Journal of Bacteriology and Virology 2013. Vol. 43, No. 4 p.290 – 296
http://dx.doi.org/10.4167/jbv.2013.43.4.290

Original Article

A Metaviromic Analysis of Viral Communities in the Feces of Unexplained Acute Gastroenteritis

In Hyuk Baek, Song ah Kim, Jong-Hwa Kim, Hee Kuk Park and Wonyong Kim*

Department of Microbiology, Chung-Ang University College of Medicine, Seoul, Korea

Although viruses are the most common cause of acute gastroenteritis (AGE) in humans, details about the causative viruses in AGE are largely unknown because many causative viruses are unable to be cultured by current culture techniques. In our study, fecal samples from 10 children under five years of age with unexplained AGE and 10 healthy children were investigated for RNA viruses using random priming (RP)-mediated sequence-independent single primer amplification (SISPA). The causative viruses in cases of cryptogenic diarrhea were then assessed for their potential diagnostic value. Of the 1,129 viral clones identified, rotavirus was most commonly associated with AGE (125 sequences, 22.4%). In contrast, bacteriophage was most common (43 sequences, 13.6%) in healthy children. The remaining 515 viral clones were unidentifiable. These findings suggest that investigation of cases or outbreaks of unexplained diarrhea using a metaviromic strategy is a new avenue for diagnosis.

Key Words: Acute gastroenteritis, Metavirome, SISPA, Rotavirus, Parechovirus

INTRODUCTION

Traditionally, the identification of viruses has been dependent on the ability to culture the viruses in host cells in order to propagate and isolate enough pure virions for characterization. However, the large majority of viruses, including enteric viruses, cannot be cultivated using standard techniques (1). The only way to obtain adequate viral particles for characterization of these enteric viruses has been to feed volunteer human adults stool filtrates derived from patients with the disease (2). Therefore, culture-based methods are insufficient for large scale characterization of a viral community in the human gastrointestinal tract. In addition, viruses do not have ubiquitously conserved genetic elements, like the bacterial ribosomal RNA gene, that can be used to quantify diversity and evolutionary distance amongst different viruses (3).

Recently, metaviromic analyses have emerged, and they offer an opportunity to directly characterize mixed genomes of uncultured microorganisms from a variety of ecological communities (4). In the case of viruses, metaviromic analysis has focused on searching for viral contents such as DNA viruses or bacteriophages in the environment (5), animals (6), insects (7), fermented foods (8), human oral cavities (9), and feces (3). There has been very little investigation of the composition of human viral communities in both diseased and healthy patients.

Received: November 4, 2013, 2013/ Revised: November 13, 2013/ Accepted: November 15, 2013
*Corresponding author: Prof. Wonyong Kim. Department of Microbiology, Chung-Ang University College of Medicine, 221 Heukseok-dong, Dongjak-ku, Seoul 156-756, Korea.
Phone: +82-2-820-5685, Fax: +82-2-822-5685, e-mail: kimwy@cau.ac.kr
**This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) (2013R1A1A2A10012148) and the Brain Korea 21 PLUS Program funded by the Ministry of Education.

Acute gastroenteritis (AGE) is a diarrheal disease of rapid onset, with or without accompanying symptoms such as nausea, vomiting, fever, and abdominal pain. Many viral pathogens, such as rotavirus, norovirus, astrovirus, calicivirus, hepatitis E virus, coronavirus, torovirus, and enteric adenovirus serotypes 40 and 41 are associated with AGE, and most of them (except adenovirus) have an RNA genome (10, 11). In this study, we assessed the potential application of metaviromic techniques to the detection and diagnosis of viral causes of unexplained AGE.

MATERIALS AND METHODS

Specimen collection

Fecal samples from children under five years of age with unexplained watery diarrhea were collected at the Department of Pediatrics, Chung-Ang University Hospital, South Korea, between April 2010 and January 2011. Fecal samples from healthy children under five years of age were collected at the Hansarang nursery in Incheon. All participants provided written informed consent. All fecal samples were analyzed using protocols approved by the Chung-Ang University College of Medicine Institutional Review Board (Protocol #2009-13).

Stool specimens

Fecal samples from 10 patients with unexplained AGE were obtained and prediagnosed as negative for rotavirus, adenovirus, astrovirus, norovirus, and other enteropathogenic bacterial infections by latex agglutination (LAT), enzyme immunoassay (EIA), and immunochromatographic methods (ICG) and RT-PCR analyses. Fecal samples from 10 healthy infants were collected and used as references for comparison by metaviromic analysis. Immediately after collection, all 20 fecal samples were transported to the Department of Microbiology, Chung-Ang University College of Medicine from the university hospital. All samples were diluted ten-fold with phosphate buffered saline (PBS; pH 7.4) and clarified via centrifugation at 10,000 g for 10 minutes. The supernatants were stored at -20℃ before use.

RNA extraction

Viral RNA was prepared following the guidelines detailed in Djikeng *et al*. (12) with some modifications. The PBS-diluted stool supernatant was filtered using 0.45 μm and 0.2 μm sterile syringe filters (Corning Costar, Corning, NY, USA) to remove bacteria, eukaryotes, and other large particles. The filtrates were then centrifuged at 12,000 g for 10 minutes using centrifugal filters (Pall, Centrifugal devices, 100 kDa, USA) to concentrate viral particles. The viral suspensions were treated with 50 units of DNase I and RNase A and incubated at 37℃ for one hour. RNA extraction was performed with a Viral RNA extraction Kit (QIAamp® Viral RNA Mini Kit; QIAGEN, Westburg, Germany) according to the manufacturer's instructions.

RP-SISPA and sequencing

The extracted RNA was then processed for random priming-mediated sequence-independent single primer amplification (RP-SISPA) (11). A 10 μl RNA sample was added to FR26RV-N primer (5'-GCCGGAGCTCTGCAG-ATATCNNNNNN-3') at a concentration of 1 μM. Before cDNA synthesis, the sample was subjected to 1 cycle of denaturation at 94℃ for five minutes. Double stranded cDNA was synthesized by a Klenow reaction (New England Biolabs, Ipswich, MA, USA) in the presence of random hexamer primers (Promega, Madison, WI USA). PCR amplification used Taq DNA polymerase with an FR20RV primer (5'-GCCGGAGCTCTGCAGATATC-3'), followed by 40 cycles of PCR (94℃ for 1 minute, 65℃ for 1 minute, and 72℃ for 2 minutes), and a final 10 minute incubation at 72℃.

Resultant analysis

The PCR amplicons were inserted into a pCR 2.1 cloning vector, introduced into *E. coli* TOP 10F' (Invitrogen, Carlsbad, CA, USA) by transformation, and sequenced using a BigDye terminator cycle sequencing kit and automatic DNA sequencer (Model 3730; Applied Biosystems). M13 forward and reverse primers were used for sequence analysis. The resultant gene and deduced amino acid

sequences were compared to the GenBank database using BLAST (BLASTn/BLASTp/BLASTx).

RESULTS

Clinical features

Amongst the 10 children with AGE of unknown origin, two patients were male and eight were female, with a mean age of 2.8 years (range 0 to 7 years). The patients with AGE were negative for rotavirus by latex agglutination (LAT) (Slidex Rota-kit 2, bioMerieux Vitek, Mercy-I'Etoile, France) or enzyme immunoassay (EIA) (VIDAS Rotavirus, bioMerieux Vitek), and adenovirus, astrovirus, and norovirus by reverse-transcription PCR. The patients with AGE of unknown origin exhibited no bacterial (*Salmonella* or *Shigella*) growth on stool culture. The main clinical features in the children with AGE were watery diarrhea (eight cases, mean duration of 4.1 days) with underlying disease, fever (eight cases, mean temperature of 38.5℃), vomiting (seven cases, mean duration of two days) and abdominal pain (five cases). Greater than two symptoms were present in nine (90%) patients. Complete blood cell counts demonstrated leukocytosis in four cases with a mean elevated white blood cell count of 12,782/mm^3, and leukopenia in one case with a white blood cell count of 2,820/mm^3. Liver function tests showed an elevated serum aspartate aminotransferase (AST) in three cases with a mean elevated AST of 62 IU/l, and a normal serum alanine aminotransferase (ALT) in all cases with an overall mean ALT of 22 IU/l. Serum c-reactive protein (CRP) levels were normal in all but one case with an overall mean value of 0.21 mg/dl. Three patients had an elevated serum erythrocyte sedimentation rate (ESR) with a mean elevated ESR of 50 mm.

Sequence data overview

A total of 1,129 clonal sequences (average length of 332.6 bases) were generated from the viral particle-enriched nucleic acids from all patients. These sequences were taxonomically assigned based on their best BLAST score (e-value, 0.001) related to sequences in GenBank, a non-redundant genomic database. Summaries of the classification of viral proteins in the human gut are shown in Fig. 1A and 1B. The most abundant viral sequences in healthy and hospitalized children were eukaryotic viruses, which accounted for 14 sequences (2.4%) and 177 sequences (31.8%), respectively. The most abundant overall sequences were bacterial sequences, which accounted for 257 sequences (44.9%) in healthy children and 84 sequences (15.1%) in hospitalized children. Bacteriophage sequences accounted for 43 sequences (7.5%) in healthy children and 28 sequences (5.0%) in hospitalized children, and eukaryotic, non-viral sequences accounted for 5 sequences (0.9%) in healthy children and 10 sequences (1.8%) in hospitalized children. Overall, 257 sequences (44.9%) from healthy children and 258 sequences (46.3%) from hospitalized children had no significant similarity to any sequences in GenBank (e-value, 0.001).

Bacteriophages in the pediatric gut

The sequences with similarities to bacteriophages made up 7.5% of all sequences in healthy children and 5.0% of sequences in hospitalized children. In healthy children, the 43 bacteriophage sequences were composed of 22 uncultured Microviridae phage sequences (3.9%, n=5), 14 Chlamydia phage sequences (2.5%, n=3), four Bdellovibrio phage sequences (0.7%, n=1), one Spiroplasma phage sequence (0.2%, n=1), one Clostridium phage sequence (0.2%, n=1), and one Trichomonas phage sequence (0.2%, n=1) (Table 1). In the hospitalized children, the bacteriophage sequences were composed of nine uncultured Microviridae phage sequences (1.6%, n=1), six Chlamydia phage sequences (1.0%, n=2), five Lactococcus phage sequences (0.9%, n=1), two Enterococcus phage sequences (0.4%, n=1), two Streptococcus phage sequences (0.4%, n=1), two Bacillus phage sequences (0.4%, n=1), one Ajellomyces phage sequence (0.2%, n=1), and one Bdellovibrio phage sequence (0.2%, n=1) (Table 1).

Eukaryotic viruses in the pediatric gut

The families of the eukaryotic viruses that were found in the pediatric fecal samples are shown in Fig. 1 and Table 1. RNA viruses capable of infecting eukaryotes were a minority

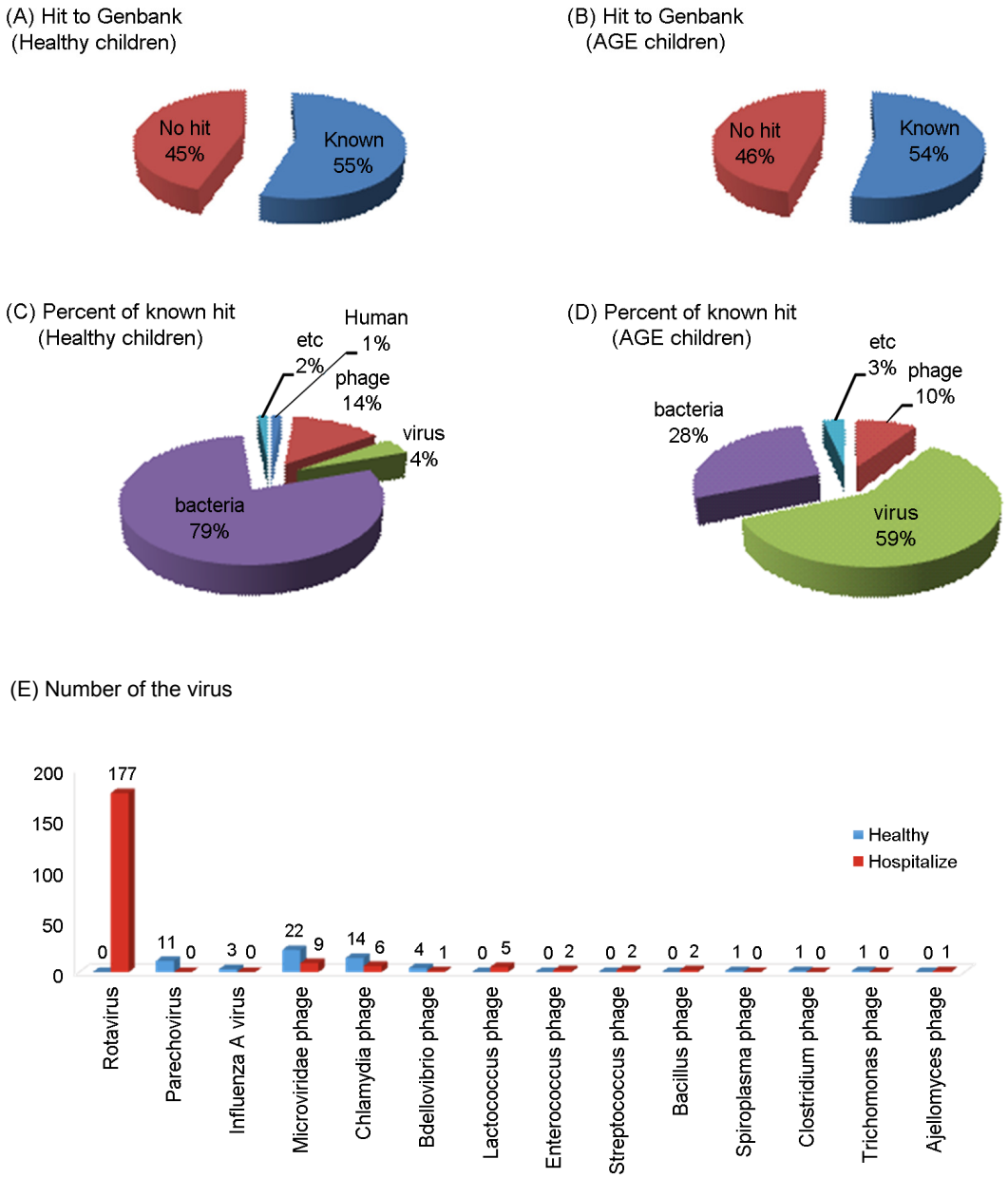

Figure 1. Genomic overview of the human fecal RNA viruses based on BLAST sequence similarities. Numbers of sequences with significant matches (e values of 0.001) in GenBank (A and C). Distribution of significant matches among known sequences of biological entities (B and D). Distribution of the number of viruses detected in both healthy and hospitalized children (E).

of all gut sequences, accounting for 31.8% of sequences in hospitalized children and 2.4% of sequences in healthy children. The RNA viruses capable of eukaryotic infection were primarily double-stranded RNA (dsRNA) viruses from the Reoviridae family. All proteins encoded by the dsRNA, eukaryotic virus-like sequences were identical to viral proteins found in hospitalized infants. The viral sequences related to the rotavirus strains known to infect humans may have been the causative agents of the AGE seen in children less than five years old. These potentially causative viral sequences included rotavirus-associated sequences (177 sequences, 31.7%). In contrast to hospitalized children, the

Table 1. Genes observed among the virus and phage sequences based on GenBank annotation

Group	Organism	Virus	No. of tested	No. of sequences
Patients	Animal virus	Rotavirus	5	177
	Bacteriophage	Uncultured Microviridae	2	9
		Chlamydia	2	6
		Lactococcus	1	5
		Enterococcus	1	2
		Streptococcus	1	2
		Bacillus	1	2
		Ajellomyces	1	1
		Bdellovibrio	1	1
Normal	Animal virus	Parechovirus	2	11
		Influenza A virus	1	3
	Bacteriophage	Uncultured Microviridae	5	22
		Chlamydia	3	14
		Bdellovibrio	1	4
		Spiroplasma	1	1
		Clostridium	1	1
		Trichomonas	1	1

fecal samples from healthy children exhibited sequences from single-stranded RNA (ssRNA) viruses, such as parechovirus (11 sequences, 1.6%) and influenza A virus (3 sequences, 0.5%). These ssRNA sequences encoded proteins that were identical to viral proteins identified in healthy children, and the ssRNA sequences were not found in hospitalized children.

DISCUSSION

It is difficult to detect and identify viral agents causing AGE because they rapidly mutate their genes (13). The vertebrate gut flora contains about 10^{14} microorganisms, but many of these microbes are difficult to isolate using current culture techniques. The vertebrate gut flora also contains bacteriophages and viruses, including known human enteric viruses and uncharacterized viruses (14). These viruses are critically involved in the regulation of the overall gut microbial community (15). Gut microbial genetic diversity is shaped by virus-mediated gene transfer, and a full knowledge of the range of potential hosts for gut viruses has been limited by difficulty in culturing potential viral hosts.

Recently, new techniques for studying microbial communities, collectively called metagenomics, have been developed to overcome the limitations of culturing in identifying gut viruses' marked genetic richness. These metagenomic techniques use culture-independent sequencing to identify microorganisms (16). In our study, many previously characterized and highly divergent eukaryotic viral sequences were detected in pediatric guts. Only 0.9~1.8% of the sequences in our analysis were non-bacterial and non-viral in origin. These non-bacterial and non-viral sequences were identified in 15 patients, five healthy and 10 with AGE. This indicated that the DNase and RNase treatments were effective in removing non-capsid-protected human host nucleic acids. Our data also revealed viral RNA genomic sequences including rotavirus, human influenza A virus, and many bacteriophages.

Despite the use of well established procedures for identifying causes of AGE in this country, many patients still stuffer from AGE of unknown etiology. In outpatient clinics, LAT, EIA, and ICG are widely used to diagnose rotavirus. However, previous studies have shown that many cases of rotavirus infection can be identified by RT-PCR of stool samples that were labeled rotavirus-negative by current testing methods (17). Indeed, one recent study showed that for the identification of rotavirus infection, RT-PCR was more sensitive and specific than other enzyme-based diagnostic methods, although RT-PCR is slower and more labor intensive (18). Based on the recently described superiority of RT-PCR testing, re-examination of stool specimens by RT-PCR for the presence of rotavirus will be worthwhile to determine the causative agent of AGE in humans (19).

Rotaviruses were found in four patients with AGE who were prediagnosed as negative for current rotavirus detection

methods such as LAT, EIA or ICG. This finding can be explained by the less-than-ideal sensitivity of the pre-diagnostic methods in detecting rotavirus. The typical limit of detection by EIA based on the VP6 group-specific antigen is >1 × 10^5 particles/ml (20). In contrast, the limit of detection by RT-PCR of viral RNA is >100 particles/ml (21) or 1.46 plaque forming units (18). An alternative explanation for the false negative rotavirus prediagnosis in these patients may have been VP6 antigenic drift in a rotavirus population in this country.

In this study, parechoviruses were detected in two healthy children. The pathogenecity of parechoviruses is currently debated, but they have been reported in a variety of disease syndromes, including respiratory and gastrointestinal illnesses, aseptic meningitis, myocarditis, encephalitis, acute flaccid paralysis and neonatal sepsis (22). However, recent reports suggest that parechoviruses are also found in healthy children (23). Our observations are consistent with the latter reports because we identified parechoviruses in fecal samples from two healthy children but no parechoviruses in children with AGE. Human influenza virus A was also detected in fecal samples from three healthy children. The influenza virus is known to cause disease when it infects the respiratory tract (9). In addition, recent studies have also reported gastrointestinal infection (24). Further research is needed to examine the effects of influenza infection in the gastrointestinal tract and to examine the usefulness of stool influenza detection in diagnosing cases of respiratory influenza.

Prior studies have shown that bacteriophage viruses constitute a significant fraction of human and equine fecal viral populations (25). In our study, the bacteriophage-associated amino acid sequences we identified had no significant similarities with any microorganismal sequences in GenBank. Unidentified sequences constituted 44.9% of all sequences in healthy children and 46.3% of all sequences in hospitalized children. This data is similar to the percentage of unclassified sequences identified in a previous metagenomic study of human stool (26). Furthermore, in our study bacteriophage sequences accounted for 5.0~7.5% of the total fecal sequences. In previous studies of fecal samples from South Asian children with non-polio acute flaccid paralysis and healthy contacts, bacteriophage sequences have accounted for 16% and 12% of total sequences, respectively (27). The predominant bacteriophage sequences we found were from uncultured Microviridae in healthy children, and uncultured Microviridae and Siphoviridae in hospitalized children. This data showed that in the presence of infective enteric viruses, the native bacteriophage community was altered.

Our study of fecal RNA viruses in children with AGE as well as healthy children revealed new information about unusual RNA viruses that can occur in children with AGE. Moreover, the metaviromic approach we used provides new potential for the discovery of unknown viruses that may be causing idiopathic infectious diseases. To fully develop a blueprint for the prevention and treatment of AGE, further studies are needed to understand the relationships between gut DNA and RNA viral microbiota and the relationships between gut viral and bacterial microbial communities.

REFERENCES

1) Chang KO, George DW, Patton JB, Green KY, Sosnovtsev SV. Leader of the capsid protein in feline calicivirus promotes replication of Norwalk virus in cell culture. J Virol 2008;82:9306-17.
2) Shim JO, Baek IH, Le VP, Ko EM, Seok WS, Uh Y, *et al*. Molecular characterization of rotavirus diarrhea among children in South Korea: detection of an unusual G11 strain. Arch Virol 2011;156:887-92.
3) Zhang T, Breitbart M, Lee WH, Run JQ, Wei CL, Soh SW, *et al*. RNA viral community in human feces: prevalence of plant pathogenic viruses. PLoS Biol 2006; 4:e3.
4) Kim W. Application of metagenomic techniques: understanding the unrevealed human microbiota and explaining in the clinical infectious diseases. J Bacteriol Virol 2012;42:263-75.
5) Bench SR, Hanson TE, Williamson KE, Ghosh D, Radosovich M, Wang K, *et al*. Metagenomic characterization of Chesapeake Bay virioplankton. Appl Environ Microbiol 2007;73:7629-41.

6) Day JM, Ballard LL, Duke MV, Scheffler BE, Zsak L. Metagenomic analysis of the turkey gut RNA virus community. Virol J 2010;7:313.
7) Ng TF, Willner DL, Lim YW, Schmieder R, Chau B, Nilsson C, *et al*. Broad Surveys of DNA Viral Diversity Obtained through Viral Metagenomics of Mosquitoes. PLoS One 2011;6:e20579.
8) Park EJ, Kim KH, Abell GC, Kim MS, Roh SW, Bae JW. Metagenomic analysis of the viral communities in fermented foods. Appl Environ Microbiol 2011;77: 1284-91.
9) Greninger AL, Chen EC, Sittler T, Scheinerman A, Roubinian N, Yu G, *et al*. A metagenomic analysis of pandemic influenza A (2009 H1N1) infection in patients from North America. PLoS One 2010;5:e13381.
10) Ramani S, Kang G. Viruses causing childhood diarrhoea in the developing world. Curr Opin Infect Dis 2009; 22:477-82.
11) Lee JI, Park SH, Kim MS, Oh YH, Yu IS, Choi BH, *et al*. Surveillance of Acute Gastroenteritis in Seoul, Korea, During May 2004 and June 2007. J Bacteriol Virol 2009;39:363-71.
12) Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, *et al*. Viral genome sequencing by random priming methods. BMC Genomics 2008;9:5.
13) Wei CJ, Boyington JC, Dai K, Houser KV, Pearce MB, Kong WP, *et al*. Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design. Sci Transl Med 2010;2:24ra1.
14) Nicholson JK, Holmes E, Wilson ID. Gut micro-organisms, mammalian metabolism and personalized health care. Nat Rev Microbiol 2005;3:431-8.
15) Xu J, Gordon JI. Honor thy symbionts. Proc Natl Acad Sci U S A 2003;100:10452-9.
16) Streit WR, Schmitz RA. Metagenomics--the key to the uncultured microbes. Curr Opin Microbiol 2004;7:492 -8.
17) Kittigul L, Pombubpa K, Taweekate Y, Diraphat P, Sujirarat D, Khamrin P, *et al*. Norovirus GII-4 2006b variant circulating in patients with acute gastroenteritis in Thailand during a 2006-2007 study. J Med Virol 2010;82:854-60.
18) Kittigul L, Ekchaloemkiet S, Utrarachkij F, Siripanichgon K, Sujirarat D, Pungchitton S, *et al*. An efficient virus concentration method and RT-nested PCR for detection of rotaviruses in environmental water samples. J Virol Methods 2005;124:117-22.
19) Harada S, Okada M, Yahiro S, Nishimura K, Matsuo S, Miyasaka J, *et al*. Surveillance of pathogens in outpatients with gastroenteritis and characterization of sapovirus strains between 2002 and 2007 in Kumamoto Prefecture, Japan. J Med Virol 2009;81:1117-27.
20) r-biopharm. RIDASCREEN Rotavirus Enzyme immunoassay for the detection of Rotavirus. R-Biopharm AG, Darmstadt. 2004:C 0901.
21) Chinsangaram J, Akita GY, Castro AE, Osburn BI. PCR detection of group A bovine rotaviruses in feces. J Vet Diagn Invest 1993;5:516-21.
22) Nix WA, Maher K, Pallansch MA, Oberste MS. Parechovirus typing in clinical specimens by nested or semi-nested PCR coupled with sequencing. J Clin Virol 2010;48:202-7.
23) Li L, Victoria J, Kapoor A, Naeem A, Shaukat S, Sharif S, *et al*. Genomic characterization of novel human parechovirus type. Emerg Infect Dis 2009;15:288-91.
24) Shu Y, Li CK, Li Z, Gao R, Liang Q, Zhang Y, *et al*. Avian influenza A(H5N1) viruses can directly infect and replicate in human gut tissues. J Infect Dis 2010; 201:1173-7.
25) Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, *et al*. Viral diversity and dynamics in an infant gut. Res Microbiol 2008;159:367-73.
26) Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, *et al*. The human gut virome: inter-individual variation and dynamic response to diet. Genome Res 2011;21:1616-25.
27) Victoria JG, Kapoor A, Li L, Blinkova O, Slikas B, Wang C, *et al*. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. J Virol 2009;83:4642-51.