ORIGINAL ARTICLE

16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis

Yunwei Wang[1,6], Jeanette D Hoenig[2,6], Kathryn J Malin[2], Sanaa Qamar[2], Elaine O Petrof[3], Jun Sun[4], Dionysios A Antonopoulos[5], Eugene B Chang[1] and Erika C Claud[1,2]

[1]*Department of Medicine, University of Chicago, Chicago, IL, USA;* [2]*Department of Pediatrics, University of Chicago, Chicago, IL, USA;* [3]*Department of Medicine, Queen's University, Kingston, Ontario, Canada;* [4]*Department of Medicine, University of Rochester, Rochester, NY, USA and* [5]*Biosciences Division, Institute for Genomics and Systems Biology, Argonne National Laboratory, Argonne, IL, USA*

Neonatal necrotizing enterocolitis (NEC) is an inflammatory intestinal disorder affecting preterm infants. Intestinal bacteria have an important function; however no causative pathogen has been identified. The purpose of this study was to determine if there are differences in microbial patterns that may be critical to the development of this disease. Fecal samples from 20 preterm infants, 10 with NEC and 10 matched controls (including 4 twin pairs) were obtained from patients in a single site level III neonatal intensive care unit. Bacterial DNA from individual fecal samples was PCR-amplified and subjected to terminal restriction fragment length polymorphism analysis and library sequencing of the 16S rRNA gene to characterize diversity and structure of the enteric microbiota. The distribution of samples from NEC patients distinctly clustered separately from controls. Intestinal bacterial colonization in all preterm infants was notable for low diversity. Patients with NEC had even less diversity, an increase in abundance of *Gammaproteobacteria*, a decrease in other bacteria species, and had received a higher mean number of previous days of antibiotics. Our results suggest that NEC is associated with severe lack of microbiota diversity that may accentuate the impact of single dominant microorganisms favored by empiric and widespread use of antibiotics.

Introduction

Necrotizing enterocolitis (NEC) is an acquired gastrointestinal disease that primarily affects premature babies. It is the most common gastrointestinal emergency in the newborn and may lead to death in severely affected infants. The incidence is 1–3 per 1000 live births, occurring in 10% of infants born under 1500 g and representing 2–5% of neonatal intensive care unit (NICU) admissions (Kosloske, 1994). Mortality has been reported to range from 9% to 28% (Stoll, 1994). Although the exact cause is unknown, the critical elements are thought to be prematurity, enteral feeding, an

inappropriate pro-inflammatory response and bacterial colonization (Claud and Walker, 2001). Because the risk factors for NEC are issues common to all premature infants, it is impossible to predict which infants will develop this devastating disease.

As opposed to the adult intestinal microbiota that comprise more than 10^{13} microorganisms, the newborn gut is sterile at birth. Although the full-term neonate becomes rapidly colonized with a diverse flora, preterm infants have a delayed colonization, fewer bacterial species present and are more susceptible to colonization with pathogenic bacteria (Goldmann and JLeclair Macone, 1978; Gewolb et al., 1999; Schwiertz et al., 2003).

Bacteria are believed to be important in the pathogenesis of NEC; however no specific pathogen has been identified. Previous studies, however, have been limited by the inability of conventional microbiological cultivation techniques to thoroughly characterize the human gastrointestinal microbiota. It has been reported that 80% of the

Correspondence: EC Claud, Department of Pediatrics, University of Chicago, 5841 S Maryland Ave, MC6060, Chicago, IL 60637, USA.
E-mail: eclaud@peds.bsd.uchicago.edu
[6]Both are Co-first authors.

human colonic microbiota are not detected by conventional culture methods (Eckburg et al., 2005). Taxonomical classification and noncultivation molecular profiling of microbiomes are now possible by sequencing the highly conserved 16S small subunit bacterial ribosomal RNA gene (rRNA), allowing identification of previously undetectable microbes (Relman, 1999). This approach therefore provides a more complete picture of the composition of human intestinal microbiota. For this study, we hypothesized that there are differences in molecular microbial profiles of preterm infants without and with NEC, which may be critical in the etiopathogenesis of disease.

Materials and methods

Patient characteristics and sample collection

Subjects were recruited from a single level III NICU at the University of Chicago. All patients underwent routine NICU care as determined by the managing service. A total of 21 fecal samples from 20 patients were obtained including 10 infants with NEC and 10 infants without NEC as control. The study set included four sets of twins. In these twin pairs, one twin developed NEC (N1-N4) and the other was spared (C1-C4) thus serving as an ideal control. Controls (C5-C10) were matched to NEC (N5-N10) by gestational age and day of life at NEC diagnosis. Control sample 9 was collected from an infant who later went on to develop NEC; however no follow-up sample after NEC diagnosis was available for this infant due to parent refusal. Only patients with definite or advanced NEC, corresponding to Bell stages II and III were included in the NEC cohort (Bell et al., 1978). For patient N10 with NEC, an additional sample was analyzed (CN10), which had been prospectively obtained 3 days before disease diagnosis.

Included subjects were born between 25 and 32 weeks gestation and the patient age when the fecal sample was obtained ranged from 4 to 49 days. Thirteen male and seven female infants were included. Seventeen patients were delivered by cesarean section and three vaginally. Ten of the infants were exclusively formula-fed and ten received breast milk. Study patient characteristics were not significantly different between NEC and control groups (Table 1).

The study was approved by institutional review board for human studies of the University of Chicago and informed consent was obtained from patient parents. Fecal samples for NEC patients were collected as soon as possible after the diagnosis was made, with an average collection time of <1 day after the diagnosis of NEC. Nurses collected the fecal sample directly from the diaper into the collection tube using the wooden end of a sterile cotton swab. The sample was immediately frozen. All samples were stored at $-80\,^{\circ}$C until processed.

Table 1 Clinical characteristics of the hospitalized preterm infants in this study

Characteristics	NEC infants (n = 10)	Control infants (n = 10)
Male/Female	6/4	7/3
Gestational age (weeks)	25–32	26–32
Sample collection age (days)	5–49	4–49
Delivery: C-section/vaginal	8/2	9/1
Feeding: breast milk/formula	4/6	6/4

Abbreviation: NEC, necrotizing enterocolitis.

DNA extraction

For DNA extraction, 50 mg frozen fecal sample was dissolved in 1 ml extraction buffer (50 mM Tris (pH 7.4), 100 mM EDTA (pH 8.0), 400 mM NaCl, 0.5% SDS) containing 20 µl proteinase K (20 mg ml^{-1}). A slurry (500 µl) of 0.1-mm diameter zirconia/silica beads (BioSpec Products, Bartlesville, OK, USA) was added into the extraction tubes and a Mini-Beadbeater-8K Cell Disrupter (BioSpec Products) was used to lyse the microbial cells. After overnight incubation at 55 °C, standard DNA extraction with phenol/chloroform/isoamyl alcohol, and precipitation with ethanol were performed. Isolated DNA was dissolved in TE buffer and stored at $-80\,^{\circ}$C.

PCR and terminal restriction fragment length polymorphism analysis

16S rRNA gene sequences were amplified from DNA samples using broad-range primers 8F (5′-AGAG TTTGATCCTGGCTCAG-3′) labeled with 6′-carboxyfluorescein (6-FAM) and 1492R (5′-GGTTACCTTGT TACGACTT-3′) for the conserved 16S bacterial domain. PCR reactions were performed for 30 cycles using Takara high-fidelity Ex Taq (Takara Mirus Bio, Madison, WI, USA) with an annealing temperature of 58 °C. Replicate PCR products were pooled and verified by electrophoresis and purified by precipitation. FAM-labeled PCR products were then digested by restriction enzyme MspI (New England Biolabs Inc., Ipswich, MA, USA), mixed with GeneScan-500 Size Standard (Applied Biosystems, Carlsbad, CA, USA) and sequenced by capillary electrophoresis.

Clone library and sequencing

Unlabeled PCR products were purified using the QIAquick gel extraction kit (Qiagen, Valencia, CA, USA) and cloned into pCR-2.1-TOPO vectors using the TOPO-TA cloning kit (Invitrogen, Carlsbad, CA, USA) for sequencing as per manufacturer's instructions. From each library, 288 colonies were picked randomly and processed for sequencing. The 5′ end

of the plasmid inserts was sequenced unidirectionally using 8F forward primers on an ABI 3730 automated fluorescent DNA sequencer (Applied Biosystems). The raw sequence data were processed using the RDP pipeline server at the Ribosomal Database Project II (RDP-II) website (http://rdp.cme.msu.edu/pipeline) by base-calling, quality-trimming and alignment. Potential chimeric sequences were checked using the SimRank 2.7 package available through the RDP and excluded (Huber et al., 2004). The classifier analysis tool of RDP-II and NCBI BLAST tool were used to assign 16S rRNA sequences to the taxonomical hierarchy at different levels. The program DOTUR with the furthest neighbor algorithm was used to group sequences into operational taxonomical units (OTUs) or phylotypes that represented the number of 16S rRNA sequence similarity groupings (Schloss and Handelsman, 2005). A 98% cutoff value was used so that sequences with more than 98% similarity were considered the same. Library coverage was calculated using Good's coverage formula as $(1-(n/N)) \times 100$ (n is the number of single clone OTUs and N is the total number of sequences for the analyzed sample) to demonstrate how completely the sequences obtained represented all sequences present in the sample (Good, 1953). The Shannon diversity index, which accounts for both abundance and evenness of the species present, was used to compare diversity of the OTUs in the libraries (Kuehl et al., 2005). The representative sequences obtained were searched using NCBI BLAST to find the nearest matched species. These sequences were deposited in the GenBank nucleotide sequence databases under accession numbers FJ557111–FJ557192 for sequences from control infants, and FJ557193–FJ557235 for sequences from NEC infants. A phylogenetic tree was constructed by the neighbor-joining distance matrix method in the ClustalX program (Thompson et al., 1997) with 1000 bootstrap replicates and displayed using the Molecular Evolutionary Genetics Analysis package (MEGA3) (Kumar et al., 2008).

Statistical analysis
For terminal restriction fragment length polymorphism (T-RFLP) analysis, fragment length and abundance were determined using GeneMapper software (Applied Biosystems). The called peaks were filtered and binned by the algorithm proposed by Abdo et al. (2006) using their R script available at http://www.ibest.uidaho.edu/tools/trflp_stats/download.php. Based on the normalized T-RFLP profile, we treated the number and height of peaks as number and abundance of bacterial phylotypes represented in samples. Diversity indices for richness (distinctive peak numbers) and Shannon diversity were calculated using formulas as described by Kuehl et al. (2005). For principal coordinate analysis (PCA), all 16S rRNA

gene sequences were imported into the ARB software package and aligned into a phylogenetic tree with sequences most similar to each other in closest proximity and connected by branch points indicating which sequences were phylogenetically related. This tree was annotated to indicate the sample origin for each sequence and used to measure the difference between bacterial communities in samples using the UniFrac metric (Lozupone and Knight 2005; Lozupone et al., 2006). The P-Test was used in the UniFrac metric to determine whether each sample was significantly different from others. A distance matrix between environments was produced by calculating values for all possible pairs of environments in the tree using PCA. Distances were converted into points in space with a number of dimensions one less than the number of samples. The principal coordinates, in descending order, describe how much of the variation is explained by each of the axes in this new space. The first principal coordinate is the axis that separates out the data to the greatest degree; the second principal coordinate provides the next most separation and so forth (Lozupone and Knight, 2005; Lozupone et al., 2006). The UniFrac interface returned information on all principal coordinate axes in the data table, and then allowed visualization of the data in a scatter-plot that compared all pairs of the first three principal coordinates. For comparison analysis, all data were expressed as mean ± s.d. Student's paired t-test was used to test differences between two samples or groups. $P < 0.05$ was considered statistically significant.

Results

Bacterial community diversity by T-RFLP analysis
Fecal samples were first examined by T-RFLP analysis in which the 16S rRNA gene is amplified from bacterial DNA and then cut with known restriction digest enzymes yielding DNA fragments of different lengths which are characteristic for different bacteria. These fragments can then be aligned by length, resulting in a graph with peaks illustrating the richness (number of peaks) and abundance (height of peaks) of the organisms in the sampled microbial community. As judged by T-RFLP analysis, the microbial profile of NEC infants was different from that of controls. Even for twin pairs, the T-RFLP profile for the NEC infant was different from its matched control healthy infant (representative profiles shown in Figures 1a and b). The absolute richness (the number of distinct terminal restriction fragments in each sample) of the community of NEC infants (12.8 ± 7.3) was observed to be significantly lower than that of controls (25.2 ± 9.8, $P < 0.05$). The Shannon diversity index, which accounts for both abundance and evenness of the species present, as calculated for NEC infants based on presence and distribution of
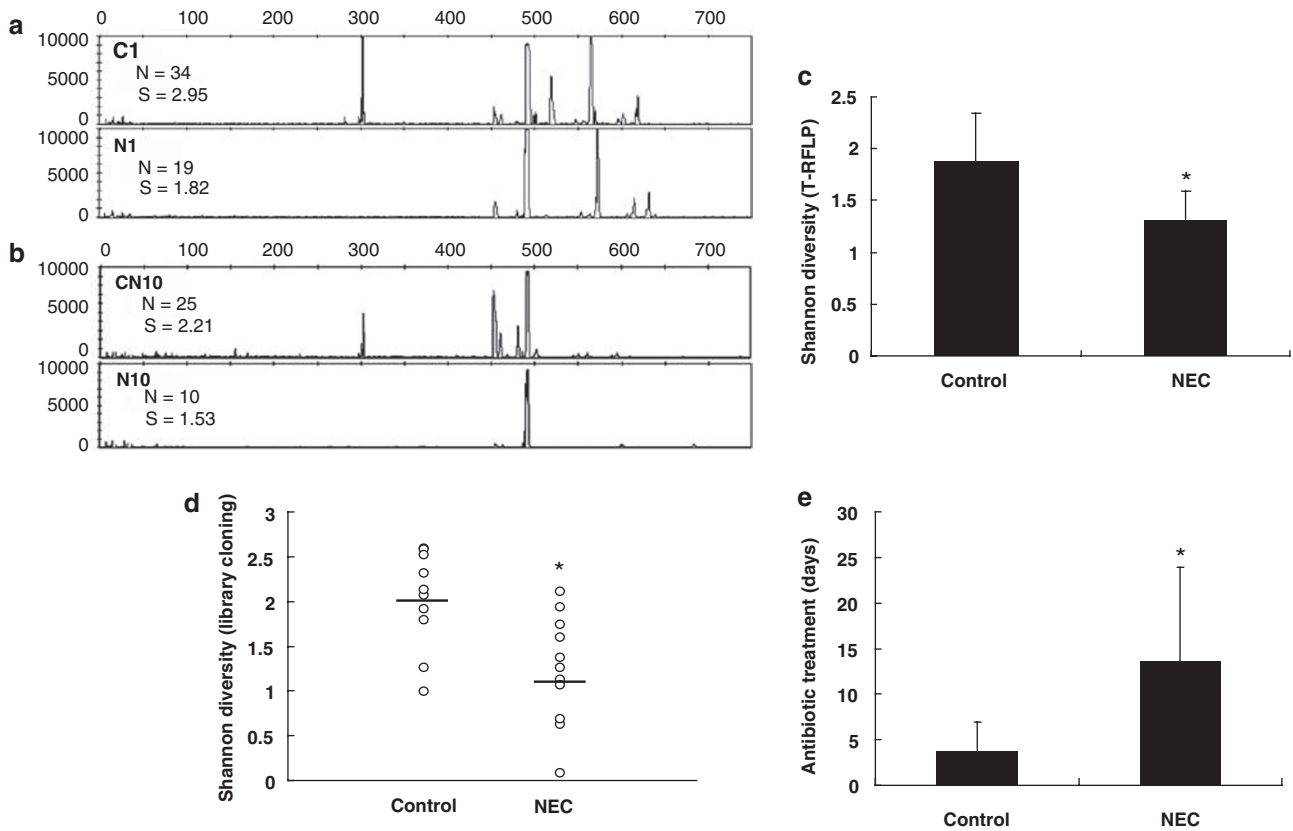
Figure 1 Diversity comparison between preterm infants with and without NEC by terminal restriction fragment length polymorphism (T-RFLP) and clone library analysis. (**a**, **b**) Representative T-RFLP profiles of intestinal bacteria in infants with and without NEC. N1 (with NEC) and C1 (without NEC) were samples collected from a genetically identical twin pair. CN10 and N10 were samples collected from the same infant before and after the development of NEC. For each profile, the number of peaks (*N*) and Shannon diversity index (S) are shown. (**c**) Overall Shannon diversity index between NEC and control groups was calculated from T-RFLP data based on peak number and distribution in T-RFLP profiles from each sample and shows decreased diversity in the NEC patients. (**d**) Overall diversity by library cloning and sequencing analysis in preterm infants with and without NEC again demonstrates decreased diversity in NEC patients. (**e**) Comparison of antibiotic exposure time between control and NEC groups. In all panels, asterisk indicates statistical significance at $P < 0.05$. NEC, necrotizing enterocolitis.

peaks was 1.13, which was also significantly lower than that of control infants without NEC (1.88, $P = 0.035$; Figure 1c).

Diversity comparison based on clone libraries
To gain insight into the specific bacterial phylotypes that explain the differences in microbial diversity between infants with and without NEC, 16S rRNA-encoding gene libraries were constructed for each sample. A total of 6048 clones were randomly picked from the 21 sample libraries and sequenced. After quality control and sequence assembly, 5354 clones yielded 700 bp of partial 16S rRNA gene sequences that were used for analysis. An average of 255 clone sequences per sample was analyzed. 16S rRNA gene sequences were assigned into OTUs or phylotypes at a similar cutoff value of 98%, so that sequences that were more than 98% similar were considered the same. Library coverage was calculated by Good's formula, and ranged from 95% to 100%, indicating that the 16S rRNA gene sequences from each sample encompassed the majority of the

microbiota in each infant. Limited diversity was found in all preterm patients, with an average of only 16 OTUs in contrast to the more than 200 phylotypes reported for human adults by similar methods (Eckburg *et al.*, 2005). NEC infants had significantly lower numbers of OTUs (10.4 ± 6.1) compared to controls (19 ± 6.7, $P = 0.008$). The Shannon diversity index was again calculated based on library cloning sequences, and was used to compare the libraries. As shown in Figure 1d, in addition to decreased numbers of OTUs, library cloning data also demonstrate that NEC infants had a significantly lower diversity of intestinal microbiota than controls (1.19 ± 0.62 vs 1.99 ± 0.55, $P = 0.005$). The low diversity of the microbial population in NEC infants was reproduced for the entire group by assigning clones to OTUs. Analysis of the 2518 clones for the control group identified 82 OTUs. However, only 43 OTUs were identified from 2575 clones in NEC infants (Supplementary Tables 1 and 2).

To investigate a possible mechanism for decreased diversity in NEC patients, clinical characteristics

were investigated between the two groups. No difference in gestational age, age at time of sample collection or type of feed was found. Patients in both groups were on comparable amounts of feeds at the time the sample was obtained, with NEC patients at an average of 79% of full feeds at the time of sample collection and control patients at an average of 84% of full feeds ($P = 0.74$). Full feeds were defined as all enteral nutrition at 130–150 cc per kg per day with no administration of hyperalimentation or intravenous fluid. Notably however, increased previous antibiotic exposure time was noted in the NEC group. Infants with NEC had a history of significantly more days of antibiotics (13.7 ± 10.2 days compared to control infants (3.7 ± 3.0 days), $P = 0.005$; Figure 1e). Antibiotic days counted were

for days before the diagnosis of NEC, thus antibiotics for NEC treatment were not included in the count.

Taxonomic structure analysis
To further analyze the composition of the microbiota, frequencies of sequences from each sample were grouped according to public database NCBI and RDP-II. Consistent with the demonstrated difference in diversity between control and NEC infants, for control infants four phyla were present: *Proteobacteria*, (34.97% relative abundance), *Firmicutes* (57.79%), *Bacteroidetes* (2.45%) and *Fusobacteria* (0.54%) with 4.25% unclassified bacteria. However, NEC patients had only two phyla, *Proteobacteria* (90.72%) and *Firmicutes* (9.12%) with
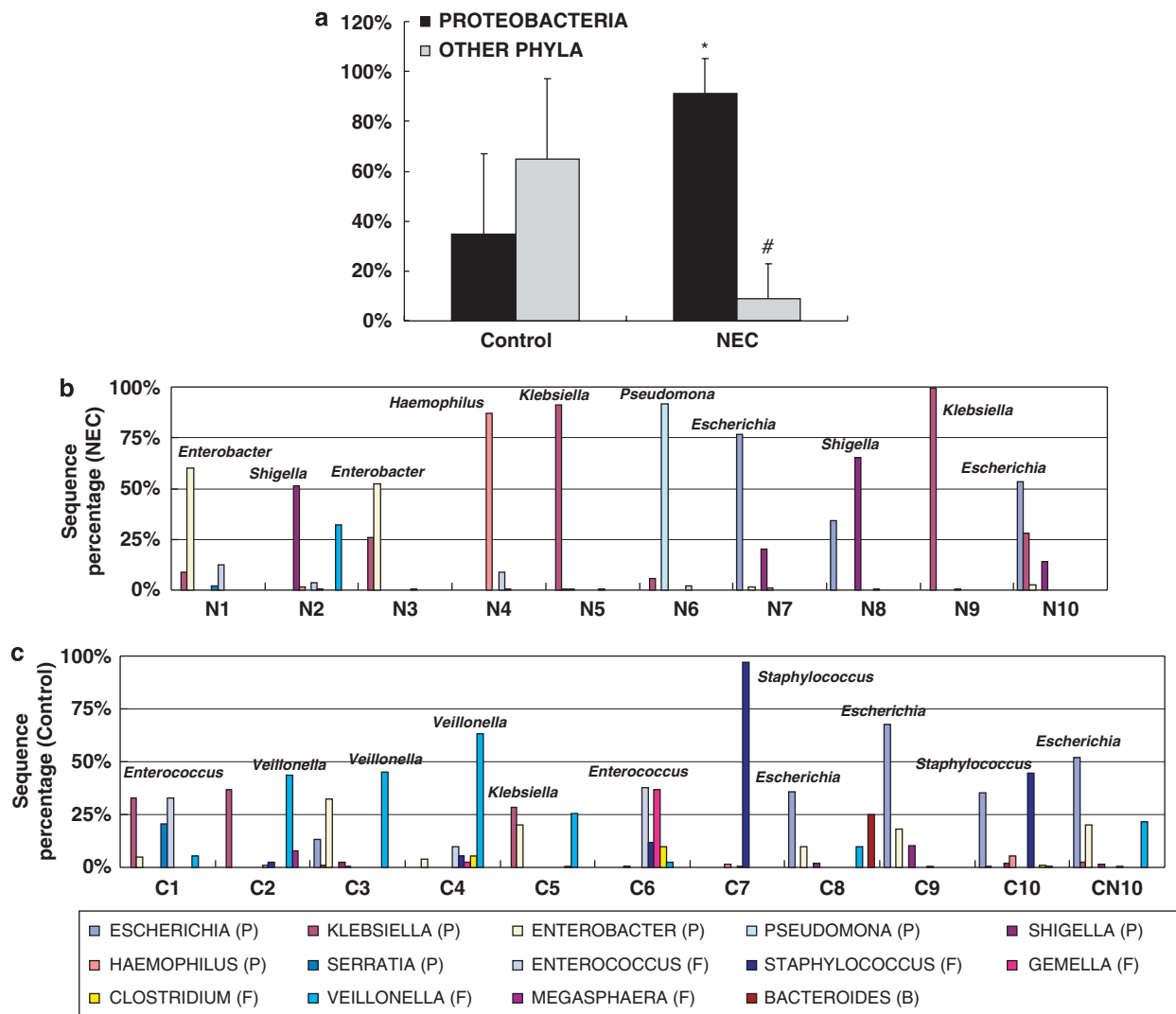
Figure 2 Overview of bacterial composition of all samples by clone library analysis. (**a**) Relative abundance of *Proteobacteria* in preterm infants. *Comparison of *Proteobacteria* between two groups, $P < 0.01$; #Comparison of other phyla (*Firmicutes* for NEC; *Firmicutes*, *Bacteroidetes* and *Fusobacteria* for control) between the two groups, $P < 0.01$. (**b**, **c**) The genus-level composition of gut microbiota. Genera whose abundance was more than 10% in any sample and its corresponding phyla (P, *Proteobacteria*; F, *Firmicutes*; B, *Bacteroidetes*) are shown as bars (NEC samples top panel, Control samples bottom panel). The dominant genus in each sample is labeled. Only samples from NEC infants had dominance of a single genus of the *Proteobacteria* phylum (more than 50%). Samples C9 and CN10 are notable for *Escherichia* dominance, and developed NEC shortly after samples were obtained. NEC, necrotizing enterocolitis.

0.16% unclassified bacteria. The average proportion of *Proteobacteria* was significantly increased and the average proportion of *Firmicutes* was significantly decreased compared to controls ($P = 0.001$; Figure 2a). Furthermore, at level 2, *Gammaproteobacteria* was the major detected class of the phylum *Proteobacteria* in both control and NEC infants. Further exploration showed that the high abundance of *Proteobacteria* in NEC infants included dominance (more than 50%) of one genera of *Proteobacteria* in each individual patient. In control patients, there was more even bacterial distribution with no *Proteobacteria* genus more than 40% of total bacteria. The exceptions were samples C9 and CN10 that were both collected from patients that later developed NEC. These samples had a predominance of *Escherichia* (Figures 2b and c).

To identify potential bacterial pathogens related to the pathogenesis of NEC, the representative OTU sequences from NEC and control groups were aligned with the GenBank database to find the closest match by BLAST search (Supplementary Tables 1 and 2). The frequency of each representative OTU was also calculated. In samples collected from the control group, the most frequently detected representative OTUs were ECC226e09 (six infants), ECC237b02 (four infants), ECC247c07 (six infants), ECC289a06 (five infants) and ECC227b03 (five infants), which exhibited the highest levels of similarity (99–100%) to *Veillonella* sp., *E. coli*, *Enterococcus* sp., *Staphylococcus* sp. and *Enterobacter aerogenes*, respectively. These phylotypes covered almost 90% of clones from analyzed libraries. However, for samples from NEC patients, the most frequently detected representative OTUs were ECN224a01 (six patients), ECN232e08 (six patients), ECN224d08 (four patients) and ECN291c02 (four patients), which were closely related (99–100%) to *Klebsiella pneumoniae*, *Shigella dysenteriae*, *Enterobacter hormaechei* and *E. coli.* These four phylotypes represented 1755 clones and one of these four was dominant (more than 50%) in the NEC samples, except for samples N4 and N6 that were dominantly colonized by *Haemophilus parainfluenzae* and *Pseudomonas* sp., which covered about 70% of library clones in these NEC patients.

To show the phylogenetic relationships among the phylotypes identified in control and NEC infants, a phylogenetic tree of all representative clones in the library was constructed. As shown in Figure 3, for both control and NEC infants, phylotypes that were affiliated with the class *Gammaproteobacteria* constructed the largest phylogenetic clade in the tree. Control infants had a more diverse gut population colonization pattern compared to NEC infants. The phylogenetic tree indicates that for most of the OTUs of *Gammaproteobacteria* identified in NEC infants, the same phylotypes or their closely related sequences could also be found in control infants. This was further illustrated by comparison between twin

pairs in which one twin developed NEC and the other was spared, representing ideal matched controls. Certain phylotypes such as *K. pneumoniae*, *S. dysenteriae* and *Enterobacter hormaechei*, could be identified in both the NEC and healthy twins. However, one of these organisms had predominant colonization in the NEC infant whereas only a small percentage of the same organism was found in the healthy control. For example, the *Enterobacter hormaechei* phylotype was detected in both samples C1 and N1. But only 3 sequences belonged to this phylotype out of 261 clones of library constructed from sample C1 compared with 160 sequences belonging to this phylotype out of 267 clones of library from sample N1. Comparative analysis of just the twin samples also highlights what was shown for the whole cohort. Individual pairings had decreased diversity and increased percentage of *Proteobacteria* in the twin with NEC (N1-N4) compared to the twin control (C1-C4; Figures 4a and b).

Multivariate analyses of clone libraries

The bacterial composition and profile in each individual sample was compared to explore the potential similarity and relationship for NEC and control infants. As revealed by library cloning and sequencing data, high interindividual variability in intestinal bacterial composition was found in the infants. Using the UniFrac program to compare the sequence composition between libraries, significant differences were observed between all pairings of infant samples ($P < 0.01$; Supplementary Table 3). Even for the four pairs of twins, intestinal microbiota varied significantly between samples.

Furthermore, the limited diversity and *Proteobacteria* dominance in NEC patients represented a unique bacterial pattern associated with NEC by PCA. As shown in Figure 5, samples from each patient were represented as discrete points without overlap, the control samples (blue squares) were clearly separated from NEC samples (red circles) by grouping into one side of the PCA plot. This suggests that the bacterial populations of the NEC infants share some characteristics that differentiate them from healthy controls. Moreover, both CN10 and C9 samples, the only control samples that grouped with the NEC patients, were obtained from patients who developed NEC 3 and 10 days later, respectively. Clustering analysis using only twin pairs again also demonstrated that sequences from NEC patients clustered separately (Figure 4c). Thus samples from NEC patients (red) were more similar to each other than were samples between twins, even genetically identical twins (sample N1/C1).

Discussion

The newborn gut is initially sterile. Thereafter, a rapid colonization by a wide variety of organisms
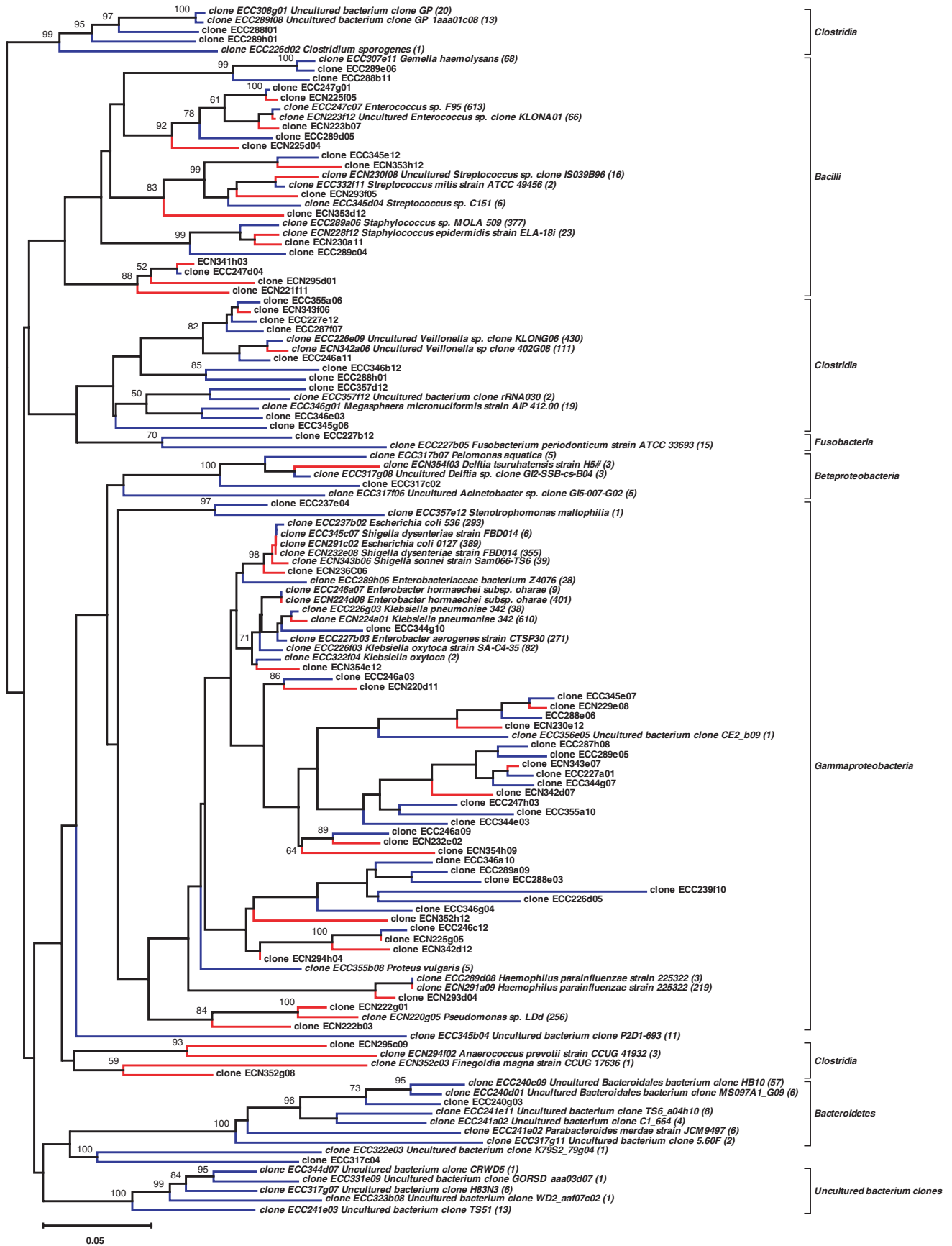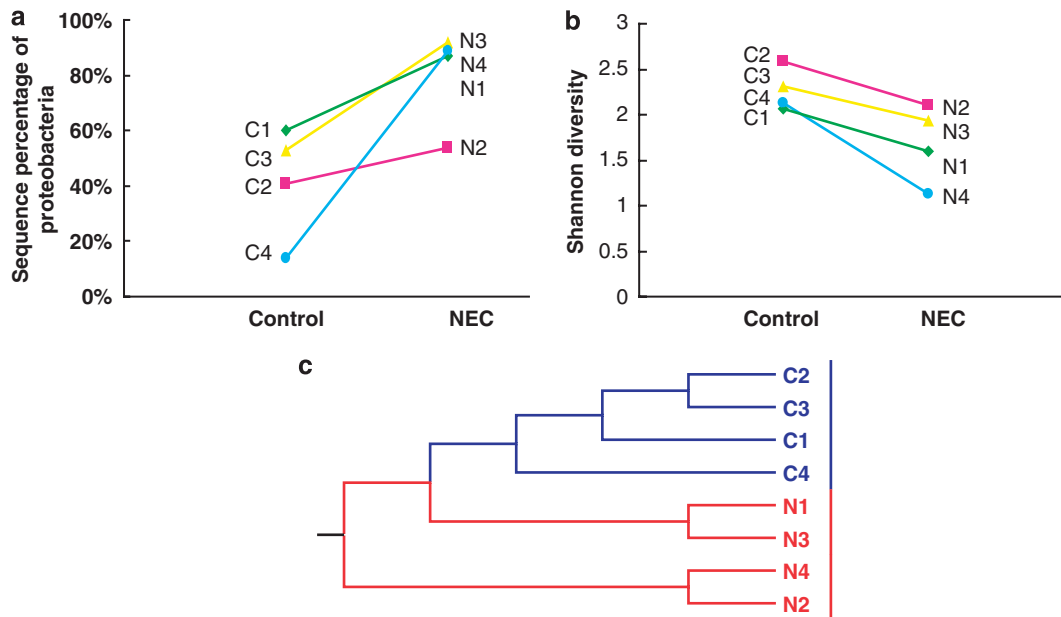
Figure 4 Increased abundance of *Proteobacteria* (**a**) and decreased diversity (**b**) are demonstrated between samples collected from infants with and without NEC in samples from four sets of twins. (**c**) Dendrogram demonstrating that even for twin pairs, NEC patients (N1-N4, red) clustered together and distinctly separated from their healthy cohorts (C1-C4, blue) by UniFrac analysis. NEC, necrotizing enterocolitis.

is a normal part of development important for gut maturation, containment of pathogenic organisms and normal metabolism (Hooper *et al.*, 2001; Stappenbeck *et al.*, 2002). The initial colonization patterns in normal full-term infants are dependent on environmental influences such as mode of delivery and type of feeding (Gronlund *et al.*, 1999; Harmsen *et al.*, 2000; Penders *et al.*, 2006; Palmer *et al.*, 2007). Full-term infants are described as having relative abundance of 46% *Proteobacteria* and 35% *Firmicutes* converging to an 'adult-like' microbiota by 1 year of age (Palmer *et al.*, 2007). Adult samples demonstrate a dominance of *Firmicutes* and *Bacteroidetes* phyla and small numbers of *Proteobacteria* (Eckburg *et al.*, 2005).

In this study we used molecular techniques to explore the composition of the fecal microbiota of preterm infants. Our data demonstrate limited diversity of the intestinal flora of all preterm infants (average 16 OTU) compared to full-term infants or adults (>200 OTU) (Eckburg *et al.*, 2005). Furthermore, we document that all preterm infants with NEC have a *Proteobacteria* predominance. This limited diversity and pathogenic bacteria dominance may explain the susceptibility of preterm infants to NEC.

Colonization of the blank slate of the preterm intestine is influenced by iatrogenic manipulations in the NICU. This includes a hospital environment; frequent use of broad-spectrum antibiotics, opioids and H2 blockers; and instrumentation with endotracheal, feeding and suctioning tubes. Resulting altered microbial flora may have significant implications for development of the immature preterm gut and susceptibility to NEC.

Previous studies using culture-based techniques have failed to document a causative pathogen for NEC. Rather, isolated studies using culture-based techniques have noted associations with various organisms including *Enterobacteriaceae* (Bell *et al.*, 1979a; Millar *et al.*, 1992), *Clostridium* (Sturm *et al.*, 1980), *E. coli* and *K. Pneumoniae* (Bell *et al.*, 1979b). However, results from bacterial cultures do not necessarily represent the most dominant or functionally relevant organisms. Molecular 16S rRNA gene cloning techniques have the advantage of giving a more complete picture of the microbial community of interest based on present bacterial DNA. Using UniFrac analysis our data demonstrate that each infant has a unique individual gut microbiota, even for twins. Consistent with other studies, our data do not implicate a specific

Figure 3 Phylogenetic relationships among the OTUs detected in fecal samples both from control and NEC infants. The representative 125 clone numbers are listed in the tree. Clones labeled with ECC on blue branches were obtained from control infants. Clones labeled with ECN on red branches were obtained from NEC infants. Sequences with a more than 99% match are given the name of the matched species in GenBank and labeled with the number of identified clones in the libraries. The accession number, the nearest neighbor, the level of similarity and the number of clones in the clone library for each sequence are shown in Supplementary Tables 1 and 2. Bootstrap values are based on 1000 replications and values above 50% are labeled. OTUs, operational taxonomical units; NEC, necrotizing enterocolitis.
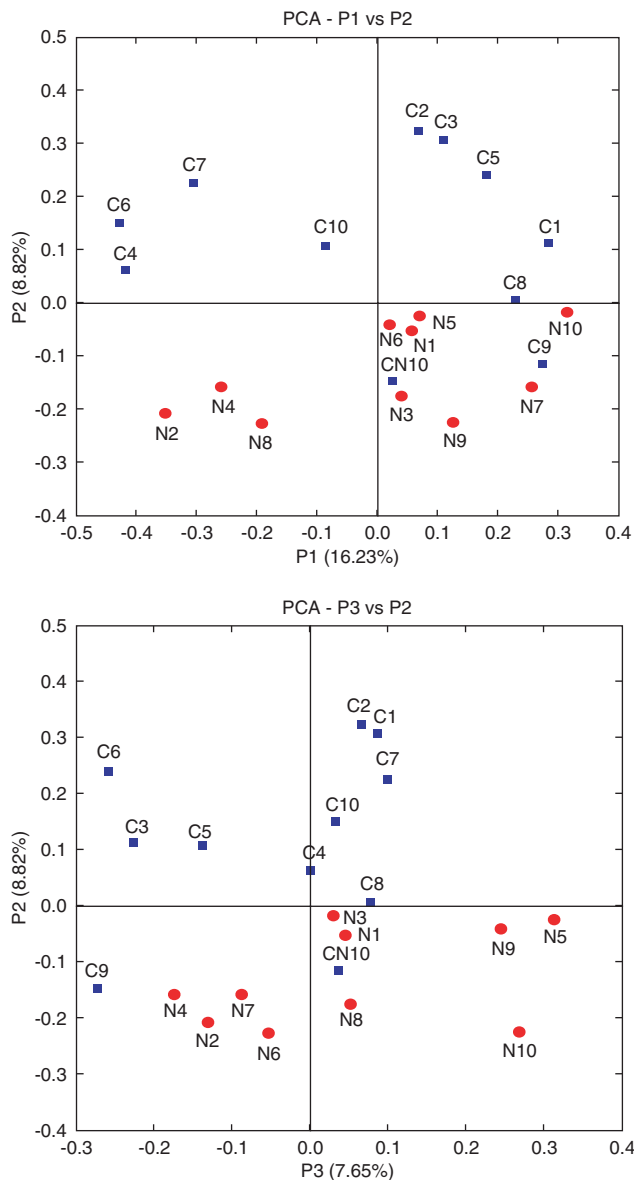
Figure 5 Principal coordinate analysis (PCA) of sequence libraries. Samples collected from infants without NEC (control) are represented by squares (blue); samples from infants with NEC are represented by circles (red). Distribution of samples collected from infants with NEC was distinct from that collected from control infants. Samples C9 and CN10, which were collected from control patients who later developed NEC, notably clustered with the NEC group. NEC, necrotizing enterocolitis.

pathogen. Rather, we find that the microbial community structure in NEC patients is distinct, based on a further decrease in diversity and increase in *Proteobacteria* dominance compared to other preterm infants, and with a bloom of a single genus of *Proteobacteria* to >50% of the overall bacterial composition. This decreased diversity is consistent with patterns found in other inflammatory bowel diseases such as Crohn's disease (Dicksved *et al.*, 2008). These findings were emphasized by the isolated analysis of four twin pairs in which one twin developed NEC and the other was spared.

These pairings represent ideal matched control sets and again demonstrated the decreased diversity and *Proteobacteria* dominance in the patient with NEC. Notably, the only two control patients with microbial community patterns which clustered with the NEC patients, went on to develop NEC days after these samples were obtained.

It has been speculated that the initial colonization of intestinal microbiota in infants is determined by the opportunistic exposure of the specific bacteria (Palmer *et al.*, 2007). Sequence matching demonstrated that many of the same organisms were found in both NEC and control infants. However, in the NEC patients a specific organism was present at a much higher percentage than the same organism in the control patients. This investigation focused on differences in bacterial colonization patterns between infants with and without NEC. We did not analyze all variations in clinical practice that could contribute to this difference. However, we did find that patients with NEC had received significantly more days of antibiotics during their neonatal course before the onset of this disease (13.7 ± 10.2 vs 3.7 ± 3.0 days, $P = 0.005$). Although it is possible that antibiotics are merely a marker for ill infants at risk for NEC or an infectious indication that increases susceptibility to NEC, our data are consistent with previous studies demonstrating an inverse correlation between days of previous antibiotic treatment and number of bacterial species in preterm infants (Gewolb *et al.*, 1999). The concept that prolonged antibiotic administration decreases intestinal microbial diversity and predisposes to intestinal disease has been demonstrated for *Clostridium difficile* colitis, and antibiotic-associated diarrhea (Young and Schmidt, 2004; Chang *et al.*, 2008). A recent study has specifically correlated prolonged duration of initial antibiotic therapy with increased rates of NEC, demonstrating that each antibiotic treatment day was associated with increased odds of NEC (Cotten *et al.*, 2009). It is possible that early antibiotic use interferes with establishment of non-*Proteobacteria* species, limiting competitive containment of these species. Diversity in any ecosystem is important in establishing stability and preventing dominance by single species. We speculate that frequent broad-spectrum antibiotic use decreases the microbial diversity of an already susceptible developing gut, allowing the bloom of a single pathogenic *Proteobacteria* species.

Although bacteria are often thought of as primarily harmful, a large body of literature demonstrates that in the gut certain commensal bacteria have critical beneficial functions. Bacteria are responsible for provision of essential nutrients such as vitamin K, vitamin B12 and short-chain fatty acids such as butyrate (Hooper *et al.*, 1998). Bacteria are important in the metabolism of polysaccharides (Salyers *et al.* 1977; Vercellotti *et al.*, 1977). In addition, certain bacteria provide bacterial interference resulting in competitive colonization against pathogenic organ-

isms. Normal commensal flora compete for host binding sites, stimulate host defense mechanisms, compete for nutrients and trigger cell-signaling events that limit the production of virulence factors; thus the growth of competitive nonpathogenic strains of bacteria may protect the infant. Bacteria are also necessary for maturation of the intestine and appropriate containment of inflammatory responses (Hooper et al., 2001). As interindividual variations in microbial composition are known to exist, it is possible that a difference in the overall pathogen vs beneficial bacterial balance predisposes to disease. A limitation of the 16S rRNA gene-based method is that the function of the identified bacteria is unknown. Microbial metagenomics is a high-throughput sequencing technique applied to whole-microbial-community DNA, which provides information about the functional diversity of a microbial community (Turnbaugh et al., 2006). Metagenomic analysis to evaluate differences in functional biochemical attributes of the bacterial populations in NEC vs control patients will be useful for future studies to gain further insight into the role of different microbial populations in the development of NEC.

These data represent patients from a single nursery, with the advantage of consistent clinical care patterns and environment. However, our study is limited by sample size, reliance on stool samples and potential bias of NICU-specific practices. For these studies we have used stool samples as a surrogate for the intestinal microbial community that would be directly influencing the incidence of NEC. Stool samples may not represent the bacteria locally at the point of intestinal injury (Eckburg et al., 2005; Lepage et al., 2005). Previous studies have suggested that bacterial adherence may be important in NEC, and studies in patients with other inflammatory bowel diseases have demonstrated increased bacterial attachment to the intestinal epithelium surface compared to controls (Panigrahi et al., 1994; Conte et al., 2006). In adult patients it has been shown that significant variability exists between stool and adherent intestinal mucosal samples, but similar studies have not been performed in preterm infants as intestinal samples are only possible from surgical resections rather than being available from endoscopic procedures (Eckburg et al., 2005). Our conclusions suggest that a decrease in diversity and shift in the balance of microbial flora rather than any specific organisms may be important in the development of NEC. Infants in different nurseries may harbor different specific bacterial strains, yet patients with NEC still have decreased diversity and increased Gammaproteobacteria dominance. Further evaluation of stool and intestinal samples from infants from multiple nurseries could yield additional important information.

Although differences in microbiota between preterm and full-term infants and adults could be developmentally based, it is likely that factors in the hospital environment or NICU practice influence microbiota development. Furthermore, limited microbiota diversity among preterm infants may render the microbiota more susceptible to perturbations introduced by clinical care or environmental conditions. Our results raise the possibility that decreased microbial diversity caused by administration of antibiotics predisposes to NEC. Breast milk feeding and probiotics have been shown to alter intestinal flora and decrease the incidence of NEC (Lucas and Cole, 1990; Hoyos, 1999; Harmsen et al., 2000; Bin-Nun et al., 2005; Lin et al., 2005). Given the clear separation of NEC and non-NEC microbial 16S rRNA gene profiles, we suggest that this approach should be studied further to determine its utility as a prognostication tool for identification of preterm infants at highest risk for development of NEC who could then be treated with targeted therapy (for example, to Gammaproteobacterias), fresh breast milk or probiotic agents.

Acknowledgements

References

Abdo Z, Schuette UM, Bent SJ, Williams CJ, Forney LJ, Joyce P (2006). Statistical methods for characterizing diversity of microbial communities by analysis of terminal restriction fragment length polymorphisms of 16S rRNA genes. Environ Microbiol 8: 929–938.

Bell MJ, Bell MJ, Shackelford P, Feigin RD, Ternberg JL, Brotherton T. (1979a). Epidemiologic and bacteriologic evaluation of neonatal necrotizing enterocolitis. J Pediatr Surg 14: 1–4.

Bell MJ, Feigin RD, Ternberg JL. (1979b). Changes in the incidence of necrotizing enterocolitis associated with variation of the gastrointestinal microflora in neonates. Am J Surg 138: 629–631.

Bell MJ, Ternberg JL, Feigin RD, Keating JP, Marshall R, Barton L et al. (1978). Neonatal necrotizing enterocolitis. Therapeutic decisions based upon clinical staging. Ann Surg 187: 1–7.

Bin-Nun A, Bromiker R, Wilschanski M, Kaplan M, Rudensky B, Caplan M et al. (2005). Oral probiotics prevent necrotizing enterocolitis in very low birth weight neonates. J Pediatr 147: 192–196.

Chang JY, Antonopoulos DA, Kalra A, Tonelli A, Khalife WT, Schmidt TM. (2008). Decreased diversity of the fecal Microbiome in recurrent Clostridium difficile-associated diarrhea. J Infect Dis 197: 435–438.

Claud EC, Walker WA. (2001). Hypothesis: inappropriate colonization of the premature intestine can cause neonatal necrotizing enterocolitis. FASEB J 15: 1398–1403.

Conte MP, Sschippa S, Zamboni I, Penta M, Chiarini F, Seganti L et al. (2006). Gut-associated bacterial microbiota in paediatric patients with inflammatory bowel disease. Gut 55: 1760–1767.

Cotten CM, taylor S, Stoll B, Goldberg RN, Hansen NI, Sanchez PJ et al. (2009). Prolonged duration of initial

954

empirical antibiotic treatment is associated with increased rates of necrotizing enterocolitis and death for extremely low birth weight infants. *Pediatrics* **123**: 58–66.

Dicksved J, Halvarson J, Rosenquist M, Jarnerot G, Tysk C, Apajalahti J et al. (2008). Molecular analysis of the gut microbiota of identical twins with Crohn's disease. *ISME J* **2**: 716–727.

Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M et al. (2005). Diversity of the human intestinal microbial flora. *Science* **308**: 1635–1638.

Gewolb IH, Schwalbe RS, Taciak VL, Harrison TS, Panigrahi P. (1999). Stool microflora in extremely low birthweight infants. *Arch Dis Child Fetal Neonatal Ed* **80**: F167–F173.

Goldmann DA, JLeclair Macone A. (1978). Bacterial colonization of neonates admitted to an intensive care environment. *J Pediatr* **93**: 288–293.

Good IJ. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, 237–264.

Gronlund MM, Lehtonen OP, Eerola E, Kero P. (1999). Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery. *J Pediatr Gastroenterol Nutr* **28**: 19–25.

Harmsen HJ, Wildeboer-Veloo AC, Raangs GC, Wagendorp AA, Klijn N, Bindels JG. (2000). Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. *J Pediatr Gastroenterol Nutr* **30**: 61–67.

Hooper LV, Bry L, Falk PG, Gordon JI. (1998). Host–microbial symbiosis in the mammalian intestine: exploring an internal ecosystem. *Bioessays* **20**: 336–343.

Hooper LV, Wong MH, Thelin A, Hansson L, Falk PG, Gordon JI. (2001). Molecular analysis of commensal host–microbial relationships in the intestine. *Science* **291**: 881–884.

Hoyos AB. (1999). Reduced incidence of necrotizing enterocolitis associated with enteral administration of *Lactobacillus acidophilus* and *Bifidobacterium infantis* to neonates in an intensive care unit. *Int J Infect Dis* **3**: 197–202.

Huber T, Faulkner G, Hugenholtz P. (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**: 2317–2319.

Kosloske AM. (1994). Epidemiology of necrotizing enterocolitis. *Acta Paediatr Suppl* **396**: 2–7.

Kuehl CJ, Wood HD, Marsh TL, Schmidt TM, Young VB. (2005). Colonization of the cecal mucosa by *Helicobacter hepaticus* impacts the diversity of the indigenous microbiota. *Infect Immun* **73**: 6952–6961.

Kumar S, Nei M, Dudley J, Tamura K. (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* **9**: 299–306.

Lepage P, Seksik P, Sutren M, de la Cochetiere MF, Jian R, Marteau P. (2005). Biodiversity of the mucosa-associated microbiota is stable along the distal digestive tract in healthy individuals and patients with IBD. *Inflamm Bowel Dis* **11**: 473–480.

Lin HC, Su BH, Chen AC, Lin TW, Tsai CH, Yeh TF. (2005). Oral probiotics reduce the incidence and severity of necrotizing enterocolitis in very low birth weight infants. *Pediatrics* **115**: 1–4.

Lozupone C, Hamady M, Knight R. (2006). UniFrac—an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* **7**: 371.

Lozupone C, Knight R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Apl Environ Microbiol* **71**: 8228–8235.

Lucas A, Cole TJ. (1990). Breast milk and neonatal necrotising enterocolitis. *Lancet* **336**: 1519–1523.

Millar MR, MacKay P, Levene M, Langdale V, Martin C. (1992). *Enterobacteriaceae* and neonatal necrotising enterocolitis. *Arch Dis Child* **67**(1 Spec No): 53–56.

Palmer C, EMBik Digiulio DB, Relman DA, Brown PO. (2007). Development of the human infant intestinal microbiota. *PLoS Biol* **5**: e177.

Panigrahi P, Sgupta S, Gewolb IH, Morris Jr JG. (1994). Occurrence of necrotizing enterocolitis may be dependent on patterns of bacterial adherence and intestinal colonization: studies in Caco-2 tissue culture and weanling rabbit models. *Pediatr Res* **36**(1 Part 1): 115–121.

Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I et al. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118**: 511–521.

Relman DA. (1999). The search for unrecognized pathogens. *Science* **284**: 1308–1310.

Salyers AA, West SE, Vercellotti JR, Wilkins TD. (1977). Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. *Appl Environ Microbiol* **34**: 529–533.

Schloss PD, Handelsman J. (2005). Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**: 1501–1506.

Schwiertz A, Gruhl B, Lobnitz M, Michel P, Radke M, Blaut M. (2003). Development of the intestinal bacterial composition in hospitalized preterm infants in comparison with breast-fed, full-term infants. *Pediatr Res* **54**: 393–399.

Stappenbeck TS, Hooper LV, Gordon JI. (2002). Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *Proc Natl Acad Sci U S A* **99**: 15451–15455.

Stoll BJ. (1994). Epidemiology of necrotizing enterocolitis. *Clin Perinatol* **21**: 205–218.

Sturm R, Staneck JL, Stauffer LR, Neblett WW. (1980). Neonatal necrotizing enterocolitis associated with penicillin-resistant, toxigenic *Clostridium butyricum*. *Pediatrics* **66**: 928–931.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**: 4876–4882.

Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**: 1027–1031.

Vercellotti JR, Salyers AA, Bullard WS, Wilkins D. (1977). Breakdown of mucin and plant polysaccharides in the human colon. *Can J Biochem* **55**: 1190–1196.

Young VB, Schmidt TM. (2004). Antibiotic-associated diarrhea accompanied by large-scale alterations in the composition of the fecal microbiota. *J Clin Microbiol* **42**: 1203–1206.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)