolation of bacteria resistant to 3rd generation cephalosporins	-41.009	0.0073412	4.9916	0.044	0.49958	0.49958	5

Note: Prop. = explanatory proportion, Cumul. = cumulative explanatory proportion, Res.df = residual degree of freedom, (+) = inclusion, PICC = Peripherally inserted central catheter. Only elements with significant effect are included.

Supplemental Information Note

Additional files:

1. File name:

Supplementary Figures

File format: .docx

Title of data:

Supplementary Figure 1. Alpha diversity of the faecal microbiota community between (A) term and preterm infants and (B) time points based on Pielou’s evenness, Simpson’s, and Shannon’s indices.

Supplementary Figure 2. Taxonomical classification and diversity of gut microbiome of term and preterm infants.

Supplementary Figure 3. Covariance plots showing the colour-coded significance of stool metabolite profiles calculated using permutation test between term and preterm infants from (A) meconium, (B) week 1, (C) week 2, (D) month 6 and (E) month 12.

2. File name:

Supplementary Tables

File format: .docx

Title of data:

Supplementary Table 1. Permutational multivariate analysis of variance (PERMANOVA) of the association between faecal bacterial community and metabolites compositions of term and preterm infants and time points.

Supplementary Table 2. Pairwise PERMANOVA of the association between faecal bacterial community composition of term and preterm infants and time points.

Supplementary Table 3. BLAST results of differentially abundant operational taxonomic units (OTUs).

Figures

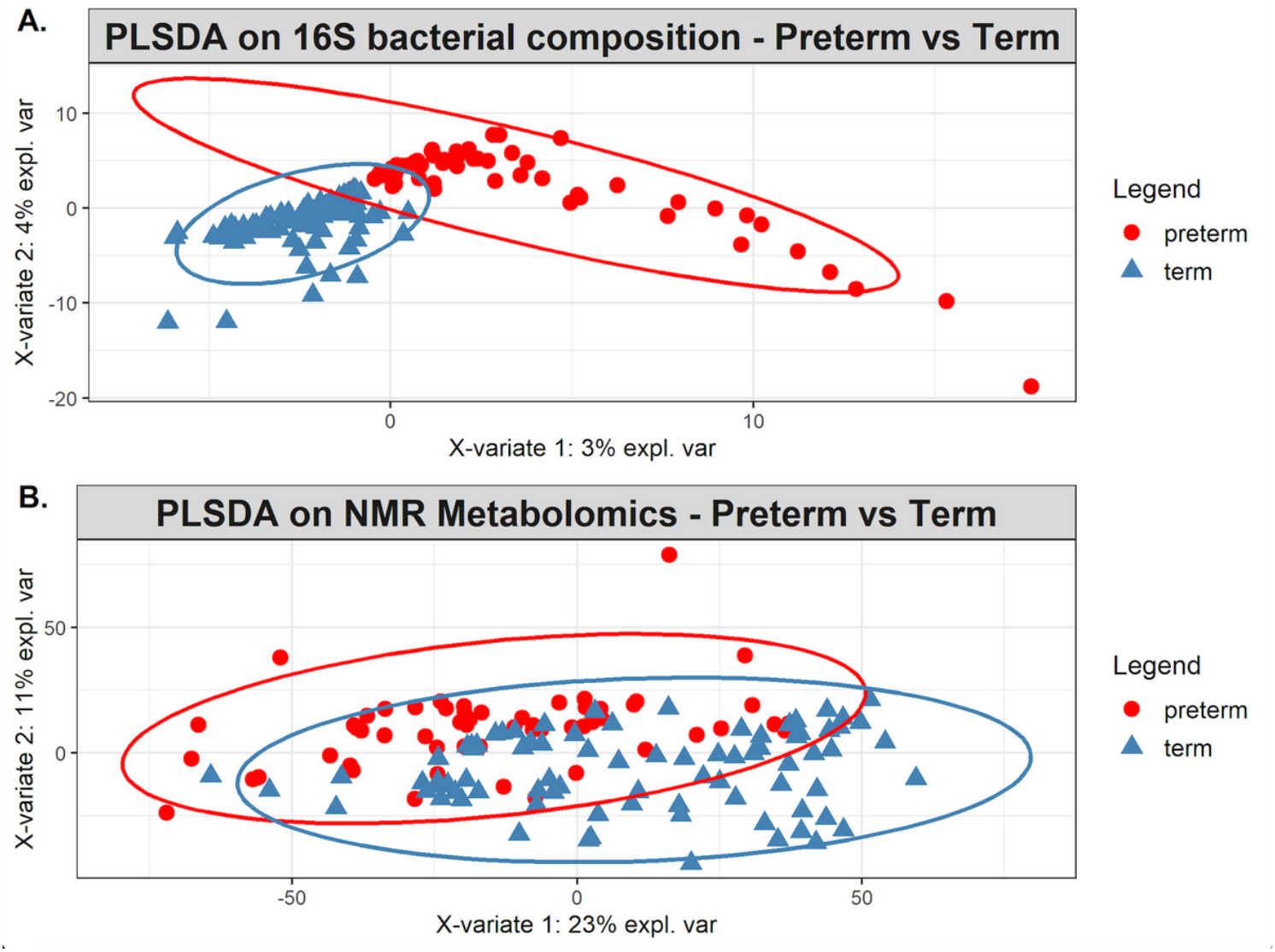

Figure 1

PLS-DA models depicting inter-subject distances based on the (A) 16S bacterial composition and (B) NMR metabolomics. Both validated with 5 fold cross-validation. Prevalence uncovers distinct clusters

among the subject group. Eclipses for each group are calculated based on a 95% confidence level.

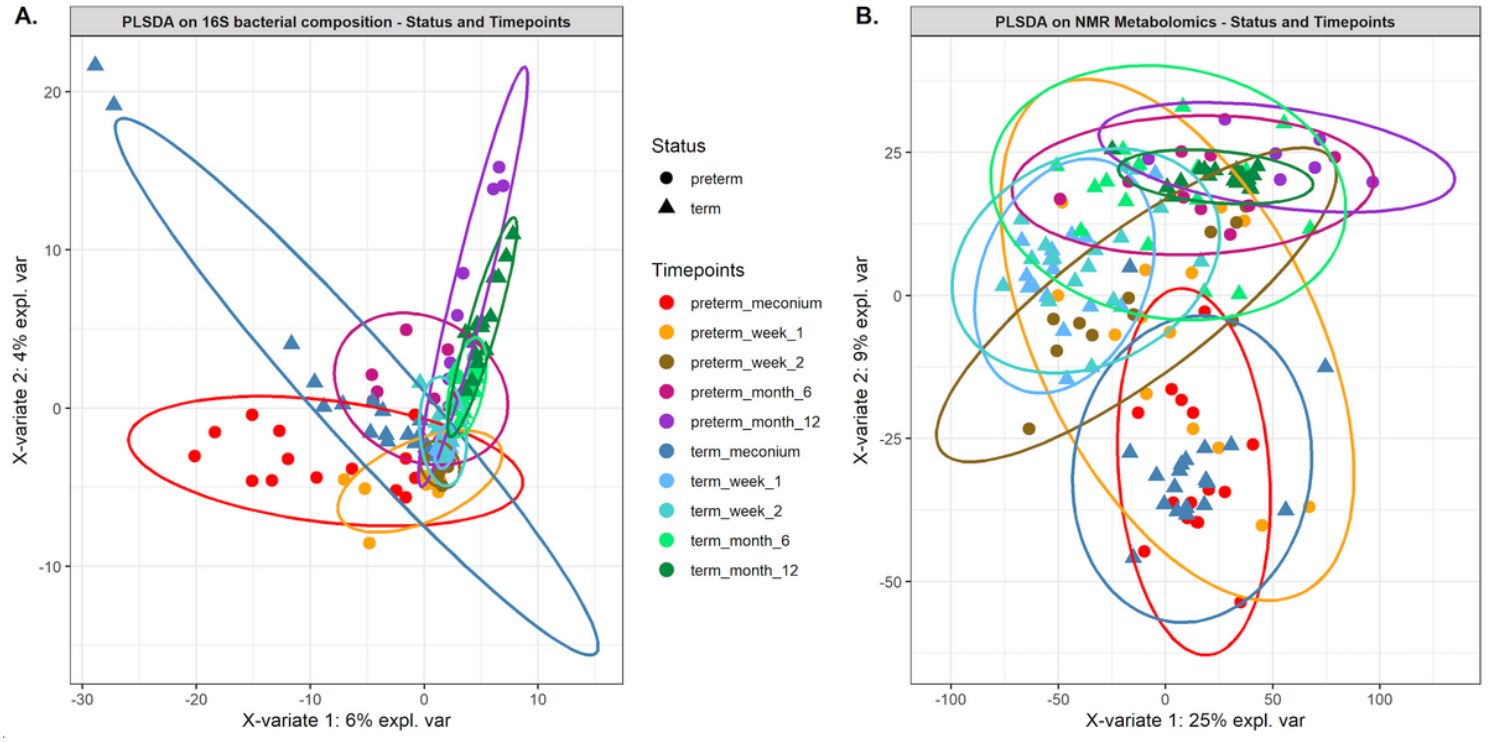

Figure 2

PLS-DA models of (A) 16S bacterial composition and (B) NMR metabolomics build on inter-subject vs the five collection time points and colour coded by group with 95% confidence eclipses. Both validated with 5 fold cross-validation.

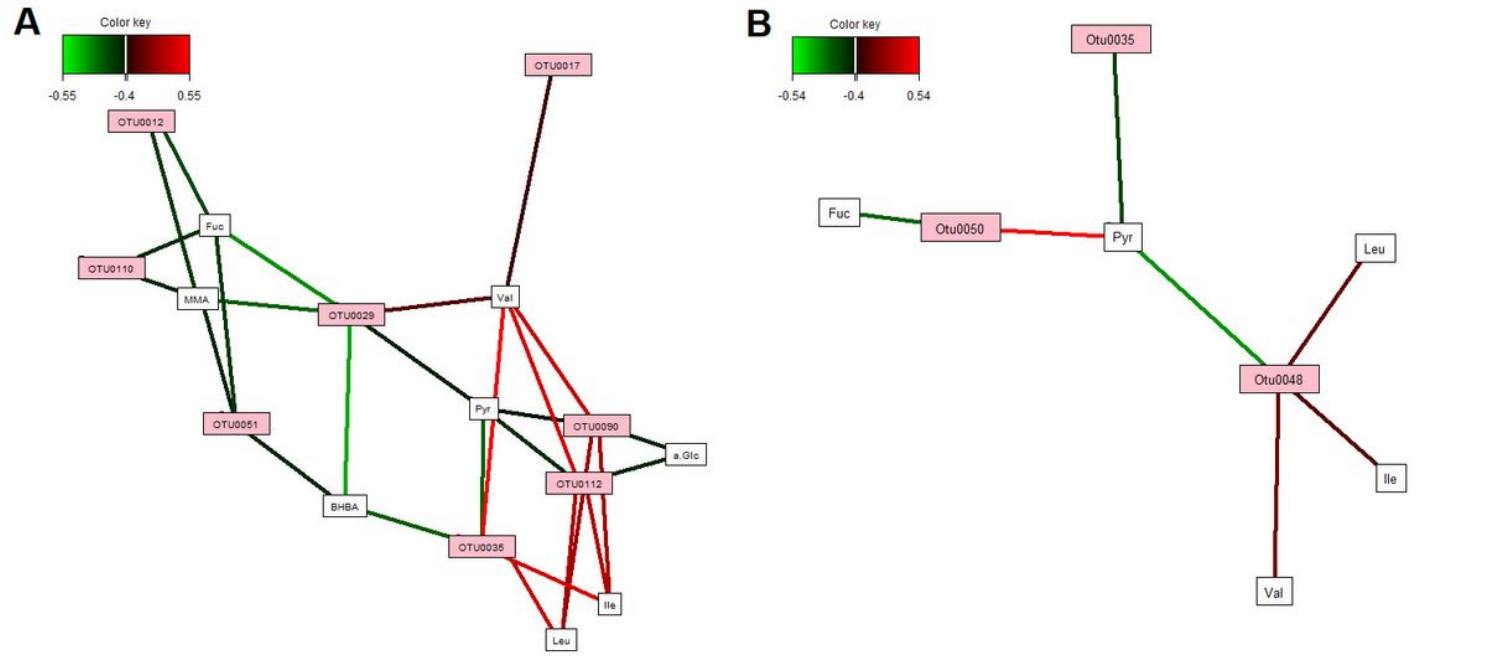

Figure 3

Spares partial least squared correlations (sPLS) between differentially expressed operational taxonomic units (OTUs) and identified significant metabolites in (A) preterm and (B) term groups. sPLS in regression mode (predict Y from X) to model a causal relationship between the OTUs and metabolites. The network is displayed graphically as nodes (OTUs and metabolites) and edges (biological relationship between nodes), with the edge colour intensity indicating the level of the association: red, positive, and green, negative. Only the strongest pairwise associations were displayed, with a cut-off threshold of 0.4 (positive and negative). a.Glu, α-glucose; BHBA, β-hydroxybutyric acid; Fuc, fucose; Ile, isoleucine; Leu, leucine; MMA, methylmalonic acid; Pyr, pyruvatoxime; Val, valine.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.docx
- SupplementaryFigures.docx