ELSEVIER

INSTITUT PASTEUR

Research in Microbiology 159 (2008) 367–373

www.elsevier.com/locate/resmic

Viral diversity and dynamics in an infant gut

Mya Breitbart [a,*], Matthew Haynes [b], Scott Kelley [b], Florent Angly [b,c], Robert A. Edwards [d,e], Ben Felts [c], Joseph M. Mahaffy [c], Jennifer Mueller [b,c], James Nulton [c], Steve Rayhawk [b,c], Beltran Rodriguez-Brito [b,c], Peter Salamon [c], Forest Rohwer [b]

[a] *College of Marine Sciences, University of South Florida, 140 Seventh Avenue South, Saint Petersburg, FL 33701, USA*
[b] *Department of Biology, LS301, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA*
[c] *Department of Mathematical Sciences, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA*
[d] *Department of Computer Science, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA*
[e] *Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Ave, Argonne, IL 60439, USA*

Received 25 February 2008; accepted 21 April 2008
Available online 1 May 2008

Abstract

Metagenomic sequencing of DNA viruses from the feces of a healthy week-old infant revealed a viral community with extremely low diversity. The identifiable sequences were dominated by phages, which likely influence the diversity and abundance of co-occurring microbes. The most abundant fecal viral sequences did not originate from breast milk or formula, suggesting a non-dietary initial source of viruses. Certain sequences were stable in the infant's gut over the first 3 months of life, but microarray experiments demonstrated that the overall viral community composition changed dramatically between 1 and 2 weeks of age.

Keywords: Infant; Feces; Virus; Phage; Metagenomics; Microarray; Gut

1. Introduction

Infants are born with sterile gastrointestinal tracts [24]. During and after birth, infants are rapidly colonized by microbes from the mother and surrounding environment [13,14,24,39,41]. In the beginning of life, the fecal microbiota is relatively simple, although there is a large amount of variation between individuals [22,32]. Microbial diversity changes very rapidly during the first few days after birth, with the community gradually becoming more diverse and stable over time [13,14,20,32,33,41,43,50]. The guts of infants are first colonized by aerobes and facultative anaerobes, which are subsequently replaced by strict anaerobes [17,32,42].

The initial acquisition and development of intestinal microbes appears to play a critical role in the health of infants. Colonizing bacteria assist in gut maturation, modulate the mucosal immune system and facilitate nutrient absorption and carbohydrate assimilation [42]. The microbiota of the developing infant gut is affected by the mode of delivery, diet, maternal microbiota, antibiotic treatment, the microbial load of the surrounding environment and characteristics of the individual's gut environment (pH, immune responses) [7,10,13,21,34,38,46].

Viruses likely play important roles in the gastrointestinal ecology of the developing infant. Adults contain an abundant and diverse community of both DNA and RNA viruses [5,8,15,53]. The majority of the DNA viruses in adult fecal samples are phages (i.e., viruses that infect bacteria), while the majority of the RNA viruses are plant viruses. Specific PCR has detected Adenovirus, Astrovirus, Bocavirus, Enterovirus, Norovirus, Picobirnavirus, Rotavirus, Sapovirus, and Torovirus in the feces of infants [9,18,19,23,29,36,47–49,51]. However, no studies to date have examined the total viral populations in healthy infants.

Here we present the first description of the DNA viral community within an infant's gut using metagenomic sequencing.

* Corresponding author. Tel.: +1 727 553 3520; fax: +1 727 553 1189.
E-mail address: mya@marine.usf.edu (M. Breitbart).

0923-2508/$ - see front matter © 2008 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.resmic.2008.04.006

At 1 week of age, the diversity of the gut viral community was extremely low. Most of the sequences were not similar to previously reported sequences, suggesting that much of the infant gut viral diversity is novel. The majority of the identifiable sequences were similar to phages, viruses that infect bacteria. These phages likely influence the diversity and structure of microbial communities in the developing infant gut. Changes in the relative abundance of different viral sequences between 1 and 2 weeks of age suggested that the infant gut viral community was dynamic.

2. Materials and methods

2.1. Sample collection and processing

Approximately 4 g of fresh fecal matter was collected from the sterile diaper of a 1-week-old male infant. The infant was full-term, delivered vaginally, had been fed a mixture of breast milk and formula from birth and had no antibiotic treatments. The feces was resuspended in 15 ml of sterile 1× SM buffer (100 nM NaCl, 10 mM $MgSO_4$, 50 mM Tris–HCl, pH 7.5) and shaken vigorously. The supernatant was filtered through a 100-μm pore size Nitex filter and then centrifuged at 3000 × *g* for 10 min and filtered through a 0.2 μm Sterivex filter to remove the bacteria and large particles. The fecal viruses were then loaded onto a cesium chloride (CsCl) step gradient and ultracentrifuged at 55,000 × *g* for 2 h at 4 °C. This step separates the viruses from any free nucleic acids present in the sample. Viruses in the 1.2–1.5 g ml^{-1} fraction were collected and used for microscopy and DNA extraction. This density range would recover all major groups of viruses [28], and this fraction contained the vast majority of the viral particles as determined by epifluorescent microscopy.

2.2. Microscopy

Epifluorescent microscopy was used to ensure that the purified viral community used for DNA extraction, metagenomic sequencing and microarray hybridization was not contaminated by any microbial cells [30]. The sample was stained with 1× SYBR Gold (Invitrogen; Carlsbad, CA) for 10 min in the dark and visualized using epifluorescent microscopy.

For electron microscopy, the CsCl-purified fecal viruses were adsorbed to glow-discharged formvar-coated 200-mesh copper grids (Ted Pella; Redding, CA) for 10 min. The grids were stained for 20 s in 1% uranyl acetate. Viruses were observed using the FEI Tecnai 12 (FEI; Hillsboro, OR) and images were recorded digitally using a Tietz TemCam-F214 (TVIPS; Gauting, Germany).

2.3. DNA extraction and shotgun library construction

Total DNA was extracted from the purified fecal viruses using formamide and cetyltrimethylammonium bromide (CTAB) extractions [40]. Lucigen Corporation (Middleton, WI) constructed a linker-amplified shotgun library (LASL) from the infant viral DNA as described previously [4–6]. Briefly, the total viral DNA was randomly sheared into 1–2 kb fragments using a HydroShear (GenMachines; San Carlos, CA). The fragments were end-repaired, short double-stranded DNA (dsDNA) linkers were ligated to the fragment ends and the fragments were PCR-amplified using primers to the linkers. The linker-amplified fragments were cloned into the pSMART vector (Lucigen; Middleton, WI).

2.4. Sequence analyses

A total of 477 clones were sequenced by SymBio Corporation (Menlo Park, CA) using the AmpL2 forward primer (GenBank accession numbers: ED651217–ED651693). Quality scores were assigned to the sequences using Phred [11,12] and the sequences were trimmed using Sequencher 4.0 (Gene Codes, Ann Arbor, MI). Sequences were trimmed until there were not more than 3 bases with confidence scores below 25 within 25 bases of either end. The trimmed sequences were compared against the GenBank non-redundant database using TBLASTX [1,2]. Sequences with significant hits (*E*-value <0.001) were classified as phages, viruses, mobile elements, Bacteria, Archaea or Eukarya based on GenBank annotation. Sequences were preferentially classified as phages or viruses if these occurred in the top 5 hits. Bacterial hits were examined manually to identify prophages within bacterial genomes.

2.5. Community structure modeling

The sequences were assembled using Sequencher 4.0 (Gene Codes, Ann Arbor, MI) to identify contigs (i.e., contiguous or overlapping sequences) based on a minimum overlap of 20 bp with a minimal match percentage of 98%. These assembly parameters allow differentiation between closely related phages, as determined previously [6]. The structure of the uncultured infant fecal viral community was mathematically modeled using the PHAge communities from the Contig Spectrum (PHACCS) program [3]. For these calculations, an average fragment size of 505 bases was used, and the average genome size was assumed to be 50 kb.

2.6. PCR for specific viruses

Primer 3 [37] was used to design primers based on contigs assembled from the week 1 sequences. Viruses were purified from the food sources of the child (breast milk from the subject's mother, and formula), as well as fecal samples at weeks 1, 2, 5 and 14 of age using the methods outlined above. Additional samples included viral DNA prepared as described above from two fecal viral DNA samples from an unrelated adult separated by a year. These samples were screened by PCR for three viral sequences. The PCR reaction (50 μl total volume) contained 1× FideliTaq PCR Master Mix (USB; Cleveland, OH), 1 μM each primer, and 1 μl of target DNA. The PCR conditions were 5 min at 95 °C; 35 cycles of 1 min at 94 °C, 1 min at 60 °C, 0.5°/cycle, and 2 min at 72 °C; followed by 10 min at 72 °C. Positive PCR products were sequenced for confirmation.

2.7. Microarray experiments

CustomArray 12k arrays (CombiMatrix Corporation; Mukilteo, WA) were constructed using probes designed from sequences from the infant fecal viral community present at 1 week of age. The arrays included two replicates of each probe, generally 35 nucleotides in length, as well as positive and negative control sequences.

To prepare the DNA for hybridizing against the microarray, total viral DNA isolated from the infant at 1 week and 2 weeks of age was amplified using Genomiphi (Amersham Biosciences; Piscataway, NJ). The amplified DNA was labeled by incorporation of 5-aha-dUTP, followed by coupling of the reactive dyes, AlexaFluor 555 or AlexaFluor 647 (Invitrogen, Carlsbad, CA). For the data presented here, the viral DNA from the infant at week 1 was labeled with AlexaFluor 555 and the viral DNA from the infant at week 2 was labeled with AlexaFluor 647. Dye swap experiments and repeated hybridization with these same arrays verified repeatability of the assay.

Hybridizations were performed at 50 °C for 16–20 h, under conditions recommended by CombiMatrix. After washing, the arrays were scanned at 532 nm and 635 nm using an Axon 4000B Microarray scanner with GenePix Pro 6.0 software (Molecular Devices, Sunnyvale, CA).

The local background at each wavelength was subtracted from the median intensity of each spot. The two replicate array spots for each probe were averaged, and the intensities were normalized by dividing the average intensity for each probe by the maximum probe intensity recorded for that wavelength. ''High'' signal was defined as $\geq$50% of the maximum intensity, ''medium'' was between 10 and 50% of the maximum intensity, ''low'' was between 1 and 10% of the maximum intensity, and any spots with intensities $<$1% of the maximum intensity were assumed to be background.

3. Results

3.1. Composition and diversity of infant viral community

Virus particles were purified from the feces of a 1-week-old infant and a shotgun library was constructed as described previously [5]. Direct counts of viruses with epifluorescent microscopy showed that the purified viral community was free of contamination by microbial cells (Fig. 1A). The majority of isolated particles appeared to be phages when viewed under the electron microscope (Fig. 1B).

The majority of sequences in the shotgun library (66%) had no significant hits (*E*-value $<$0.001) to any sequences in the GenBank non-redundant database (Fig. 2A), suggesting that the fecal viral community was mostly novel. The most common known matches were to phages (Fig. 2B). Only one sequence with significant similarity to a eukaryotic virus was observed. The closest GenBank hit for this sequence was lymphocystis disease virus, which infects fish. Sequences with significant similarities to the major double-stranded DNA phage groups, Podo-, Sipho-, and Myoviruses, were observed in the metagenomic library. Among the sequences with similarities to phages, the majority (72%) were similar to Siphoviruses and prophages (Fig. 2C). The most common hits were to phages that infect *Bacillus*, *Lactobacillus*, *Lactococcus*, *Bacteroides*, *Listonella*, and *Staphylococcus*. More than half of the phage hits were similar to phages known to infect Gram-positive hosts. Seventy-five percent of the phage hits were to known proteins, with endonucleases and terminases being the most common (Table 1). Several sequences were similar to genes involved in lysogeny, including integrases, and antirepressors. In addition, numerous similarities to mobile elements such as transposons and plasmids were observed.

The diversity and community structure of the infant fecal viruses was mathematically modeled using the PHAge Communities from Contig Spectrum (PHACCS) program [3]. First, the shotgun sequences were assembled using Sequencher 4.0 with a 20 bp overlap and 98% sequence identity. Among the 477 sequences, there were 138 sequences that did not overlap with any other sequences, 31 2-contigs, 11 3-contigs, eight 4-contigs, nine 5-contigs, three 6-contigs, one 7-contig, six 8-contigs, one 9-contig, three 12-contigs, one 15-contig, one 16-contig and one 18-contig. Mathematical modeling using the average fragment size of 505 bases, an average genome size of 50 kb, and a power law fit for the rank-abundance distribution, predicted low diversity of the infant gut viral community (Table 2). The infant fecal viral community present at 1 week of age consisted of 8 viral genotypes, with the most abundant genotype comprising 43.6% of the total viral community. The Shannon–Wiener Index for the infant viral community was 1.69 nats.

Specific primers were designed to amplify a portion of three viral sequences based on the largest contigs assembled in Sequencher (Table 3). A portion of one of these contigs (primer set 2) was identical to a viral sequence recovered from the feces of an unrelated adult [5]. Viral DNA purified from breast milk, formula, and infant fecal samples from 1, 2, 5, and 14 weeks of age was screened by PCR for the three viral sequences (Fig. 3). In addition, viral DNA isolated from the fecal viral community of an unrelated adult (on two separate occasions separated by a year) was also analyzed for the presence of these sequences. Although this assay is not quantitative, it is evident that these viral sequences were present in the infant's gut throughout the first 3 months of life. In addition, these sequences, which predominate in the infant's gut at week 1, do not appear to originate from breast milk or formula. One of the viral sequences was also present in two fecal samples collected from an unrelated adult. This sequence was most similar to a Bacteroides integrase.

3.2. Monitoring temporal changes in an infant's viral community using microarrays

To further describe the temporal variation of the infant gut viral community, viral DNA samples from week 1 and week 2 were hybridized against a microarray containing probes constructed from the week 1 viral sequences. A ''high'' signal was a spot with $\geq$50% of the maximum intensity, ''medium''

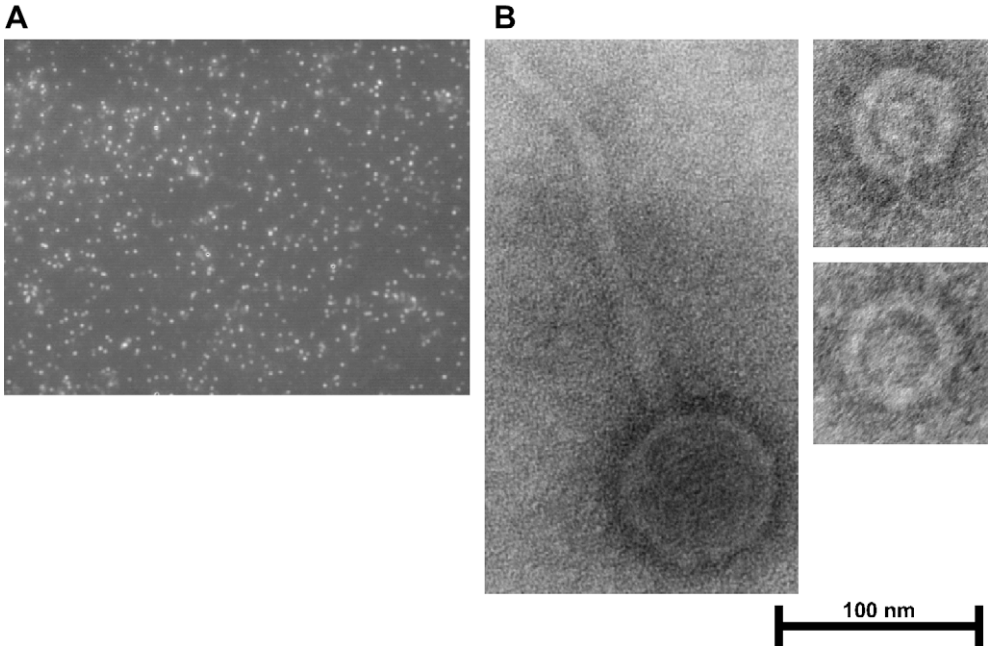

Fig. 1. Microscopy of viruses in infant feces at 1 week of age. (A) Purified infant fecal viral concentrate stained with SYBR Gold and viewed under epifluorescence microscopy. (B) Representative electron micrographs of viruses observed in the week 1 infant fecal sample.

was defined as 10–50% of the maximum intensity, ''low'' was between 1 and 10% of the maximum intensity, and any spots with intensities <1% of the maximum intensity were considered to be ''background''. The majority of the sequences were present at medium levels at week 1 (73%), but then became low (23%) or background (46%) at week 2 (Fig. 4). Over half (56%) of the sequences were present at week 1 but not week 2. All of the most abundant sequences at week 2 were of medium abundance at week 1. No sequences were present at high levels at both week 1 and week 2, suggesting that the infant viral community was dynamic over this short time period, with turnover of the most abundant viral sequences.

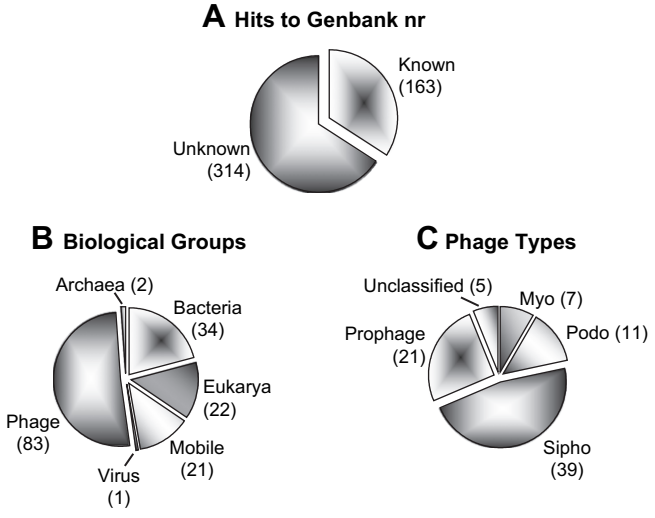

Fig. 2. TBLASTX sequence similarities from the uncultured viral community from infant feces collected at 1 week of age. GenBank accession numbers: ED651217–ED651693. (A) Sequences with significant similarity (*E*-value <0.001) to the GenBank nr database. (B) Similarities to major biological groups within the ''known'' hits. (C) Major phage types identified in the infant fecal sample.

4. Discussion

Direct epifluorescent microscope counts did not detect any viral particles in the meconium (i.e., the earliest infant stool samples). By the end of the first week, however, there were $\sim 10^8$ virus particles per gram wet weight of feces. High background and clumping in the direct counts precluded the determination of exact viral numbers in these samples. However, it is evident that viruses are an abundant and dynamic component of the microbiota of the developing infant gut.

Table 1
Genes observed amongst the phage hits based on GenBank annotation

Gene	Number of sequences
Unknown	21
Endonuclease	11
Terminase	11
Methylase	7
Integrase	7
Antirepressor	6
Methyltransferase	6
Recombination protein	4
Helicase	3
Transcriptional regulator	3
DNA polymerase	1
dCMP deaminase	1
Enterotoxin	1
Tail-sheath stabilizer	1

Table 2
Predictions for viral community structure and diversity in human feces as determined by mathematical modeling of contig spectra using a power law rank-abundance distribution and an average viral genome length of 50,000 bp

	Baby	Adult
Total no. of viral genotypes	8	1930
% Most abundant virus	43.6	6.3
Shannon–Wiener Index (H_{nats})	1.69	6.43

The initial source of viruses in the infant human gut is unknown, though inoculation from dietary, environmental, and maternal sources is possible. Additionally, the earliest infant fecal viruses may be the result of induction of prophages from the newly colonized microbial flora. Over 25% of the phage-like sequences were most similar to phages that infect lactic acid bacteria, including *Lactobacillus*, *Lactococcus*, *Streptococcus*, and *Enterococcus*, which are known to be abundant in breast milk [26]. However, in this study, the most abundant viral sequences at 1 week of age were not detected in either the breast milk or formula.

Metagenomic sequences from both the infant and a previously described adult fecal viral community were mostly unknown. Both infant and adult viral communities were dominated by Siphophages and prophages, supporting the idea that the gut environment might be dominated by phages with the ability to be temperate. Many hits to phages that infect Gram-positive hosts were seen in both the infant and adult fecal samples, which correlates with the abundance of Gram-positive hosts in the gut environment. Interestingly, one identical viral sequence was recovered from both the feces of the infant and an unrelated adult. This sequence was most similar to the *Bacteroides thetaiotamicron* integrase gene.

The most striking difference between the adult and infant viral communities was that diversity of the infant fecal viral community was much lower than that of the adult fecal viral community (Table 2). Based on mathematical modeling of sequence data, the adult fecal viral community contained approximately 2000 genotypes, while the infant viral community only contained 8 genotypes. In the infant fecal sample, the most abundant viral genotype comprised over 40% of the total community, while the adult community had no dominant viral genotype (most abundant virus ~6% of the total community). The low diversity in the viral community correlates with low microbial diversity in infant guts [13,14,24,25].

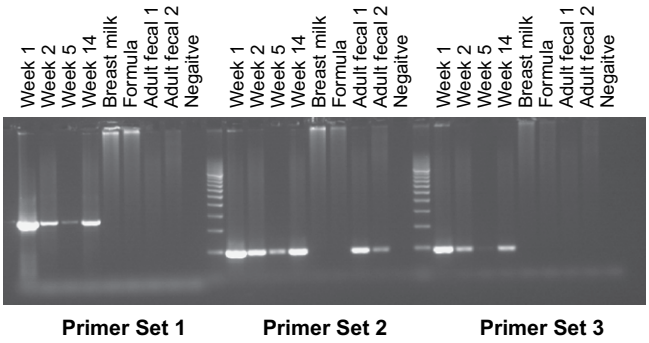

Fig. 3. PCR for three specific viral sequences originally identified in the week 1 infant fecal virus community. Samples included infant fecal virus samples over a 3-month period, viruses purified from breast milk and infant formula, as well as two fecal samples (separated by a year) from an unrelated adult.

The human gastrointestinal tract has been regarded as "one of the fiercest competitive ecological niches found in nature" [21]. In addition to competition for nutrients and space, the microbial populations may also be regulated by bacteriocins [21] or predators. Phages are an important force in controlling the abundance and composition of microbial communities [52]. The presence of phages may be especially important during gut colonization and succession, by infecting the dominant hosts and creating the opportunity for another strain to become abundant [45]. This model of predator–prey dynamics, termed "kill the winner" [44], has not previously been suggested for gut microflora. In this model, blooms of a specific microbial species lead to blooms of their corresponding phage, followed by decreases in abundance of both. A prediction of this model is that the most abundant phage genotype will not be the same at different time points. Although some of the viral sequences were stable in the infant gut over the first 3 months of life, the microarray experiments showed dramatic changes in the overall composition of the fecal viral community between 1 and 2 weeks of age. The microbial community in developing infant guts over this time period has also been shown to be extremely dynamic [43,50].

Future studies should examine the ecological roles of phages in the development of normal gut microflora in infants. The microarray approaches used here are an example of how these studies might be done. However, there are a number of improvements that should be made. First, we suggest that the viral communities from multiple infants be sampled through time. These communities can then be pooled,

Table 3
Primer sequences used in this study

Primer set	Primer sequence (5′–3′)	Product size (bp)	BLAST hit of amplified sequence
Baby1F	GTGGTTGTTTGCATCCACTG	1430	*Streptococcus pyogenes* phage terminase, large subunit
Baby1R	ACCAAAAGCACCGTGGTAAG		
Baby2F	GCACATGAGACAGCTCCGTA	495	*Bacteroides thetaiotamicron* integrase
Baby2R	GACGGCTCACATGGTAGGTT		
Baby3F	TTGTGTTCGCTTCATTCTCG	475	*Bacillus thuringiensis* phage MZTP02
Baby3R	AAACAAGGTGGGACTGTTCG		

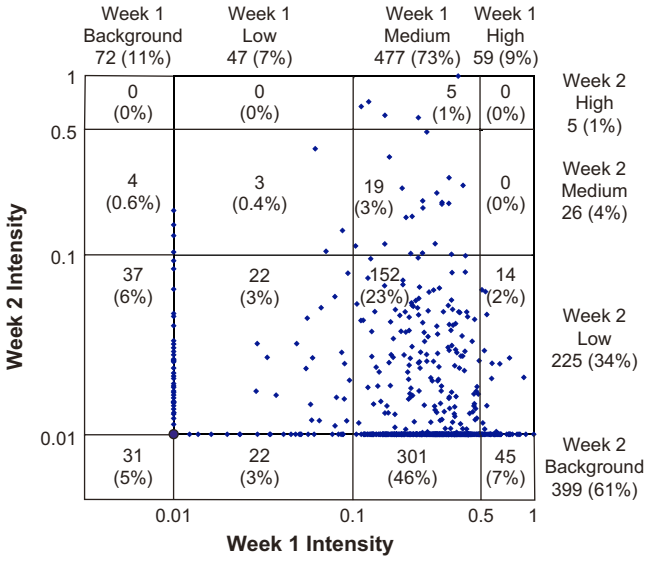

Fig. 4. Viral DNA samples from fecal samples collected at week 1 and week 2 of age were hybridized against a microarray containing probes designed from week 1 metagenomic viral sequences. After subtracting the local background and averaging replicate array features, the intensities were normalized by dividing the average intensity for each probe by the maximum probe intensity recorded for that wavelength. Spots with intensities <1% of the maximum intensity were designated as "background", "low" signal was between 1 and 1% of the maximum intensity, "medium" was between 10 and 50% of the maximum intensity and a "high" signal was defined as ≥50% of the maximum intensity.

sequenced, and used to make a combined metagenome microarray. Individual samples can then be probed against the microarray to monitor individual phage population dynamics. Multiple individuals need to be sampled in order to recognize general trends in viral dynamics during colonization. These studies would be further improved by concurrent sampling of the microbial community. Recent advances in examining the microbial communities of the human gastrointestinal tract with microarrays make studies of this nature immediately possible (e.g., [31]). Understanding the viral population in the infant gut will become increasingly important as prebiotics and probiotics are being more used for manipulation of the developing intestinal microbiota [16,24,27,35,38].

Acknowledgements

The authors thank Linda Wegley and Anca Segall for insightful discussions and Steve Barlow at the SDSU Microscopy Facility for help with microscopy. This project was supported by the Gordon and Betty Moore Foundation (grant to FR). The electron microscope was supported by NSF instrumentation grant DBI-0308029.

References

[1] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

[2] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

[3] Angly, F., Rodriguez-Brito, B., Bangor, D., McNairnie, P., Breitbart, M., Salamon, P., Felts, B., Nulton, J., Mahaffy, J., Rohwer, F. (2005) PHACCS, an online tool for estimating the structure and diversity of uncultured viral communities using metagenomic information. BMC Bioinformatics 6, 41.

[4] Breitbart, M., Felts, B., Kelley, S., Mahaffy, J., Nulton, J., Salamon, P., Rohwer, F. (2004) Diversity and population structure of a nearshore marine sediment viral community. Proc. R. Soc. B 271, 565–574.

[5] Breitbart, M., Hewson, I., Felts, B., Mahaffy, J., Nulton, J., Salamon, P., Rohwer, F. (2003) Metagenomic analyses of an uncultured viral community from human feces. J. Bacteriol. 85, 6220–6223.

[6] Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., Azam, F., Rohwer, F. (2002) Genomic analysis of uncultured marine viral communities. Proc. Natl. Acad. Sci. USA 99, 14250–14255.

[7] Brunser, O., Figueroa, G., Gotteland, M., Haschke-Becher, E., Magliola, C., Rochat, F., Cruchet, S., Palframan, R., Gibson, G., Chauffard, F., Haschke, F. (2006) Effects of probiotic or prebiotic supplemented milk formulas on fecal microbiota composition of infants. Asia-Pacific. J. Clin. Nutr. 15, 368–376.

[8] Cornax, R., Morinigo, M., Gonzalez-Jaen, F., Alonso, M., Borrego, J. (1994) Bacteriophages presence in human faeces of healthy subjects and patients with gastrointestinal disturbances. Zentralbl. Bakteriol. 281, 214–224.

[9] Cruz, J., Bartlett, A., Herrmann, J., Caceres, P., Blacklow, N., Cano, F. (1992) Astrovirus-associated diarrhea among Guatemalan ambulatory rural children. J. Clin. Microbiol. 30, 1140–1144.

[10] Dicksved, J., Floistrup, H., Bergstrom, A., Rosenquist, M., Pershagen, G., Scheynius, A., Roos, S., Alm, J.S., Engstrand, L., Braun-Fahrlander, C., von Mutius, E., Jansson, J.K. (2007) Molecular fingerprinting of the fecal microbiota of children raised according to different lifestyles. Appl. Environ. Microbiol. 73, 2284–2289.

[11] Ewing, B., Green, P. (1998) Basecalling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8, 186–194.

[12] Ewing, B., Hillier, L., Wendl, M., Green, P. (1998) Basecalling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8, 175–185.

[13] Fanaro, S., Chierici, R., Guerrini, P., Vigi, V. (2003) Intestinal microflora in early infancy: composition and development. Acta Paediatr. Suppl. 441, 48–55.

[14] Favier, C., Vaughan, E., DeVos, W., Akkermans, A. (2002) Molecular monitoring of succession of bacterial communities in human neonates. Appl. Environ. Microbiol. 68, 219–226.

[15] Furuse, K., Osawa, S., Kawashiro, J., Tanaka, R., Osawa, Z., Sawamura, S., Yanagawa, Y., Nagao, T., Watanabe, I. (1983) Bacteriophage distribution in human faeces: continuous survey of healthy subjects and patients with internal and leukemic diseases. J. Gen. Virol. 64, 2039–2043.

[16] Haarman, M., Knol, J. (2005) Quantitative real-time PCR assays to identify and quantify fecal *Bifidobacterium* species in infants receiving a prebiotic infant formula. Appl. Environ. Microbiol. 71, 2318–2324.

[17] Hopkins, M., Macfarlane, G., Furrie, E., Fite, A., Macfarlane, S. (2005) Characterisation of intestinal bacteria in infant stools using real-time PCR and northern hybridisation analyses. FEMS Microbiol. Ecol. 54, 77–85.

[18] Isakbaeva, E., Bulens, S., Beard, R., Adams, S., Monroe, S., Chaves, S., Widdowson, M., Glass, R. (2005) Norovirus and child care: challenges in outbreak control. Pediatr. Infect. Dis. J. 24, 561–563.

[19] Khamrin, P., Maneekarn, N., Peerakome, S., Tonusin, S., Malasao, R., Mizuguchi, M., Okitsu, S., Ushijima, H. (2007) Genetic diversity of noroviruses and sapoviruses in children hospitalized with acute gastroenteritis in Chiang Mai. Thailand. J. Med. Virol. 79, 1921–1926.

[20] Klaassens, E.S., de Vos, W.M., Vaughan, E.E. (2007) Metaproteomics approach to study the functionality of the microbiota in the human infant gastrointestinal tract. Appl. Environ. Microbiol. 73, 1388–1392.

[21] Klijn, A., Mercenier, A., Arigoni, F. (2005) Lessons from the genomes of bifidobacteria. FEMS Microbiol. Rev. 29, 491–509.

[22] Kurokawa, K., Itoh, T., Kuwahara, T., Oshima, K., Toh, H., Toyoda, A., Takami, H., Morita, H., Sharma, V.K., Srivastava, T.P., Taylor, T.D., Noguchi, H., Mori, H., Ogura, Y., Ehrlich, D.S., Itoh, K., Takagi, T., Sakaki, Y., Hayashi, T., Hattori, M. (2007) Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. DNA Res. 14, 169–181.

[23] Lodha, A., De Silva, N., Petric, M., Moore, A. (2005) Human torovirus: a new virus associated with neonatal necrotizing enterocolitis. Acta Paediatr. 94, 1085–1088.

[24] Mackie, R., Sghir, A., Gaskins, H. (1999) Developmental microbial ecology of the neonatal gastrointestinal tract. Am. J. Clin. Nutr. 69, 1035S–1045S.

[25] Magne, F., Abely, M., Boyer, F., Morville, P., Pochart, P., Suau, A. (2006) Low species diversity and high interindividual variability in faeces of preterm infants as revealed by sequences of 16S rRNA genes and PCR-temporal temperature gradient gel electrophoresis profiles. FEMS Microbiol. Ecol. 57, 128–138.

[26] Martin, R., Langa, S., Reviriego, C., Jimenez, E., Marin, M.L., Xaus, J., Fernandez, L., Rodriguez, J.M. (2003) Human milk is a source of lactic acid bacteria for the infant gut. J. Pediatr. 143, 754–758.

[27] Moro, G., Arslanoglu, S. (2005) Reproducing the bifidogenic effect of human milk in formula-fed infants: why and how? Acta Paediatr. 94, 14–17.

[28] Murphy, F.A., Fauquet, C.M., Bishop, D.H.L., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A., & Summers, M.D. (Eds.). (1995). Virus Taxonomy: Sixth Report of the International Committee on Taxonomy of Viruses. New York: Springer-Verlag. Vol. Supplement 10.

[29] Nguyen, T.A., Yagyu, F., Okame, M., Phan, T.G., Trinh, Q.D., Yan, H., Hoang, K.T., Cao, A.T.H., Le Hoang, P., Okitsu, S., Ushijima, H. (2007) Diversity of viruses associated with acute gastroenteritis in children hospitalized with diarrhea in Ho Chi Minh City, Vietnam. J. Med. Virol. 79, 582–590.

[30] Noble, R.T., Fuhrman, J.A. (1998) Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. Aquat. Microb. Ecol. 14, 113–118.

[31] Palmer, C., Bik, E., Eisen, M., Eckburg, P., Sana, T., Wolber, P., Relman, D., Brown, P. (2006) Rapid quantitative profiling of complex microbial populations. Nucleic Acids Res. 34. e5.

[32] Palmer, C., Bik, E.M., DiGiulio, D.B., Relman, D.A., Brown, P.O. (2007) Development of the human infant intestinal microbiota. PLoS Biol. 5, 1556–1573.

[33] Park, H., Shim, S., Kim, S., Park, J., Park, S., Kim, H., Kang, B., Kim, C. (2005) Molecular analysis of colonized bacteria in a human newborn infant gut. J. Microbiol. 43, 345–353.

[34] Penders, J., Thijs, C., Vink, C., Stelma, F.F., Snijders, B., Kummeling, I., van den Brandt, P.A., Stobberingh, E.E. (2006) Factors influencing the composition of the intestinal microbiota in early infancy. Pediatrics 118, 511–521.

[35] Rinne, M., Gueimonde, M., Kalliomaki, M., Hoppu, U., Salminen, S., Isolauri, E. (2005) Similar bifidogenic effects of prebiotic-supplemented partially hydrolyzed infant formula and breastfeeding on infant gut microbiota. FEMS Immunol. Med. Microbiol. 43, 59–65.

[36] Rosenlew, M., Stenvik, M., Roivainen, M., Jarvenpaa, A., Hovi, T. (1999) A population-based prospective survey of newborn infants with suspected systemic infection: occurrence of sporadic enterovirus and adenovirus infections. J. Clin. Virol. 12, 211–219.

[37] Rozen, S., Skaletsky, H. (2000) Primer3 on the WWW for General Users and for Biologist Programmers. In S. Krawetz, & S. Misener (Eds.), Bioinformatics Methods and Protocols: Methods in Molecular Biology. Totowa, NJ: Humana Press.

[38] Rubaltelli, F., Biadaioli, R., Pecile, P., Nicoletti, P. (1998) Intestinal flora in breast- and bottle-fed infants. J. Perinat. Med. 26, 186–191.

[39] Sakata, S., Tonooka, T., Ishizeki, S., Takada, M., Sakamoto, M., Fukuyama, M., Benno, Y. (2005) Culture-independent analysis of fecal microbiota in infants, with special reference to *Bifidobacterium* species. FEMS Microbiol. Lett. 243, 417–423.

[40] Sambrook, J., Fritsch, E., Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

[41] 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.

[42] Songjinda, P., Nakayama, J., Kuroki, Y., Tanaka, S., Fukuda, S., Kiyohara, C., Yamamoto, T., Izuchi, K., Shirakawa, T., Sonomoto, K. (2005) Molecular monitoring of the developmental bacterial community in the gastrointestinal tract of Japanese infants. Biosci. Biotechnol. Biochem. 69, 638–641.

[43] Tapiainen, T., Ylitalo, S., Eerola, E., Uhari, M. (2006) Dynamics of gut colonization and source of intestinal flora in healthy newborn infants. Apmis. 114, 812–817.

[44] Thingstad, T. (2000) Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnol. Oceanogr. 45, 1320–1328.

[45] Thingstad, T.F., Lignell, R. (1997) Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. Aquat. Microb. Ecol. 13, 19–27.

[46] Thompson-Chagoyan, O.C., Maldonado, J., Gil, A. (2007) Colonization and impact of disease and other factors on intestinal microbiota. Digest Dis. Sci. 52, 2069–2077.

[47] Valentini, S., Lustri, W., Falcao, D. (1992) Rotavirus and adenovirus in diarrheic and non-diarrheic feces of children in Araraquara, SP, Brazil. Rev. Microbiol. 23, 24–27.

[48] Vicente, D., Cilia, G., Montes, M., Perez-Yarza, E.G., Perez-Trallero, E. (2007) Human bocavirus, a respiratory and enteric virus. Emerg. Infect. Dis. 13, 636–637.

[49] Wakuda, M., Pongsuwanna, Y., Taniquchi, K. (2005) Complete nucleotide sequences of two RNA segments of human picobirnavirus. J. Virol. Methods 126, 165–169.

[50] Wang, M., Ahrne, S., Antonsson, M., Molin, G. (2004) T-RFLP combined with principal component analysis and 16S rRNA gene sequencing: an effective strategy for comparison of fecal microbiota in infants of different ages. J. Microbiol. Methods 59, 53–69.

[51] Wang, Y.H., Kobayashi, N., Zhou, D.J., Yang, Z.Q., Zhou, X., Peng, J.S., Zhu, Z.R., Zhao, D.F., Uu, M.Q., Gong, J. (2007) Molecular epidemiologic analysis of group A rotaviruses in adults and children with diarrhea in Wuhan city, China, 2000–2006. Arch. Virol. 152, 669–685.

[52] Weinbauer, M. (2004) Ecology of prokaryotic viruses. FEMS Microbiol. Rev. 28, 127–181.

[53] Zhang, T., Breitbart, M., Lee, W., Run, J.-Q., Wei, C., Soh, S., Hibberd, M., Liu, E., Rohwer, F., Ruan, Y. (2006) Prevalence of plant viruses in the RNA viral community of human feces. PLoS Biol. 4. e3.