



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Human Bocaviruses Are Highly Diverse, Dispersed, Recombination Prone, and Prevalent in Enteric Infections

**Citation for published version:**

Kapoor, A, Simmonds, P, Slikas, E, Li, L, Bodhidatta, L, Sethabutr, O, Triki, H, Bahri, O, Oderinde, BS, Baba, MM, Bukbuk, DN, Besser, J, Bartkus, J & Delwart, E 2010, 'Human Bocaviruses Are Highly Diverse, Dispersed, Recombination Prone, and Prevalent in Enteric Infections', *The Journal of Infectious Diseases*, vol. 201, no. 11, pp. 1633-1643. <https://doi.org/10.1086/652416>

**Digital Object Identifier (DOI):**

[10.1086/652416](https://doi.org/10.1086/652416)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

The Journal of Infectious Diseases

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.





Published in final edited form as:

*J Infect Dis.* 2010 June 1; 201(11): 1633–1643. doi:10.1086/652416.

## Human bocaviruses are highly diverse, dispersed, recombination prone, and prevalent enteric infections

A Kapoor<sup>1,2</sup>, P. Simmonds<sup>3</sup>, B. Slikas<sup>1,2</sup>, Linlin Li<sup>1,2</sup>, L. Bodhidatta<sup>4</sup>, O. Sethabutr<sup>4</sup>, H. Triki<sup>5</sup>, Olfa Bahri<sup>5</sup>, B. Oderinde<sup>6</sup>, M. Baba<sup>6</sup>, D. Bukbuk<sup>7</sup>, J. Besser<sup>7</sup>, J. Bartkus<sup>7</sup>, and E. Delwart<sup>1,2,\*</sup>

<sup>1</sup>Blood Systems Research Institute, San Francisco, CA, USA. <sup>2</sup>Dept of Laboratory Medicine, University of California San Francisco, San Francisco, CA, USA. <sup>3</sup>Center for Infectious Diseases, University of Edinburgh, Edinburgh, United Kingdom. <sup>4</sup>Department of Enteric Diseases, Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand. <sup>5</sup>Laboratory of Clinical Virology, WHO Regional Reference Laboratory on Poliomyelitis, Institut Pasteur, Tunis, Tunisia. <sup>6</sup>WHO National Polio Laboratory, University of Maiduguri Teaching Hospital, Borno State, Nigeria. <sup>7</sup>Molecular Epidemiology Unit, Public Health Laboratory Division, Minnesota Dept. of Health, St Paul, MN, USA.

### Abstract

A new species of parvovirus tentatively named human bocavirus 4 (HBoV4) was genetically characterized. Among 641 feces samples from children and adults the most commonly detected bocaviruses species were HBoV2>HBoV3>HBoV4>HBoV1 with HBoV2 prevalence of 21% and 26% in Nigerian and Tunisian children. HBoV3 and HBoV4 species combined were found in 12/192 cases of non-polio acute flaccid paralysis (AFP) from Tunisia and Nigeria and 0/96 healthy Tunisian contacts (p=0.01). Evidence of extensive recombination at the NP1 and VP1 gene boundary between and within species was found. The multiple species and high degree of genetic diversity seen among the human bocaviruses found in feces relative to the highly homogeneous HBoV1 suggest that this world-wide distributed respiratory pathogen may have recently evolved from an enteric bocavirus, perhaps after acquiring an expanded tropism favoring the respiratory track. Elucidating the possible role of the newly identified enteric bocaviruses in human diseases including AFP and diarrhea will require further epidemiological studies.

### INTRODUCTION

Parvoviruses are small, non-lipid enveloped, environmentally resistant viral particles with single-stranded linear DNA genomes that frequently infect animals through the fecal–oral route (1). Parvoviruses can cause numerous symptoms in animals, including enteritis, panleukopenia, hepatitis, erythrocyte aplasia, and cerebellar ataxia (1). Efficacious vaccines against animal parvovirus infections are widely employed (2–4).

Following the identification of the human parvovirus B19 in 1976 several other human parvoviruses have recently been genetically characterized. PARV4 was found in the blood of

**Reprints or correspondence:** Eric Delwart, Blood Systems Research Institute, 270 Masonic Ave., San Francisco, CA 94118. Tel: 415-9235763, Efax: 419-7914220, delwarte@medicine.ucsf.edu.

\*Communicating author: Blood Systems Research Institute, 270 Masonic Ave, San Francisco CA 94118. delwarte@medicine.ucsf.edu Tel: 415-9235763 Fax: 415-5675899

**Conflict of interest:** A provisional patent application has been filed for HBoV2-3-4.

an adult intravenous drug user (5), HBoV1 in the nasopharyngeal secretion of a child with respiratory problems (6), HBoV2 in the stool of children with non-polio acute flaccid paralysis (7), and HBoV3 in the stool of Australian children with diarrhea (8). In contrast to PARV4 or B19, bocaviruses contain a third open reading frame of unknown function (10)(11).

The first bocavirus identified was in cows (12), and the name of the genus is derived from its first known hosts (bovine-canine). Animal bocaviruses can cause both respiratory and gastrointestinal diseases and embryonic and fetal death (10).

HBoV1 infection has been linked with mild to severe, primarily lower respiratory tract infections in children, frequently in association with other viral infections (11)(13)(14)(15)(17)(18)(19)(20)(21)(22)(23)(25)(26)(27)(28). HBoV1 has also been detected at low frequency in stool samples although association with diarrhea appears weaker than with respiratory symptoms (11)(29)(30)(31)(32)(33)(34). HBoV1 strains show a very low degree of genetic variability world-wide (35,36). HBoV2 was recently identified in stool samples of Pakistani children (7). Lower frequencies of HBoV2 were detected in the stool of Scottish adults and children (7). A third bocavirus species, HBoV3, as well as HBoV2, were recently found in stool samples from Australian children with diarrhea (8).

In this study, pan-bocavirus PCR primers were designed and used on human stool samples from several countries. All three recently identified bocaviruses plus a fourth species, HBoV4, were identified. A high degree of genetic diversity, relative to that seen for HBoV1, was seen among human bocaviruses in feces.

## MATERIALS AND METHODS

### Samples

Stool samples were collected as part of previous clinical studies and anonymized. This study was approved by the UCSF CHR. Samples from Nigeria and Tunisia were collected as part of the WHO's poliovirus eradication program from children with non-polio AFP between the ages of 4 months and 15 years. Stool samples from healthy contacts of cases from Tunisia were matched for age. Stool samples from Nepal were from adult travelers and resident expatriates with diarrhea with no known enteric pathogens detected following standard microbiology for enteric bacteria, EIA for rotavirus, adenovirus, astrovirus, Giardia and Cryptosporidium and RT-PCR for norovirus. Stool samples from healthy, asymptomatic controls were collected from the same population. Stool samples from the Minnesota Department of Health are from individuals with diarrhea and healthy individuals, matched for age and residential area code.

### PCR amplification of bocaviruses

We used nested PCR targeting the VP1/2 region of both HBoV1 and HBoV2 (nucleotide positions 3233 to 3808 numbered here and subsequently using the HBoV2 prototype sequence GenBank entry FJ170278). Nucleic acids (DNA and RNA) was extracted (QIAamp Viral RNA mini kit) from 140 µl of clarified stool supernatant and eluted into 60 µl of water. First-round PCR primers were AK-VP-F1:(CGCCGTGGCTCCTGCTCT) and AK-VP-R1:(TGTTCCGCATCACAAAAGATGTG) and second-round primers were AK-VP-F2:(GGCTCCTGCTCTAGGAAATAAAGAG) and AK-VP-R2:(CCTGCTGTTAGGTCGTTGTTGTATGT). PCR contained 2.5U of Taq DNA polymerase (NEB) in 1.1×Thermopol reaction buffer with MgCl (2.0mM), 20pmol (each) of forward and reverse primers, and 2.5 µl of nucleic acids (first round) or 1µl of the first-round PCR product (second round) as template in 50ul PCRs. First-round conditions were 10 cycles of 95°C for 35 sec, 58°C for 1 min, and 72°C for 1 min, with a decrease of 0.5°C in annealing temperature each cycle; 30 cycles of at 95°C for 30 sec, 54°C for 45 sec, and 72°C for 45 sec; a final

extension at 72°C for 10 min. Similar conditions were used for the second round, except that the initial annealing temperatures were 60°C and 58° in the first and second group of PCR cycles, respectively. Amplicons of the appropriate size detected by agarose gel electrophoresis were directly sequenced. The PCR products whose sequences are reported here produced unambiguous dideoxy-sequencing electrophoregram peaks indicating the predominance (>90%) of a single bocavirus variant. The sensitivity of the pan-bocavirus nested PCR was determined using dilution of plasmids containing the target sequences from each of the four bocavirus species and was estimated at 10–100 plasmid copies for each species.

### Complete genome sequencing

Complete genomes were amplified using PCR primers designed from alignments of HBoV1 and HBoV2 genomes and then directly sequenced by primer walking (7). The terminal sequences were acquired by a modified protocol for rapid amplification of cDNA ends (5). The terminal sequences are incomplete because of extensive hairpin structures preventing extensions to the viral 5' and 3' extremities.

### Distance measurements and phylogenetic analyses

Phylogenetic relationships were evaluated by using Mega 4.1 (<http://www.megasoftware.net/mega41.html>) (37). Neighbor-joining trees were inferred using a matrix of pairwise maximum-likelihood distances computed from a nucleotide alignment including the genomes obtained in this study and in Genbank plus a matrix of PAM distances computed from the inferred amino acid alignment.

### Recombination analyses

Similarity values based on Jukes-Cantor corrected nucleotide distances between full-length sequences were calculated using the program SequenceDist in the Simmonic2005 v1.6 Sequence Editor Package (38). To assess similarity across the genomes, sequence scans were performed using a fragment length of 300 bases and an increment of 9 bases between fragments. For sequence comparisons with HBoV-1 and HBoV2, a mean pairwise distance was computed using a set of 14 HBoV1 and 7 HBoV2A sequences (as shown in Fig. 2).

### Nucleotide sequence accession numbers

The near full genomes and partial VP1 gene sequences have been deposited in GenBank under accession numbers FJ973558-FJ973563 and GQ506558-GQ506661

## RESULTS

### Wide spread geographic distribution of human bocaviruses and identification of new species

Pan-bocavirus PCR primers were designed that could anneal to both HBoV1 and HBoV2 amplifying a fragment approximately 576 nt long of the VP1 capsid gene (see Materials and Methods). Nucleic acids from human stools collected from Nigeria, Tunisia, Nepal and the United States were then analyzed using nested PCR (Table 1). Of the 641 samples tested, 101 (16%) were confirmed positive by PCR sequencing, with the highest prevalence in Tunisian children with AFP (33%). To determine the phylogenetic relationship of these strains, the sequences were aligned with available sequences of HBoV1, 2, and 3. Only 4/101 strains (from Nigeria and Tunisia) grouped with HBoV1. The remaining strains shared a more recent common ancestry with HBoV2, although the sequences clustered into four distinct genetic lineages, labeled HBoV2A, HBoV2B, HBoV3, HBoV4 (Fig. 1).

The HBoV2A clade included all three published HBoV2 genotypes (7), plus four new strains from Nigeria and Tunisia and the recently reported W153 strain from Australia (8). Because

of the high degree of genetic diversity observed and to prevent excessive splitting of HBoV2 into a multitude of genotypes, all strains in that cluster were re-classified into a new, now more diverse genotype A (HBoV2A). The nearly full genome of a new HBoV2A variant was sequenced (TU-C-114-06) and showed very high protein identities (>98%) to the Pakistani HBoV2 prototype PK5510 (FJ170278).

The HBoV2B cluster included 76 of the 101 bocavirus strains reported here. Two near full genomes were sequenced showing pair-wise amino acid divergence in VP1 of 0.45% (Table 2B)(NI-213 and NI-327)(Genbank accession: FJ973560 and FJ973559). Greater divergence was observed when these strains were compared to the VP1 of HBoV2A strains (average 3.9% range 3.3%–4.5%)(Table 2C). The nucleotide and amino acid distances at the other loci are shown (Table 2C). Using the amino acid divergence between VP1 proteins as the criterion for genotype level classification, this cluster was therefore labeled as HBoV2 genotype B (HBoV2B). The large majority (76/101) of HBoV variants identified in this study belonged to the HBoV2B genotype (Table 1).

The HBoV3 cluster included 11 strains, and the complete genomes of two representative strains, NI-374 and TU-A-210-07, were acquired (Genbank accession: FJ973563 and FJ973562). When compared to HBoV2, the NS1 region of both strains showed an average amino acid divergence of 26% in the NS1 region and 9% in the VP1 region. Compared to HBoV1, the HBoV3 strains showed an average amino acid divergence of 12% in the NS1 region and 20% in the VP1 region (Table 2C). All 11 strains were classified as members of the new HBoV3 species recently described by Arthur et al (8). A distantly related variant of HBoV3 was also recently identified in a US sewage sample (Fig 1 HBoV3B-CA-1-C1)(39). To confirm that this variant represents a second genotype of HBoV3 will require full genome sequencing.

The HBoV4 cluster included six strains. The complete genome of one representative strain (NI-385) was obtained (Genbank accession: FJ973561). The NS1 protein of HBoV4 (NI-385) showed an average of 11 % divergence to HBoV2 and 25–27 % divergence to HBoV1 or HBoV3 (Table 2C). For the VP1 protein, the relative divergences were reversed with an average divergence of 8.5% relative to HBoV3, 9.5% compared to HBoV2, and 19% compared to HBoV1. NI-385 therefore also appears to be a recombinant with the 5' genes NS1 and NP1 most related to HBoV2 (particularly genotype A), while the 3' VP1 slightly more similar to HBoV3 than HBoV2 (Table 2C). According to ICTV species demarcation criteria in the genus *Bocavirus*, members of different species must show greater than 5% divergence in their non-structural gene nucleotide sequences (40). The genetic distance of NI-385 to its closest relatives in the NS1 gene (HBoV2A) was 10.8% (range 6.8–12.6), indicating that NI-385 qualifies, pending ICTV review, as the prototype of a fourth HBoV species (HBoV4). A distant VP1 (partial) variant of HBoV4 was also detected in the US (Fig.1 US-MN-964-05). To determine if the latter variant represents a second HBoV4 genotype will require full genome sequencing.

### Nearly complete genomes and phylogenetic analysis of new bocavirus species

In a manner similar to HBoV1 and 2, all the new genomes of HBoV2, 3, and 4 encoded three large ORFs (Figure 5). The left ORF encodes the non-structural protein NS, the middle ORF encodes NP1 and the right ORF encodes overlapping VP1/VP2 capsid proteins. Conserved motifs associated with rolling circle replication, helicase and ATPase were identified within the NS protein. NP1 is a highly phosphorylated protein of currently undetermined function (41); NP1 differed in length between species, ranging from 214 to 219 aa. Situated within the VP1-unique (VP1u) region the phospholipase A2 motifs required for parvovirus infectivity were found in all 6 genomes, together with the presence of the calcium-binding loop and catalytic residues.

### Further evidence of recombination in human bocaviruses

The six new HBoV genomes from this study were aligned with those of HBoV1, 2, and 3 from prior studies. To determine the relationship between HBoV2 and other members of the *Bocavirus* genus, phylogenetic analyses of NS1, NP1 and VP1/VP2 were performed, by use of both nucleotide sequences and deduced protein sequences (Fig 2). NS1 and NP1 genes of HBoV3 clustered with HBoV1 while their VP1/2 gene clustered with HBoV2. The incongruence in phylogenetic association between loci provided further evidence that HBoV3 originated from a recombination event bringing together the NS1/NP1 gene of HBoV1 and the VP1/2 gene of HBoV2 (8). The likely recombinant origin of HBoV4, clustering with HBoV2 in the NS1/NP1 but with HBoV3 in the VP1, is also shown.

A scan of sequence divergence between complete genome sequences further supported the hypothesis of past recombination between HBoV1 and 2 in the generation of HBoV3 and of HBoV2 and HBoV3 in the generation of HBoV4, with both recombination points near the NP1 and VP1 junction (Fig 3). When different HBoV2 variants were similarly analyzed for recombination, intraspecies HBoV2 recombinants were also detected (data not shown and (7))

### Diversity amongst respiratory HBoV1 and enteric HBoV2-4

We compared the intra-species diversity of HBoV1 with that of HBoV2 using the partial VP1 sequence data generated with the pan-bocavirus PCR primers available for the greatest number of HBoV2, 3, 4 variants (Fig 4A and 4B). A very low average pair-wise difference was seen for HBoV1 collected world-wide (Fig. 4A). HBoV2, including both genotypes, was more diversified than HBoV1 although HBoV2B alone showed a low level of diversity comparable to that of HBoV1 (HBoV2B generated the large low divergence peak of figure 4B). The homogeneity of HBoV1 and HBoV2B can also be visualized in the small branches length in figure 1. When the pair-wise distances of all the enteric (non-HBoV1) sequences were plotted the distribution was much larger than that for HBoV1 (Fig 4C). The low level of intra-species diversity of HBoV1 relative to the other species is also reflected in Table 1B.

### Splicing in NS

The genomic organization of HBoV species are remarkably similar to those of animal bocaviruses, except that all HBoV NS ORFs encode a shorter NS2 protein (630 to 650 aa) compared to animal bocaviruses NS1 (716 to 726 aa) (Fig 5). We noticed in all four HBoV species the presence of a stretch of encoded amino acids similar to the C-terminus of the longer NS1 of animal bocaviruses overlapping the NP ORF but in a different frame (Fig 5). Genomes of all HBoV species were aligned to determine the presence of conserved potential RNA splicing signals near the end of the smaller NS2 ORF and the putative second exon encoding the C-terminal region of NS1 (Fig 5). The putative NS1 resulting from such a spliced transcript encoded a 750–780 amino acid long protein whose carboxy-termini showed significant similarity to that of the canine and bovine bocaviruses NS1 (Fig 5)(42,43). A recent publication using RT-PCR for the detection of HBoV1 viral transcript in human lung epithelial cells failed to detect the NS splicing proposed here (44). On the other hand the proposed NS1 RNA splicing and NS1 protein expression itself were detected using Northern blots and NS1 C-termini specific sera respectively in 293 and human epithelial cells transfected with plasmids expressing HBoV1 transcripts (Jianming Qiu, personal communication).

### Human bocavirus 2–4 disease association

Most positive stool samples contained HBoV2B (76/101) making this genotype the most commonly detected enteric human bocavirus (Table 1). HBoV3 was identified in 11 stool samples and HBoV4 in 6 making them the second and third most common enteric human

bocaviruses respectively in the regions analyzed here. Only four of the 101 positive samples belonged to HBoV1. Stools from AFP cases from Nigeria and Tunisia both showed a high prevalence of HBoV2A+B (21–26%). Since both AFP cases and healthy contacts from Tunisia showed a comparable prevalence of HBoV2 no association was seen between AFP and HBoV2 shedding. The prevalence of HBoV2 in adults with diarrhea from the US and Nepal were also compared to those in healthy matched subjects. No associations were observed between HBoV2 shedding and diarrhea. While the numbers of HBoV3 and HBoV4 detected were relatively small it was noticed that the eight HBoV3 and four HBoV4 were found only amongst 192 AFP samples while none was found among 96 healthy Tunisian matched contacts. Comparing the combined Nigerian and Tunisia AFP infected with HBoV3 and HBoV4 to Tunisian controls yielded a 2 tail Fisher's exact p value of 0.01 while comparing only the Tunisian AFP to Tunisian controls gave a p value of 0.059. HBoV3 was also found from one case each of diarrhea from Nepal and the US and from one healthy person. HBoV4 was also found in a diarrhea case and one healthy US adult. Further testing will be needed to confirm this trend of an association of HBoV3/4 with AFP in children.

## DISCUSSION

We report on a previously uncharacterized species of bocavirus we tentatively named HBoV4. HBoV1 to HBoV3 were also detected by this pan-PCR approach. HBoV2 was recently described in the stool of Pakistani children (7) and HBoV3 in stool samples of Australian children (8) and both were recognized as recombinant viruses. The newly reported Australian HBoV3 (EU918736) appears closely related to a Tunisian strain (TA-210-07)(Fig 2). A highly prevalent genotype of HBoV2 (HBoV2B), together with partial genomic support for second genotypes of HBoV3 and HBoV4, were also identified. The availability of novel bocavirus genomes will allow the design of species specific PCR or microarray oligonucleotides for their detection and disease association studies now required for the 3 recently characterized enteric human bocavirus species (HBoV2,3,4). Based on the phylogenetic clustering observed for a large number of partial VP1 sequences (Fig 1) and the distances among full genomes (Table 2) we propose for future classification that HBoV strains showing >8% protein and >10% nucleotide difference in complete VP1 gene be considered different species while those showing >1.5% protein and >5% nucleotide difference be considered as different genotypes. Such VP1 based classification would retain the four proposed human bocavirus species. The VP1 locus was selected since it is likely to strongly influence tissue tropism and potentially pathogenesis (45).

HBoV1 is primarily, although not exclusively (29–34), a respiratory virus. We show here a higher prevalence of HBoV2 (particularly HBoV2B) than HBoV1 in stool samples from children of different countries. A study testing for HBoV1 and HBoV2 DNA using species specific nested PCRs failed to detect any HBoV2 in >6500 respiratory secretion samples from Edinburgh and Bangkok while HBoV1 was found in 3% and 14% of these respiratory samples respectively (46). Another study found 5/212 (2%) nasopharyngeal samples from Korean children with acute lower respiratory track infections to be positive for HBoV2 DNA while unexpectedly no HBoV1 DNA was found (47). Analyzing for HBoV1 and HBoV2 in both respiratory secretion and stool samples collected from the same individuals will be required to confirm whether the tropism of HBoV2 favors the digestive track and is distinct from that of the largely respiratory HBoV1.

Extensive evidence for recombination was observed through full genome analyses including the likely recombinant origin of HBoV3 and HBoV4 and the high level of intra-species recombination between HBoV2 variants (7,8). The high prevalence of bocavirus infection provides the opportunity for co-infections, the first step in generating recombinant viruses.

Indeed a HBoV3 and HBoV4 co-infection was detected based on the pattern of mixed bases in one directly sequenced PCR product (data not shown).

The frequent detection of HBoVs in stool from both healthy children and adults support the likelihood of long periods of viral shedding and/or frequent re-infections. Whether symptoms such as diarrhea were associated with acute infection in subsequently healthy viral shedders will require quantitative viral load measurements and analysis of longitudinally collected samples. Whether prior infection provides any protection against re-infection with the same or different genotype or species is also unknown.

A borderline statistical association of HBoV2 shedding with diarrhea was reported in Australian children (8) requiring validation by larger studies. In both diarrhea sample sets analyzed here, consisting mostly of adults, no association with HBoV2 shedding was detected. If HBoV2 causes diarrhea, it may do so in only a small subset of infected children, possibly those without passively transferred maternal antibodies or protective immune responses from prior infections. Co-infections with other enteric viruses may also exacerbate symptoms. Given the very large number of infections (viral shedding prevalence of >20% in some countries) even low virulence could translate into a large disease burden.

A trend of an association of HBoV3/HBoV4 detection with AFP was detected. The small numbers of cases will require independent confirmation. The neurological damages caused by bovine and canine bocaviruses in their animal hosts provide a precedent for infant nervous system pathogenicity.

The totality of HBoV1 sequences collected worldwide by multiple groups show very low protein and nucleotide sequence diversity (35,36)(Table 2 and Fig 4). In contrast this single study found substantial diversity among HBoV2 and HBoV3, a fourth species (HBoV4) and extensive viral recombination. Assuming comparable rates of evolution, the genetically homogeneous and largely respiratory HBoV1 therefore appears to be the more recently evolved species relative to the more diverse HBoV2, 3, and 4 found predominantly in feces. We propose that HBoV1 evolved from an enteric bocavirus ancestor that acquired, through mutation and/or recombination, enhanced respiratory tract tropism. Single stranded DNA parvoviruses have been shown to have a mutation rate approaching that of RNA viruses and recombination among animal parvoviruses have been reported (48–51). Parvoviruses also have the demonstrated capacity to rapidly expand their host species tropism resulting in a recent pandemic in dogs (45,52,53). A recent study showed that HBoV1 could replicate in differentiated human airway epithelial cells (44). Whether HBoV2, 3, and 4 show an in vitro tropism and in vivo distribution that is more biased towards digestive track cells will require further studies.

## Acknowledgments

we thank Dr Michael P. Busch and Blood Systems Research Institute for sustained support, Hope Biswas for statistical assistance, and NHLBI grant R01HL083254 to ELD.

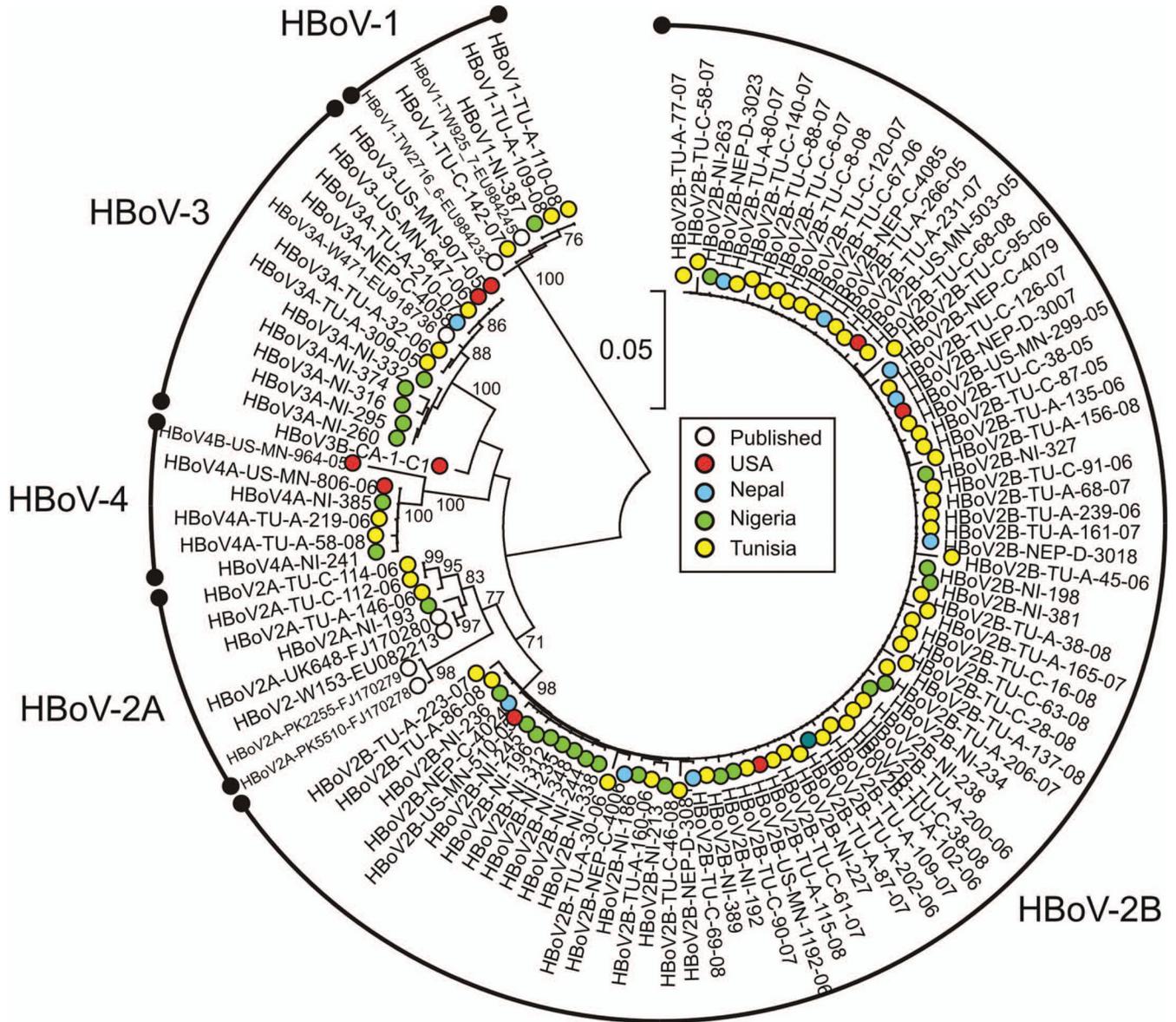
## BIBLIOGRAPHY

1. Bloom, Marshall E.; Young, Neal S. Parvoviruses. Knippe, DM.; Howley, PM., editors. Philadelphia, PA: Lippincott Williams and Wilkins; 2001.
2. Schultz RD. Duration of immunity for canine and feline vaccines: a review. *Vet Microbiol* 2006;117:75–79. [PubMed: 16707236]
3. Patel JR, Heldens JG. Review of companion animal viral diseases and immunoprophylaxis. *Vaccine* 2009;27:491–504. [PubMed: 19041354]
4. Appel MJ. Forty years of canine vaccination. *Adv Vet Med* 1999;41:309–324. [PubMed: 9890024]

5. Jones MS, Kapoor A, Lukashov VV, Simmonds P, Hecht F, Delwart E. New DNA viruses identified in patients with acute viral infection syndrome. *J Virol* 2005;79:8230–8236. [PubMed: 15956568]
6. Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, Andersson B. Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proc Natl Acad Sci U S A* 2005;102:12891–12896. [PubMed: 16118271]
7. Kapoor A, Slikas E, Simmonds P, Chieochansin T, Naeem A, et al. A Newly Identified Bocavirus Species in Human Stool. *J Infect Dis* 2009;199:196–200. [PubMed: 19072716]
8. Arthur JL, Higgins GD, Davidson GP, Givney RC, Ratcliff RM. A novel bocavirus associated with acute gastroenteritis in Australian children. *PLoS Pathog* 2009;5:e1000391. [PubMed: 19381259]
9. Manteufel J, Truyen U. Animal bocaviruses: a brief review. *Intervirology* 2008;51:328–334. [PubMed: 19023216]
10. Allander T. Human bocavirus. *J Clin Virol* 2008;41:29–33. [PubMed: 18055252]
11. Chen KC, Shull BC, Moses EA, Lederman M, Stout ER, Bates RC. Complete nucleotide sequence and genome organization of bovine parvovirus. *J Virol* 1986;60:1085–1097. [PubMed: 3783814]
12. Choi EH, Lee HJ, Kim SJ, Eun BW, Kim NH, et al. The association of newly identified respiratory viruses with lower respiratory tract infections in Korean children, 2000–2005. *Clin Infect Dis* 2006;43:585–592. [PubMed: 16886150]
13. Bastien N, Brandt K, Dust K, Ward D, Li Y. Human Bocavirus infection, Canada. *Emerg Infect Dis* 2006;12:848–850. [PubMed: 16704852]
14. Foulongne V, Rodiere M, Segondy M. Human Bocavirus in children. *Emerg Infect Dis* 2006;12:862–863. [PubMed: 16710957]
15. Kupfer B, Vehreschild J, Cornely O, Kaiser R, Plum G, et al. Severe pneumonia and human bocavirus in adult. *Emerg Infect Dis* 2006;12:1614–1616. [PubMed: 17176591]
16. Ma X, Endo R, Ishiguro N, Ebihara T, Ishiko H, et al. Detection of human bocavirus in Japanese children with lower respiratory tract infections. *J Clin Microbiol* 2006;44:1132–1134. [PubMed: 16517912]
17. Sloots TP, McErlean P, Speicher DJ, Arden KE, Nissen MD, Mackay IM. Evidence of human coronavirus HKU1 and human bocavirus in Australian children. *J Clin Virol* 2006;35:99–102. [PubMed: 16257260]
18. McIntosh K. Human bocavirus: developing evidence for pathogenicity. *J Infect Dis* 2006;194:1197–1199. [PubMed: 17041844]
19. Kesebir D, Vazquez M, Weibel C, Shapiro ED, Ferguson D, et al. Human bocavirus infection in young children in the United States: molecular epidemiological profile and clinical characteristics of a newly emerging respiratory virus. *J Infect Dis* 2006;194:1276–1282. [PubMed: 17041854]
20. Manning A, Russell V, Eastick K, Leadbetter GH, Hallam N, et al. Epidemiological profile and clinical associations of human bocavirus and other human parvoviruses. *J Infect Dis* 2006;194:1283–1290. [PubMed: 17041855]
21. Arden KE, McErlean P, Nissen MD, Sloots TP, Mackay IM. Frequent detection of human rhinoviruses, paramyxoviruses, coronaviruses, and bocavirus during acute respiratory tract infections. *J Med Virol* 2006;78:1232–1240. [PubMed: 16847968]
22. Kahn J. Human bocavirus: clinical significance and implications. *Curr Opin Pediatr* 2008;20:62–66. [PubMed: 18197041]
23. Kantola K, Hedman L, Allander T, Jartti T, Lehtinen P, et al. Serodiagnosis of human bocavirus infection. *Clin Infect Dis* 2008;46:540–546. [PubMed: 18199037]
24. Schildgen O, Muller A, Allander T, Mackay IM, Volz S, et al. Human bocavirus: passenger or pathogen in acute respiratory tract infections? *Clin Microbiol Rev* 2008;21:291–304. table of contents. [PubMed: 18400798]
25. Sloots TP, Whitley DM, Lambert SB, Nissen MD. Emerging respiratory agents: New viruses for old diseases? *J Clin Virol*. 2008
26. Albuquerque MC, Rocha LN, Benati FJ, Soares CC, Maranhao AG, et al. Human bocavirus infection in children with gastroenteritis, Brazil. *Emerg Infect Dis* 2007;13:1756–1758. [PubMed: 18217564]
27. Cheng WX, Jin Y, Duan ZJ, Xu ZQ, Qi HM, et al. Human bocavirus in children hospitalized for acute gastroenteritis: a case-control study. *Clin Infect Dis* 2008;47:161–167. [PubMed: 18532891]

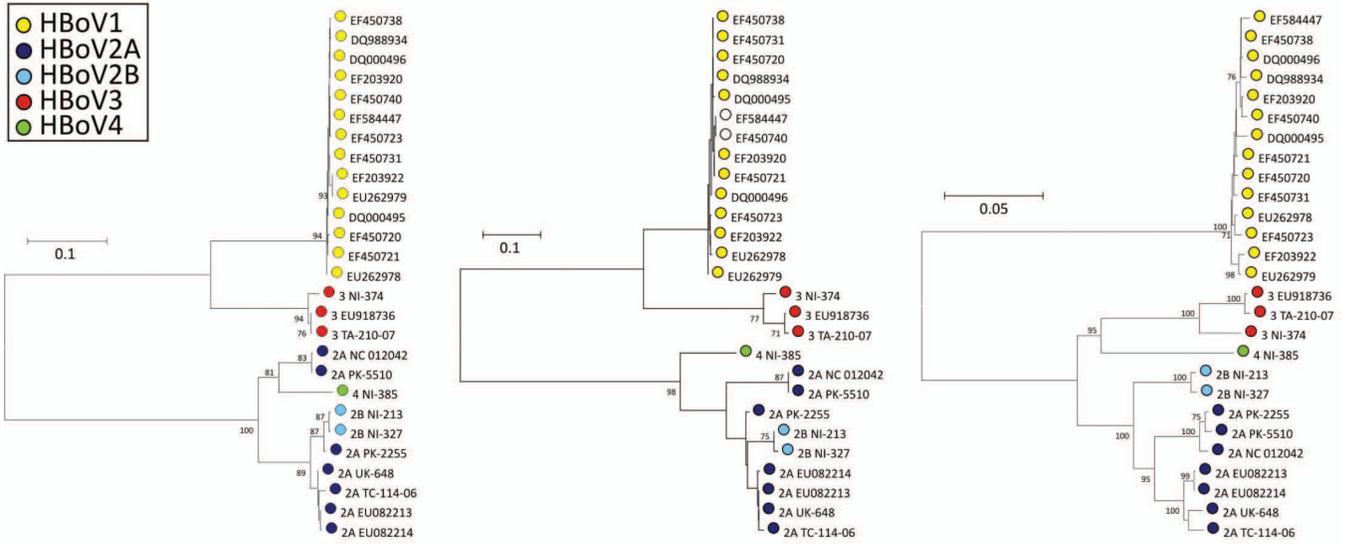
28. Lau SK, Yip CC, Que TL, Lee RA, Au-Yeung RK, et al. Clinical and molecular epidemiology of human bocavirus in respiratory and fecal samples from children in Hong Kong. *J Infect Dis* 2007;196:986–993. [PubMed: 17763318]
29. Lee JI, Chung JY, Han TH, Song MO, Hwang ES. Detection of human bocavirus in children hospitalized because of acute gastroenteritis. *J Infect Dis* 2007;196:994–997. [PubMed: 17763319]
30. Vicente D, Cilla G, Montes M, Perez-Yarza EG, Perez-Trallero E. Human bocavirus, a respiratory and enteric virus. *Emerg Infect Dis* 2007;13:636–637. [PubMed: 17553287]
31. Yu JM, Li DD, Xu ZQ, Cheng WX, Zhang Q, et al. Human bocavirus infection in children hospitalized with acute gastroenteritis in China. *J Clin Virol*. 2008
32. Chieochansin T, Chutinimitkul S, Payungporn S, Hiranras T, Samransamruajkit R, et al. Complete coding sequences and phylogenetic analysis of Human Bocavirus (HBoV). *Virus Res* 2007;129:54–57. [PubMed: 17532505]
33. Ditt V, Viazov S, Tillmann R, Schildgen V, Schildgen O. Genotyping of human bocavirus using a restriction length polymorphism. *Virus Genes* 2008;36:67–69. [PubMed: 18071891]
34. Kumar S, Nei M, Dudley J, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* 2008;9:299–306. [PubMed: 18417537]
35. Simmonds P, Smith DB. Structural constraints on RNA virus evolution. *J Virol* 1999;73:5787–5794. [PubMed: 10364330]
36. Blinkova O, Rosario K, Li L, Kapoor A, Slikas B, et al. Frequent detection of highly diverse Cardioviruses, Cosaviruses, Bocaviruses and Circoviruses in US sewage. *J Clin Microbiol*. 2009
37. Fauquet, CM.; Mayo, MA.; Maniloff, J.; Desselberger, U.; Ball, LA. 8th reports of the International Committee on Taxonomy of Viruses. Academic Press; 2005. Virus taxonomy.
38. Lederman M, Patton JT, Stout ER, Bates RC. Virally coded noncapsid protein associated with bovine parvovirus infection. *J Virol* 1984;49:315–318. [PubMed: 6319731]
39. Sun Y, Chen AY, Cheng F, Guan W, Johnson FB, Qiu J. Molecular characterization of infectious clones of the minute virus of canines reveals unique features of bocaviruses. *J Virol* 2009;83:3956–3967. [PubMed: 19211770]
40. Qiu J, Cheng F, Johnson FB, Pintel D. The transcription profile of the bocavirus bovine parvovirus is unlike those of previously characterized parvoviruses. *J Virol* 2007;81:12080–12085. [PubMed: 17715221]
41. Dijkman R, Koekkoek SM, Molenkamp R, Schildgen O, van der Hoek L. Human bocavirus can be cultured in differentiated human airway epithelial cells. *Virol* 2009;83:7739–7748.
42. Parrish CR, Kawaoka Y. The origins of new pandemic viruses: the acquisition of new host ranges by canine parvovirus and influenza A viruses. *Annu Rev Microbiol* 2005;59:553–586. [PubMed: 16153179]
43. Chieochansin T, Kapoor A, Delwart E, Poovorawan Y, Simmonds P. Absence of detectable replication of the human bocavirus species 2 in the respiratory tract despite frequent detection in faecal samples. *Emerg Infect Dis* 2009;15:1503–1505. [PubMed: 19788826]
44. Han T-H, Chung J-Y, Hwang E-S. Human Bocavirus 2 in Children, South Korea. *Emerg Infect Dis* 2009;15:1698–1699. [PubMed: 19861084]
45. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet* 2008;9:267–276. [PubMed: 18319742]
46. Norja P, Eis-Hubinger AM, Soderlund-Venermo M, Hedman K, Simmonds P. Rapid sequence change and geographical spread of human parvovirus B19; comparison of B19 evolution in acute and persistent infections. *J Virol*. 2008
47. Shackelton LA, Hoelzer K, Parrish CR, Holmes EC. Comparative analysis reveals frequent recombination in the parvoviruses. *J Gen Virol* 2007;88:3294–3301. [PubMed: 18024898]
48. Lukashov VV, Goudsmit J. Evolutionary relationships among parvoviruses: virus-host coevolution among autonomous primate parvoviruses and links between adeno-associated and avian parvoviruses. *J Virol* 2001;75:2729–2740. [PubMed: 11222696]
49. Hoelzer K, Shackelton LA, Parrish CR, Holmes EC. Phylogenetic analysis reveals the emergence, evolution and dispersal of carnivore parvoviruses. *J Gen Virol* 2008;89:2280–2289. [PubMed: 18753238]

50. Shackelton LA, Parrish CR, Truyen U, Holmes EC. High rate of viral evolution associated with the emergence of carnivore parvovirus. *Proc Natl Acad Sci U S A* 2005;102:379–384. [PubMed: 15626758]

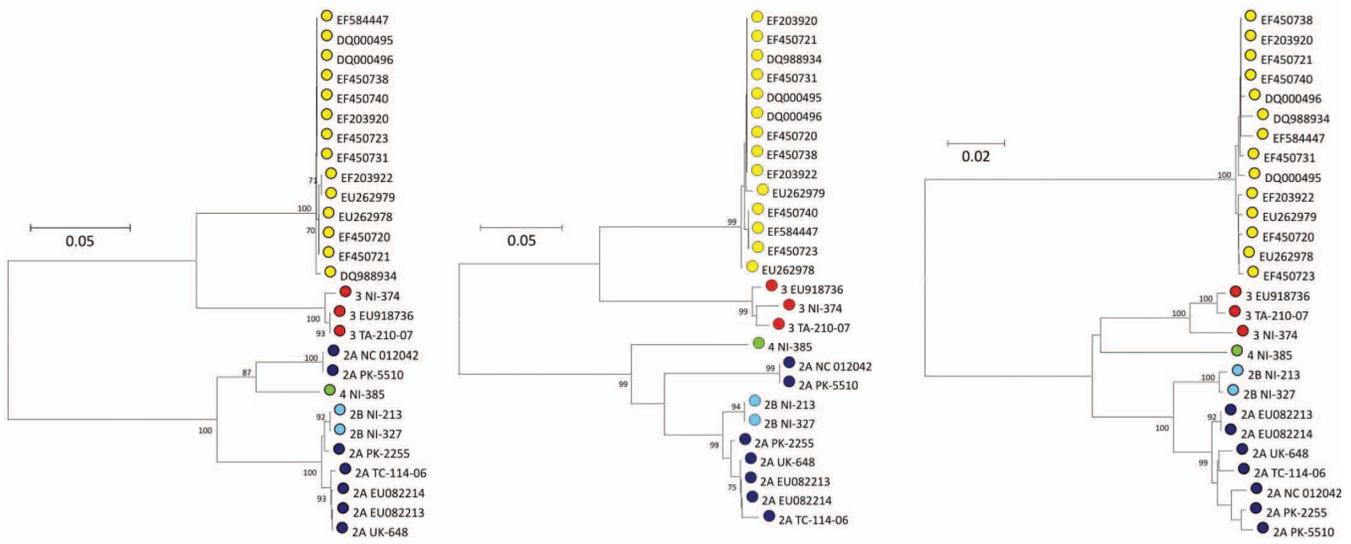


**Figure 1.** Phylogenetic analysis of partial VP1 pan-bocavirus PCR amplicons of different species of human bocaviruses. Tree constructed by neighbour-joining of pairwise maximum composite likelihood distances between nucleotide sequences; bootstrap values  $\geq 70\%$  shown.

Nucleotides, MCL distances



Amino acid, PAM distances

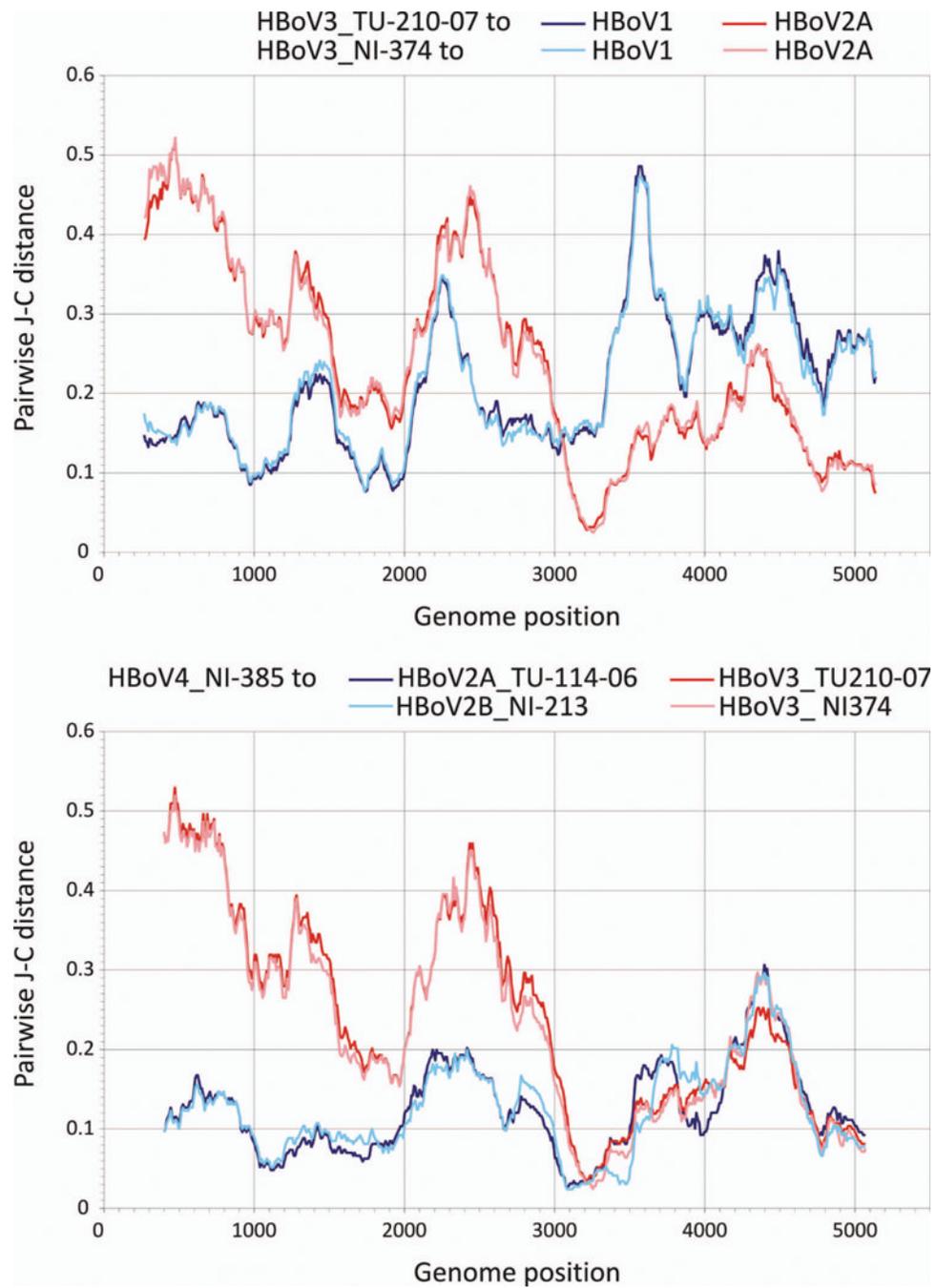


NS1 (253-2598)

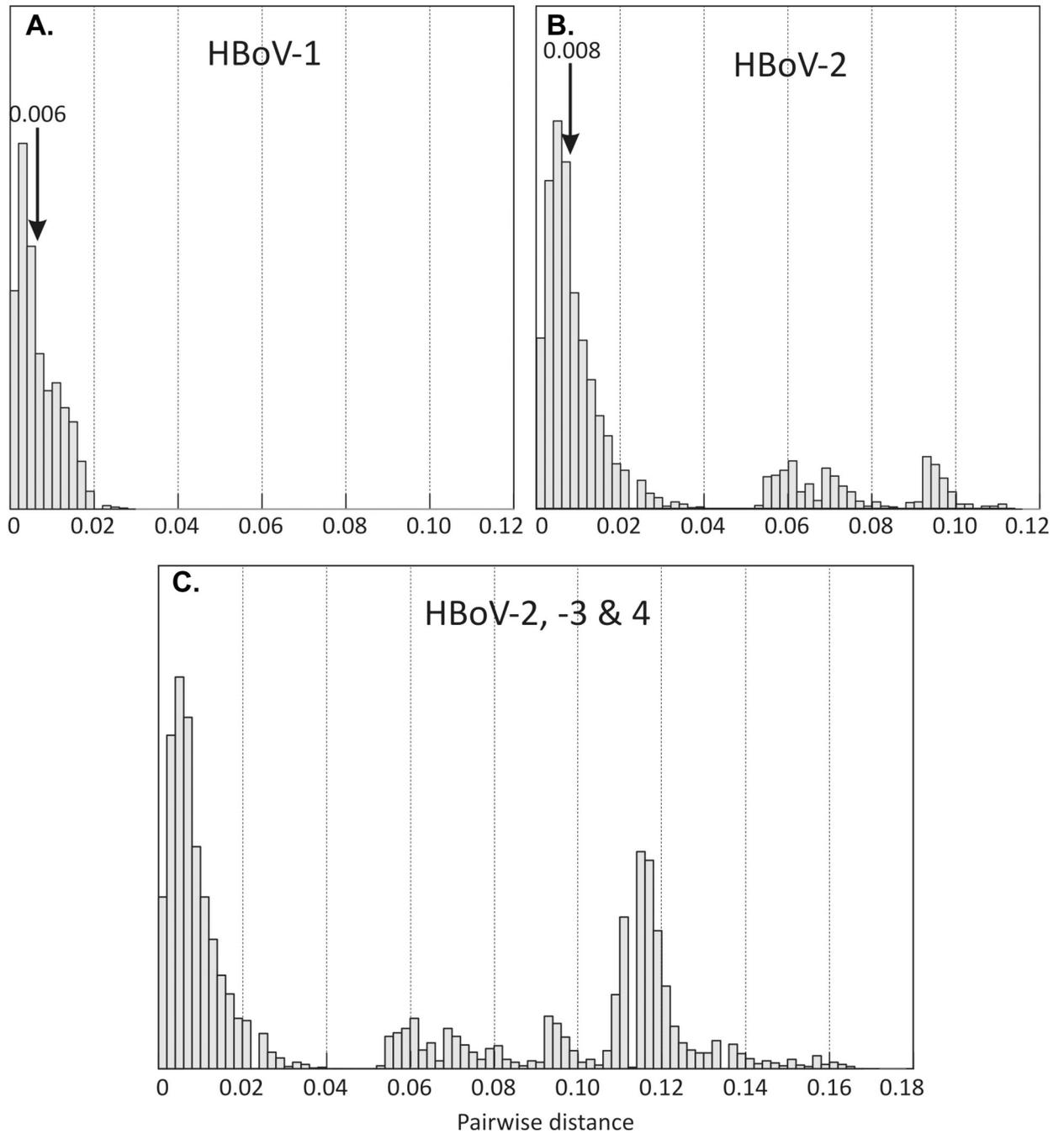
NP1 (2410-3069)

VP1(3056-5071)

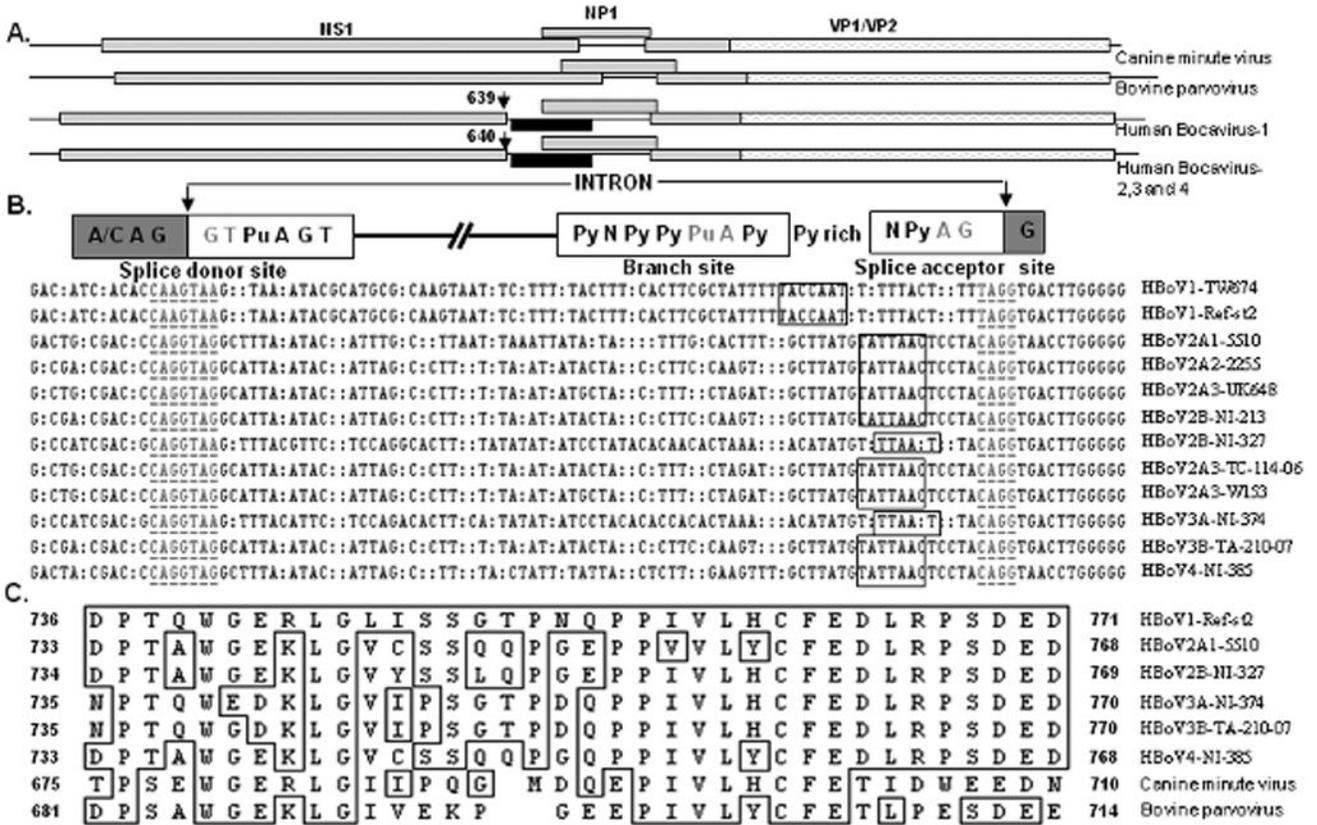
**Figure 2.** Phylogenetic analyses of nucleotide (upper panel) and inferred amino acid sequences of the three principal open reading frames of HBoV1-4; bootstrap values  $\geq 70\%$  shown. The NS1 trees used HBoV sequences spliced as described.



**Figure 3.** HBoV3 and HBoV4 are recombinant genomes. Pair-wise Jukes-Cantor corrected distance scans of HBoV3 and HBoV4 sequences against presrepresentative sequences of HBoV1 and 2A, or individual sequences as indicated in legend.



**Figure 4.** Representation of genetic diversity within and between different bocavirus species. Distribution of pair-wise uncorrected p-distances in the partial VP1 region amplified using pan-bocavirus PCR among A) HBoV1, B) HBoV2 and C) HBoV2-4 combined.



**Figure 5.** Evidence for RNA splicing in HBoVs genome to generate longer NS1 proteins. (A) proposed genomic organization of HBoVs relative to that of animal bocaviruses. Black boxes represent exon 2 of NS1 protein. (B) Canonical sequence of splice donor, branch site and splice acceptor (Pu=A/G, Py=T/C). All HBoVs species were aligned to show presence of putative RNA splicing elements in the NS exon. (c) region of highest similarity between extended carboxy-termini of NS1 of HBoV1-4 species and termini of animal bocaviruses NS1. The amino acid positions used in alignment are shown.

**Table 1**

Distribution of Bocavirus species and genotypes in different countries and cohorts

Country Symptoms Age group	Nigeria AFP Children	Tunisia AFP Children	Tunisia Healthy contacts Children	Nepal Diarrhea Adults	Nepal Healthy Adults	US Diarrhea Adults	US Healthy Adults	TOTAL
total # tested	96	96	96	96	96	87	74	641
total # positive (%)	28 (29%)	32 (33%)	24 (25%)	4 (4%)	5 (5%)	6 (7%)	2 (3%)	101
# HBov1	1	2	1	0	0	0	0	4
# HBov2A	1	1	2	0	0	0	0	4
# HBov2B	19	24	21	4	4	4	0	76
# HBov3	5	3	0	0	1	1	1	11
# HBov4	2	2	0	0	0	1	1	6

TABLE 2

Genetic distances between and within bocavirus species and genotypes.

A Number of pairwise comparisons						
	HBov1	HBov2	HBov2A	HBov2B	HBov3	HBov4
HBov1	91					
HBov2	126					
HBov2A	98	21				
HBov2B	28	14	1			
HBov3	42	27	6	3		
HBov4	14	9	2	3	0	

B % divergence within species and genotype for each gene [mean (min-max)]						
Nucleotide	Amino acid			Amino acid top, nucleotide below.		
	NS1	NP1	VP1/2	NS1	NP1	VP1/2
HBov1	0.33 (0.04-0.64)	0.38 (0-0.91)	0.84 (0.1-1.64)	0.14 (0-0.51)	0.29 (0-0.91)	0.43 (0-0.89)
HBov2A	4.61 (0.04-8.59)	3.92 (0-7.44)	3.19 (0.05-5.19)	4.45 (0.13-9.14)	6.09 (0-12.33)	1.445 (0-2.25)
HBov2B	0.26 (0.26-0.26)	0.15 (0.15-0.15)	0.55 (0.55-0.55)	0 (0-0)	0 (0-0)	0.45 (0.45-0.45)
HBov3	0.89 (0.13-1.29)	2.44 (0.46-3.5)	3.09 (0.4-4.43)	0.6 (0-0.9)	2.12 (1.36-2.73)	2.09 (0.6-2.84)
HBov4						

C % divergence between species and genotype for each gene [mean (min-max)]. Amino acid top, nucleotide below.						
Nucleotide	Amino acid			Amino acid top, nucleotide below.		
	NS1	NP1	VP1/2	NS1	NP1	VP1/2
NS1			25.5 (25.2-25.9)	25.6 (025.5-25.8)	12.0 (11.8-12.3)	25.3 (25.3-25.5)
HBov1						
HBov2	26.7 (26.3-26.9)				25.5 (24.8-26.2)	11.4 (6.8-12.6)
HBov2A	26.6 (26.3-26.9)			34.7 (0.8-9.3)	25.4 (24.8-26.2)	10.8 (6.8-12.6)
HBov2B	26.8 (26.6-26.9)		4.0 (0.8-8.6)		25.5 (25.3-25.8)	12.1 (12.1-12.1)
HBov3	13.9 (13.6-14.4)	16.2 (25.7-26.6)	25.9 (25.7-26.3)	26.5 (26.4-26.6)		26.8 (26.7-26.9)
HBov4	26.9 (26.8-27.0)	9.8 (6.7-10.4)	9.3 (6.7-10.4)	10.3 (10.3-10.3)	26.6 (26.6-26.6)	
NP1			32.1 (30.9-35.2)	31.3 (30.9-31.3)	19.0 (18.2-20.0)	31.4 (31.0-31.5)
HBov2	23.5 (22.5-24.5)				31.4 (30.6-32.9)	16.7 (14.9-17.7)

C % divergence between species and genotype for each gene [mean (min-max)]. Amino acid top, nucleotide below.

	HBov2A	23.5 (22.5–24.5)					6.9 (3.7–14.4)	31.6 (30.6–32.9)	16.7 (14.9–17.7)	
	HBov2B	23.5 (23.0–23.9)				5.5 (3.3–9.8)		31.3 (30.9–31.8)	16.7 (16.7–16.7)	
	HBov3	13.8 (13.4–14.2)	24.8 (23.6–25.6)			24.3 (23.6–25.0)	25.4 (25.3–25.6)		34.1 (33.8–34.3)	
	HBov4	22.7 (22.3–22.9)	12.1 (9.6–13.0)			11.3 (9.6–12.1)	12.9 (12.9–13.0)	24.7 (24.1–25.1)		
VP1/2	HBov1		19.7 (19.3–20.7)			19.9 (19.5–20.7)	19.5 (19.3–19.8)	19.7 (19.5–20.0)	19.4 (19.3–19.8)	
	HBov2	21.8 (21.4–22.3)						9.2 (8.8–9.9)	9.5 (9.1–10.5)	
	HBov2A	21.9 (21.4–22.3)					3.9 (3.3–4.5)	9.4 (8.8–9.9)	9.8 (9.3–10.5)	
	HBov2B	21.7 (21.6–22.0)				6.2 (5.7–7.1)		9.0 (8.8–9.3)	9.2 (9.1–9.3)	
	HBov3	22.9 (22.5–23.3)	12.4 (11.6–13.5)			12.9 (12.3–13.5)	11.8 (11.6–12.2)		8.5 (8.1–9.3)	
	HBov4	22.3 (22.2–22.5)	12.3 (11.9–12.9)			12.6 (12.3–12.9)	12.0 (11.9–12.0)	11.9 (11.6–12.4)		
		HBov1		HBov2		HBov2A		HBov2B		HBov3
										HBov4