Infection, Genetics and Evolution 14 (2013) 265–274

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

Evolutionary time-scale of primate bocaviruses

Igor V. Babkin [a,*], Alexander I. Tyumentsev [b], Artem Yu. Tikunov [a], Alexander M. Kurilshikov [a], Elena I. Ryabchikova [a], Elena V. Zhirakovskaya [a], Sergei V. Netesov [b], Nina V. Tikunova [a]

[a] Laboratory of Molecular Microbiology, Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, Russia
[b] Laboratory of Nanobiotechnologies, Novosibirsk State University, Novosibirsk, Russia

ARTICLE INFO

Article history:
Received 2 October 2012
Received in revised form 16 December 2012
Accepted 17 December 2012
Available online 8 January 2013

Keywords:
Human bocavirus
Genotyping
Complete genome
Virus evolution
Recombination

ABSTRACT

Human bocavirus (HBoV) is associated with acute gastroenteritis in humans, occurring mostly in young children and elderly people. Four bocavirus genotypes (HBoV1–HBoV4) have been found so far. Since there were no data on the contribution of HBoV to gastroenteritis in Russia, 1781 fecal samples collected from infants hospitalized with acute gastroenteritis in Novosibirsk, Russia during one year were tested for the presence of nucleic acids from HBoV and three major gastrointestinal viruses (rotavirus A, norovirus II, and astrovirus). HBoV was detected only in 1.9% of the samples: HBoV1 was detected in 0.6% and HBoV2, in 1.3%. Complete genome sequencing of three Novosibirsk isolates was performed. An evolutionary analysis of these sequences and the available sequences of human and great apes bocaviruses demonstrated that the current HBoV genotypes diverged comparatively recently, about 60–300 years ago. The independent evolution of bocaviruses from chimpanzees and gorillas commenced at the same time period. This suggests that these isolates of great apes bocaviruses belong to separate genotypes within the species of human bocavirus, which is actually the primate bocavirus. The rate of mutation accumulation in the genome of primate bocaviruses has been estimated as approximately 9×10^{-4} substitutions/site/year. It has been demonstrated that HBoV1 diverged from the ancestor common with chimpanzee bocavirus approximately 60–80 years ago, while HBoV4 separated from great apes bocaviruses about 200–300 years ago. The hypothesis postulating independent evolution of HBoV1 and HBoV4 genotypes from primate bocaviruses has been proposed.

1. Introduction

Human bocavirus (HBoV) is a recently identified virus of the family *Parvoviridae*. HBoV was initially reported in Sweden in the nasopharyngeal aspirates of children with acute infections of the upper and lower respiratory tracts (Allander et al., 2005). A metagenomic study of the feces of Australian children with acute gastroenteritis detected the human bocavirus of genotype 2 (HBoV2) (Arthur et al., 2009). HBoV3 was detected in the same year (Kapoor et al., 2009), and HBoV4 was found in 2010 (Kapoor et al., 2010a). The high genetic diversity observed in HBoV2 suggested that this genotype should be subdivided into two subgenotypes HBoV2A and HBoV2B (Kapoor et al., 2010a). Lau et al. first documented association HBoV with gastroenteritis in early study (Lau et al., 2007). HBoV1 is primarily, although not exclusively a respiratory virus. Unlike HBoV1, HBoV2–HBoV4 genotypes seem to occur mainly in human stool (Jartti et al., 2012; Malecki et al., 2011).

HBoV is a small (18–26 nm) nonenveloped icosahedral virion (Gurda et al., 2010). The HBoV genome is represented by a linear single-stranded DNA, which is a coding strand; the genome length is about 5300 nucleotides (Schildgen et al., 2012). The genome putatively encodes two forms of the nonstructural protein NS1; nuclear phosphoprotein NP1, unique for bocaviruses; and two major structural proteins, VP1 and VP2. The non-coding regions in the genomic termini contain palindromic sequences, commonly known as inverted terminal repeats, which play a vital role in the viral replication (Chen et al., 2010; Dijkman et al., 2009; Kapoor et al., 2011; Sun et al., 2009). The HBoV DNA replication and virus assembly take place in the cell nucleus. Homology of the protein amino acid sequences of the HBoVs belonging to different genotypes is rather low, amounting to 70–80%. The similarity between the HBoV NP1 amino acid sequences to other members of the genus *Bocavirus*—bovine parvovirus and minute virus of canines—does not exceed 47% (Kapoor et al., 2009, 2010a, 2011).

The HBoVs are characterized by rapid evolution. As has been demonstrated, the rate of mutation accumulation in the genome of this virus is comparable to that of RNA viruses (Malecki et al., 2011; Zehender et al., 2010). In addition, it has been found that recombination between different HBoV strains make a significant contribution to emergence of new genetic variants of this virus.

* Corresponding author. Address: Laboratory of Molecular Microbiology, Institute of Chemical Biology and Fundamental Medicine, pr. Lavrentieva 8, Novosibirsk 630090, Russia. Tel.: +7 383 363 5157; fax: +7 383 363 5153.
E-mail addresses: i_babkin@mail.ru, babkin@niboch.nsc.ru (I.V. Babkin).

1567-1348/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2012.12.023

Some researchers propose that HBoV2 appeared as a result of recombination between HBoV1 and HBoV4 with the additional contribution of an intra-genotype recombination among HBoV2 variants (Fu et al., 2011; Kapoor et al., 2009, 2010a; Song et al., 2010). It is also believed that HBoV3 is a recombinant of other HBoV genotypes (Kapoor et al., 2010a; Cheng et al., 2011). However, this issue is to be further clarified. For example, it is still unclear which particular genotypes have been involved in the formation of HBoV3: HboV1 and HBoV2 (Kapoor et al., 2010a) or HBoV1 and HBoV4 (Cheng et al., 2011). In several cases, the Shimodaira–Hasegawa test fails to confirm recombination events in HBoV2 (Cheng et al., 2011). All these questions require further studies.

So far, it has been shown that HBoV is present in almost all regions of the world, including Europe (Allander et al., 2005; Zehender et al., 2010; Kapoor et al., 2009), Asia (Cheng et al., 2011; Song et al., 2010; Kapoor et al., 2009), America (Santos et al., 2010; Kapoor et al., 2011), Africa (Kapoor et al., 2010a) and Australia (Arthur et al., 2009). Three types of HBoV were detected in southern Ireland (Cashman and O'Shea, 2012). Since there were no data on the HBoV contribution to the gastroenteritis in Russia, we have searched for this agent in the children hospitalized with acute gastroenteritis in Novosibirsk (Russia) during March 2010–February 2011 and conducted molecular genetic analysis of the detected isolates. In addition, we determined the genomic sequences of one HBoV1 isolate and two HBoV2 isolates and used this data for dating the molecular history of primate bocaviruses.

2. Materials and methods

2.1. Samples

Stool samples were collected from 1781 children up to 3 years of age hospitalized with a diagnosis of acute gastroenteritis in Infectious Department of the 3rd Novosibirsk Child Hospital during March 01, 2010–February 28, 2011. This project was considered and approved by the Ethical Committee with the State Research Center of Virology and Biotechnology Vector (IRB0001360).

2.2. HBoV detection and sequencing

Viral DNA was isolated from the feces clarified with PBS–glycerin buffer by affinity sorption using a Ribo-Sorb (Interlabservis, Russia) extraction kit. The samples were assayed for the presence of bocavirus DNA by polymerase chain reaction (PCR) using the primers HBoV2-sf2 (5′-TGCTTCAACAGGCAAAACAA-3′) and HBoV2-sr2 (5′-TCCAAGAGGAAATGAGTTTGG-3′) according to Kapoor et al. (2009). The primers were annealed at 52 °C; in total 41 amplification cycles were performed. The amplification products were analyzed by electrophoresis in 1.5% agarose gel. The size of the target PCR product was 495 bp.

The amplified fragments were purified with a Silica Bead DNA Gel Extraction kit (Fermentas) and sequenced using a BigDye® Terminator v. 3.1 Cycling Sequencing kit (Applied Biosystems). The reaction products were purified on CentriSep (Applied Biosystems) columns and assayed by electrophoresis using an ABI PRISM 3100 (Applied Biosystems) genetic analyzer. The sequences were analyzed with the help of the MEGA 5.0 software (Tamura et al., 2011).

2.3. Sequence analysis

The complete genomic sequences, used for similarity analysis, were extracted from the GenBank database (http://www.ncbi.nih.gov). Sequence alignment and pairwise nucleotide identities were performed using the programs BioEdit v. 7.0 (Hall, 1999) and ClustalX v. 1.8 (Thompson et al., 1997). Phylogenetic analysis was performed using the neighbor joining and maximum likelihood methods with the help of the MEGA v. 5.0 software (Tamura et al., 2011).

Positive and negative selections were estimated using two approaches: single likelihood ancestor counting (SLAC) (Kosakovsky Pond and Frost, 2005) and PARRIS test for positive selection (Scheffler et al., 2006).

In order to identify the recombinant strains and exclude them from analysis, we used the RDP3 program implementing five different methods: GENECONV, RDP, SiScan, Bootscan, and MaxChi, with a highest acceptable p value of 0.05.

Following each search, a likelihood ratio test, LRT (Huelsenbeck and Rannala, 1997), was performed to test for departure from a clocklike evolution in the molecular sequence data (test for rate constancy). The tests were performed on the trees found by maximum likelihood searches. The significance of the likelihood ratio test statistics was approximated using a χ^2 distribution.

The rates of nucleotide substitution per site and the times to the most recent common ancestor were estimated using the Bayesian Markov chain Monte Carlo (MCMC) method available in the BEAST package (Drummond and Rambaut, 2007). A relaxed (uncorrelated lognormal) molecular clock was utilized for each dataset (Lemey et al., 2010). Sufficient MCMC chains were run to ensure convergence with an initial 10% of the MCMC chains discarded as burn-in. Statistical uncertainty around the mean estimates was provided by the 95% highest probability density values. Finally, the BEAST analysis also enabled us to infer maximum clade credibility trees for each dataset.

2.4. Electron microscopy

The stool samples containing HBoV DNA detected by PCR and sequencing were examined by electron microscopy. A drop of stool sample was adsorbed for 30 s on a copper grid covered with formvar film; then the excess fluid was removed, and the grid was placed on a drop of 1% uranyl acetate for 15 s. The grids were examined with a JEM 1400 transmission electron microscope at an accelerating voltage of 80 kV. Virus images were collected by a Veleta digital camera.

3. Results

3.1. HBoV detection in diarrheic stool samples

In this study, we tested a set of 1781 feces samples from infants hospitalized with acute gastroenteritis during March 2010–February 2011. The HBoV DNA was detected by PCR using primers that can detect four HBoV genotypes (Cheng et al., 2011; Kapoor et al., 2009). During this period, HBoV isolates were detected in 34 clinical samples (1.9% of all the analyzed specimens). Note that all the collected samples were assayed in parallel by RT-PCR for the presence of RNA of rotaviruses A (HRVA), genogroup II noroviruses (HNoVGII), and astroviruses (HAstV) (Fig. 1).

During the observation period, all the listed viruses were detected in 707 of the assayed 1781 cases, accounting for 39.7% of all examined children. Only HRVA was present in 393 (22.1%) samples; HNoVGII, in 164 (9.2%) samples; HAstV, in 33 (1.9%) samples; and 81 (4.5%) samples contained two or more different viruses (HRVA, HNoVGII, and/or HAstV). In 1073 (60.3%) samples, gastroenteritis were determined by other causes (Fig. 1). Since each sample was assayed for the presence of all the listed pathogens, we could assess the HBoV abundance as a monoinfection and mixed infections of combined etiologies. Only HBoV DNA was present in 19 (1.1%) samples and HBoV DNA was detectable as a mixed