Journal of Clinical Virology 60 (2014) 161–164

Contents lists available at ScienceDirect

Journal of Clinical Virology

journal homepage: www.elsevier.com/locate/jcv

Short Communication

New astrovirus in human feces from Burkina Faso

Tung Gia Phan[a,b], Johan Nordgren[c], Djeneba Ouermi[d], Jacques Simpore[d], Leon W. Nitiema[d], Xutao Deng[a], Eric Delwart[a,b,*]

[a] Blood Systems Research Institute, San Francisco, CA 94118, USA
[b] Department of Laboratory Medicine, University of California at San Francisco, San Francisco, CA 94118, USA
[c] Division of Molecular Virology, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden
[d] Centre de Recherche Biomoléculaire Pietro Annigoni Saint Camille CERBA/LABIOGENE, Université de Ouagadougou, Ouagadougou, Burkina Faso

ARTICLE INFO

Article history:
Received 20 January 2014
Received in revised form 28 March 2014
Accepted 31 March 2014

Keywords:
Astrovirus
Species
Diarrhea
Metagenomics
Children

ABSTRACT

Background: A significant fraction of cases of diarrhea, a leading cause of childhood mortality worldwide, remain unexplained.
Objectives: To identify viruses in unexplained cases of diarrhea using an unbiased metagenomics approach.
Study design: Viral nucleic acids were enriched from the feces from 48 cases of unexplained diarrhea from Burkina Faso, sequenced, and compared against all known viral genomes.
Results: The full genome of a highly divergent astrovirus was sequenced in a sample co-infected with parechovirus 1. RT-PCR identified a single astrovirus infection in these 48 patients indicating a low prevalence. Human astrovirus-BF34 was most closely related to mamastrovirus species 8 and 9 also found in human with which it shared 62%, 74%, and 57% amino acid identities over its protease, RNA dependent RNA polymerase and capsid proteins, respectively.
Conclusions: Burkina Faso astrovirus is proposed as prototype for a novel species in the genus *Mamastrovirus*, here tentatively called *Mamastrovirus* 20, representing the fifth human astrovirus species.

1. Background

Astroviruses are small non-enveloped viruses with a characteristic star-like structure whose RNA genomes are 6.4–7.7 Kb in size and contains three ORFs designated ORF1a, 1b, and 2, coding for protease, protease-RdRp fusion, and capsid proteins, respectively. The first astrovirus was identified by electron microscopy in human fecal samples in 1975 [1]. The family *Astroviridae* is currently classified into two distinct genera: *Mamastrovirus* and *Avastrovirus* infecting mammals and birds, respectively.

According to the ICTV different strains of the same astrovirus species should share >75% identity in their capsid proteins. The genus *Mamastrovirus* currently consists of at least 19 species. The prototypic astrovirus species from human is highly diverse consisting of 8 genotypes which together form the species HAstV or mamastrovirus 1. The second species of human astrovirus, prototype strain MLB1, was characterized by Finkbeiner et al. in 2008 [2] is now called mamastrovirus 6. The third human astrovirus species (mamastrovirus 8) was reported in 2009 with strains VA2 [3] and HMO-A [4]. The fourth human astrovirus species (mamastrovirus 9), also described in 2009, includes strains VA1 [5], VA3 [6], HMO-B and HMO-C [4]. HAstV-VA1 was reported as the causative agent for a diarrhea outbreak in a child care center in Virginia [5]. Serological studies showed HAstV-HMO-C to be a highly prevalent human infection [7] with approximately 65% of adult in the US showing antibody reactivity. Recently HAstV-VA4 was discovered in diarrhea from Nepalese children with a prevalence of 1% (2/196) [6]. The HAstV-VA4's capsid shared the highest identity of 77.5% to HAstV-VA2 [6], suggesting that it also belonged to *Mamastrovirus* 8.

2. Objectives

To analyze the fecal virome in unexplained cases of diarrhea using an unbiased metagenomics approach and genetically characterize novel viruses.

3. Study design

Feces from Burkina Faso children with unexplained acute gastroenteritis were analyzed by deep sequencing. These specimens

* Corresponding author at: Blood Systems Research Institute, San Francisco, CA 94118, USA. Tel.: +1 415 923 5762; fax: +1 415 567 5899.
E-mail address: delwarte@medicine.ucsf.edu (E. Delwart).

http://dx.doi.org/10.1016/j.jcv.2014.03.024
1386-6532/

had been previously screened for rotavirus, norovirus, pathogenic bacteria, parasites, and yeasts and non-reactive samples were selected here for further viral metagenomics analysis [8,9]. Fecal viral particles from 48 such patients were enriched by filtration of stool supernatants and nuclease treatment of the filtrate to reduce the concentration of naked host and bacteria nucleic acids. Briefly, the samples were clarified by 15,000 × *g* centrifugation for 10 min. A total of 200 μl of supernatants was filtered through a 0.45-μm filter (Millipore) to remove bacterium-sized particles. The filtrate was then treated with a cocktail of DNases (Turbo DNase from Ambion, Baseline-ZERO from Epicenter, and Benzonase from Novagen) and with RNase (Fermentas) to digest unprotected nucleic acids. Nuclease resistant nucleic acids were then extracted [10]. A DNA library was constructed using ScriptSeq™ v2 RNA-Seq Library Preparation Kit (Epicenter) according to the manufacturer's instructions, and sequenced using the Illumina MiSeq platform. The Illumina kit generated sequences were compared to the GenBank nonredundant protein databases using BLASTx.

4. Results

All 48 samples were analyzed in 5 pools using one Illumina MiSeq run of 250 bases paired-end reads. One pool showed a single read encoding an astrovirus-like protein segment with a best BLASTx E-score of 10^{-11} to human astrovirus HMO-A (GenBank NC_013443). The specific fecal sample within the pool containing this sequence was identified using RT-PCR with primers complementary to this sequence and then re-analyzed individually using the same metagenomics approach. A total of 127 unique sequences covering 57% of the astrovirus genome were identified. Also detected were 233 unique sequences covering 82% of a parechovirus 1 genome indicating a viral co-infection. The complete genome of the astrovirus was then determined by filling gaps between sequence fragments by RT-PCR, and the RNA genome extremities were amplified by 5′ and 3′ RACE. PCR amplicons were directly Sanger sequenced by primer walking.

The complete genome of the human astrovirus named Burkina Faso 34 (HAstV-BF34, GenBank: KF859964) is 6561-bp in length excluding the polyadenylated tract, with a GC content of 43%. HAstV-BF34 has a 5′ UTR of 43 bases, three overlapping open reading frames (ORFs), and a 3′ UTR of 119 bases. ORF1a encoded the nonstructural polyprotein (888-aa) with a conserved trypsin-like serine protease domain with best BLASTx match to that of mamastrovirus 8, sharing 60% aa-identities. ORF1b encoded RdRp (523-aa) expressed via a ribosomal frameshift caused by a conserved "slippery heptamer" sequence [AAAAAAC]. Pair-wise analysis showed that the RdRp region had the highest identity of 73–74% to those of mamastrovirus 8 strains. NCBI ORF finder and protein alignments revealed that HAstV-BF34 had an alternative ORF2 start codon **M^{1}**GNS upstream of the putative start codon of **M^{63}**AGKQ seen in other genetically-related human astroviruses (Fig. 1A). An unusual long N-terminus for ORF2 was also observed in porcine and dog astroviruses sharing the same **M^{1}**GNS with HAstV-BF34 (Fig. 1A). The consensus promoter initiating ORF2 subgenomic RNA synthesis in HAstV-BF34 was identified as CUUUGGAGGGGAGGACCAAAGCGUGGUGAUGGC (**M^{63}** start codon underlined) [11]. We selected the second methionine codon (**M^{63}**AGKQ) as the putative start codon for this analysis. The shorter ORF2 encoded a viral capsid precursor protein 739-aa in size. It was most closely related to those of HAstV-HMO-A, HAstV-VA2, and HAstV-VA4 (mamastrovirus 8) with 57% aa-identity and to HAstV-HMO-C (mamastrovirus 9) with 52% aa-identity. Because the identity over the capsid protein is less than 75% to those of previously reported astroviruses, HAstV-BF34 is proposed as prototype for a novel species in the genus *Mamastrovirus*, here tentatively called *Mamastrovirus* 20 pending ICTV review, representing the fifth human astrovirus species.

Sequence alignment was performed using CLUSTAL X with the default settings [12]. Phylogenetic tree with 100 bootstrap resamples of the alignment data sets was generated using MEGA version 5 [13]. Bootstrap values (based on 100 replicates) for each node are shown if >70 (Fig. 1B). Phylogenetic analysis showed that HAstV-BF34 was most closely related to *Mamastrovirus* 8 and 9 infecting humans (Fig. 1B).

A nested PCR assay was used to determine the prevalence of this virus in the 48 diarrhea feces from Burkina Faso children. Primers BF34-F1 (5′-GTC CTG AAG ATT ACA GCA AGT CCT-3′) and BF34-R1 (5′-GAC CCA TCC GAG TGA GTG TG-3′) were used for the first round of PCR, and primers BF34-F2 (5′-GAA CCA TTG ACT AAC ATA AAA GCC A-3′) and BF34-R2 (5′-TCC TCA AAA ACA CAG CCT ATT CT-3′) for the second round of PCR, resulting in an expected amplicon of ~350 bp. The PCR conditions was as follows: denaturation at 95 °C for 5 min, 35 cycles of 95 °C for 30 s, 51 °C or 49 °C (for the first or second round, respectively) for 30 s and 72 °C for 1 min, a final extension at 72 °C for 10 min, and then held at 4 °C. No other samples except the one initially detected by deep sequencing were PCR positive yielding a low prevalence of detection of <2% (1/48) in this population.

5. Discussions

The classic human astrovirus (mamastrovirus 1) has been associated with acute gastroenteritis in humans worldwide [11,14–17]. Chronic infections of astrovirus were reported in immunodeficient patients [18,19]. Mamastrovirus 9 (strain HAstV-PS) was also identified as the causative agent in a fatal case of encephalitis in a child with agammaglobulinemia [20] and HAstV-MLB2 (whose capsid is 74% identical to mamastrovirus 6 strain MLB1 and may therefore also qualify as a distinct species) has also been reported in the plasma of a febrile child [21]. Astroviruses were also found in the brain of minks with a shaking symptoms [22] and in the brain and spinal cord of cows with neurologic symptoms [23] indicating that astroviruses are not exclusively enteric infections but can also affect other organ systems. Varied astroviruses have also been detected in feces from a wide variety of farm, wild, and laboratory aminals [11] indicating that the known diversity of this viral family is likely to continue its rapid expansion.

In this study, fecal shedding of a previously uncharacterized astrovirus was detected in an unexplained case of diarrhea. The 14-month old female patient has significant malnutrition indicators including severe underweight, severe wasting, and moderate stunting. This patient showed typical symptoms of acute gastroenteritis including fever (38 °C), vomiting, abdominal pain, liquid feces (4 times/day) and moderate dehydration. The simultaneous detection of parechovirus 1 in this fecal sample provides an alternative explanation for the patient's symptoms although a primary or aggravating role for the astrovirus is also conceivable. Because the parechovirus and astrovirus genomes both consist of ssRNA of similar length their relative viral loads are likely reflected by the respective number of sequence reads (233/127) or nearly twice as many parchovirus as astrovirus RNA genomes. The absence of detected recombination between mamastrovirus 8, 9, and proposed mamastrovirus 20 genomes, using SimPlot (data not shown), would support the presence of different astrovirus species. The phylogenetic clustering of now three species of human astrovirus (mamastrovirus 8, 9, and proposed 20) to the exclusion of animal astroviruses may also reflect their monophyletic origin and these species could potentially be subsumed into a single, more diverse viral group. It may also be concluded that astrovirus evolution