Longitudinal investigation of the faecal microbiota of healthy full-term infants using fluorescence *in situ* hybridization and denaturing gradient gel electrophoresis

Laure C. Roger and Anne L. McCartney

Microbial Ecology and Health Group, Food Microbial Sciences Unit, Department of Food and Nutritional Sciences, University of Reading, Whiteknights Campus, PO Box 226, Reading RG6 6AP, UK

From birth onwards, the gastrointestinal (GI) tract of infants progressively acquires a complex range of micro-organisms. It is thought that by 2 years of age the GI microbial population has stabilized. Within the developmental period of the infant GI microbiota, weaning is considered to be most critical, as the infant switches from a milk-based diet (breast and/or formula) to a variety of food components. Longitudinal analysis of the biological succession of the infant GI/faecal microbiota is lacking. In this study, faecal samples were obtained regularly from 14 infants from 1 month to 18 months of age. Seven of the infants (including a set of twins) were exclusively breast-fed and seven were exclusively formula-fed prior to weaning, with 175 and 154 faecal samples, respectively, obtained from each group. Diversity and dynamics of the infant faecal microbiota were analysed by using fluorescence *in situ* hybridization and denaturing gradient gel electrophoresis. Overall, the data demonstrated large inter- and intra-individual differences in the faecal microbiological profiles during the study period. However, the infant faecal microbiota merged with time towards a climax community within and between feeding groups. Data from the twins showed the highest degree of similarity both quantitatively and qualitatively. Inter-individual variation was evident within the infant faecal microbiota and its development, even within exclusively formula-fed infants receiving the same diet. These data can be of help to future clinical trials (e.g. targeted weaning products) to organize protocols and obtain a more accurate outline of the changes and dynamics of the infant GI microbiota.

INTRODUCTION

From birth onwards, infant health can be easily affected by the ever-changing, developing and complex microbiota that colonizes the human gastrointestinal (GI) tract (Mackie et al., 1999; Fanaro et al., 2003). Microbial colonization of the infant GI tract has been divided into different stages (Tannock et al., 1990; Langhendries, 2006), although recent data indicate *in utero* microbiological exposure is possible (Martín et al., 2004; Satokari et al., 2009). The initial stage of colonization includes birth and the first hours of life (acquisition and initial colonization/development), while the second stage has been defined as the period of exclusive milk-feeding (Langhendries, 2006; Tannock et al., 1990). The development of the microbiota has been shown to be heavily dependent on the type of milk given in the first weeks/months of life [breast milk and/or formula milk, as well as the type of formula (standard or supplemented)] (Harmsen et al., 2000a; Mountzouris et al., 2002). The third stage of microbial colonization of the infant gut begins with the introduction of solid food into the infant’s diet (weaning). In general, it is thought that the infant GI microbiota stabilizes at around 2 years of age (Hentges, 1993). The acquisition of a beneficial GI microbiota is essential for host health. Commensals are known to occupy space and utilize nutrients available in the GI tract and thereby act as competitors to more harmful micro-organisms (Mitsuoka, 1996). The GI microbiota scavenges energy and enhances the bioavailability of certain nutrients and minerals for the host. It also stimulates epithelial development and maturation of the immune system (Cebra, 1999; Bourlioux et al., 2003). Indeed, certain microbial populations, such as *Bifidobacterium breve* and *Lactobacillus rhamnosus*, have been shown to reduce the risk of...
developing asthma or atopic dermatitis (Gorbach, 1990; Kirjavainen et al., 2001). Commensals can also control colonization by competing for adhesion sites and/or substrates, or by altering the environment (e.g. by generating a lower pH or producing inhibitory compounds) (Wold & Adlerberth, 2000). Overall, the main factors contributing to bacterial colonization of the GI tract of the newborn are environmental conditions, GI flow, competitive microbial interactions, presence of mucosal immunity and diet. The complex mixture of carbohydrates found in infant milk that reaches the infant colon allows beneficial bacteria, which are able to convert substrates into energy, to increase in numbers, colonize the GI tract and allow healthy immunological and physiological development (Vandenplas, 2002).

Only a small proportion of the bacterial population inhabiting the GI tract has been isolated and characterized, but it is thought to include at least 1000 different species in adult humans (Bäckhed et al., 2005; Zoetendal et al., 2008). It is more difficult to estimate a relative number of bacterial species in the infant GI microbiota, as this is dependent on both age and diet. Palmer et al. (2007) detected around 2150 taxonomic groups and species from infants, parents, siblings and milk samples. The main phyla corresponded to Bacteroidetes, Proteobacteria, Firmicutes and Actinobacteria. For many years, scientists have tried to examine the acquisition, complexity and succession of the infant GI microbiota. Most studies have employed cultivation techniques, which lack accuracy and sensitivity, and generally underestimate bacterial numbers and diversity. However, a number of studies based on molecular techniques have been undertaken recently (Harmsen et al., 2000a; Favier et al., 2002, 2003; Palmer et al., 2007), and a more detailed understanding of the infant GI microbiota is emerging. A number of pilot studies have used molecular-based techniques to follow the infant faecal microbiota over time and/or in relation to diet. Most either are longitudinal studies (involving small numbers of subjects) or follow high numbers of subjects for a limited period of time. A recent study used microarrays, together with the construction of clone libraries and sequencing, to investigate the succession of the GI microbiota of 14 infants (including a set of twins) during their first year of life (Palmer et al., 2007). The study followed seven exclusively breast-fed and seven mixed-fed infants, with more regular collection of samples during early infancy (first 3 months, i.e. pre-weaning). Six of the infants received a course of antibiotics during the study. Introduction of solid food was investigated; however, information related to the timing of weaning was only available for six infants (four breast-fed and two mixed-fed infants). Overall, each infant had its own distinguishable microbiota and no common developmental pattern was identified for either feeding group. The microbiota of all infants tended to resemble that of an adult by the end of the study (i.e. 12 months of age).

Prior to weaning, the development of the microbiota of infants tends to follow a general trend of colonization at the phylum level. It is generally accepted that the microbiota of breast-fed infants is less complex than that of formula-fed infants and consists of mainly (in order of predominance) bifidobacteria (up to 90% of the total bacterial population) and members of the family Enterobacteriaceae (Orghage & Nord, 1999; Fanaro et al., 2003). Formula-fed infants tend to be colonized by a more ‘adult-like’ microbiota, with higher levels and frequency of Enterobacteriaceae and enterococci compared with breast-fed infants, although a microbiota dominated by similar levels of bifidobacteria, bacteroides, clostridia and streptococci is seen in formula-fed infants prior to weaning (Harmsen et al., 2000a; Palmer et al., 2007). The start of weaning is generally linked to an increase in bacterial diversity. In general, the Bifidobacterium population is gradually superseded by a more complex ecosystem, with the main organisms being Bacteroides, Clostridium and Eubacterium spp. However, it is difficult to identify a consensus microbial pattern, as the type of diet and environment of the infant is likely to affect the pattern of colonization (Alm et al., 2002). Weaning is a crucial step in the development of the infant GI microbiota; however, the majority of studies to date have included low numbers of infants and/or collected insufficient samples from each infant to be able to fully elucidate the effect of weaning on bacterial succession. Magne et al. (2006) examined the impact of introducing formula milk to the diet of 11 exclusively breast-fed infants on the faecal microbiota (referring to said supplementation as ‘weaning’). Commencement of mixed milk-feeding was defined as when mothers decided to begin weaning, between 6 and 18 weeks of age. However, the World Health Organisation recommends not to begin weaning (introduction of solid food) until at least 6 months (24 weeks). No information regarding the introduction of solid food to the infants’ diets was provided by Magne et al. (2006) and post-weaning was defined as the cessation of breast-feeding (which occurred at 7 weeks of age for at least one infant). Therefore, one questions how such a study design affords investigation of the infant faecal microbiota during weaning (i.e. the gradual introduction of solid foods to the diet).

The overall aim of this work was to investigate the bacterial succession (in terms of diversity and dynamics) of the faecal microbiota of breast- and formula-fed infants during exclusive milk-feeding and through weaning and post-weaning, to provide essential information on the microbiological challenges encountered by infants during the developmental stages of gut microbiology. Quantitative analysis was performed to enumerate the predominant bacterial groups (e.g. bifidobacteria, eubacteria, certain clostridia, atopobia, bacteroides, lactic acid bacteria and Escherichia coli) using fluorescence in situ hybridization (FISH). Qualitative analysis was performed using denaturing gradient gel electrophoresis (DGGE), which provided an overview of the overall diversity and dynamics of the infant faecal microbiota.
METHODS

Study design and recruitment of volunteers. Fourteen infants were recruited over the course of this study (between 2005 and 2007); seven were exclusively breast-fed and seven were exclusively formula-fed. Faecal samples were obtained from infants at specific intervals during the first 18 months of life. For inclusion in the study, infants had to meet the following criteria: have been born at full-term (>37 weeks gestation); be of normal birth weight (>2.5 kg); be <5 weeks old and generally healthy; and be exclusively breast-fed or exclusively formula-fed SMA Gold or SMA White (Wyeth Pharmaceuticals), to avoid supplemented formulae and to keep consistency within the formula group. The mothers of the breast-fed infants had not consumed any antibiotics within the 3 months prior to the study and had not taken any prebiotics and/or probiotics. The infants enrolled in the study were not required to consume any test product and no restrictions on their lifestyles and feeding patterns were given. Ethical approval was obtained from the University of Reading Ethics Committee. To perform this longitudinal study, NHS healthcare workers were contacted; leaflets and posters were produced and distributed in local doctors’ surgeries, in postnatal clinics, to numerous healthcare workers and in the maternity wards of the Royal Berkshire Hospital and the Wokingham Hospital. Signed informed consent forms were obtained from the parents of each infant. Dates of vaccinations were not recorded during this study. Table 1 gives details of the infants enrolled in the study. Withdrawal of breast milk and/or addition of formula milk into the diet of exclusively breast-fed infants did not happen until the start of weaning. Infants BF1, BF2, BF4 and BF7 were prescribed a course of antibiotics at some point during the study (Table 1).

Sample collection. Faecal samples were collected from each infant at 1 month of age and then every 2 weeks until their parents introduced solid food into the diet [Fig. 1; pre-weaning (PW) period]. During the early weaning (W) period, samples were collected weekly for 10 successive weeks. Samples were subsequently collected monthly (M) until each infant was 12 months old and then at 15 and 18 months old. Sample collection began during the fourth week of life to allow the initial microbiota to settle and stabilize and for practical reasons (i.e. size of faecal sample and choice of artificial milk). An unforeseen difficulty arose during the collection of the faecal samples. Each infant did not produce a faecal sample at similar times; therefore, each infant had an individual collection timetable (Table 2). Furthermore, the parents decided to introduce solid food at different ages; therefore, the start of weaning was different between infants. Recruitment of formula-fed infants was more problematic than for breast-fed infants, and the first sample was collected during the fifth week of life for three of the seven formula-fed infants. In order to compare data, samples from all infants in each feeding regime were collated in chronological order within the three dietary phases (e.g. PW1, PW2, PW3, ...; W1, W2, W3, ...; M1, M2, ...). This overcame the limitation of inter-individual variation in sampling times (age of infant) and age at which weaning began.

Parents were supplied with an Anaerocult jar (Merck) at each sampling time. A freshly soiled nappy was placed in the jar and an Anaerogen 2.5 l sachet (Oxoid) was used (following the manufacturer’s instructions) to generate an anaerobic atmosphere in the jar. The investigator was then contacted and the sample was collected from the parents’ house, returned to the laboratory and processed immediately (all samples were processed within 2 h of defecation).

Processing of faecal samples. Faecal material was transferred from the soiled nappy to a stomacher bag using a sterile spatula. Pre-reduced 1 × PBS (0.1 M, pH 7.2; Oxoid) was added to provide a 1 in 10 dilution (w/v), and the sample was homogenized by massaging the sample by hand. The faecal homogenate was used to prepare the following samples. Duplicate aliquots (2 × 1 ml) of the homogenate were pipetted into 1.5 ml microcentrifuge tubes and stored at –20 °C until required for DGGE. Aliquots (3 × 375 μl) of the homogenate were processed for FISH analysis (Waldram et al., 2009).

FISH. Faecal bacterial populations were assessed by FISH analysis as described by Waldram et al. (2009), using a collection of 5’ Cy3-labelled oligonucleotide probes (commercially synthesized by MWG Biotech): Bif164, Bac303, Chis150, Erec482, Ato291, EC1531 and...
Lab158 (Table 3). The nucleic acid stain 4',6-diamidino-2-phenylindole (DAPI; 500 ng ml⁻¹) was used to enumerate the total bacterial load of samples. The following formula was used to calculate the number of bacteria in each sample:

\[
C = \frac{n \times 155.6 \times 14873.74 \times (1000/q)}{155.6 \times 14873.74 \times (1000/q)};
\]

\(C\), cells (g wet weight faeces); \(n\), average number of cells counted per field of view; \(q\), volume of hybridization mixture used; 14873.74, the magnification constant (which takes into account the filter area and microscope grid counting area); 155.6, the dilution factor.

**Principal component analysis (PCA).** Multivariate analysis was employed to analyse the FISH data using SIMCA P+ 11.0 (Umetrics). ‘SIMCA-P+ Batch modelling’ was used for the interpretation and comparison of the dataset for both feeding groups.

**Statistical analyses.** These were performed using the software SPSS 15.0.1. A statistical model for repeated measurements could not be used as there were not enough degrees of freedom (too many repeated samples for too few individuals). All data were transformed into logarithm values (log₁₀) and aggregated into mean values for each infant. Normal distributions were determined by visual inspection of normal probability plots (Q-Q plots), and one value that was outside of the 95% confidence interval was removed. Significant differences

---

**Table 2. Organization of sample collection for each infant**

<table>
<thead>
<tr>
<th>Sample</th>
<th>BF1</th>
<th>BF2</th>
<th>BF3*</th>
<th>BF4*</th>
<th>BF5</th>
<th>BF6</th>
<th>BF7</th>
<th>FF1</th>
<th>FF2</th>
<th>FF3</th>
<th>FF4</th>
<th>FF5</th>
<th>FF6</th>
<th>FF7</th>
</tr>
</thead>
<tbody>
<tr>
<td>PW1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>10</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>PW2</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>PW3</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>16</td>
<td>12</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>PW4</td>
<td>12</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>18</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PW5</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PW6</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PW7</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PW8</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PW9</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PW10</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W2</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W3</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W4</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W5</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W6</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W7</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W8</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W9</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W10</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Identical twins.*
were determined using the analysis of variance (ANOVA, General Linear model) and post-analyses were performed by using the independent-samples Student’s t test (for differences observed between feeding regimes). Statistical significance was defined as $P<0.05$.

**DNA extraction from faecal samples.** When needed, samples were thawed on ice and centrifuged at 13 792 $g$ (rotor no. 3753 28RS; Heraeus SEPATECH) at 4 °C for 5 min in a sterile 2 ml micro-centrifuge tube. Supernatants were discarded and the DNA was extracted from the pellets by using the QIamp DNA Stool Mini kit (Qiagen) following the manufacturer’s instructions. Step 3 of the protocol was done at both 70 and 95 °C initially (using duplicate samples), to determine which gave the better profiles. The higher temperature was chosen as Gram-positive bacteria have thicker cell walls and this temperature lysed the cells better; it also gave more complex DGGE profiles (data not shown). The quantity of DNA in each sample was assessed by running 5 $l$ aliquots of the extracts alongside a MassRuler-DNA Ladder Mix (Helena Biosciences Europe) in 0.8% ultraPURE agarose (GibcoBRL) gels containing ethidium bromide (0.5 ng ml$^{-1}$); bands were visualized using UV light.

**PCR for DGGE.** Fragments (233 bp) derived from the variable V3 region of the 16S RNA gene were amplified from samples by using the universal primers p2 and p3 (synthesized by MWG Biotech; Muyzer et al., 1993). PCR mixtures comprised 5 $l$ 10× MgCl$_2$-free buffer (Promega), 5 $l$ dNTPs (12.5 mM each; Promega), 1 $l$ each of primers p2 and p3 (both 20 pmol), 6 $l$ MgCl$_2$ (25 mM; Promega), 30 $l$ sterile H$_2$O, 1 $l$ Taq (1.25 U DNA polymerase; Promega) and 1 $l$ template DNA, to give a final volume of 50 $l$. PCR were carried out as described by Muyzer et al. (1993). Amplification products were examined by using agarose gel electrophoresis (1.5% agarose containing 0.5 ng ethidium bromide ml$^{-1}$) and products were stored at $-20$ °C until required.

**DGGE.** This was done as described by Waldram et al. (2009) but with a minor modification: gels comprised 30-65% gradients. After electrophoresis, gels were silver-stained according to the method described by Sanguinetti et al. (1994) with minor modifications. Gels were scanned using CanonScan (Canon) and imported into the TotalLab TL120 v2006f software (NonLinear Dynamics) for further analysis.

**DGGE ladder.** A DGGE ladder (or reference profile) was required to allow normalization of profiles during gel analysis (Marker 1). A number of bacterial strains were randomly selected from the laboratory collection or purchased from type culture collections. These included *Acidaminococcus fermentans* ALM2 (band h), *Bacteroides fragilis* ALM13 (band f), *Bifidobacterium bifidum* DSM 20456$^T$ (band j), *Clostridium ramosum* ALM43 (band g), *Coriobacterium* spp. ALM55 (band k) and *Lactobacillus casei* ALM74 (band i). A second ladder (Marker 2) was also developed to allow an overview of the Bifidobacterium diversity using the universal primers. In this case, five *Bifidobacterium* type strains, obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen, were employed (*Bifidobacterium angulatum* DSM 20098$^T$, band c; *Bifidobacterium breve* DSM 20213$^T$, band a; *Bifidobacterium longum* subsp. *infantis* DSM 20088$^T$, band b; *B. longum* subsp. *longum* DSM 20219$^T$, band b; *Bifidobacterium pseudolongum* subsp. *pseudolongum* DSM 20099$^T$, band d).

**DGGE analysis.** Distortion of gels was corrected automatically by using TotalLab TL120 v2006f. Calculations to obtain dendrograms were performed based on the Jaccard similarity coefficient. In addition to having Markers 1 and 2 as reference profiles, two DGGE gels were made with two random samples (one from the PW and the other from the M phase) from all infants. We were able to give the same reference number to bands on each infant DGGE profile and comparisons could be made between gels. Dendrograms did not show any major clustering; therefore, comparisons of DGGE profiles were based on similarity matrices. Lanes sharing more than 60% similarity were entered into the calculations.

**RESULTS**

Faecal samples from seven exclusively breast-fed infants and seven exclusively formula-fed infants were collected on...
a regular basis from 1 month to 18 months of age. The number of samples obtained from the breast-fed group during the PW period was higher than that obtained from the formula-fed group, since breast-fed infants started weaning at a later age than formula-fed infants (20.6 and 17 weeks old, respectively). Both quantitative and qualitative data were obtained from all samples. Data from the quantitative (FISH) analysis are presented by feeding group; this allows the changes in each bacterial group to be monitored throughout the study for breast- and formula-fed infants. Data from the qualitative (DGGE) analysis are presented by dietary phase; this allows the overall variation of the microbial fingerprints to be examined throughout the study.

Quantitative analysis

Breast-fed infants. Total bacteria (DAPI) counts remained relatively stable (between log_{10} 10.5 and log_{10} 10.9) throughout the study for both breast-fed and formula-fed infants (Fig. 2). According to FISH, bacteria detected by Bif164 (bifidobacteria) predominated in the breast-fed group and stayed relatively stable (between log_{10} 9.9 and log_{10} 10.6; mean 56.3 and 54.3 % of total counts, ranges 4.8–100 % and 5.3–90.1 %) during the PW and W phases but decreased significantly (P<0.05) at the start of the M phase to approximately log_{10} 9.5 (Fig. 2). Sample PW6 had a large standard error due to the undetectable levels of bifidobacteria from infant BF4 (who had had antibiotics immediately prior to this sample). The mean levels of bacteria detected by Erec482 started to increase (without interruption) just prior to W phase (although Erec482 counts made up <1 % of total counts during PW) and continued to increase throughout the remainder of the study (mean 0.2, 2.5 and 18.4 % for PW, W and M phases, respectively). The increase seen at PW9 could not be linked to change in diet, as samples were taken when the infants were still exclusively breast-fed (however, a PW9 sample was only available for four of the seven breast-fed infants). It could be speculated that the nutritional changes in the breast milk and/or age of infant (thus GI tract development) might have affected the microbiota. Erec482 counts increased significantly (P=0.001) during the W and M phases, with significantly higher levels than were seen during PW phase (P<0.001). Bac303 counts were detected from the start of the study (PW1), but extensive inter-individual variation was observed. Nevertheless, Bac303 levels were significantly (P<0.05) lower during PW and W phases compared with M phase (Fig. 2), with mean proportions of 3.5, 4.4 and 13.4 % for PW, W and M phases, respectively. Bacterial counts detected by probes Ato291 (mean <5 % of microbe), EC1531 (mean <1 %) and Chis150 (mean <1 %) did not change drastically during the study, although PW levels of Ato291 (mean 1.5 %) were significantly lower than those during the M phase (P<0.01) (mean 3.1 %). Bacteria targeted by Lab158 were detected throughout the study (mean <1 %) and initial weaning did not appear to have an impact on their counts (Fig. 2).

Bacteria detected by Erec482 were the population that showed the most significant changes due to dietary phase in breast-fed infants. Bif164 levels generally started to decrease around 8.4 months old (mean age of infants when M1 was collected), while Bac303 levels began to increase around the same time (Fig. 2). Interestingly, the four infants that received antibiotics during this study were all breast-fed. Antibiotic administration only dramatically affected the bifidobacteria of BF4, with their level reduced to below the detection limit in the subsequent sample (PW6).

Formula-fed infants. Unlike the breast-fed group, there were no drastic shifts in the bacterial populations of the formula-fed group during the study. The bifidobacteria (Bif164) levels fluctuated between log_{10} 8.6 and log_{10} 9.3 during the PW phase (mean 25.4 %) and increased slightly at the start of the W phase (Fig. 2). Levels continued to fluctuate between log_{10} 9.1 and log_{10} 10.1 until the end of the study, although the variability between infants clearly decreased. A similar pattern occurred for the Erec482 counts, with fluctuations between log_{10} 9 and log_{10} 9.7 during the PW phase (mean 15.0 %) and relative stability over the rest of the study (between log_{10} 9.6 and log_{10} 10; mean 24.0 and 29.8 % for W and M phases, respectively). There was a slight (non-significant) increase in the Chis150 levels at the start of the M phase (Fig. 2). Ato291 levels increased from log_{10} 7.2 at PW1 to log_{10} 8.6 at M7, although inter-individual variation meant that this was not statistically significant. EC1531 levels decreased from log_{10} 8.7 at PW1 to log_{10} 7.6 at M7 (P<0.01). A significant decrease was also observed in EC1531 levels between the W and M phases (P<0.05). EC1531 and Ato291 counts appeared to follow opposite trends; when EC1531 decreased, Ato291 increased over time. A marked increase in Bac303 counts was detected throughout the study (log_{10} 7.2 at PW1 and log_{10} 9.7 at M7). Similar to the breast-fed group, the inter-individual variation in Bac303 levels was relatively high. Nevertheless, levels were significantly different (P<0.05) between PW and M phases (with mean proportions of 10, 17.8 and 23.7 % for PW, W and M, respectively). Lab158 counts were detected more often during the PW and W phase than the M phase, although detectable Lab158 levels [>10^{6} cells (g wet weight faeces)^{-1}] were obtained for at least one infant at each sample time. Overall, no major changes were observed in bacterial levels at the start of the W phase in formula-fed infants.

Breast-fed versus formula-fed infants. Comparison of the bacterial levels harboured by breast-fed and formula-fed infants during each dietary phase (PW, W and M) demonstrated that formula-fed infants had significantly higher Erec482 (P<0.001) and Chis150 (P<0.01) counts than their breast-fed counterparts throughout the study (Fig. 2). Bif164 levels were significantly lower in the
Fig. 2. Investigation of the bacterial succession of the faecal microbiota of breast- and formula-fed infants detected by FISH [mean log_{10} cells (g wet weight faeces)^{-1} ± 2 se]. PW, Pre-weaning samples (exclusive milk-feeding); W, initial weaning samples (obtained for 10 successive weeks after introduction of solid food); M, monthly samples. Top: ■ DAP; ○, Bif164; △, Ato291. Middle: ×, Bac303; ■, Lab158; ○, Erec482. Bottom: □, Chis150; ●, EC1531. The detection limit of FISH is ~10^6 cells (g wet weight faeces)^{-1}. 

The infant faecal microbiota: diversity and stability
formula-fed group during PW and W phases of the study, compared with those of the breast-fed group (P<0.05). Statistical analysis within each feeding group (i.e. between the different phases) demonstrated greater changes in the breast-fed group than in the formula-fed cohort across the study. In particular, the Erec482 population of the breast-fed group increased significantly between the PW and W phases, as well as between the W and M phases. As the diet was becoming more complex (together with the cessation of breast-feeding), Erec482 counts increased progressively. Erec482 levels did not fluctuate significantly in the formula-fed group, but were approximately the same level during the entire study. In the breast-fed group, Bif164 counts decreased significantly between the start and the end of the study. It was only toward late weaning (the end of the W phase) that their levels were significantly lower. Although Bif164 counts were lower in the formula-fed group, they did not change significantly during the study. The EC1531 population experienced a significant decrease between the PW phase and the M phase in the formula-fed group. This indicates that EC1531 levels were significantly higher while the infants were exclusively formula-fed compared with when the infants were having a more complex diet. It is not clear from the data whether the faecal microbiota of infants had stabilized by 18 months of age.

PCA of the entire dataset (from both feeding groups) detected similarities between the two feeding groups, while demonstrating the overall succession of the faecal microbiota over time and in relation to diet (Fig. 3). The PCA model was explained by three principal components (PCs). PC1 explained 32% of the variation within the data, whereas PC3 explained 45% of the variation. No clear clusters appeared at first glance. However, the positions occupied by the PW and W data of the breast-fed group were relatively distinct from other data points (Fig. 3). It was also apparent that the faecal microbiotas during the PW and W phases of the breast-fed infants and formula-fed infants were different. The M phase samples from breast-fed infants appeared to cluster together with most of the W phase and all of the M phase samples of formula-fed infants. All but two samples from the PW phase of the breast-fed group were positioned on the left-hand side of the plot (namely sample PW5 of BF5 and sample PW9 of BF1). Both correlated with higher Bac303, Erec482 and Ato291 counts. Overall, all data from the PW phase of the breast-fed group correlated with higher Bif164 and DAPI counts together with lower Erec482, Chis150 and Ato291 levels. The PW phase data from the formula-fed infants were also spread out, but were not confined within a specific quadrant of the plot. The data that clustered most tightly together were those from the M phase of the formula-fed group. However, they did not cluster separately from other samples. Overall, the bacterial population of the breast-fed group evolved in the score plot during W and M phases, migrating towards the upper right-hand side of the plot (essentially demonstrating modulation of the bacterial profile to be more like that of the formula-fed group). The microbiota of the formula-fed group also changed during the course of the study, shifting towards the lower right-hand side of the plot, linked to higher levels of Chis150, Ato291, Bac303 and Erec482 together with lower levels of Bif164, DAPI, Lab158 and EC1531.

Fig. 3. PCA showing comparison of breast- and formula-fed faecal microbiotas from different dietary phases, as determined by FISH analysis. (a) Score plot of the PCA model for comparison of all samples, with PC1 (x axis) versus PC3 (y axis). Squares, PW samples; circles, W samples; triangles, M samples. Red, breast-fed; blue, formula-fed. (b) Loading plot of the PCA model for the comparison of all samples, with PC1 (x axis) versus PC3 (y axis).
Qualitative analysis

In addition to FISH analysis, each faecal sample was subjected to DGGE analysis (Fig. 4). Faecal bacterial DNA was extracted and the variable V3 region of the 16S rRNA gene was amplified. Each DGGE profile was subjected to clustering analysis and a dendrogram was obtained for each gel (containing all samples obtained for each infant). The high complexity and inter-individual differences in the DGGE profiles were such that no obvious grouping of profiles was seen (although clustering according to dietary phase was clear for all infants). To compare the different profiles more efficiently, a similarity matrix was obtained for each dietary phase of each feeding group. A cut-off value of 60% similarity between profiles was accepted for demonstration that the two DGGE profiles were comparable and/or relatively similar.

Three similarity matrices were generated to show the similarity among profiles during each dietary phase: (i) of breast-fed infants, (ii) of formula-fed infants and (iii) across feeding regimes. Overall, the DGGE results were consistent with the PCA findings from FISH analysis. PW phase DGGE profiles were shown to be unique to each infant, with few profiles sharing >60% similarity (either within or across feeding regimes), and the twins’ profiles displayed the greatest similarity (i.e. more instances of >60% similarity between profiles). Intra-individual variation was common among infant PW phase DGGE profiles. A shift in banding pattern was seen for infant BF4 due to the administration of antibiotics (in accord with the quantitative data).

Greater similarity was seen between W profiles from infants within each dietary group than for PW profile comparisons, with the DGGE profiles from the twins again displaying the highest degree of resemblance. Greater resemblance was generally found between W profiles of formula-fed infants than was seen among W profiles of breast-fed infants. Intra-individual variation was less for breast-fed infants’ W profiles compared with that seen for their PW profiles. Comparison of W profiles between breast-fed and formula-fed infants demonstrated that profile similarity was higher than that seen during PW phase. Overall, the similarity analysis highlighted that a shift in banding profiles occurred within each infant as a result of weaning and that the DGGE profiles began to converge across the two feeding groups. In other words, weaning appeared to elicit changes in the infant faecal microbiota that effectively reduced inter-individual variation.

The similarity matrix of M profiles for each breast-fed infant suggested that changes were still occurring in the faecal microbiota of some infants even between 15 and 18 months of age. With the exception of FF5, chronological comparisons of the formula-fed infants’ M profiles suggested relative stability of the faecal microbiota. Overall, less inter-individual variation was seen within each feeding group between M profiles (compared with PW and W phases) and greater similarity was observed between M profiles across the feeding groups than during PW and W phases.

Overall, universal DGGE profiles from the PW phase displayed the greatest intra-individual variation within each feeding group (compared with W and M phases). Although the infants were on an exclusively milk-fed diet, both inter- and intra-individual variation was evident.

Fig. 4. Investigation into the bacterial succession of the faecal microbiota of a breast-fed infant (BF3), using universal PCR-DGGE profiling. DNA was obtained from faecal samples and PCR-DGGE was performed using the universal primer set p2/p3 (Muyzer et al., 1993). L1, Marker 1 containing five dominant bands (a–e); and L2, Marker 2 containing seven dominant bands (f–k). The markers were used to normalize banding profiles across and between gels. Asterisks correspond to bands that migrated to the same position as a band in one of the markers.
Interestingly, the similarity between the DGGE profiles of the twins was greater during the W phase than either PW or M phases. However, the greatest similarity between the universal DGGE profiles of all infants was generally seen during M phase. Although the DGGE profiles changed with time and tended to become more similar to one another as the study progressed, unique profiles were observed for each infant.

**DISCUSSION**

The results of the current study corroborate the global observation of previous studies investigating the infant gut microbiota; namely inter-individual variability initially, with convergence towards a more adult-like faecal microbiota after 1 year of age. No obvious clustering was observed in relation to gender, mode of delivery, antibiotic usage, presence of animals and/or siblings in the household, or the season the infant was born into, although this may reflect the relatively small subject cohort (14 infants). The data obtained from the twins showed relatively high similarity. Both individual PCA models showed a similar pattern (with the exception of sample PW6 of BF4, which was after antibiotic administration). Moreover, when the DGGE data were subjected to cluster analysis, separation according to dietary phase and not to infants was evident for the twins. This correlates with microarray analysis by Palmer et al. (2007), which showed that the environment and genetics play a role in determining microbial acquisition.

The major discrepancy between our results and those of Palmer et al. (2007) was in the detection of bifidobacteria. FISH analysis identified bifidobacteria as one of the predominant bacterial groups within the infant faecal microbiota, yet Palmer et al. (2007) did not detect bifidobacteria within the predominant microbial populations (even when they used *Bifidobacterium*-specific qPCR to confirm their result). This is most likely due to differences in the collection, handling and processing of faecal samples between the two studies. In the former study, samples were collected by parents and stored in stool collection vials (with no addition of glycerol) in the home freezer until transfer to the laboratory (2 weeks, 3 months and 6 months from the start of the study). This, together with the DNA extraction procedure and PCR protocol employed, may have implications on the relative efficacy of monitoring different members of the mixed faecal microbiota. Indeed, during the course of our study, we demonstrated that the incubation temperature used during DNA extraction (QiAmp DNA Stool Mini kit; the same as used by Palmer et al., 2007) affected the diversity observed based on DGGE banding profiles. It is unclear whether Palmer and colleagues used 70 or 95 °C for DNA extraction during their study (as both temperatures are mentioned in the manufacturer’s instructions). Another observation from our work (corroborated by Dr Anna Snelling, University of Bradford, UK, personal communication) was the gradual degradation of DNA obtained by this method during repeated freeze–thaw cycles. This may explain the relatively low estimates of abundance obtained by Palmer and colleagues during the *Bifidobacterium*-specific qPCR. Different infant cohorts were used in the two studies, which may also explain the discrepancies. However, we consider that the potential bias from each stage of sample collection, handling and processing (including the molecular protocols employed) are the more likely contributors, and that polyphasic strategies should be employed for studies investigating mixed microbial ecosystems. Palmer et al. (2007) may have used more than one molecular method, but all were performed on the same DNA samples.

Data obtained by molecular techniques also need to be carefully interpreted. FISH allows enumeration of important bacterial groups; however, this technique has a detection limit of ~10⁶ bacterial cells (g wet weight faeces)⁻¹. Therefore, when BF4 had undetectable levels of bifidobacteria at week 14, we do not know whether this was due to complete loss of this bacterial group or if levels were simply below the detection limit of the method. Furthermore, FISH does not facilitate investigation of changes in terms of species variation. In general, FISH allows detection of important trends that warrant further investigation. DGGE analysis showed that the diversity of the infant faecal microbiota increased over time and/or with change in diet. Universal DGGE did not show any major clustering or specific banding patterns for the different feeding groups or dietary stages.

Data obtained in this study were from formula-fed infants that received standard formula. However, an increasing number of formulae supplemented with prebiotics are available. Therefore, it would be interesting to see if the transition between an exclusively milk-fed diet (with added prebiotics) and the start of weaning showed significant differences between dietary phases (as observed in the breast-fed group) or corresponded more to the formula-fed microbial development seen herein. Furthermore, what is the impact of mixed-milk feeding (i.e. breast and formula milk) on the development and complexity of the infant faecal microbiota? And does this depend on the ratio of formula to breast milk?

The bacterial population targeted by probe Erec482 appears to be an important comparison criterion between feeding group and dietary phase due to the obvious changes observed throughout the current study. Probe Erec482 detects members of a number of different genera, including *Blautia, Clostridium, Eubacterium, Roseburia* and *Ruminococcus* (Supplementary Fig. S1, available with the online version of this paper). Members of these genera have been detected in the infant microbiota. However, *Roseburia* spp. have only been identified in two studies to date. Duncan et al. (2002) detected five strains of *Roseburia intestinalis* in infant faeces, while Barcenilla et al. (2000) detected *Roseburia ceccocola* in the faeces of a pre-term


