DNA Microarray for Detection of Gastrointestinal Viruses

DNA microarray for detection of gastrointestinal viruses

Miguel A. Martínez[1], María de los Dolores Soto-del Río[1*], Rosa María Gutiérrez[2], Charles Y. Chiu[3,4], Alexander L. Greninger[5], Juan Francisco Contreras[6], Susana López[1], Carlos F. Arias[1], Pavel Isa[1#].

1. Departamento de Genética del Desarrollo y Fisiología Molecular, Instituto de Biotecnología UNAM, Cuernavaca, Morelos, Mexico, 2. Departamento de Microbiología Molecular, Instituto de Biotecnología UNAM, Cuernavaca, Morelos, Mexico, 3. Department of Laboratory Medicine, University of California San Francisco, San Francisco, CA, 4. UCSF-Abbott Viral Diagnostics and Discovery Center, University of California San Francisco, CA, 5. Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, CA, 6. Departamento de Microbiología e Inmunología, Universidad Autónoma de Nuevo León, Monterrey, Nuevo León, Mexico.

* Present address: Department of Animal Pathology, University of Turin, Grugliasco, Turin, Italy

Corresponding author: PhD Pavel Isa, pavel@ibt.unam.mx Instituto de Biotecnología, UNAM. Avenida Universidad 2001, Col Chamilpa, Cuernavaca, Morelos 62210, Mexico.

Phone: (52) (777) 329 1612. Fax (52) (777) 317 2388

Key words: gastroenteritis, virus, diagnostic, microarrays

Running title:
DNA microarray for detection of viruses found in stools

ABSTRACT

Gastroenteritis is a clinical illness of humans and other animals characterized by vomiting and diarrhea, caused by a variety of pathogens including viruses. An increasing number of viral species have been associated with gastroenteritis or have been found in stool samples as new molecular tools are developed. In this work, a DNA microarray capable in theory of parallel detection of more than 100 viral species was developed and tested. Initial validation was done with 10 different virus species and an additional 5 species were validated using clinical samples. Detection limits of $1x10^3$ virus particles of *human adenovirus C* (HAdV), *human astrovirus* (HAstV) and group A *rotavirus* (RV-A) were established. Furthermore, when exogenous RNA was added, the limit for RV-A detection decreased in one log. In a small group of clinical samples from children with gastroenteritis (n = 76), the microarray detected at least one viral species in 92% of the samples. Single infection was identified in 63 samples (83%) and more than one virus was identified in 7 samples (9%). The most abundant virus species were RV-A (58%), followed by *anellovirus* (15.8%), HAstV (6.6%), HAdV (5.3%), *Norwalk virus* (6.6%), *human enterovirus* (HEV; 9.2%), *human parechovirus* (1.3%), *Sapporo virus* (1.3%) and *human bocavirus* (1.3%). To further test the specificity and sensitivity of the microarray, the results were verified by RT-PCR detection of 5 gastrointestinal viruses. The RT- PCR assay detected a virus in 59 samples (78%). The microarray showed good performance for detection of RV-A, HAstV and calicivirus, while the sensitivity for HAdV and HEV was low. Furthermore some discrepancies in detection of mixed infections were observed, and were addressed by RT-qPCR of the viruses involved. It was observed that differences in the amount of genetic material favored the detection of the most abundant virus. The microarray described in this work should help to understand the etiology of gastroenteritis in humans and animals.

INTRODUCTION

Gastroenteritis stands among the five principal causes of mortality by disease and morbidity at all ages worldwide. The most affected population is children under 5 years of age, where it accounts for the second cause of post-neonatal death with approximately 2.6 millions deceased per year [1]. Although the majority of deaths occur in developing countries, diarrheal disease is among the most common causes of illness worldwide, with approximately 4,620 millions episodes annually [1]. Besides humans, all vertebrate species suffer from enteric diseases. Infections in farm animals can lead to big economical losses, while household pets, like dogs and cats, are also affected. On the other hand, wild animals, like deer, monkeys, bats, foxes, wolves, and boars, among others, can act as potential reservoirs for pathogens [2]. Gastrointestinal (GI) infections are caused by a variety of pathogens including parasites, bacteria and viruses. The characterization of pathogens causing GI infections of viral etiology have led to the establishment of a main group of pathogens [*rotavirus A* (RV-A), *Norwalk virus* (NV), *human astrovirus* (HAstV), and *human adenovirus F* (HAdV-F)] [3] for which specific diagnostic tests were developed [4]. Tests for secondary or rare virus are available but usually restricted to experimental use. Routine diagnostic methods for viral gastroenteritis are nowadays based on the detection of virus components by immunoassays or by molecular methods [5,6,7,8], with the majority of these tests designed to evaluate only a single pathogen at a time.

The use of two or more specific primer sets (multiplexing) in polymerase chain reaction (PCR) allows the amplification of several targets in one test. Although multiplex tests are available for diverse viruses [9,10,11,12,13], facilitating rapid and sensitive detection of the main GI agents, these assays are still limited in number of viruses detected, and the results can be affected by mutations at primer binding sites. On the other hand, DNA microarrays represent an alternative to detect hundreds to thousands of potential pathogens in a single assay. Microarray detection is based on solid phase hybridization, in which specific probes are deposited on a surface and react with a mixture of labeled

nucleic acids. So far, different microarrays have been developed to detect causative infectious agents associated with a number of diseases: respiratory [14,15,16], hemorrhagic [17], blood borne [18,19], and central nervous system syndromes [20]. Other broad microarrays have been developed for virus discovery [21], however, diagnostic microarray specific for viruses found in GI tract is missing. Given the recent rise in the number of new viral species [22,23,24,25,26], diagnostic DNA microarrays represent the possibility to test their clinical importance and impact in human and animal health.

In this work, the development and validation of a DNA microarray designed to detect in principle more than 100 viral species associated with GI tract in vertebrates is presented. This microarray was successfully used to identify viruses in a small set of gastroenteritis clinical samples.

MATERIALS AND METHODS

Cells, viruses, and clinical samples

Viruses were either present in our laboratory or kindly provided by different partner laboratories (Table 1). Clinical samples from children presenting gastroenteritis during the winter season 2004-2005 were obtained in Monterrey, Mexico, with the written consent of parent or guardian. Analysis of human clinical samples was approved by the Bioethics Committee of the Instituto de Biotecnologia. The initial screening of samples for RV-A was performed in Monterrey by silver staining of RV-A segmented double-stranded RNAs separated by SDS-PAGE electrophoresis**.** No previous screening for bacterial or parasitic agents was performed on group of samples. Triple layered particles of RV-A strain RRV were purified by cesium chloride density gradient as described previously [27].

Microarray probe design

All virus species that have been either associated with gastroenteritis or found in the gastrointestinal tract were identified by an extensive review of published literature, and selected to be

included in microarray. All available full-length genomes or complete gene sequences of the selected virus species were obtained from GenBank (up to February 2009) and the proper databases were created. For each virus species, sequence redundancy was eliminated according to a sequence similarity with cutoff values 95-99% using CD-HIT software [28]. One sequence for each species was selected as source for probes production and was processed as described [29]. Specifically, sequences were consecutively split into 70-mers with a shifting window of 3 nucleotides, with each 70-mer corresponding a potential probe. The 70-mer length probes have sufficient size to allow for stringent hybridization conditions while allowing for certain degree of mismatches, but are small enough to maintain species specificity [30,31,32]. Target probes were selected to be included in the microarray by analysis of BLAST results and calculation of hybridization thermodynamics (ΔG) calculated by the nearest-neighbor method [33]. To be considered good candidates for microarray, the ΔG was required to be at least -70 kcal/mol for homologous sequence and higher than -40 kcal/mol for heterologous sequences. A minimum of 6 non-overlapping probes from conserved regions in virus genomes were selected for each virus, and each available genome sequence in target database for given species was recognized by at least two probes. When necessary, due to variability within a species, two or more source sequences were chosen and each single sequence was processed as described above.

Microarray probes *in-silico* analysis

Hybridization thermodynamics of RV-A selected probes was evaluated *in-silico* with VP1, VP2 and NSP5 segments of RV-A strains representing all full-genome G and P genotypes available. Hybridization ΔG (kcal/mol) between probe and target was calculated by nearest-neighbor method. Best probe-target ΔG was plotted in heatmap using R. Detection of a target is considered with ΔG < -50 kcal/mol.

Microarray production

Selected 70-mer probes were synthesized by Illumina Oligator (Illumina Inc, CA, USA). Oligonucleotides were resuspended to 400 pmol in 3X SSC buffer (0.45 M NaCl, 45 mM sodium citrate pH 7.0), and spotted onto epoxide-coated glass slides in the Microarray Facility of the Prostate Centre at Vancouver General Hospital, Vancouver, British Columbia, Canada. Each spot contained one specific probe to detect one virus species, and 4 pmol of spike70 (a 70-mer without biological complementary known sequence) [34] used to precisely identify probe spot locations on the microarray. Slides were maintained in humidity-free chamber until their use.

Nucleic acid extraction, amplification and labeling

Genetic material from virus lysates (cell culture supernatants from reference strains) was extracted by PureLink Viral RNA/DNA kit according to the manufacturer's instructions (Invitrogen, USA). For clinical samples, Norwalk and Sapporo virus positive controls, 100 μg of stool was added to conical screw cap tubes containing 100 mg of 150-212 μm glass beads (Sigma, USA), chloroform (100 μl) and PBS up to 1ml. Samples were homogenized in a bead beater (Biospec Products, USA). After 10 min centrifugation at 2000 x g supernatants were recovered and filtered in Spin-X 22 μm pore filters (Costar, NY) at 5000 x g during 10-20 min. Filtered samples were treated with Turbo DNAse (Ambion, USA) and RNAse (Sigma, USA) for 30 min at 37°C and immediately chilled on ice. Nucleic acids were then extracted from 200 μl using PureLink Viral RNA/DNA kit according to manufacturer's instructions (Invitrogen, USA). Nucleic acids from virus lysates or clinical samples were eluted in nuclease-free water, aliquoted, quantified in NanoDrop ND-1000 (NanoDrop Technologies, DE) and stored at -70°C until further use.

Sample processing and random-amplification of nucleic acids was performed essentially as described previously [21,35,36]. Briefly, reverse transcription was done using SuperScript III Reverse

Transcriptase (Invitrogen, USA) and primer-A (5'-GTTTCCCAGTAGGTCTCN$_9$-3'). The complementary DNA (cDNA) strand was generated by two rounds of synthesis with Sequenase 2.0 (USB, USA). The obtained cDNA was then amplified with KlenTaq polymerase (Sigma, USA) or Taq Polymerase (New England Biolabs, USA) using the primer-B (5'-GTTTCCCAGTAGGTCTC-3') by 30 cycles of the following program: 30 sec at 94°C, 1 min at 50°C, 1 min at 72°C. As a last step, nucleotide-analogue aminoallyl-dUTP (TriLink, USA) in a 7:3 ratio with dTTP was incorporated during an additional 20 cycles of PCR using the same conditions described above and 5 µl of product from the previous PCR as starting material. The amplified products were purified with DNA Clean & Concentrator-5 kit (Zymo Research, USA). Coupling reactions of sample DNA with Cy3 and probe-70 (70-mer complementary to spike-70) with Cy5 dyes (GE HealthCare, USA) were done as described elsewhere [31]. Fluorophore-labeled DNA was purified by the Zymo DNA Clean & Concentrator-5 kit and label incorporation was quantified by NanoDrop.

Slides preparation, hybridization and scanning

Microarray slides were treated just before their use with an ethanolamine wash solution (50 mM ethanolamine, 0.1 % SDS, 0.1M Tris pH 9) for 15 min at 50°C, followed by two washes in distilled water, and they were then dried by centrifugation for 5 min at 500 rpm. Processed slides were loaded with 30 µl of a combination of Cy3 and Cy5 labeled DNA in 3xSSC buffer, and the hybridization was left to proceed in a sealed chamber submerged in a water bath at 65°C for 8-12 h. After incubation the slides were washed consecutively in the following solutions: 2xSSC (65°C), 2xSSC, 1xSSC, 0.2xSSC and dried for 5 min at 500 rpm. Hybridization images were acquired with an Axon GenePix 4000B scanner (Molecular Devices, USA) synchronized with GenePix Pro 6.0 software to detect and measure spot intensities.

Data Analysis

Hybridization spot intensities were first filtered by the following spot-quality control parameters: spot size and shape (denoted as good/bad/absent), channel 532 foreground (F532) signal saturation (% F532 saturated < 5), and F532 signal proportion over channel 532 background (B532) signal [(% > B532 + 2 standard deviations) > 50%]. Spots showing good quality were used to generate microarray level background values. Normalization of intensity values was done with the formula: (F532i/F532m)-(B532i-B5532m) where F532i and B532i stands for foreground and background signal of spot "i" respectively, and F532m and B532m stands for the sum of all foreground or background spots respectively.

The statistical significance of probe intensities in the reference samples was obtained by the rank products algorithm [37] using a minimum of three technical replicates. Rank values from negative control samples were recorded and used to generate a 'spot rank value' included in subsequent spot-quality analysis. For clinical samples, z-score transformed intensities and their p-values were analyzed with fdr tool package [38] in R [39]. Positive virus species were defined as having at least two probes with p-values < 0.05 and a false discovery rate < 0.01.

Limits of detection assays

In order to determine the amount of virus particles detectable by the microarray, three reference viruses with different genome types were assayed: dsRNA RV-A, positive single stranded RNA (ssRNA+) HAstV and double stranded DNA (dsDNA) HadV-C. RNA was extracted from purified RV-A strain RRV, and MA104 cells. The RV-A genome molecular mass was calculated according to the following formula: [genome length (bp) x 325] / 6.022×10^{23} [40]. Decreasing dilutions of RV-A RNA corresponding from 1×10^{8} to 10 particles; were analyzed alone or mixed with an excess of MA104 cells RNA (50 ng). Similarly, decreasing dilutions of focus forming units titrated cell lysates of HAstV or

196 HAdV-C, corresponding from $1x10^7$ to 100 virus particles were extracted, amplified, labeled, and
197 processed using the full microarray protocol as described above.
198
199 **Conventional diagnostic or confirmatory RT-PCR**
200 Nucleic acids extracted from clinical samples were used to perform diagnostic RT-PCR using
201 Qiagen's One-Step RT-PCR kit (Qiagen, USA) or Super Script III One-step RT-PCR with Platinum
202 Taq (Invitrogen, USA). For confirmatory RT-PCR, cDNA was generated with SuperScript III Reverse
203 Transcriptase (Invitrogen, USA), and Taq Polymerase (New England Biolabs) was used for PCR
204 reactions following the manufacturer's instructions. Oligonucleotide primers used in diagnostic or
205 confirmatory RT-PCR are listed in Supplementary Table 1. PCR reactions for RV-A detection
206 included a 5 min boiling followed by immediate ice-chilling step just before RT-PCR. Amplification
207 conditions for RV-A , HAstV and calicivirus (CV) were: 30 min at 50°C, 15 min at 95°C; 40 cycles of
208 30 sec at 95°C, 30 sec at 50°C, 1 min at 72°C, with final extension of 5 min at 72°C. RT-PCR
209 conditions for human adenovirus (HAdV) were: 30 min at 50°C, 15 min at 95°C; 40 cycles of 30 sec at
210 95°C, 30 sec at 55°C, 1 min at 72°C, with final extension for 5 min at 72°C. Human enterovirus (HEV)
211 amplification program was: 30 min at 50°C, 15 min at 95°C; 40 cycles of 30 sec at 95°C, 30 sec at
212 50°C, 30 sec at 72°C, with final extension for 5 min at 72°C. *Human parechovirus* (HPeV)
213 amplification was: 30 min at 50°C, 15 min at 95°C; 35 cycles of 1 min at 95°C, 1 min at 48°C, 1 min at
214 72°C, with final extension for 5 min at 72°C. *Anellovirus* (TTV) confirmation was performed as semi-
215 nested PCR, conditions for the first round were: 2 min at 94°C; 35 cycles of 30 sec at 94°C, 30 sec at
216 55°C, 30 sec at 72 °C, with final extension for 5 min at 72°C. The second round used the same program
217 but only with 30 cycles. *Human bocavirus* (HBoV) was detected by Seeplex RV15 OneStep ACE
218 Detection (Seegene, USA). PCR products were visualized in 2.0% agarose gels except for HEV, which
219 required 3.5% gels due to small amplicon size.
220

Semi quantitative RT-PCR/PCR detection of viruses

One-step real-time RT-PCR and real-time PCR were performed using primers targeting conserved genomic regions (Supplementary Table 1). RV-A detection required previous sample boiling for 5 minutes and immediate ice-chill. For the RNA viruses (RV-A, HAstV, NV and HEV), detection was performed as a two-step process. First, 3 µl of RNA (5 ng) were reverse transcribed with 0.125 µl (50 U/µl)_SuperScript III Reverse Transcriptase (Invitrogen, USA), 0.25 µl of RNAse inhibitor (20 U/µl), 12.5 µl of SYBR Green Master Mix 2X (Applied Biosystems, USA), 1 µl of the primer and DEPC treated water in a 24 µl final volume. Samples were incubated for 30 min at 48°C followed by enzyme inactivation for 10 min at 90 °C. In the second step 1 µl of second primer was added and PCR conditions were carried out as follows: HAstV and RV-A amplification program consisted of 10 min at 95°C, and 40 cycles of 15 sec at 95°C and 1 min at 60°C. NV amplification program was 5 min at 95°C, 45 cycles of 10 sec at 95°C, 20 sec at 48°C, 45 sec at 60°C. HEV program was 10 min at 95°C, 45 cycles of 20 sec at 95°C, 20 sec at 55°C, 1 min at 72°C and final extension of 5 min at 72°C. In the case of HEV, both specific primers were added before PCR, since the RT step was performed using random hexamers. HAdV amplification reactions consisted in 3 µl (5 ng) of DNA, 12.5 µl of SYBR Green Master Mix 2X and 1µl of each corresponding primer in a 25 µl volume reaction. Conditions were 95°C for 8 min, 45 cycles of 30 sec at 95°C, 20 sec at 55°C, 20 sec at 72°C, final extension of 5 min at 72°C. Amplifications were carried out in an ABI Prism 7500 Sequence Detector System (Applied Biosystems). Dissociation curves were evaluated for non-specific products. Ct values corresponding to detection of specific virus sequences were obtained from triplicates of selected samples presenting co-infections and compared for the viruses detected. PCR primer sets for detection of CV, HAdV and HEV were designed to recognize the target at the genus level [5,6,41].

RESULTS

Selection of viruses related to gastrointestinal infections

An advantage of the microarray technology is the capacity to test hundreds, and even thousands of targets in a single assay. The main goal of this study was to develop an assay for detection of all viruses that have been found in stool samples from vertebrate, associated or not with gastroenteritis, what should facilitate clinical and epidemiological studies in humans and animals. A deep search of the scientific literature available in public databases resulted in a list of 128 viral species of viruses reported to be present in the gastrointestinal tract, representing 55 genera that belonged to 17 viral families (Supplementary Table 2). The list of virus species includes the well known human gastroenteritis viruses (calicivirus group, rotaviruses, human astroviruses, enteric adenoviruses), together with some recently described human viruses [*human adenovirus G* [23], *human bocavirus* [42], *cosavirus* [24], *Saffold virus* [43] and *salivirus A* [25,44]]. Classical, non-human gastrointestinal viruses (coronavirus, circovirus, and pestiviruses), and other new discovered viral agents (at the time of the microarray design) from different animal species, such as animal anelloviruses [45,46], bat astroviruses [47], and bovine kobuviruses [48], whose participation as pathogens is not well understood, are also included in the microarray. Thus, the virus species of interest comprehended a variety of viruses with different characteristics, such as RNA and DNA genomes, enveloped/non-enveloped virions, segmented or non-segmented genomes, single- or double-stranded genomes. All available complete gene or genome sequences were retrieved from a public database (GenBank) and were organized in a taxonomic hierarchical database following the ICTV classification at the date the microarray probes were design (ICTV, 2009) or, for novel species, as suggested by the discoverer.

Probe selection and microarray validation

A set of 1,256 70-mer microarray probes were selected from conserved regions and designed to identify 128 viral species associated with the GI tract, with at least 6 probes designed for each viral species and at least 2 probes corresponding to each sequenced viral genomes. To maintain stringent

experimental conditions (hybridization at 65^{o}C) while allowing certain number of sequence variability, the probes were designed as 70-mers. The highest number of probes covered RV-A (42 probes), *alphacoronavirus* (28 probes), and mammalian *orthoreovirus* (25 probes) (Supplementary Tables 2 and 3). For some viruses, the design of a complete set of 6 oligonucleotides was not possible due to the lack of enough complete sequences, nevertheless available probes were included for each viral species.

Reference strains for 10 viral species were available for probe validation. These species represent 6 viral families and include 4 main human pathogens (HAstV, NV, SV, RV-A), other human viruses (mammalian *orthoreovirus* , HAdV-C, *dengue virus 4*), and three non-human viruses (*feline calicivirus, bovine viral diarrhea virus 1,* and *bovine parainfluenza virus 3*) (Table 1). All reference strains tested were detected as expected, including four different RV-A strains (human strain Wa, simian strain RRV, porcine virus TFR-41, and bovine strain UK) and two different mammalian *orthoreovirus* strains (T1L and T3D) (Table 1). To test *in-silico* capacity of probes to recognize different and variable strains, 42 probes specific for rotavirus were analyzed with panel of all available G and P genotypes (Supplementary Figure 1). The only genotype that the microarray would not probably detect was G22P[35] belonging to a turkey rotavirus strain.

Sensitivity and specificity of the assay

To determine the sensitivity limits of the DNA microarray, the virus genetic material was extracted from lysates of HAstV- or HAdV-C-infected cells or from CsCl-purified simian strain RRV particles. In a series of cell lysate dilutions (corresponding from 10^2 to 10^7 viral particles), the microarray was able to detect as low as 10^3 HAdV-C or HAstV virus particles. Similarly, RV-A RNA (corresponding from 10 to 10^8 viral particles) was amplified before or after addition of a constant amount of cellular RNA (50 ng). In the absence of cellular RNA the detection limit of viral RNA was 10^3 genome copies, however, when the complexity of the sample was augmented by adding cellular RNA, the detection limit was one logarithm lower, detecting 10^4 genome copies.

To evaluate the probe specificity, a rank products algorithm [37] was applied to the results obtained from technical replicates of reference viruses and mock-infected cell controls (MA104 cells, A549 cells, and C6/36 cells). Based on the false discovery rate test (FDR) included in the software, 16 probes were identified as presenting nonspecific behavior (marked with * in Supplementary Table 3). When analyzed, these nonspecific probes did not show any common feature, although some presented high GC content (>70%). In the following experiments the results obtained with these probes were excluded from analysis.

Analysis of clinical samples

To further test the capacity of the microarray to detect viruses, 76 samples from children under 5 years of age, collected during the winter season 2004-2005 in Monterrey, Mexico, was analyzed. The collection of samples was originally screened for RV-A by polyacrylamide gel electrophoresis (SDS-PAGE), and stored at -70°C. Using the microarray developed in this study, a viral agent was detected in 70 out of 76 (92 %) samples tested; a single virus was found in 63 (83%) samples, while two or more viral species were detected in 7 (9%) samples (Fig. 1). Among the viruses detected, the most common was RV-A (44 samples), followed by TTV (12 positives), HEV (7), caliciviruses (6 positives – 5 NV and 1 SV), HAstV (5), HAdV (4 positive samples- 3 HAdV-F and 1 HAdV-A), HPeV (2) and one HBoV (Fig. 1). It is important to mention that only 6 (8%) samples remained negative after microarray detection, and that not all viruses found are known to be pathogenic. As mentioned above, after collection all samples were screened for the presence of RV-A by SDS-PAGE. Additionally, as described below, all samples tested with the microarray were tested for selected viruses by diagnostic RT-PCR, including RV-A. In thirty-four samples RV-A was identified by the three methods tested; 5 additional samples were found positive by microarray and RT-PCR tests (Fig. 2). Another 8 were found positive either by microarray (n=5) or by RT-PCR (n=3) (Fig 2). Notably, the 3 samples that were only positive for RV-A by RT-PCR were found in mixed-infection samples.

To compare the results of the microarray with those of a routine diagnostic method for viral gastroenteritis, RT-PCR detection for a panel of 5 viruses [RV-A, HAstV, HAdV, CV (NV and SV) and HEV] was performed in all clinical samples. It is important to point out that the primer sets for HAdV, CV, and HEV are designed to recognize its target at the genus level [5,6,41].

The RT-PCR panel detected at least one virus in 59 samples (78%; Fig. 1B), a lower detection rate when compared to the DNA microarray when analyzing only these 5 viruses (n= 65, 85%). At the individual virus level, the RT-PCR panel confirmed the microarray results in all HAdV-positive samples (1 HAdV-A and 3 HAdV-F), having a positive predictive value (PPV) of 100%, all CV (5 NV and 1 SV) (PPV, 100%) , and 39 of 44 RV-A positive samples (PPV, 89%), while PPVs were lower for HAstV, with 3 of 5 positive samples identified (PPV 60%), and 5 of 7 HEV positive samples identified by microarray (PPV 71%) (Fig. 3).

Detection of viruses in mixed infections

The RT-PCR screening resulted in the identification of 16 mixed infections (MI), while the microarray identified only 7 MI (Fig. 1). The microarray detected up to 4 different viruses within one sample, with TTV found in all MI samples. The following viral combinations were found by microarray: HEV B/TTV (3 samples); and one sample of each NV/TTV; HEV-B/HAstV/TTV; RV-A/HPeV/TTV; and SV/HEV-B/HPeV/TTV (Fig. 1). Of interest, *human parechovirus* and *Sapporo virus* were detected only in co-infection. The MI combinations observed in RT-PCR were: RV-A/HAdV (8), RV-A/HEV (5), HAstV/HEV (1), RV-A/CV (1), and one triple infection HAdV/CV/HEV (Fig. 1B). Examining these 16 samples, we observed that RV-A was the only virus identified by microarray in all samples with RV-A/HAdV co-infection (n= 8), and in 4 out of 5 RV-A/HEV samples, while HAstV was the only virus identified in sample with HAstV/HEV co-infection (Table 2). In one sample NV was identified as the sole species by microarray, while RT-PCR results showed CV/HAdV/HEV triple co-infection (Table 2). Thus, in all of these 16 samples a single virus

345 was identified by the microarray, while at least two viral species were detected by RT-PCR.

346 One possible explanation for the discrepancies in the identification of mixed infections using
347 microarrays and RT-PCR could be the variability in the relative amount of each virus genetic material
348 in clinical samples, as it has been observed that individuals infected with some viruses, for example
349 RV-A and NV, can shed large amount of viral particles in the acute stage of infection [49,50,51]. To
350 explore this possibility, the amount of viral genetic material in selected samples with mixed infection
351 was quantified by real-time RT-PCR. The use of equal quantities of starting material allowed us to
352 compare directly the amplification Ct's of two viruses within a sample. The results showed that the
353 single virus detected by microarray had, in most cases, a lower Ct value as compared to the second
354 virus detected by qRT-PCR, with the only exception being the combination of RV-A/HEV, where RV-
355 A was the only virus identified by microarray despite the fact that HEV had lower Ct values (Table 2).
356 This indicates that MIs presenting large differences in the amount of the genetic material of the viral
357 agents involved are prone to result in single virus detection by the microarray, generally, the one
358 present more abundantly.

359 As consequence, when comparing the sensitivity and specificity of he microarray, as compared
360 with the panel of individual diagnostic RT-PCR, the most prevalent or most frequently found viruses in
361 single infections, such as RV-A, HAstV, and CV, showed good sensitivity and specificity (from 85-
362 100%), while the sensitivity for viruses like HAdV and HEV was low, ranging from 30 to 42%, being
363 clearly affected by other viruses present in the sample (Fig. 3, Table 2). For example, 4 samples that
364 presented only HAdV were found positive by both microarray and RT-PCR, while in the remaining 9
365 samples, which presented HAdV co-infection with RV-A (8 samples) and one with CV, only the
366 second virus was identified by microarray (Table 2). It should be pointed out that most of these
367 samples contained a low level of HAdV genetic material, with Ct values close to non-template control
368 value (Ct 44.5) (Table 2).

369

Detection of not common GI viruses

Of note, the microarray found 3 viruses that usually are not evaluated in gastroenteritis samples. Two samples presented HPeV, both in co-infection (one with RV-A/ TTV, another with SV/ HEV B/ TTV). An additional sample containing HBoV was identified (RV-A was identified by RT-PCR in this sample), and 12 samples presented TTV, 5 samples as single infection and 7 in co-infection with other viruses. As reference samples for these viruses were not available, confirmation RT-PCR coupled with capillary sequencing were performed, and the viruses detected by the microarray were confirmed in all these samples (results not shown). Fact that TTV single virus positive samples were found is not an indicator of causativeness.

DISCUSSION

Current routine viral testing is designed to detect only the most prevalent viruses, leaving frequently a 30-50% of cases without an agent identified [52]. In recent years, advances in molecular biology and the implementation of next generation sequencing has allowed the identification of several new viruses in intestinal samples [53,54,55,56,57]. The role of most of these viruses in diarrheal disease remains unclear (*Aichi virus*, *anellovirus*, *human bocavirus*, *human parechovirus*, *human picobirnavirus*, and some enteroviruses, among others) raising the need to study in detail their epidemiology. In order to gather information on GI virus diversity, proper tools are required for their monitoring. In this work a comprehensive and sensitive DNA microarray was developed and tested, which allows in principle the parallel detection of more than 100 gastrointestinal associated virus species.

Implementation of the microarray for detection of viruses is not an easy task. Design of probes and experimental conditions are two important parameters to consider. Resequencing microarrays permit identification of mutations but require high numbers of probes for a single agent, increasing its cost [58]. Arrays for sub-typing use less and shorter probes but are often designed for only one viral

species [59,60,61,62,63]. Microarrays used for virus discovery have proven very useful when usual suspects are discarded or in rare disease cases, but identification is not clear and require complex analysis [34].

Several DNA microarrays have been previously reported for identification of the main known gastrointestinal pathogenic viruses [59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75], however, they were mostly oriented to the identification or sub-typification of one viral species, and none had specifically addressed the list of viruses that can be found in stool samples.

The microarray platform described in this work has been validated with 14 reference viral strains, representing 10 different virus species. Importantly, 5 other viral species were identified using the microarray when analyzing clinical samples: HAdV-F, HAdV-A, HPeV , HBoV, and several TTV. The capacity of the microarray to identify correctly viruses whose probes were not validated in this work with cultured reference strains confirms that the methodology used to design probes is adequate, and increases the probability that the remaining probes will be also capable to identify their target viruses, this is additionally supported by *in-silico* detection of wide variety of RV-A strains using probes obtained from conserved genes; however, testing with other reference strains would be necessary. During the validation experiments some probes were found to react non-specifically with the amplified labeled DNA, regardless of its origin, in other words, they were found to be 'sticky', and they were excluded from further analysis. No common characteristic was found between these probes that could account for their non-specific binding behavior.

One of the critical parameters in virus detection is the sensitivity of the assay. There are several factors that can affect the sensitivity. In the case of a microarray, sample nucleic acids are generally processed by random primed amplification prior to hybridization, to ensure amplification of a wide variety of viruses. The product of random-PCR could be lower than that of specific PCR, decreasing the sensitivity of the assay as all genetic material is amplified, diluting the positive signal [76]. Limit of detection for three viruses with different genome types (dsDNA, dsRNA and ssRNA+) was established

420 in 10^3 virus particles, suggesting that genome nature does not affect the sensitivity of the assay.
421 Moreover, testing the sensitivity of the microarray with purified RV-A RNA, we observed that addition
422 of 50 ng of cellular RNA as a non-specific diluting RNA, decreased the sensitivity of detection ten-
423 fold. To try to solve the sensitivity problem in complex clinical samples, agent-specific primers have
424 been included in previous reports, together with random primers during amplification of the genetic
425 material [14,15], with the disadvantage of narrowing the scope of targets for the microarray assay.
426 We subsequently analyzed a group of clinical samples collected from children with diarrhea.
427 Initially, the clinical samples were screened by SDS-PAGE, which led to the identification of 34 RV-A
428 positive samples, while the microarray presented in this work identified 44, suggesting that the
429 microarray platform has a higher sensitivity than traditional methods. A similar sensitivity was
430 obtained by RT-PCR, as 42 samples were found RV-A-positive. Even though our results indicate that
431 the limit of detection of purified virus ($1x10^3$ viral particles) is similar to that reached with PCR assays
432 [8], the microarray had a higher number of positive results when clinical samples were tested, possibly
433 due to the natural genetic variation in primer binding regions of viruses found in sample viruses.
434 Although multiplexed assays are being developed, their use in routine testing is not generally
435 implemented, and most studies use single pathogen tests. When RT-PCR screening for the most
436 common viruses is performed, the percentage of clinical samples without a virus identified remains
437 around 30-50% [13,77,78], while the microarray presented in this study detected a virus in 92% of the
438 samples. This high detection rate could have been influenced by the time of sampling, since winter is a
439 high season for viral gastroenteritis in the region and no pre selection for pathogens was performed. An
440 additional advantage of the microarray test as compared to a set of different RT-PCR assays, is the
441 capacity to identify viruses that are not commonly tested for, like those previously associated with
442 diarrhea (like HPeV) and those of unclear clinical significance in GI disease (HBoV and TTV). In this
443 work we found a wide range of circulating not typical viruses among children, similarly as observed in
444 other studies [79,80] and their continuous surveillance should be considered. To our knowledge, this is

445 first report of description of HPeV, HBoV, and TTV in Mexican children.

446 As a consequence of the limited number of virus species routinely tested, prevalence of co-
447 infections is a poorly explored issue. Usually, when a panel of up to 5 viruses is used, co-infection rates
448 between 4-18% are observed, being the most common combination RV-A/NV [2,13,77,81,82,83].
449 More recently, wide scope ranging metagenomic studies have shown that mixed infections are more
450 common than previously thought [4,80], even in healthy individuals [79]. The analysis of the small set
451 of clinical samples analyzed in this work showed that 30% (23 out of 76) contained more than 1
452 gastrointestinal virus. The identification of individual viruses in co-infections presented some
453 discrepancies when comparing the results from microarray and RT-PCR tests. Of seven samples with
454 mixed infections identified by the microarray, five were confirmed by RT-PCR, while in 16 mixed
455 infection identified by RT-PCR, a single virus was identified by the microarray, suggesting that the
456 microarray may be less sensitive than RT-PCR for detection of mixed infections. To address this
457 inconsistency, real time RT-PCR was implemented for the principal combinations of viruses that were
458 missed by the microarray. This platform showed certain advantage for detection of RV-A over HAdV
459 and HEV, as RV-A was identified even when the HEV genome was present in larger amounts. HEV
460 was identified by microarray in samples co-infected with RV-A only when RV-A RNA was present in
461 low amounts, close to negative control levels (Table 3). Preferential identification of RV-A by the
462 microarray could be due to the large amount of virus particles excreted during the acute phase of
463 infection, and by the large number of probes selected (42 oligonucleotides), as compared to 5 and 17
464 probes for HEV A and HEV C, respectively, and 17 probes for HAdV. On the other hand, the two
465 HEV samples positive for microarray that were missed by RT-PCR correspond to mixed infections,
466 with HAstV/TTV and SV/HpeV/TTV, respectively. Despite several attempts to identify HEV in these
467 samples by RT-PCR resulted in negative results, thus the possibility of a microarray false positive
468 result cannot be discarded.

469 The number of virus species identified has increased considerably in the last decade with the

application of emerging genomic technologies such as microarrays and unbiased next-generation sequencing in studies of fatal or rare cases of disease in humans, wild, and domestic animals [25,56,84,85,86]. Adequate tools that allow detection of well-known pathogenic viruses while are capable of detecting the new or rare viruses in a single assay will contribute with useful epidemiological information about both kinds of viruses. This microarray includes viruses of different host origin in order to extend the range of use to veterinary studies. The oligonucleotide probes selected should allow the identification of target viruses despite the sequence variations that will occur in the following years, however, it will be important to update the microarray design on a regular basis to maintain the capacity to broadly detect pathogenic viruses and to include newly found viral species.

Parallel detection of gastroenterical viruses beyond the most common viruses should facilitate a better understanding of virus etiology, as it increases the rate of positive cases, closing the diagnostic gap, and allows inspection for mixed infections where secondary viral agents could represent an important factor. Adding data from case control studies and inclusion of other host parameters, as serological data, will help to provide evidence of virus pathogenicity. Furthermore, adequate and comprehensive epidemiological studies in wild and domestic animals should be considered.

Acknowledgments

The computational analysis was performed using the cluster of the Instituto de Biotecnología, UNAM, with the assistance of M.C. Jerome Verleyen. Authors thank Paul Gaytán Colín, Eugenio López-Bustos, and Santiago Becerra Ramírez from Unidad de Síntesis y Secuenciación de DNA, Instituto de Biotecnología, UNAM, for synthesis of the oligonucleotides used in RT-PCR assays. This work was supported by grant S0008-111593 from CONACYT. Miguel A. Martínez and María de los Dolores Soto-del Río were supported by a scholarship from CONACYT-Mexico.

REFERENCES

1. WHO. 2011. World Health Statistics 2011. Geneva: World Health Organization.
2. Simpson, V.R. 2002. Wild animals as reservoirs of infectious diseases in the UK. Veterinary journal 163: 128-146.
3. Wilhelmi, I., Roman, E., and Sanchez-Fauquier, A. 2003. Viruses causing gastroenteritis. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases 9: 247-262.
4. Rovida, F., Campanini, G., Sarasini, A., Adzasehoun, K.M., Piralla, A., and Baldanti, F. 2013. Comparison of immunologic and molecular assays for the diagnosis of gastrointestinal viral infections. Diagnostic microbiology and infectious disease 75: 110-111.
5. Allard, A., Albinsson, B., and Wadell, G. 2001. Rapid typing of human adenoviruses by a general PCR combined with restriction endonuclease analysis. Journal of clinical microbiology 39: 498-505.
6. Farkas, T., Zhong, W.M., Jing, Y., Huang, P.W., Espinosa, S.M., Martinez, N., Morrow, A.L., Ruiz-Palacios, G.M., Pickering, L.K., and Jiang, X. 2004. Genetic diversity among sapoviruses. Archives of virology 149: 1309-1323.
7. Royuela, E., Negredo, A., and Sanchez-Fauquier, A. 2006. Development of a one step real-time RT-PCR method for sensitive detection of human astrovirus. Journal of virological methods 133: 14-19.
8. Schwarz, B.A., Bange, R., Vahlenkamp, T.W., Johne, R., and Muller, H. 2002. Detection and quantitation of group A rotaviruses by competitive and real-time reverse transcription-polymerase chain reaction. Journal of virological methods 105: 277-285.
9. Coupland, L.J., McElarney, I., Meader, E., Cowley, K., Alcock, L., Naunton, J., and Gray, J. 2013. Simultaneous detection of viral and bacterial enteric pathogens using the Seeplex(R) Diarrhea ACE detection system. Epidemiology and infection 141: 2111-2121.
10. Khamrin, P., Okame, M., Thongprachum, A., Nantachit, N., Nishimura, S., Okitsu, S., Maneekarn, N., and Ushijima, H. 2011. A single-tube multiplex PCR for rapid detection in feces of 10 viruses causing diarrhea. Journal of virological methods 173: 390-393.
11. Rohayem, J., Berger, S., Juretzek, T., Herchenroder, O., Mogel, M., Poppe, M., Henker, J., and Rethwilm, A. 2004. A simple and rapid single-step multiplex RT-PCR to detect Norovirus, Astrovirus and Adenovirus in clinical stool samples. Journal of virological methods 118: 49-59.
12. Svraka, S., van der Veer, B., Duizer, E., Dekkers, J., Koopmans, M., and Vennema, H. 2009. Novel approach for detection of enteric viruses to enable syndrome surveillance of acute viral gastroenteritis. Journal of clinical microbiology 47: 1674-1679.
13. van Maarseveen, N.M., Wessels, E., de Brouwer, C.S., Vossen, A.C., and Claas, E.C. 2010. Diagnosis of viral gastroenteritis by simultaneous detection of Adenovirus group F, Astrovirus, Rotavirus group A, Norovirus genogroups I and II, and Sapovirus in two internally controlled multiplex real-time PCR assays. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology 49: 205-210.
14. Palacios, G., Quan, P.L., Jabado, O.J., Conlan, S., Hirschberg, D.L., Liu, Y., Zhai, J., Renwick, N., Hui, J., Hegyi, H., Grolla, A., Strong, J.E., Towner, J.S., Geisbert, T.W., Jahrling, P.B., Buchen-Osmond, C., Ellerbrok, H., Sanchez-Seco, M.P., Lussier, Y., Formenty, P., Nichol, M.S., Feldmann, H., Briese, T., and Lipkin, W.I. 2007. Panmicrobial oligonucleotide array for diagnosis of infectious diseases. Emerging infectious diseases 13: 73-81.
15. Quan, P.L., Palacios, G., Jabado, O.J., Conlan, S., Hirschberg, D.L., Pozo, F., Jack, P.J., Cisterna, D., Renwick, N., Hui, J., Drysdale, A., Amos-Ritchie, R., Baumeister, E., Savy, V., Langer, K.M., Rucht, J.A., Boyle, D.B., Garcia-Sastre, A., Casas, I., Perez-Brena, P., Briese, T., and

Lipkin, W.I. 2007. Detection of respiratory viruses and subtype identification of influenza A viruses by GreeneChipResp oligonucleotide microarray. Journal of clinical microbiology 45: 2359-2364.
16. Sengupta, S., Onodera, K., Lai, A., and Melcher, U. 2003. Molecular detection and identification of influenza viruses by oligonucleotide microarray hybridization. Journal of clinical microbiology 41: 4542-4550.
17. Xiao-Ping, K., Yong-Qiang, L., Qing-Ge, S., Hong, L., Qing-Yu, Z., and Yin-Hui, Y. 2009. Development of a consensus microarray method for identification of some highly pathogenic viruses. Journal of medical virology 81: 1945-1950.
18. Striebel, H.M., Birch-Hirschfeld, E., Egerer, R., Foldes-Papp, Z., Tilz, G.P., and Stelzner, A. 2004.Enhancing sensitivity of human herpes virus diagnosis with DNA microarrays using dendrimers. Experimental and molecular pathology 77: 89-97.
19. Zhaohui, S., Wenling, Z., Bao, Z., Rong, S., and Wenli, M. 2004. Microarrays for the detection of HBV and HDV. Journal of biochemistry and molecular biology 37: 546-551.
20. Boriskin, Y.S., Rice, P.S., Stabler, R.A., Hinds, J., Al-Ghusein, H., Vass, K., and Butcher, P.D. 2004. DNA microarrays for virus detection in cases of central nervous system infection. Journal of clinical microbiology 42: 5811-5818.
21. Wang, D., Coscoy, L., Zylberberg, M., Avila, P.C., Boushey, H.A., Ganem, D., and DeRisi, J.L. 2002.Microarray-based detection and genotyping of viral pathogens. Proceedings of the National Academy of Sciences of the United States of America 99: 15687-15692.
22. Holtz, L.R., Finkbeiner, S.R., Zhao, G., Kirkwood, C.D., Girones, R., Pipas, J.M., and Wang, D. 2009. Klassevirus 1, a previously undescribed member of the family Picornaviridae, is globally widespread. Virology journal 6: 86.
23. Jones, M.S., 2nd, Harrach, B., Ganac, R.D., Gozum, M.M., Dela Cruz, W.P., Riedel, B., Pan, C., Delwart, E.L., and Schnurr, D.P. 2007. New adenovirus species found in a patient presenting with gastroenteritis. Journal of virology 81: 5978-5984.
24. Kapoor, A., Victoria, J., Simmonds, P., Wang, C., Shafer, R.W., Nims, R., Nielsen, O., and Delwart, E. 2008. A highly divergent picornavirus in a marine mammal. Journal of virology 82: 311-320.
25. Li, L., Victoria, J., Kapoor, A., Blinkova, O., Wang, C., Babrzadeh, F., Mason, C.J., Pandey, P., Triki, H., Bahri, O., Oderinde, B.S., Baba, M.M., Bukbuk, D.N., Besses, J.M., Bartkus, J.M, and Delwart, E.L. 2009. A novel picornavirus associated with gastroenteritis. Journal of virology 83: 12002-12006.
26. Rockett, R.J., Sloots, T.P., Bowes, S., O'Neill, N., Ye, S., Robson, J., Whiley, D.M., Lambert, S.B., Wang, D., Nissen, M.D., and Bialasiewicz, S. 2013. Detection of novel polyomaviruses, TSPyV, HPyV6, HPyV7, HPyV9 and MWPyV in feces, urine, blood, respiratory swabs and cerebrospinal fluid. PloS one 8: e62764.
27. Zarate, S., Espinosa, R., Romero, P., Guerrero, C.A., Arias, C.F., and Lopez, S. 2000. Integrin alpha2beta1 mediates the cell attachment of the rotavirus neuraminidase-resistant variant nar3. Virology 278: 50-54.
28. Li, W., and Godzik, A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658-1659.
29. Paulin, L.F., Soto-Del Rio Mde, L., Sanchez, I., Hernandez, J., Gutierrez-Rios, R.M., Lopez-Martinez, I., Wong-Chew, R.M., Parissi-Crivelli, A., Isa, P., Lopez, S., and Arias, C.F. 2014. PhyloFlu, a DNA Microarray for Determining the Phylogenetic Origin of Influenza A Virus Gene Segments and the Genomic Fingerprint of Viral Strains. Journal of clinical microbiology 52: 803-813.
30. He, Z., Wu, L., Fields, M.W., and Zhou, J. 2005. Use of microarrays with different probe sizes for monitoring gene expression. Applied and environmental microbiology 71: 5154-5162.

31. Hughes, T.R., Mao, M., Jones, A.R., Burchard, J., Marton, M.J., Shannon, K.W., Lefkowitz, S.M., Ziman, M., Schelter, J.M., Meyer, M.R., Kobayashi, S., Davis, C., Dai, H., He, Y.D., Stephaniants, S.B., Cavet, G., Walker, W.L., West, A., Cofrey, E., Shoemaker, D.D., Stoughton, R., Blanchard, A.P., Friend, S.H., and Linsey, P.S. 2001. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. Nature biotechnology 19: 342-347.
32. Wang, H.Y., Malek, R.L., Kwitek, A.E., Greene, A.S., Luu, T.V., Behbahani, B., Frank, B., Quackenbush, J., and Lee, N.H. 2003. Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. Genome biology 4: R5.
33. SantaLucia, J., Jr. 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proceedings of the National Academy of Sciences of the United States of America 95: 1460-1465.
34. Urisman, A., Fischer, K.F., Chiu, C.Y., Kistler, A.L., Beck, S., Wang, D., and DeRisi, J.L. 2005. E-Predict: a computational strategy for species identification based on observed DNA microarray hybridization patterns. Genome biology 6: R78.
35. Bohlander, S.K., Espinosa, R., 3rd, Le Beau, M.M., Rowley, J.D., and Diaz, M.O. 1992. A method for the rapid sequence-independent amplification of microdissected chromosomal material. Genomics 13: 1322-1324.
36. Chen, E.C., Miller, S.A., DeRisi, J.L., and Chiu, C.Y. 2011. Using a pan-viral microarray assay (Virochip) to screen clinical samples for viral pathogens. Journal of visualized experiments : JoVE.
37. Breitling, R., Armengaud, P., Amtmann, A., and Herzyk, P. 2004. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS letters 573: 83-92.
38. Strimmer, K. 2008. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. Bioinformatics 24: 1461-1462.
39. Team, R.t.C. 2013. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing
40. Teo, J., Di Pietro, P., San Biagio, F., Capozzoli, M., Deng, Y.M., Barr, I., Caldwell, N., Ong, K.L., Sato, M., Tan, R., and Lin, R. 2011. VereFlu: an integrated multiplex RT-PCR and microarray assay for rapid detection and identification of human influenza A and B viruses using lab-on-chip technology. Archives of virology 156: 1371-1378.
41. Rotbart, H.A. 1990. Enzymatic RNA amplification of the enteroviruses. Journal of clinical microbiology 28: 438-442.
42. Allander, T., Tammi, M.T., Eriksson, M., Bjerkner, A., Tiveljung-Lindell, A., and Andersson, B. 2005. Cloning of a human parvovirus by molecular screening of respiratory tract samples. Proceedings of the National Academy of Sciences of the United States of America 102: 12891-12896.
43. Nielsen, A.C., Gyhrs, M.L., Nielsen, L.P., Pedersen, C., and Bottiger, B. 2013. Gastroenteritis and the novel picornaviruses aichi virus, cosavirus, saffold virus, and salivirus in young children. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology 57: 239-242.
44. Greninger, A.L., Runckel, C., Chiu, C.Y., Haggerty, T., Parsonnet, J., Ganem, D., and DeRisi, J.L. 2009. The complete genome of klassevirus - a novel picornavirus in pediatric stool. Virology journal 6: 82.
45. Biagini, P., Uch, R., Belhouchet, M., Attoui, H., Cantaloube, J.F., Brisbarre, N., and de Micco, P. 2007. Circular genomes related to anelloviruses identified in human and animal samples by using a combined rolling-circle amplification/sequence-independent single primer amplification approach. The Journal of general virology 88: 2696-2701.

641 46. Okamoto, H., Takahashi, M., Nishizawa, T., Tawara, A., Fukai, K., Muramatsu, U., Naito, Y., and
642 Yoshikawa, A. 2002. Genomic characterization of TT viruses (TTVs) in pigs, cats and dogs and
643 their relatedness with species-specific TTVs in primates and tupaias. The Journal of general
644 virology 83: 1291-1297.
645 47. Chu, D.K., Poon, L.L., Guan, Y., and Peiris, J.S. 2008. Novel astroviruses in insectivorous bats.
646 Journal of virology 82: 9107-9114.
647 48. Khamrin, P., Maneekarn, N., Peerakome, S., Okitsu, S., Mizuguchi, M., and Ushijima, H. 2008.
648 Bovine kobuviruses from cattle with diarrhea. Emerging infectious diseases 14: 985-986.
649 49. Atmar, R.L., Opekun, A.R., Gilger, M.A., Estes, M.K., Crawford, S.E., Neill, F.H., and Graham,
650 D.Y. 2008. Norwalk virus shedding after experimental human infection. Emerging infectious
651 diseases 14: 1553-1557.
652 50. Lee, N., Chan, M.C., Wong, B., Choi, K.W., Sin, W., Lui, G., Chan, P.K., Lai, R.W., Cockram,
653 C.S., Sung, J.J., and Leung, W.K. 2007. Fecal viral concentration and diarrhea in norovirus
654 gastroenteritis. Emerging infectious diseases 13: 1399-1401.
655 51. Ramani, S., Sankaran, P., Arumugam, R., Sarkar, R., Banerjee, I., Mohanty, I., Jana, A.K.,
656 Kuruvilla, K.A., and Kang, G. 2010. Comparison of viral load and duration of virus shedding
657 in symptomatic and asymptomatic neonatal rotavirus infections. Journal of medical virology 82:
658 1803-1807.
659 52. Dennehy, P.H. 2011. Viral gastroenteritis in children. The Pediatric infectious disease journal 30:
660 63-64.
661 53. Breitbart, M., Hewson, I., Felts, B., Mahaffy, J.M., Nulton, J., Salamon, P., and Rohwer, F. 2003.
662 Metagenomic analyses of an uncultured viral community from human feces. Journal of
663 bacteriology 185: 6220-6223.
664 54. Finkbeiner, S.R., Allred, A.F., Tarr, P.I., Klein, E.J., Kirkwood, C.D., and Wang, D. 2008.
665 Metagenomic analysis of human diarrhea: viral detection and discovery. PLoS pathogens 4:
666 e1000011.
667 55. Reyes, A., Haynes, M., Hanson, N., Angly, F.E., Heath, A.C., Rohwer, F., and Gordon, J.I. 2010.
668 Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 466: 334-338.
669 56. Victoria, J.G., Kapoor, A., Li, L., Blinkova, O., Slikas, B., Wang, C., Naeem, A., Zaidi, S., and
670 Delwart, E. 2009. Metagenomic analyses of viruses in stool samples from children with acute
671 flaccid paralysis. Journal of virology 83: 4642-4651.
672 57. Zhang, T., Breitbart, M., Lee, W.H., Run, J.Q., Wei, C.L., Soh, S.W., Hibberd, M.L., Liu, E.T.,
673 Rohwer, F., and Ruan, Y. 2006. RNA viral community in human feces: prevalence of plant
674 pathogenic viruses. PLoS biology 4: e3.
675 58. Lin, B., Blaney, K.M., Malanoski, A.P., Ligler, A.G., Schnur, J.M., Metzgar, D., Russell, K.L., and
676 Stenger, D.A. 2007. Using a resequencing microarray as a multiple respiratory pathogen
677 detection assay. Journal of clinical microbiology 45: 443-452.
678 59. Brinkman, N.E., and Fout, G.S. 2009. Development and evaluation of a generic tag array to detect
679 and genotype noroviruses in water. Journal of virological methods 156: 8-18.
680 60. Brown, D.W., Gunning, K.B., Henry, D.M., Awdeh, Z.L., Brinker, J.P., Tzipori, S., and Herrmann,
681 J.E. 2008. A DNA oligonucleotide microarray for detecting human astrovirus serotypes. Journal
682 of virological methods 147: 86-92.
683 61. Chizhikov, V., Wagner, M., Ivshina, A., Hoshino, Y., Kapikian, A.Z., and Chumakov, K. 2002.
684 Detection and genotyping of human group A rotaviruses by oligonucleotide microarray
685 hybridization. Journal of clinical microbiology 40: 2398-2407.
686 62. Lovmar, L., Fock, C., Espinoza, F., Bucardo, F., Syvanen, A.C., and Bondeson, K. 2003.
687 Microarrays for genotyping human group a rotavirus by multiplex capture and type-specific
688 primer extension. Journal of clinical microbiology 41: 5153-5158.

689 63. Vinje, J., and Koopmans, M.P. 2000. Simultaneous detection and genotyping of "Norwalk-like
690 viruses" by oligonucleotide array in a reverse line blot hybridization format. Journal of clinical
691 microbiology 38: 2595-2601.
692 64. Ayodeji, M., Kulka, M., Jackson, S.A., Patel, I., Mammel, M., Cebula, T.A., and Goswami, B.B.
693 2009. A microarray based approach for the identification of common foodborne viruses. The
694 open virology journal 3: 7-20.
695 65. Chen, H., Mammel, M., Kulka, M., Patel, I., Jackson, S., and Goswami, B.B. 2011. Detection and
696 identification of common food-borne viruses with a tiling microarray. The open virology
697 journal 5: 52-59.
698 66. Chen, Q., Li, J., Deng, Z., Xiong, W., Wang, Q., and Hu, Y.Q. 2010. Comprehensive detection and
699 identification of seven animal coronaviruses and human respiratory coronavirus 229E with a
700 microarray hybridization assay. Intervirology 53: 95-104.
701 67. Honma, S., Chizhikov, V., Santos, N., Tatsumi, M., Timenetsky Mdo, C., Linhares, A.C.,
702 Mascarenhas, J.D., Ushijima, H., Armah, G.E., Gentsch, J.R., and Hoshino, Y. 2007.
703 Development and validation of DNA microarray for genotyping group A rotavirus VP4 (P[4],
704 P[6], P[8], P[9], and P[14]) and VP7 (G1 to G6, G8 to G10, and G12) genes. Journal of clinical
705 microbiology 45: 2641-2648.
706 68. Jaaskelainen, A.J., Piiparinen, H., Lappalainen, M., Koskiniemi, M., and Vaheri, A. 2006.
707 Multiplex-PCR and oligonucleotide microarray for detection of eight different herpesviruses
708 from clinical specimens. Journal of clinical virology : the official publication of the Pan
709 American Society for Clinical Virology 37: 83-90.
710 69. Kim, J.M., Kim, S.Y., Park, Y.B., Kim, H.J., Min, B.S., Cho, J.C., Yang, J.M., Cho, Y.H., and Ko,
711 G. 2012. Simultaneous detection of major enteric viruses using a combimatrix microarray.
712 Journal of microbiology 50: 970-977.
713 70. Leblanc, N., Gantelius, J., Schwenk, J.M., Stahl, K., Blomberg, J., Andersson-Svahn, H., and
714 Belak, S. 2009. Development of a magnetic bead microarray for simultaneous and simple
715 detection of four pestiviruses. Journal of virological methods 155: 1-9.
716 71. Lin, B., Vora, G.J., Thach, D., Walter, E., Metzgar, D., Tibbetts, C., and Stenger, D.A. 2004. Use of
717 oligonucleotide microarrays for rapid detection and serotyping of acute respiratory disease-
718 associated adenoviruses. Journal of clinical microbiology 42: 3232-3239.
719 72. Lopez-Campos, G., Coiras, M., Sanchez-Merino, J.P., Lopez-Huertas, M.R., Spiteri, I., Martin-
720 Sanchez, F., and Perez-Brena, P. 2007. Oligonucleotide microarray design for detection and
721 serotyping of human respiratory adenoviruses by using a virtual amplicon retrieval software.
722 Journal of virological methods 145: 127-136.
723 73. Mattison, K., Corneau, N., Berg, I., Bosch, A., Duizer, E., Gutierrez-Aguirre, I., L'Homme, Y.,
724 Lucero, Y., Luo, Z., Martyres, A., Myrmel, M., O'Ryan, M., Pagotto, F., Sano, D., Svraka, S.,
725 Urzua, U., and Bidawid, S. 2011. Development and validation of a microarray for the
726 confirmation and typing of norovirus RT-PCR products. Journal of virological methods 173:
727 233-250.
728 74. Shih, S.R., Wang, Y.W., Chen, G.W., Chang, L.Y., Lin, T.Y., Tseng, M.C., Chiang, C., Tsao, K.C.,
729 Huang, C.G., Shio, M.R., Tai, J.H., Wang, S.H., Kuo, R.L., and Liu, W.T. 2003. Serotype-
730 specific detection of enterovirus 71 in clinical specimens by DNA microchip array. Journal of
731 virological methods 111: 55-60.
732 75. Susi, P., Hattara, L., Waris, M., Luoma-Aho, T., Siitari, H., Hyypia, T., and Saviranta, P. 2009.
733 Typing of enteroviruses by use of microwell oligonucleotide arrays. Journal of clinical
734 microbiology 47: 1863-1870.
735 76. Vora, G.J., Meador, C.E., Stenger, D.A., and Andreadis, J.D. 2004. Nucleic acid amplification
736 strategies for DNA microarray-based pathogen detection. Applied and environmental
737 microbiology 70: 3047-3054.

77. Pang, X.L., Preiksaitis, J.K., and Lee, B.E. 2013. Enhanced enteric virus detection in sporadic gastroenteritis using a multi-target real-time PCR panel: A one-year study. Journal of medical virology.
78. Simpson, R., Aliyu, S., Iturriza-Gomara, M., Desselberger, U., and Gray, J. 2003. Infantile viral gastroenteritis: on the way to closing the diagnostic gap. Journal of medical virology 70: 258-262.
79. Kapusinszky, B., Minor, P., and Delwart, E. 2012. Nearly constant shedding of diverse enteric viruses by two healthy infants. Journal of clinical microbiology 50: 3427-3434.
80. Mitui, M.T., Bozdayi, G., Ahmed, S., Matsumoto, T., Nishizono, A., and Ahmed, K. 2013. Detection and molecular characterization of diarrhea causing viruses in single and mixed infections in children: A comparative study between Bangladesh and Turkey. Journal of medical virology.
81. Koh, H., Baek, S.Y., Shin, J.I., Chung, K.S., and Jee, Y.M. 2008. Coinfection of viral agents in Korean children with acute watery diarrhea. Journal of Korean medical science 23: 937-940.
82. Roman, E., Wilhelmi, I., Colomina, J., Villar, J., Cilleruelo, M.L., Nebreda, V., Del Alamo, M., and Sanchez-Fauquier, A. 2003. Acute viral gastroenteritis: proportion and clinical relevance of multiple infections in Spanish children. Journal of medical microbiology 52: 435-440.
83. Tran, A., Talmud, D., Lejeune, B., Jovenin, N., Renois, F., Payan, C., Leveque, N., and Andreoletti, L. 2010. Prevalence of rotavirus, adenovirus, norovirus, and astrovirus infections and coinfections among hospitalized children in northern France. Journal of clinical microbiology 48: 1943-1946.
84. Belak, S., Karlsson, O.E., Blomstrom, A.L., Berg, M., and Granberg, F. 2013. New viruses in veterinary medicine, detected by metagenomic approaches. Veterinary microbiology 165: 95-101.
85. Bodewes, R., Kik, M.J., Raj, V.S., Schapendonk, C.M., Haagmans, B.L., Smits, S.L., and Osterhaus, A.D. 2013. Detection of novel divergent arenaviruses in boid snakes with inclusion body disease in The Netherlands. The Journal of general virology 94: 1206-1210.
86. Bodewes, R., van de Bildt, M.W., Schapendonk, C.M., van Leeuwen, M., van Boheemen, S., de Jong, A.A., Osterhaus, A.D., Smits, S.L., and Kuiken, T. 2013. Identification and characterization of a novel adenovirus in the cloacal bursa of gulls. Virology 440: 84-88.

Figure 1. Prevalence of viruses in clinical samples. A group of 76 clinical samples from children presenting gastroenteritis was analyzed by the described microarray (A) or by diagnostic RT-PCR for 5 most common gastrointestinal pathogens (B). Samples with co-infections are shown. NEG is for negative result (no virus identified).

Figure 2. Identification of rotavirus group A. A group of 76 gastroenteritis samples was analyzed by three methods for the presence of rotavirus. These included visualization of rotavirus dsRNA on SDS-PAGE gel, RT-PCR, and the microarray designed in this work. The circles represent numbers of rotavirus positive samples identified by one, two, or three methods used .

Figure 3. Microarray diagnostic sensitivity and specificity. A panel of 5 virus groups [rotavirus group A (RV-A), human astrovirus (HAstV), human adenovirus (HAdV), calicivirus (CV), and enterovirus (HEV)] was tested by RT-PCR in all 76 samples. Results were compared to those obtained by microarray analysis. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of microarray (array), as compared to RT-PCR (PCR), for detection of particular pathogens are shown.

Supplementary Figure 1. *In-silico* hybridization of available RV-A G and P genotypes. Hybridization thermodynamic values (ΔG) were calculated for all 42 RV-A probes (labeled on the bottom) interacting "*in-silico*" with RV-A isolates representing different G and P genotype combinations. RV-A probes belong to VP1, VP2 and NSP5 segments (marked on top). G and P genotype combinations (marked on left) belong to RV-A strains isolated from different hosts (human, cow, dog, cat, simian, rhesus, pig, horse, goat, lamb, sheep, antelope, guanaco, mouse, rabbit, pigeon, chicken, turkey), marked on right, for whom complete genomes are available. From 27 G genotypes, complete genomes for 21 are available, while from 35 P genotypes only 23 are represented. Successful detection by RV-A probes is considered when ΔG value is below -50 (shades of blue). RV-A reference strains (Wa, UK, RRV) used for validation are depicted as G1P[8]-human, G6P[5]-cow and G3P[3]-simian respectively.

Figure 1.

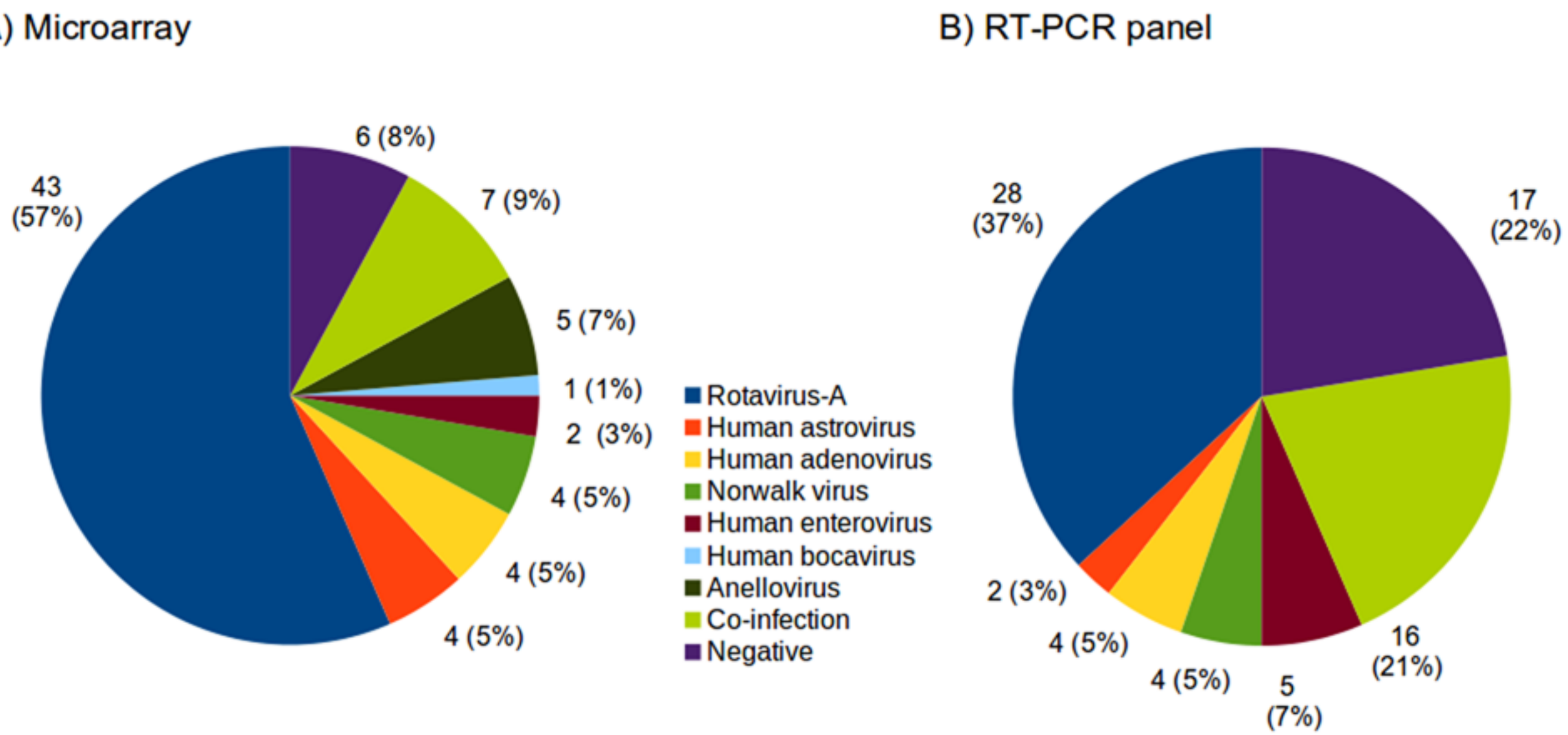
A) Microarray
6 (8%)
7 (9%)
43
(57%)
5 (7%)
1 (1%)
2 (3%)
4 (5%)
4 (5%)
4 (5%)
Rotavirus-A
Human astrovirus
Human adenovirus
Norwalk virus
Human enterovirus
Human bocavirus
Anellovirus
Co-infection
Negative
B) RT-PCR panel
28
(37%)
17
(22%)
2 (3%)
4 (5%)
4 (5%)
5
(7%)
16
(21%)

Figure 2.

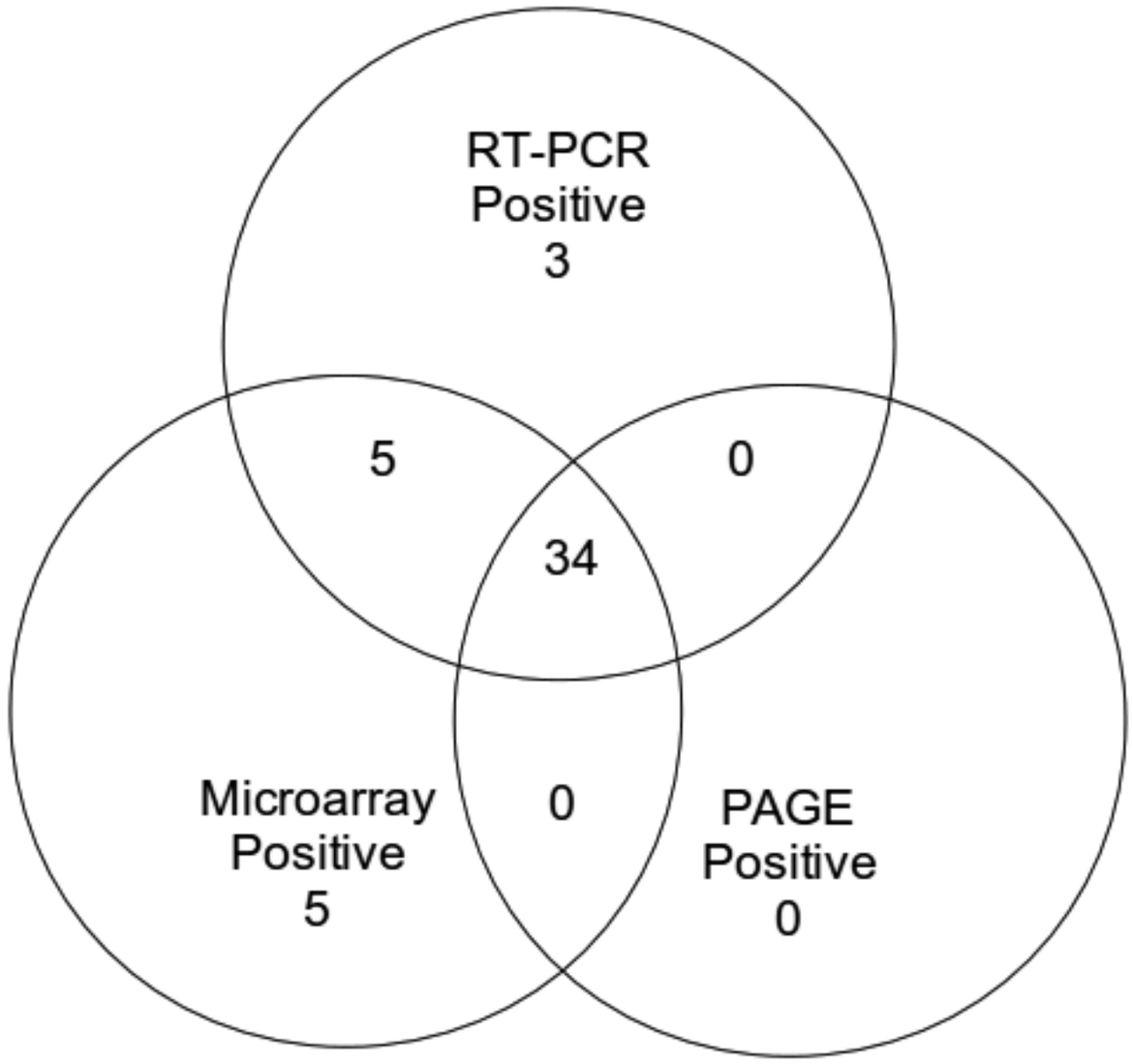
RT-PCR
Positive
3
5
0
34
0
Microarray
Positive
5
PAGE
Positive
0

Figure 3.

RV-A

	PCR+	PCR-	
array +	39	5	44
array -	3	29	32
	42	34	76

Sensitivity 93%
Specificity 85%
PPV 88.6%
NPV 90.6%

HAdV

	PCR+	PCR-	
array +	4	0	4
array -	9	63	72
	13	63	76

Sensitivity 30%
Specificity 100%
PPV 100%
NPV 87.5%

CV

	PCR+	PCR-	
array +	6	0	6
array -	0	70	70
	6	70	76

Sensitivity 100%
Specificity 100%
PPV 100%
NPV 100%

HAstV

	PCR+	PCR-	
array +	3	2	5
array -	0	71	71
	3	73	76

Sensitivity 100%
Specificity 97%
PPV 60%
NPV 100%

HEV

	PCR+	PCR-	
array +	5	2	7
array -	7	62	69
	12	64	76

Sensitivity 42%
Specificity 97%
PPV 71.6%
NPV 89.9%

Table 1. Reference virus species used in microarray validation

Family	Genus	Species	Strain[a]	probes positive/total[b]
Astroviridae	Mammastrovirus	*Human astrovirus*	Yuc8	4/4
Adenoviridae	Mastadenovirus	*Human adenovirus C*	Adv5	10/13
Caliciviridae	Vesivirus	*Feline Calicivirus*	F9	14/22
	Norovirus	*Norwalk virus*[c]	-	8/12
	Sapovirus	*Sapporo virus*[c]	-	5/14
Flaviviridae	Pestivirus	*Bovine Viral Diarrhea Virus 1*	NADL	6/6
	Flavivirus	*Dengue virus 4*	-	9/9
Paramyxoviridae	Respirovirus	*Bovine parainfluenza virus 3*	SF-4	9/9
Reoviridae	Rotavirus	*Rotavirus A*	RRV	22/42
			TFR-41	14/42
			UK	19/42
			Wa	21/42
	Orthoreovirus	*Mammalian Orthoreovirus*	T1L	11/25
			T3D	19/25

[a] Reference strains were provided by Dr. Ramon Gonzalez, FC-UAEM, (Human adenovirus C); Dra. Lorena Gutierrez, CINVESTAV-IPN, (Feline Calicivirus, Norwalk virus, Sapporo virus); Dra. Rosa E. Sarmiento, FMVZ-UNAM (Bovine Viral Diarrhea Virus 1, Bovine parainfluenza virus 3); Dra. Rosa Ma. Del Angel, CINVESTAV-IPN (Dengue virus 4); M.D. Terrence S. Dermody, Vanderbilt University School of Medicine (Mammalian orthoreovirus)

[b] number of oligonucleotide probes which recognized virus/ total number of oligonucleotide probes designed to bind viral species

[c] clinical reference samples

Table 2. Ct values of viral nucleic acid quantification in samples with co-infection

RT-PCR[a]	Microarray[b]	RV-A[c]	HAdV	HEV	NV[d]	HAstV
RV-A[e]	RV-A	21.9[f]				
HAdV	HAdV		14.7			
HEV	HEV			23.8		
*NV	NV				19.6	
HAstV	HAstV					14.8
RV-A/HAdV	RV-A	**20.5**[g]	37.6			
RV-A/HAdV	RV-A	**22.5**	28.4			
RV-A/HAdV	RV-A	**28.2**	41.3			
RV-A/HAdV	RV-A	**28.6**	44.5			
RV-A/HAdV	RV-A	**29.1**	43.8			
RV-A/HAdV	RV-A	**29.2**	43.4			
RV-A/HAdV	RV-A	**30.4**	30.6			
RV-A/HEV	RV-A	29.2		**25.4**		
RV-A/HEV	RV-A	29.2		**27.3**		
RV-A/HEV	RV-A	29.6		**28.5**		
RV-A/HEV	RV-A	30.8		**27.8**		
RV-A/HEV	HEV	38.4		**28**		
RV-A/NV	NV	30.1			**23.6**	
HAstV/HEV	HAstV			28		**23.2**
NV/HEV/HAdV	NV		34.1	28.4	**20.7**	
NTC[h]		36.4	44.5	33.7	33.7	29.9

[a] virus identified by diagnostic RT-PCR

[b] virus identified by microarray

[c] virus determined by real-time RT-PCR, RV-A (rotavirus group A), HAdV (human adenovirus), HEV (human enterovirus), NV (Norwalk virus), HAstV (human astrovirus)

[d] NV is detected at genus level as calicivirus

[e] Single infection samples were used as positive control

[f] Ct value detected by real-time RT-PCR

[g] lower Ct values are shown in bold

[h] NTC non template control