

### MICROBIOLOGY

## Development of the gut microbiota in southern Indian infants from birth to 6 months: a molecular analysis

Jayakanthan Kabeerdoss<sup>1</sup>, Shama Ferdous<sup>1</sup>, Ramadass Balamurugan<sup>1</sup>, John Mechenro<sup>1</sup>, R. Vidya<sup>1</sup>, Sridhar Santhanam<sup>2</sup>, Atanu K. Jana<sup>2</sup> and Balakrishnan S. Ramakrishna<sup>1\*</sup>

<sup>1</sup>Department of Gastrointestinal Sciences, Christian Medical College, Vellore 632004, India

<sup>2</sup>Department of Neonatology, Christian Medical College, Vellore 632004, India

(Received 6 July 2012 – Final revision received 6 January 2013 – Accepted 29 January 2013)

*Journal of Nutritional Science* (2013), vol. 3, e2, page 1 of 7

doi:10.1017/jns.2013.6

#### Abstract

Acquisition of the gastrointestinal microbiota at birth may have long-term health impacts. We longitudinally characterised major microbial communities in the faeces of a cohort of infants using molecular methods. Faecal samples were prospectively obtained at several time points after birth from eighty-three infants. Real-time PCR using SYBR green and primers targeted at 16S rRNA gene sequences were used to quantify *Bifidobacterium*, *Lactobacillus acidophilus* group, *Bacteroides-Prevotella* group, *Enterobacteriaceae*, *Enterococcus*, *Clostridium coccoides-Eubacterium rectale* group, *Clostridium leptum* group and *Staphylococcus*. Microbial community abundance was expressed relative to amplification of sequences conserved universally for domain bacteria. Faecal copy number of 16S rRNA genes increased non-significantly from a mean of  $4.1 \times 10^9$ /g on day 1 to  $1.1 \times 10^{10}$ /g on day 4. All microbial communities were detected from day 1 after birth. *Enterobacteriaceae* and lactobacilli predominated on day 1, while bifidobacteria and staphylococci increased on day 4. *Bacteroides-Prevotella* and *C. coccoides-E. rectale* increased by day 180. *C. leptum* was detected in half of the cohort at birth and in a slightly larger percentage by 6 months. Caesarean section was associated with delayed colonisation by several bacterial communities. Higher socio-economic status was associated with more abundant lactobacilli and *Bacteroides-Prevotella* at 90 and 180 d. Supplemental feeding was associated with a reduction in *Enterobacteriaceae*. Microbial colonisation of the gut was well established on the first day of birth, and relative abundance of microbial communities was influenced by mode of delivery, socio-economic status and supplemental feeding. These findings may have relevance to infant nutrition and growth.

**Key words:** Faeces: 16S rRNA gene: *Bifidobacterium*: *Lactobacillus*: *Bacteroides-Prevotella*: *Clostridium*

The gastrointestinal tract is host to a number of bacteria that reside within its lumen. The adult human harbours approximately  $10^{14}$  microbial organisms belonging to 400–500 different species in the gastrointestinal tract<sup>(1–3)</sup>. The gut is sterile while the fetus is in the uterus, and becomes colonised with microbes immediately after birth. The nature of the microbes that colonise the gut initially is thought to be important in conditioning the development of the host immune system<sup>(4,5)</sup>. The establishment of the gut microbiota after birth was studied from the late 1950s through culture of the faeces in neonates<sup>(6–9)</sup>. Culture-based studies showed that facultative

anaerobes (*Enterobacteriaceae*, streptococci and staphylococci) appear first. These establish an anaerobic luminal environment in the colon leading to the subsequent appearance of obligate anaerobes such as bifidobacteria<sup>(10)</sup>. Differences in the faecal microbiota have been noted between vaginally delivered neonates and those delivered by Caesarean section<sup>(11)</sup>. By the second year of life, the gut microbiota composition stabilises with approximately 400 species of bacteria, although studies in southern India suggest there is further remodelling of the microbiota through childhood and adolescence<sup>(12)</sup>. Traditional bacterial culture is not capable of detecting many of the fastidious

**Abbreviation:** Ct, threshold cycle.

\* **Corresponding author:** B. S. Ramakrishna, fax +91 4162282486, email [rama@cmcvellore.ac.in](mailto:rama@cmcvellore.ac.in)



obligate anaerobes present in the gut as they rapidly die on exposure to minimal amounts of oxygen. The use of molecular techniques to study the gut flora indicates that 50–70 % of the faecal microbiota is not fit for culture, and thus molecular techniques give us a more comprehensive representation of the microbiota composition. Several molecular studies have examined the development of the gut microbiota in high-income countries<sup>(1,3–15)</sup>.

The developing intestinal microbiome is important in intestinal immune conditioning and in nutrition<sup>(2,4,5)</sup>. Specific microbial communities such as lactobacilli and bifidobacteria are used as probiotics to alter the gut microbiota profile in a manner that will favour health. Neonates in low-income countries are born in a contaminated environment. Infant feeding practices in these countries differ from those in high-income countries. The only information on development of the gut microbiota in neonates in low-income countries comes from limited culture-based studies done four decades ago<sup>(7,16)</sup>. The present studies were undertaken to examine the development of the gut microbiota in neonatal life, extending through to 6 months after birth, using molecular techniques directed at the 16S rRNA gene, specifically to elucidate the pattern of development of the gut microbiota in a developing country and to ascertain differences in microbiota composition associated with mode of delivery, socio-economic status, feeding practices and birth weight.

## Methods

The study was carried out in the Departments of Gastrointestinal Sciences and Neonatology of the Christian Medical College, Vellore. Infants delivered in the inpatient services of the hospital were the primary participants. Mothers admitted to the obstetric wards for delivery were informed of the study on admission. Inclusion criteria were infants born in the hospital after 28 weeks of gestation, hailing from families residing within 30 km of the hospital. Exclusion criteria were preterm infants born at gestational age <27 weeks, birth weight <900 g, infants with lethal anomalies and critically-ill infants requiring cardio-respiratory support on the first day of life with oxygen requirement of more than 80 % at 24 h of age. Focus discussions were held with mothers prior to obtaining informed written consent. This study was conducted according to the guidelines laid down in

the Declaration of Helsinki and all procedures involving human subjects/patients were approved by the Institutional Review Board of the Christian Medical College, Vellore (approval no. 13-19-03-2008). Written informed consent was obtained from all participating mothers.

Demographic details and clinical details including mode of delivery and immediate postnatal interventions were entered into case report forms. Information regarding the infant's health and feeding was recorded. The infant's length and weight were recorded, at birth and at 3 and 6 months after birth. The socio-economic status of the family was scored using a standard scoring system based on education, possessions and income adjusted for 2007<sup>(17)</sup>. Infants were followed up to the age of 1 year. Samples of stool were obtained from each infant on postnatal days 1, 2, 4, 7, 14, 28, 90 and 180, transported immediately to the laboratory, divided into aliquots, identifier numbers allotted and stored at  $-80^{\circ}\text{C}$  until processing. DNA was extracted from stool in batches using QiaAMP Stool DNA mini-kits (Qiagen), checked for quality, quantified using Qubit dsDNA BR assay kits (Invitrogen) in a Qubit fluorometer (Invitrogen) and stored at  $-20^{\circ}\text{C}$  until analysis. The DNA was subjected to real-time PCR using primers (Table 1) which were targeted at 16S rRNA gene sequences specific to microbial communities at the species, genus or group level<sup>(18–24)</sup>. The microbial communities that were targeted included *Bifidobacterium*, *Lactobacillus acidophilus* group, *Bacteroides–Prevotella* group, *Clostridium coccooides–Eubacterium rectale* group, *Clostridium leptum* group, *Enterococcus* and *Staphylococcus*. These microbial communities were targeted because of their significance to human health. The DNA sequences of interest were amplified using SYBR green in a Chromo4 instrument (Biorad)<sup>(25)</sup>. DNA of bacteria of interest was expressed as the relative difference relative to the amplification of 'universal' DNA sequences conserved throughout domain bacteria. Melting curve analysis was always done to check the specificity of the amplification. Quantification was based on the fluorescence intensity obtained from the intercalated SYBR Green dye. The cycle number at which the signal was first detected – the threshold cycle (Ct) – correlated with the original concentration of DNA template. 16S rRNA gene copy number was calculated for total bacteria in faeces using an external set of standard DNA copies for the universal bacteria sequences. For the other bacterial communities studied,

**Table 1.** Primers used for real-time PCR amplification of the bacterial communities of interest

Primer	Specificity	Forward sequence (5'–3')	Reverse sequence (5'–3')	Reference
Universal	Domain bacteria	CGGTGAATACGTTCCCGG	TACGGCTACCTTGTACGACTT	Suzuki <i>et al.</i> <sup>(18)</sup>
<i>Bifidobacterium</i>	Genus	CGGGTGTAGTAATGCGTGACC	TGATAGGACGCGACCCCA	Sokol <i>et al.</i> <sup>(19)</sup>
<i>Lactobacillus acidophilus</i>	Group	AGCAGTAGGGAATCTTCCA	CGCCACTGGTGTTCYTCCATATA	Sokol <i>et al.</i> <sup>(19)</sup>
<i>Bacteroides–Prevotella</i>	Group	GGTGTCCGCTTAAGTGCCAT	CGGACGTAAGGGCCGTGC	Balamurugan <i>et al.</i> <sup>(20)</sup>
Enterobacteria	Genus	CATTGACGTTACCCGAGAAG	CTCTACGAGACTCAAGCTTGC	Ahmed <i>et al.</i> <sup>(21)</sup>
<i>Clostridium coccooides–Eubacterium rectale</i>	Group	CGGTACTGACTAAGAAGC	AGTTTCATTCTTGCGAAC	Rinttila <i>et al.</i> <sup>(22)</sup>
<i>Clostridium leptum</i>	Subgroup	GCACAAGCAGTGGAGT	CTTCTCCGTTTTGTCAA	Matsuki <i>et al.</i> <sup>(23)</sup>
<i>Enterococcus</i>	Genus	CCCTTATTGTTAGTTGCCATCATT	ACTCGTTGTACTIONCCATTGT	Rinttila <i>et al.</i> <sup>(22)</sup>
<i>Staphylococcus</i>	Genus	ACGGTCTTGCTGTCACTATA	TACACATATGTTCTCCCTAATAA	Matsuda <i>et al.</i> <sup>(24)</sup>



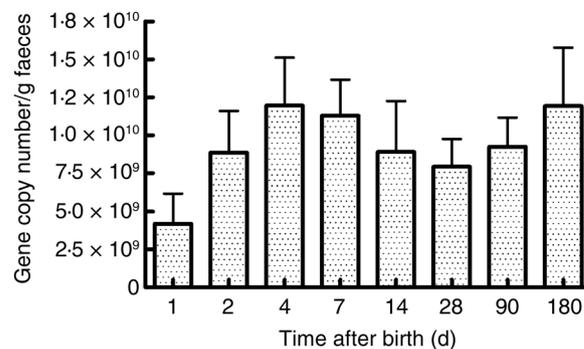
relative DNA copy abundance was calculated from the Ct at which the target DNA was detected relative to the Ct at which 'universal bacterial' DNA was detected after amplification, using the formula  $2^{-\Delta Ct}$  and expressed as relative fold difference<sup>(26)</sup>. In brief, the calculation of the relative fold difference was based on the difference between the PCR cycle number at which the 'universal bacteria' amplicon reached a set threshold (Ct) and the PCR cycle number at which the amplicon of interest reached the same threshold. Each PCR plate used in the quantitative PCR contained, besides the PCR for the amplicons of interest, PCR with primers for amplification of the 'universal domain bacteria' sequences that are strongly conserved in the 16S rRNA gene. The ratio calculated for each of the duplicate samples correlated well with the other, with an adjusted  $R^2$  in the region of 0.80.

Bacterial quantification was handled as a continuous variable. The data on bacterial quantification were expressed as the relative fold difference which is a logarithmic value, and significance of difference was evaluated using non-parametric tests. Repeated measures ANOVA (with *post hoc* Tukey's multiple comparison) was done to compare microbial abundance within the same infants at different time points. The Student's *t* test with Welch's correction was used to compare two groups at any time point.

## Results

A total of eighty-three infants (forty-one male) participated in these studies. Their mean birth weight was 2.87 (SD 0.45) kg and birth length 48.4 (SD 2.5) cm. At 3 months age, their mean weight and length were 5.9 (SD 1.0) kg and 60.5 (SD 3.6) cm, respectively. Of the infants, ten were delivered by Caesarean section while the rest were vaginally delivered (eight had forceps or suction cup used during delivery); three infants were of very low birth weight, eight were of low birth weight and the remainder were of normal birth weight. The gestation period was 32–36 weeks for seven infants, while the rest were all born after 36 weeks of gestation; delivery was post-term in 18 mothers. A total of six infants received formula milk in the first week of life, while the others had exclusive breast-feeding for periods ranging from 1 to 15 months with the majority continuing exclusive breast feeds for 6 months. A total of eight infants were weaned to solid foods in the fourth month of life, and eighteen in the fifth month of life, while weaning foods were introduced in the remaining infants at the age of 6 months or later. The mean socio-economic score was 16 (SD 5) where a score of eleven–fifteen was classified as middle lower middle class and sixteen–twenty-five as upper middle class.

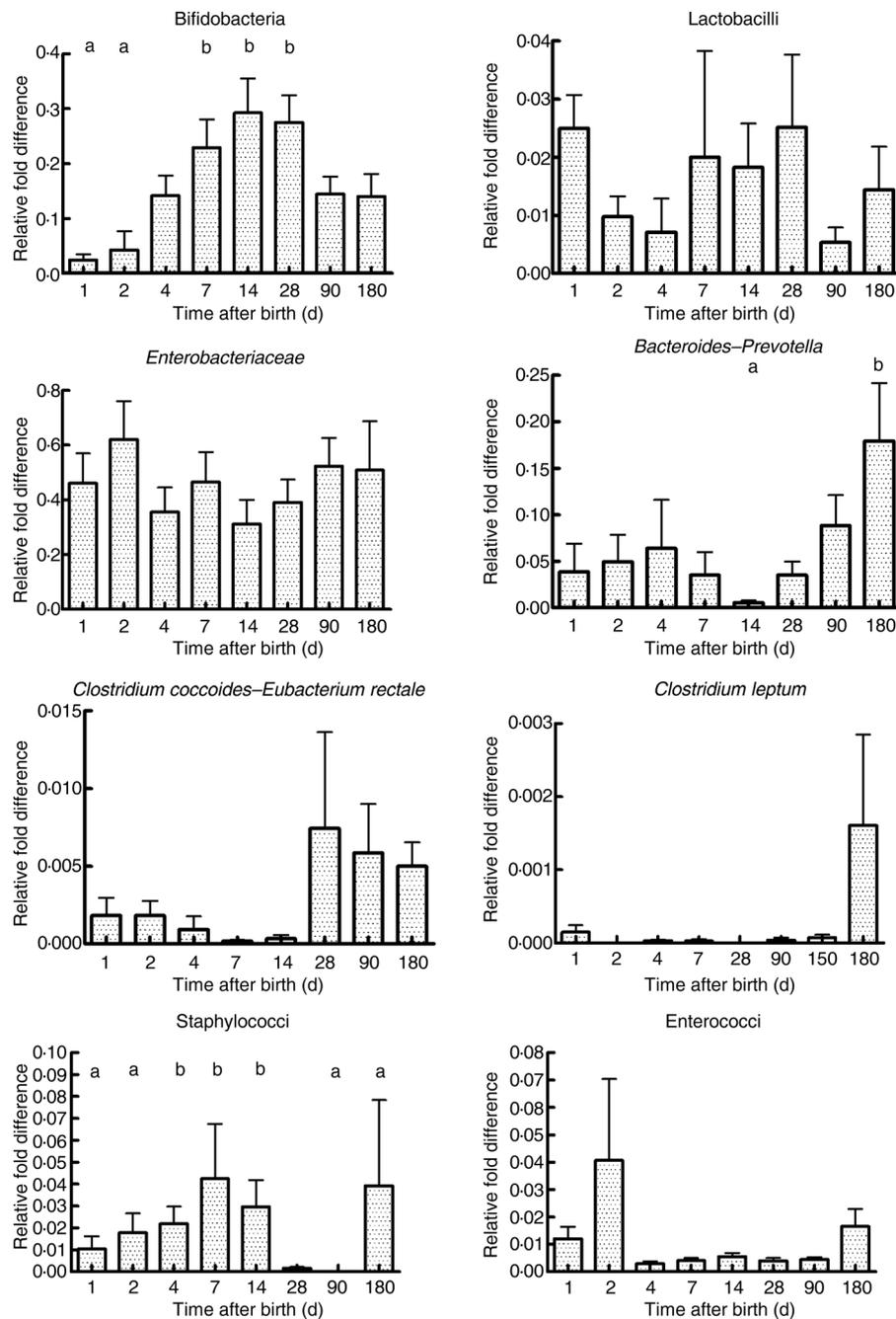
As seen in Fig. 1, there were a mean of  $4.1 \times 10^9$  (SEM  $1.9 \times 10^9$ ) bacterial rRNA gene copies/g stool on the first day of life that increased to  $1.2 \times 10^{10}$  (SEM  $3.1 \times 10^9$ ) copies/g stool by day 4 of life (NS) and remained stable at those levels for the first 6 months. All bacterial groups tested were detected in the stool from the first day of life and all could be detected until the end of 6 months of age. *Enterobacteriaceae* predominated in the first week of life (Fig. 2). Bifidobacteria were



**Fig. 1.** Total bacterial count in faeces estimated using real-time PCR amplification of universal 16S rDNA sequences conserved for domain bacteria. Quantification was done with reference to an external standard of plasmid DNA copies. Values are means, with standard errors represented by vertical bars. None of the values was statistically significantly different from the others.

low in numbers on day 1 after birth but increased 10-fold or more from day 4 and remained constant thereafter (Fig. 2). Staphylococci also increased by day 4 and then returned to earlier levels by the end of the first month. The *Bacteroides–Prevotella* group, a prominent microbial community in the adult gut, was present in small numbers after birth and increased around 6 months of age to become one of the more abundant communities in the faecal microbiota. Enterococci were less abundant constituents of the faecal microbiota, remaining at somewhat similar levels from day 1 to 6 months of age. *C. coccoides–E. rectale*, a prominent carbohydrate-fermenting bacterial community, were found in very low relative numbers in the faeces of these infants until the age of 6 months (Fig. 2). The *C. leptum* group of bacteria was not detected in over a half of infants in the study.

The mode of delivery influenced bacterial colonisation. On the first day of life, lactobacilli and *Enterobacteriaceae* were more abundant in the faeces of vaginally delivered infants compared with those delivered by Caesarean section (Fig. 3), but their abundance in subsequent days was not different between the two groups. *Bacteroides–Prevotella* were more abundant in the faeces of vaginally delivered infants compared with those delivered by Caesarean section from birth up to 90 d (Fig. 3). Bifidobacteria were more abundant in vaginally delivered infants at 90 d (Fig. 3). The socio-economic status of the family was associated with differences in the faecal microbiota at 90 and 180 d when lactobacilli and *Bacteroides–Prevotella* were noted to be more abundant in infants from families with higher socio-economic score than in those coming from a lower socio-economic stratum (Fig. 4). Birth weight did not have a significant association with microbiota composition at birth, although lactobacilli were less abundant at 90 d in low-birth-weight infants (<2.5 kg) compared with large (>3.5 kg) infants. The overwhelming majority of infants were breast-fed from birth, making it difficult to compare the effect of breast-feeding and bottle-feeding on the microbiota. Cessation of breast-feeding by day 90 was not associated with any significant alteration in abundance of these microbial communities (data not shown). Infants who received supplemental cows' milk (in addition to breast feeds) by day 90 had lower numbers of *Enterobacteriaceae* than infants who were exclusively breast fed. We did not find



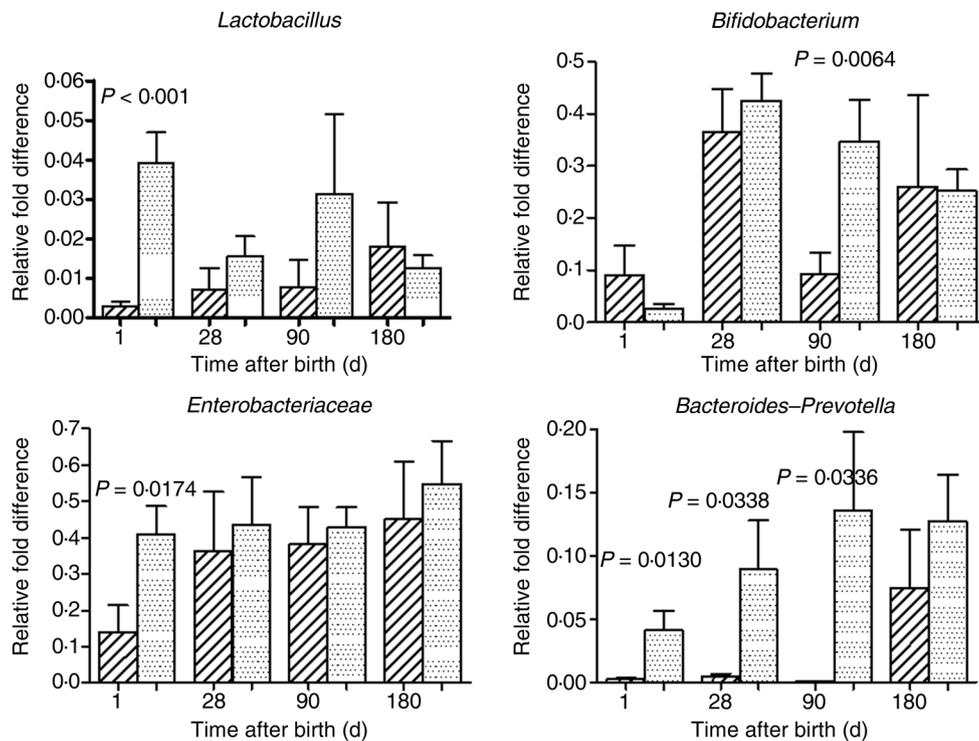
**Fig. 2.** Abundance of specific microbial communities in the faeces of neonates followed longitudinally at 6 months. Data were acquired by real-time PCR targeting specific sequences of the 16S rRNA gene and are expressed as relative difference in comparison with amplification of the 'universal' Bacteria domain sequence. Values are means, with standard errors represented by vertical bars. <sup>a,b</sup>Mean values with unlike letters were significantly different ( $P < 0.05$ ; repeated measures ANOVA with *post hoc* Tukey's test).

any association between faecal microbiota abundance at birth and length at birth or at the age of 1 year.

## Discussion

This study provides a longitudinal documentation of the faecal microbial community composition of Indian infants from birth to 6 months of age. Faecal bacterial counts reached peak numbers by the fourth day of life and remained steady thereafter in the first 6 months of life. *Enterobacteriaceae* and

lactobacilli were the most abundant bacteria in the first 2 d of life, while bifidobacteria and staphylococci increased in faeces by the fourth day, possibly accounting for the increase in faecal bacterial counts. *Bacteroides-Prevotella* group, another prominent component of the faecal microbiota, increased at the time of weaning. *C. coccoides-E. rectale* were low in abundance in the first 6 months of life, while the *C. leptum* group was not detected in significant numbers of infants. The study also confirms differences in the faecal microbiota between vaginally delivered infants and those delivered by



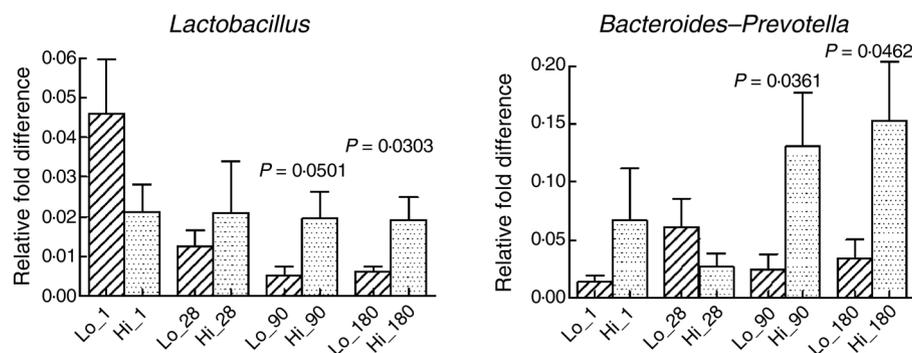
**Fig. 3.** Association of mode of delivery (normal vaginal delivery (▨); lower segment Caesarean section (▩)) and faecal microbial colonisation. Values are means, with standard errors represented by vertical bars.  $P$  values shown were calculated using the Student's  $t$  test with Welch's correction.

Caesarean section, and demonstrates associations between gut colonisation patterns and socio-economic status of the family, exclusive breast-feeding and weaning.

The faecal microbiota is composed of obligate anaerobes with a smaller contribution from facultative anaerobes. The strict anaerobes are often difficult to culture. In the last two decades, the study of the faecal microbiota has been greatly enhanced by the advent of molecular methods<sup>(27)</sup>. Most of these techniques depend upon targeting of the 16S rRNA gene, which aids in taxonomic identification of the microbes due to a large database of information available in public databases<sup>(24)</sup>. The present study used real-time PCR targeting these sequences in order to obtain quantitative abundance information of specific groups or genera of bacteria that are of

biological significance to human health. This technique is very sensitive, being able to detect 2–1000 copies/ml of the target organism or group of organisms<sup>(24)</sup>. Recently deep sequencing and next-generation sequencing techniques have been used to examine the faecal microbiota. The latter techniques are particularly useful to identify the presence of novel or previously unknown microbes and are complementary to real-time PCR. For our purposes, the use of real-time PCR was more appropriate, providing quantitative information about bacterial groups or genera of interest.

The development of the neonatal faecal microbiota has been studied by culture-dependent techniques from the 1950s onwards<sup>(6–10)</sup>. Rotimi & Duerden compared bacterial abundance in the first week of life by the percentage of children



**Fig. 4.** Association of socio-economic status and microbial colonisation. Based on the socio-economic score, mothers were classified into three equal groups, with the lowest tertile being represented as Lo\_ and the highest tertile as Hi\_. Values are means, with standard errors represented by vertical bars.  $P$  values shown were calculated using Student's  $t$  test with Welch's correction. The other microbial communities tested did not significantly differ between low and high socio-economic groups (data not shown).



from whose faeces the bacteria could be isolated<sup>(9)</sup>. These studies showed that *Escherichia coli*, enterococci, staphylococci and streptococci were the most prominent colonisers in the first week of life<sup>(10)</sup>. It was postulated that these facultative anaerobes were able to survive in the neonatal colon where the abundance of oxygen does not permit obligate anaerobes to flourish, and that their growth led to a drop in oxygen tension which then allowed strict anaerobes, such as bifidobacteria, *Bacteroides* and *Clostridium* to colonise the gut<sup>(10)</sup>. Colonisation by lactobacilli was reported to be uncommon in the first year of life<sup>(10)</sup>. DNA-based studies of microbial colonisation suggest that *Proteobacteria*, *Bacteroides*, *Firmicutes* (dominated by *Clostridium*), *Actinobacteria* (and its constituent bifidobacteria) and *Verrucomicrobia* are the dominant microbiota in the first few months of life<sup>(28)</sup>. The dominance of *Enterobacteriaceae* among the faecal microbiota in the first week of life has been described in both Western and Japanese neonates<sup>(10,29)</sup>. Our study shows that they continued to be relatively abundant in the first 6 months of life. In the present study, faecal bacterial numbers did not significantly increase after day 1 and remained constant in the first 6 months of life, indicating very rapid colonisation to peak levels immediately after birth. This may be a feature of gut microbial colonisation in low-income countries. A culture-based study compared a low socio-economic status cohort of infants in Pakistan with a cohort in Sweden and found that colonisation of the intestine was significantly more rapid in the former<sup>(30)</sup>.

In the present study lactobacilli, *Bacteroides-Prevotella* and *Enterobacteriaceae* were all significantly less abundant in faeces on the first day of life in neonates delivered by Caesarean section compared with those delivered vaginally. Our findings are corroborated by studies in which infants delivered vaginally had higher numbers of *Bacteroides* and bifidobacteria in their faecal microbiota compared with those delivered by Caesarean section<sup>(31,32)</sup>. Another study of 1-d-old infants using pyrosequencing of 16S rRNA genes suggested that vaginally delivered infants had greater numbers of *Lactobacillus*, *Prevotella* or *Sneathia* species, while faeces of neonates delivered by Caesarean section had greater abundance of *Staphylococcus*, *Corynebacterium* and *Propionibacterium*. Neonates in the latter study were not followed up after 1 d of life<sup>(33)</sup>. In the case of delivery by Caesarean section, our findings suggest that colonisation of the neonatal gut by these specific microbial communities was delayed only by a few days at most.

Breast-fed neonates have been reported to have higher faecal bacterial counts during the first week of life than bottle-fed neonates. In breast-fed infants, bifidobacteria accounted for 90 % of faecal bacteria, with lesser numbers of *Enterobacteriaceae* and enterococci<sup>(34,35)</sup>. In bottle-fed infants, on the other hand, streptococci, *Enterobacteriaceae* and *Bacteroides* tended to predominate<sup>(36)</sup>. Breast milk contains factors that promote the growth of *Bifidobacterium bifidum*, while formula feeds are considered to promote implantation and persistence of *Clostridium perfringens*<sup>(37)</sup>. Exclusive breast-feeding up to the age of 6 months is a message actively promoted to all young mothers in India. In our cohort, only six neonates received infant formula in the first week of life, in conditions where there was partial failure of breast-feeding. There was no significant difference noted in microbial

colonisation of the gut in this group compared with breast-fed infants, but this must be interpreted with caution because of the very low numbers in the non-breast-fed group. The complete cessation of breast-feeding by 3 months did not have significant influence on abundance of faecal microbial communities. However, exclusive breast-feeding (no complementary feeds at all) was associated with a higher abundance of *Enterobacteriaceae* compared with infants who received complementary milk feeds in addition to breast milk. The significance of this observation is not clear and may need further investigation.

During weaning, the diet changes from a high-fat-milk-based diet to a diet rich in carbohydrate. Bacterial oxidative metabolism leads to changes in the redox potential of the colon and results in increased abundance of strict anaerobes such as *Bacteroides* and *Clostridium*. Bifidobacteria, *C. coccoides* and *Bacteroides* were reported to be more prominent after weaning in a multi-country European study<sup>(38)</sup>. In the present study, complementary feeding was introduced in a small number of infants before the age of 6 months and did not significantly alter microbiota composition.

This study had several shortcomings that need to be considered when interpreting the data in conjunction with other data available in the literature. Molecular techniques for examination of a complex microbial ecosystem as in the human gut do not usually correct for the multiplicity of 16S rRNA genes in an individual microbe. However, the gene copy number for total bacteria in our quantitative study was in the region of  $10^9$ – $10^{10}$  copies/g wet weight faeces, which is in excellent agreement with the results obtained by Palmer *et al.* using a similar method of quantitative real-time PCR<sup>(28)</sup>. While the gene copy number is more or less constant in species, it may vary significantly across species and genera. The relative fold difference in this study thus does not give an absolute number, but since measurements were done in the same infants at different time points, the overall direction of change is an important indicator of changes in abundance of a particular microbial community. Secondly, the differences that we had anticipated due to influences such as Caesarean section, type of milk provided and socio-economic status, were relatively minor and not to the level that we had anticipated. This could partly be secondary to the small numbers of infants we had in some of the groups, but it is quite possible that they could truly represent the gut microbiota development pattern in developing countries.

These molecular studies show that the gastrointestinal tract was colonised by microbes even on the first day of life with representation of all major microbial communities. Colonisation reached peak levels almost immediately although there was a minor, non-significant increase in the next 2 d. There were differences in the proportions of the microbial communities in neonates delivered by Caesarean section and neonates born to mothers hailing from a lower socio-economic class but these differences did not persist beyond the second day of life. Introduction of supplemental feeding was associated with reduced abundance of *Enterobacteriaceae*. These studies provide an understanding of gastrointestinal colonisation between birth and weaning in low-income countries that will serve as background information for designing nutritional interventions targeted at the microbiota.



## Acknowledgements

This study was supported by *ad hoc* research grant no. 5/8-1 (223)/D/2008-ECD-II from the Indian Council of Medical Research, New Delhi. J. K. was supported by a Senior Research Fellowship from the Indian Council of Medical Research. The authors' contributions are as follows: J. K. was responsible for the laboratory analyses and statistical analyses; S. F. was responsible for participant recruitment, sample acquisition and sample preparation; R. B. was responsible for study design and initiation of the study; J. M. and R. V. were responsible for participant recruitment, sample acquisition and participant follow up; S. S. and A. K. J. were responsible for participant recruitment, study design and data acquisition; B. S. R. was responsible for study design, data analysis and presentation, and acquisition of funding. All authors read and approved the final manuscript. None of the authors has any conflict of interest to declare.

## References

- O'Hara AM & Shanahan F (2006) The gut flora as a forgotten organ. *EMBO Rep* **7**, 688–693.
- Kau AL, Ahern PP, Griffin NW, *et al.* (2011) Human nutrition, the gut microbiome and the immune system. *Nature* **474**, 327–336.
- Wallace TC, Guarner F, Madsen K, *et al.* (2011) Human gut microbiota and its relationship to health and disease. *Nutr Rev* **69**, 392–403.
- Lee YK & Mazmanian SK (2010) Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science* **330**, 1768–1773.
- Bezirtzoglou E & Stavropoulou E (2011) Immunology and probiotic impact of the newborn and young children intestinal microflora. *Anaerobe* **17**, 369–374.
- Colgan MT (1956) The bacterial flora of the intestinal tract: changes in diarrheal disease and following antimicrobial therapy. *J Pediatr* **49**, 214–228.
- Mata LJ & Urrutia JJ (1971) Intestinal colonization of breast-fed children in a rural area of low socioeconomic level. *Ann NY Acad Sci* **176**, 93–109.
- Mata LJ, Mejicanos ML & Jimenez F (1972) Studies on the indigenous gastrointestinal flora of Guatemalan children. *Am J Clin Nutr* **25**, 1380–1390.
- Rotimi VO & Duerden BI (1981) The development of the bacterial flora in normal neonates. *J Med Microbiol* **14**, 51–62.
- Adlerberth I & Wold AE (2009) Establishment of the gut microbiota in Western infants. *Acta Paediatr* **98**, 229–238.
- Biasucci G, Rubini M, Riboni S, *et al.* (2010) Mode of delivery affects the bacterial community in the newborn gut. *Early Hum Dev* **86**, Suppl. 1, 13–15.
- Balamurugan R, Janardhan HP, George S, *et al.* (2008) Bacterial succession in the colon during childhood and adolescence: molecular studies in a southern Indian village. *Am J Clin Nutr* **88**, 1643–1647.
- Eggesbø M, Moen B, Peddada S, *et al.* (2011) Development of gut microbiota in infants not exposed to medical interventions. *APMIS* **119**, 17–35.
- Karlsson CL, Molin G, Cilio CM, *et al.* (2011) The pioneer gut microbiota in human neonates vaginally born at term—a pilot study. *Pediatr Res* **70**, 282–286.
- Hascoët JM, Hubert C, Rochat F, *et al.* (2011) Effect of formula composition on the development of infant gut microbiota. *J Pediatr Gastroenterol Nutr* **52**, 756–762.
- Albert MJ, Bhat P, Rajan D, *et al.* (1978) Faecal flora of South Indian infants and young children in health and with acute gastroenteritis. *J Med Microbiol* **11**, 137–143.
- Kumar N, Shekhar C, Kumar P, *et al.* (2007) Kuppaswami's socioeconomic status scale – updating for 2007. *Indian J Pediatr* **74**, 1131–1132.
- Suzuki MT, Taylor LT & DeLong EF (2000) Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5-nuclease assays. *Appl Environ Microbiol* **66**, 4605–4614.
- Sokol H, Pigneur B, Watterlot L, *et al.* (2008) *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA* **105**, 16731–16736.
- Balamurugan R, Janardhan HP, George S, *et al.* (2008) Bacterial succession in the colon during childhood and adolescence. *Am J Clin Nutr* **88**, 1643–1647.
- Ahmed S, Macfarlane GT, Fite A, *et al.* (2007) Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. *Appl Environ Microbiol* **73**, 7435–7442.
- Rinttilä T, Kassinen A, Malinen E, *et al.* (2004) Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. *J Appl Microbiol* **97**, 1166–1177.
- Matsuki T, Watanabe K, Fujimoto J, *et al.* (2004) Use of 16S rRNA gene-targeted group-specific primers for real-time PCR analysis of predominant bacteria in human faeces. *Appl Environ Microbiol* **70**, 7220–7228.
- Ramakrishna BS (2007) The normal bacterial flora of the human intestine and its regulation. *J Clin Gastroenterol* **41**, Suppl. 1, S2–S6.
- Matsuda K, Tsuji H, Asahara T, *et al.* (2007) Sensitive quantitative detection of commensal bacteria by rRNA-targeted reverse transcription-PCR. *Appl Environ Microbiol* **73**, 32–39.
- Balamurugan R, Janardhan HP, George S, *et al.* (2008) Molecular studies of faecal anaerobic commensal bacteria in acute diarrhea in children. *J Pediatr Gastroenterol Nutr* **46**, 514–519.
- Schmittgen TD & Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **3**, 1101–1108.
- Palmer C, Bik EM, DiGiulio DB, *et al.* (2007) Development of the human infant intestinal microbiota. *PLoS Biol* **5**, e177.
- Tsuji H, Oozeer R, Matsuda K, *et al.* (2012) Molecular monitoring of the development of intestinal microbiota in Japanese infants. *Benef Microbes* **3**, 113–125.
- Adlerberth I, Carlsson B, de Man P, *et al.* (1991) Intestinal colonization with *Enterobacteriaceae* in Pakistani and Swedish hospital-delivered infants. *Acta Paediatr Scand* **80**, 602–610.
- Rotimi VO, Olowe SA & Ahmed I (1985) The development of bacterial flora of premature neonates. *J Hyg (Lond)* **94**, 309–318.
- Penders J, Thijs C, Vink C, *et al.* (2006) Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118**, 511–521.
- Dominguez-Bello MG, Costello EK, Contreras M, *et al.* (2010) Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci USA* **107**, 11971–11975.
- Stark PL & Lee A (1982) The microbial ecology of the large bowel of breast-fed and formula-fed infants during the first year of life. *J Med Microbiol* **15**, 189–203.
- Yoshioka H, Iseki K & Fujita J (1983) Development and differences of intestinal flora in the neonatal period in breastfed and bottlefed infants. *Pediatrics* **72**, 317–321.
- Benno Y, Sawada K, Mitsuoka T (1984). The intestinal microflora of infants: composition of the faecal flora in breastfed and bottlefed infants. *Microbiol Immunol* **28**, 975–986.
- Solis G, de Los Reyes-Gavilan CG, Fernández N, *et al.* (2010) Establishment and development of lactic acid bacteria and bifidobacteria microbiota in breast-milk and the infant gut. *Anaerobe* **16**, 307–310.
- Fallani M, Amarri S, Uusijarvi A, *et al.* (2011) Determinants of the human infant intestinal microbiota after the introduction of first complementary foods in infant samples from five European centres. *Microbiology* **157**, 1385–1392.