MAJOR ARTICLE

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

Amit Kapoor,[1,2] Peter Simmonds,[3] Elizabeth Slikas,[1,2] Linlin Li,[1,2] Ladaporn Bodhidatta,[4] Orntipa Sethabutr,[4] Henda Triki,[5] Olfa Bahri,[5] Bamidele S. Oderinde,[6] Marycelin M. Baba,[6] David N. Bukbuk,[7] John Besser,[7] Joanne Bartkus,[7] and Eric Delwart[1,2]

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

A new species of parvovirus, tentatively named human bocavirus 4 (HBoV4), was genetically characterized. Among 641 feces samples obtained from children and adults, the most commonly detected bocavirus species were, in descending order, HBoV2, HBoV3, HBoV4, and HBoV1, with an HBoV2 prevalence of 21% and 26% in Nigerian and Tunisian children, respectively. HBoV3 or HBoV4 species were found in 12 of 192 patients with non-polio acute flaccid paralysis in Tunisia and Nigeria and 0 of 96 healthy Tunisian contacts ($P = .01$). Evidence of extensive recombination at the NP1 and VP1 gene boundary between and within bocavirus species was found. The high degree of genetic diversity seen among the human bocaviruses found in feces specimens, relative to the highly homogeneous HBoV1, suggest that this worldwide-distributed respiratory pathogen may have recently evolved from an enteric bocavirus after acquiring an expanded tropism favoring the respiratory tract. Elucidating the possible role of the newly identified enteric bocaviruses in human diseases, including acute flaccid paralysis and diarrhea, will require further epidemiological studies.

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 a febrile 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 (AFP) [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 [9, 10].

The first bocavirus identified was in cows [10], and the name of the genus is derived from its first known hosts (bovine-canine). Animal bocaviruses can cause both respiratory and gastrointestinal diseases, as well as embryonic and fetal death [11].

HBoV1 infection has been linked with mild-to-severe, primarily lower respiratory tract infections in children, frequently in association with other viral infections [12–25]. HBoV1 has also been detected in stool

Received 23 September 2009; accepted 3 November 2009; electronically published 22 April 2010.

Financial support: Blood Systems Research Institute and the National Heart, Lung and Blood Institute (grant R01HL083254 to E.L.D.).

Potential conflicts of interest: A provisional patent application has been filed for HBoV2–3-4.

Reprints or correspondence: Eric Delwart, Blood Systems Research Institute, 270 Masonic Ave, San Francisco, CA 94118 (delwarte@medicine.ucsf.edu).

The Journal of Infectious Diseases 2010; 201(11):1633–1643

0022-1899/2010/20111-0006$15.00
DOI: 10.1086/652416

samples, although association with diarrhea appears weaker than with respiratory symptoms [26–31]. HBoV1 strains show a very low degree of genetic variability worldwide [32, 33].

In this study, pan-bocavirus polymerase chain reaction (PCR) primers were designed and used on human stool samples from several countries. All 3 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 were anonymized. This study was approved by the University of California San Francisco Committee on Human Research. Samples from Nigeria and Tunisia were collected as part of the World Health Organization'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 case patients from Tunisia were matched for age. Stool samples from Nepal were obtained from adult travelers and resident expatriates with diarrhea with no known pathogens detected by standard microbiologic testing for enteric bacteria; enzyme immunoassay for rotavirus, adenovirus, astrovirus, *Giardia* and *Cryptosporidium;* and reverse transcription PCR for norovirus. Stool samples from healthy, asymptomatic control persons 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 3233–3808, numbered here and subsequently with use of the HBoV2 prototype sequence; GenBank accession number FJ170278). Nucleic acids (both DNA and RNA) were extracted (QIAamp Viral RNA mini kit) from 140 μL of clarified stool supernatant and were eluted into 60 μL of water. First-round PCR primers were AK-VP-F1 (5′-CGCCGTGGCTCCTGCTCT-3′) and AK-VP-R1 (5′-TGTTCGCCATCACAAAAGATGTG-3′) and second-round primers were AK-VP-F2 (5′-GGCTCCTGCTCTAGGAAATAAAGAG-3′) and AK-VP-R2 (5′-CCTGCTGTTAGGTCGTTGTTGTATGT-3′). PCR reactions contained 2.5 U of Taq DNA polymerase (NEB) in 1.1× Thermopol reaction buffer with MgCl (2.0 mmol/L), 20 pmol/L (each) of forward and reverse primers, and 2.5 μL of nucleic acids (first round) or 1 μL of the first-round PCR product (second round) as template in a 50 μL total volume. First-round conditions were 10 cycles at 95°C for 35 s, 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 at 95°C for 30 s, 54°C for 45 s, and 72°C for 45 s; and a final extension at 72°C for 10 min. Similar conditions were used for the second PCR round, except that the initial annealing temperatures were 60°C and 58°C in the first and second group of PCR cycles, respectively. Amplicons of the appropriate size, as 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 4 bocavirus species and was estimated at 10–100 plasmid copies for each species.

Complete genome sequencing. Nearly complete genomes were amplified using PCR primers designed from alignments of HBoV1 and HBoV2 genomes and were then directly sequenced by primer walking [7]. The terminal sequences were acquired by a modified protocol for rapid amplification of complementary DNA ends [5]. The terminal sequences are incomplete because of extensive hairpin structures that prevented extensions to the viral 5′ and 3′ extremities.

Distance measurements and phylogenetic analyses. Phylogenetic relationships were evaluated with use of Mega 4.1 (http://www.megasoftware.net/mega41.html) [34]. Neighbor-joining trees were inferred using a matrix of pair-wise 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 [35]. 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 HBoV1 and HBoV2, a mean pair-wise distance was computed using a set of 14 HBoV1 and 7 HBoV2A sequences.

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

Widespread 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 ~576-nucleotide fragment of the VP1 capsid gene. Nucleic acids from human stools collected from Nigeria, Tunisia, Nepal, and the United States were then analyzed using this nested PCR (Table 1). Of the 641 samples tested, 101 (16%) had confirmed positive results by PCR sequencing, with the highest prevalence in Tunisian children with AFP

Table 1. Distribution of Human Bocavirus (HBoV) Species and Genotypes in Different Countries and Cohorts

Variable	Country, cohort							
	Nigeria	Tunisia cohort 1	Tunisia cohort 2	Nepal cohort 1	Nepal cohort 2	United States cohort 1	United States cohort 2	Total
Symptom	Acute flaccid paralysis	Acute flaccid paralysis	Healthy contacts	Diarrhea	Healthy	Diarrhea	Healthy	…
Age group	Children	Children	Children	Adults	Adults	Adults	Adults	…
No tested	96	96	96	96	96	87	74	641
No (%) positive for HBoV	28 (29)	32 (33)	24 (25)	4 (4)	5 (5)	6 (7)	2 (3)	101 (16)
No of patients infected								
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

(33%). To determine the phylogenetic relationship of these strains, the sequences were aligned with available sequences of HBoV1, HBoV2, and HBoV3. Only 4 of 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 4 distinct genetic lineages, labeled HBoV2A, HBoV2B, HBoV3, HBoV4 (Figure 1).

The HBoV2A clade included all published HBoV2 genotypes [7], plus 4 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 reduce splitting of HBoV2 into a multitude of genotypes, all strains in that cluster were reclassified 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 identity (>98%) to the Pakistani HBoV2 prototype PK5510 (FJ170278).

The HBoV2B cluster included 76 of the 101 bocavirus strains reported here (Figure 1). Two nearly full HBoV2B genome sequences demonstrated a pair-wise amino acid divergence in VP1 of 0.45% (Table 2) (NI-213 and NI-327; GenBank accession numbers FJ973560 and FJ973559). Greater divergence was observed when these strains were compared with the VP1 of HBoV2A strains (average divergence, 3.9% [range, 3.3%–4.5%] (Table 2). The nucleotide and amino acid distances at the other loci are shown in Table 2.

The HBoV3 cluster included 11 strains, and the genomes of 2 representative strains, NI-374 and TU-A-210–07, were sequenced (GenBank accession numbers FJ973563 and FJ973562). When compared with 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 with HBoV1, the HBoV3 strains showed an average amino acid divergence of 12% in the NS1 region and 20% in the VP1 region (Table 2). All 11 strains were classified as members of the HBoV3 species, recently described by Arthur et al [8]. A distantly related variant of HBoV3 was also recently identified in a sewage sample in the United States (HBoV3B-CA-1-C1; Figure 1) [36]. To confirm whether this variant represents a second genotype of HBoV3 will require full genome sequencing.

The HBoV4 cluster included 6 strains. The complete genome of 1 representative strain (NI-385) was obtained (GenBank accession number FJ973561). The NS1 protein of HBoV4 (NI-385) showed an average of 11% divergence with HBoV2 and 25%–27% divergence with HBoV1 or HBoV3 (Table 2). For the VP1 protein, the relative divergences were reversed, with an average divergence of 8.5% with HBoV3, 9.5% with HBoV2, and 19% with HBoV1. NI-385, therefore, also appears to be a recombinant, with the 5′ genes NS1 and NP1 closely related to HBoV2 (particularly genotype A) and the 3′ VP1 gene slightly more similar to HBoV3 than HBoV2 (Table 2). According to the International Committee on Taxonomy of Viruses species demarcation criteria in the genus *Bocavirus,* members of different species must show >5% divergence in their nonstructural gene nucleotide sequences [37]. 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 International Committee on Taxonomy of Viruses review, as the prototype of a fourth HBoV species (HBoV4). A distant VP1 (partial) variant of HBoV4 was also detected in the United States (US-MN-964–05; Figure 1). To determine whether 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 3 large open reading frames (ORF). The left ORF encodes the nonstructural 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 [38]; NP1 differed in length between species, ranging from 214 to 219 amino acids. Situated within the VP1-unique (VP1u)

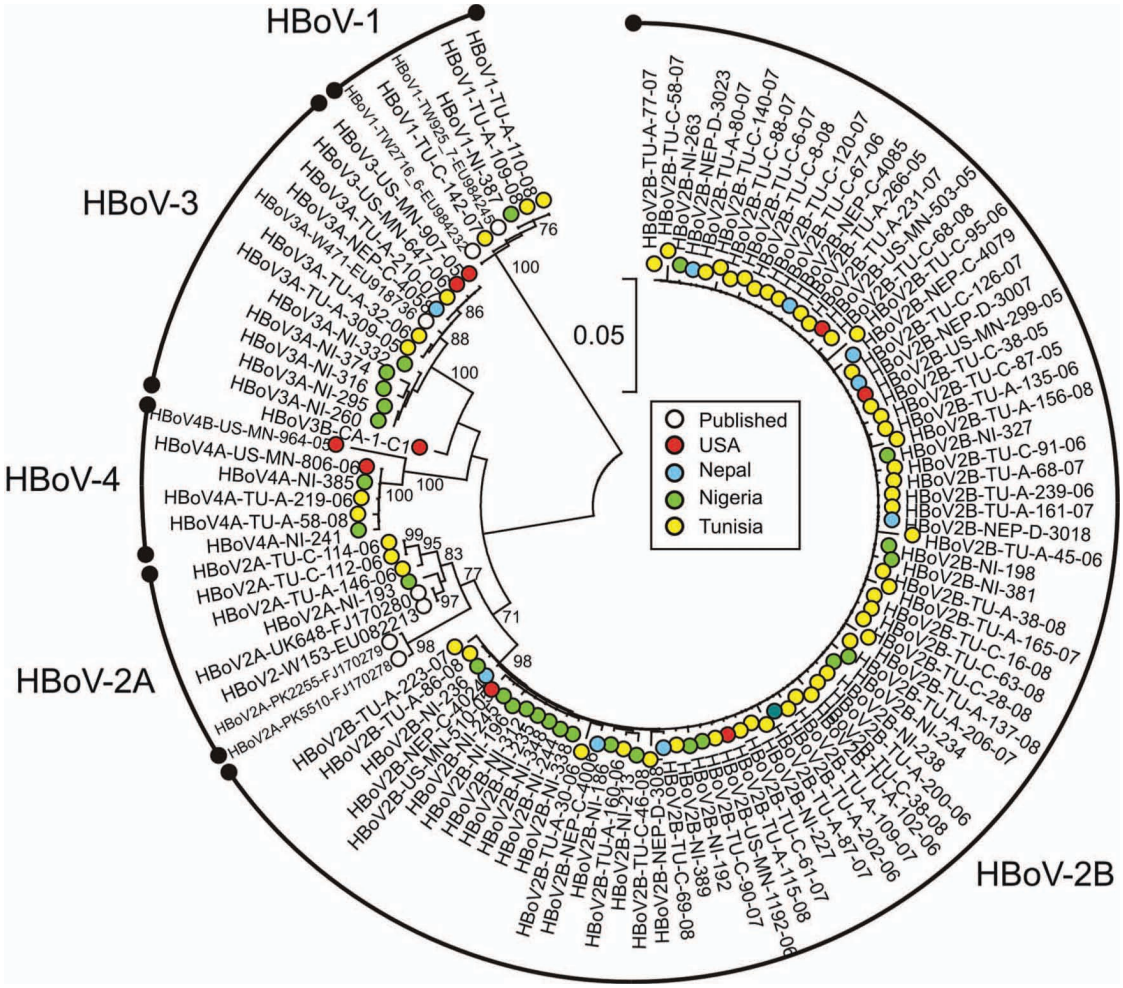

Figure 1. Phylogenetic analysis of partial VP1 pan-bocavirus polymerase chain reaction amplicons of different species of human bocaviruses. Tree constructed by neighbor-joining of pair-wise maximum composite likelihood distances between nucleotide sequences; bootstrap values ⩾70% are shown.

region, the phospholipase A2 motifs required for parvovirus infectivity were found in all 6 genomes, together with the calcium-binding loop and catalytic residues.

Further evidence of recombination in human bocaviruses. To further determine the relationship between members of the *Bocavirus* genus, phylogenetic analyses of NS1, NP1, and VP1/VP2 were performed, with use of both nucleotide sequences and deduced protein sequences (Figure 2). The NS1 and NP1 genes of HBoV3 clustered with HBoV1, whereas 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 (Figure 2).

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 recombination of HBoV2 and HBoV3 in the generation of HBoV4, with both recombination points near the NP1 and VP1 junction (Figure 3). When different HBoV2 variants were similarly analyzed for recombination, intraspecies HBoV2 recombinants were also detected (data not shown) [7].

Diversity among respiratory HBoV1 and enteric HBoV2–4. We compared the intraspecies diversity of HBoV1 with that of HBoV2 by use of the partial VP1 sequence data generated with the pan-bocavirus PCR primers available for the greatest number of HBoV2, 3, 4 variants (Figure 4*A* and 4*B*). A very low average pair-wise difference was seen for HBoV1 collected worldwide (Figure 4*A*). 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 in Figure 4*A*). The homogeneity of HBoV1 and HBoV2B can also be visualized in the small branch lengths in Figure 1. When the pair-wise dis-

Table 2. Genetic Distances Between and Within Human Bocavirus (HBoV) Species and Genotypes

	Genotype, no of pair-wise comparisons					
Genotype	HBoV1	HBoV2	HBoV2A	HBoV2B	HBoV3	HBoV4
HBoV1	91	…	…	…	…	…
HBoV2	126	…	…	…	…	…
HBoV2A	98	…	21	…	…	…
HBoV2B	28	…	14	1	…	…
HBoV3	42	27	21	6	3	…
HBoV4	14	9	7	2	3	0
	Divergence within species and genotype for each gene, mean % (minimum-maximum)					
	Nucleotide			Amino acid		
	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	…	…	…	…	…	…
	Divergence between species and genotypes for each gene, mean % (minimum-maximum)					
Gene, genotype	HBoV1	HBoV2	HBoV2A	HBoV2B	HBoV3	HBoV4
NS1						
HBoV1	…	…	25.5 (25.2–25.9)[a]	25.6 (25.5–25.8)[a]	12.0 (11.8–12.3)[a]	25.3 (25.3–25.5)[a]
HBoV2	26.7 (26.3–26.9)[b]	…	…	…	25.5 (24.8–26.2)[a]	11.4 (6.8–12.6)[a]
HBoV2A	26.6 (26.3–26.9)[b]	…	…	34.7 (0.8–9.3)[a]	25.4 (24.8–26.2)[a]	10.8 (6.8–12.6)[a]
HBoV2B	26.8 (26.6–26.9)[b]	…	4.0 (0.8–8.6)[b]	…	25.5 (25.3–25.8)[a]	12.1 (12.1–12.1)[a]
HBoV3	13.9 (13.6–14.4)[b]	16.2 (25.7–26.6)[b]	25.9 (25.7–26.3)[b]	26.5 (26.4–26.6)[b]	…	26.8 (26.7–26.9)[a]
HBoV4	26.9 (26.8–27.0)[b]	9.8 (6.7–10.4)[b]	9.3 (6.7–10.4)[b]	10.3 (10.3–10.3)[b]	26.6 (26.6–26.6)[b]	…
NP1						
HBoV1	…	32.1 (30.9–35.2)[a]	32.9 (31.2–35.2)[a]	31.3 (30.9–31.3)[a]	19.0 (18.2–20.0)[a]	31.4 (31.0–31.5)[a]
HBoV2	23.5 (22.5–24.5)[b]	…	…	…	31.4 (30.6–32.9)[a]	16.7 (14.9–17.7)[a]
HBoV2A	23.5 (22.5–24.5)[b]	…	…	6.9 (3.7–14.4)[a]	31.6 (30.6–32.9)[a]	16.7 (14.9–17.7)[a]
HBoV2B	23.5 (23.0–23.9)[b]	…	5.5 (3.3–9.8)[b]	…	31.3 (30.9–31.8)[a]	16.7 (16.7–16.7)[a]
HBoV3	13.8 (13.4–14.2)[b]	24.8 (23.6–25.6)[b]	24.3 (23.6–25.0)[b]	25.4 (25.3–25.6)[b]	…	34.1 (33.8–34.3)[a]
HBoV4	22.7 (22.3–22.9)[b]	12.1 (9.6–13.0)[b]	11.3 (9.6–12.1)[b]	12.9 (12.9–13.0)[b]	24.7 (24.1–25.1)[b]	…
VP1/2						
HBoV1	…	19.7 (19.3–20.7)[a]	19.9 (19.5–20.7)[a]	19.5 (19.3–19.8)[a]	19.7(19.5–20.0)[a]	19.4 (19.3–19.8)[a]
HBoV2	21.8 (21.4–22.3)[b]	…	…	…	9.2 (8.8–9.9)[a]	9.5 (9.1–10.5)[a]
HBoV2A	21.9 (21.4–22.3)[b]	…	…	3.9 (3.3–4.5)[a]	9.4 (8.8–9.9)[a]	9.8 (9.3–10.5)[a]
HBoV2B	21.7 (21.6–22.0)[b]	…	6.2 (5.7–7.1)[b]	…	9.0 (8.8–9.3)[a]	9.2 (9.1–9.3)[a]
HBoV3	22.9 (22.5–23.3)[b]	12.4 (11.6–13.5)[b]	12.9 (12.3–13.5)[b]	11.8 (11.6–12.2)[b]	…	8.5 (8.1–9.3)[a]
HBoV4	22.3 (22.2–22.5)[b]	12.3 (11.9–12.9)[b]	12.6 (12.3–12.9)[b]	12.0 (11.9–12.0)[b]	11.9 (11.6–12.4)[b]	…

[a] Amino acid comparison.
[b] Nucleotide comparison.

tances of all the enteric (non-HBoV1) sequences were plotted (Figure 4*C*), the distribution was much larger than that for HBoV1. The low level of intraspecies diversity of HBoV1 relative to the other species is also reflected in Table 2.

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 (630–650 amino acids), compared with the longer NS1 of animal bocaviruses (716–726 amino acids) (Figure 5*A*). We noticed in all 4 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 (Figure 5*A*). 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 (Figure

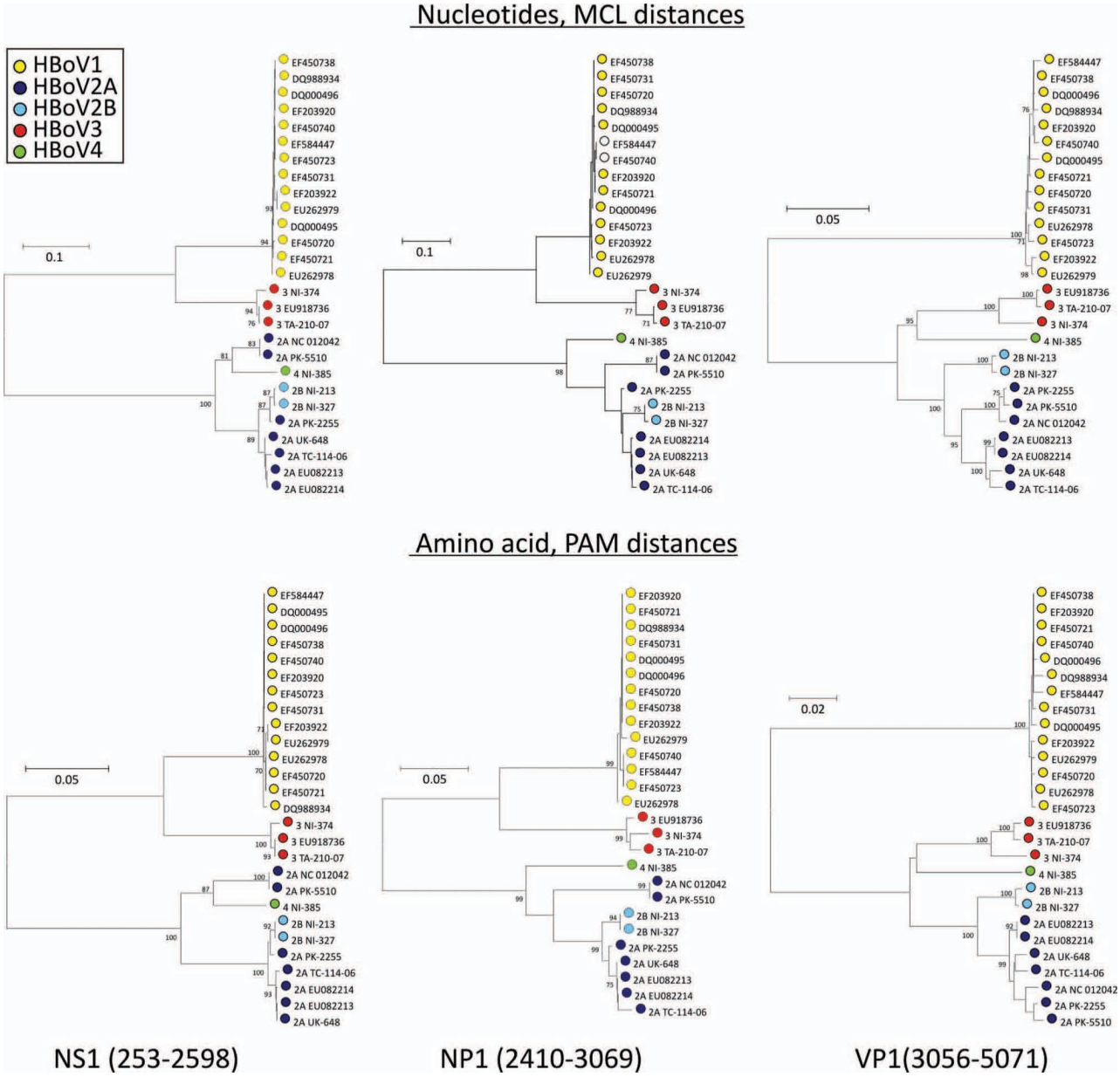

Figure 2. Phylogenetic analyses of nucleotide (*upper panel*) and inferred amino acid sequences of the 3 principal open reading frames of human bocavirus (HBoV) 1–4; bootstrap values ⩾70% are shown. The NS1 trees used HBoV sequences spliced as described in the text.

5*B*). The putative NS1 resulting from such a spliced transcript encoded a 750–780 amino acid protein with a carboxy terminus that showed significant similarity to that of the canine and bovine bocaviruses NS1 (Figure 5*C*) [39, 40]. A study using RT-PCR for the detection of HBoV1 viral transcript in human lung epithelial cells did not detect the NS splicing proposed here [41]. 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).

HBoV2–4 disease association. Most PCR-positive stool samples contained HBoV2B (76 of 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. HBoV1 was detected in only 4 of the 101 positive samples. Stool samples from patients with AFP in Nigeria and Tunisia both showed a high prevalence of HBoV2A+B (21%–26%). Because both patients with AFP and healthy contacts from Tunisia showed a comparable prevalence of HBoV2, no asso-

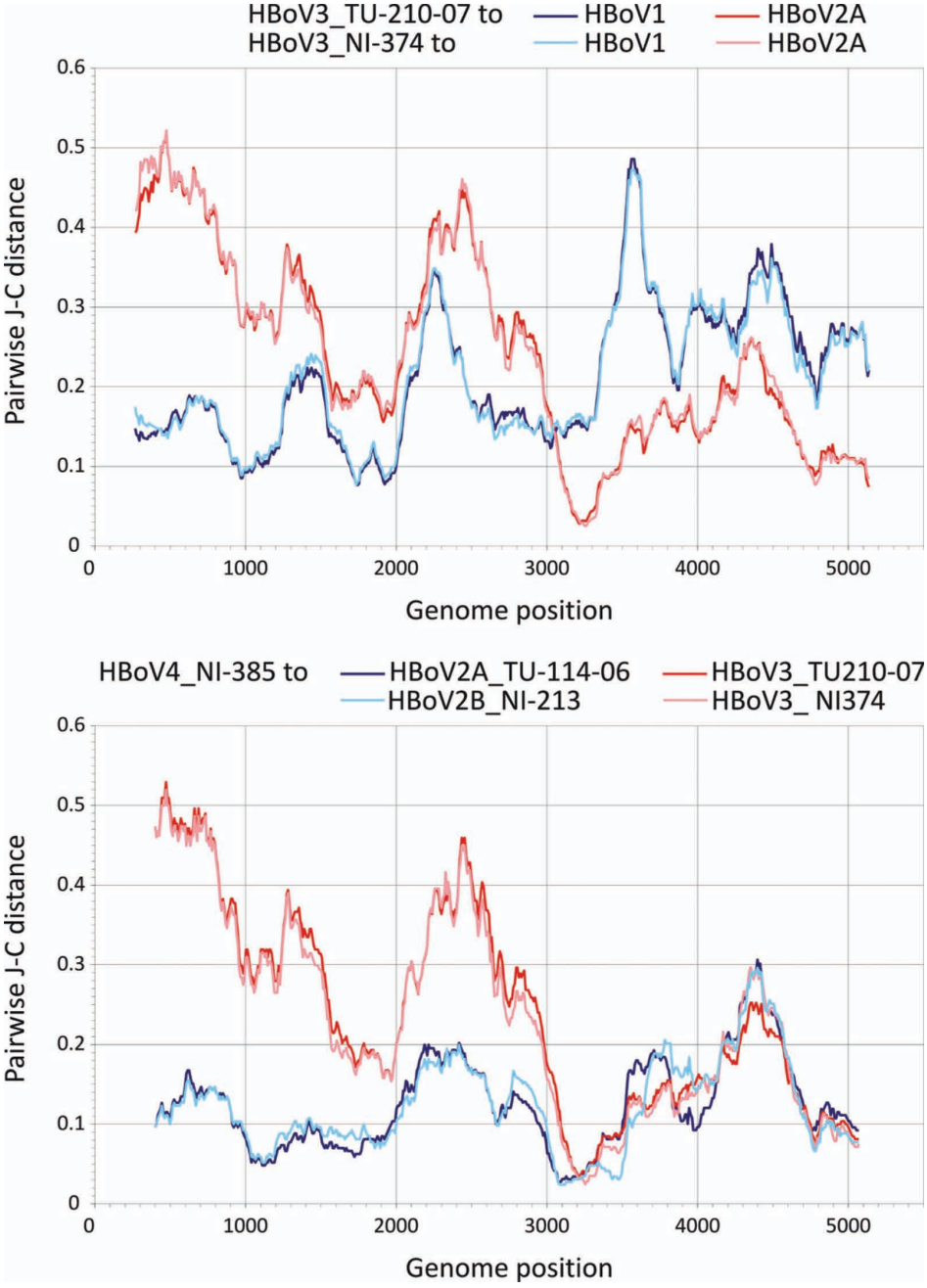

Figure 3. Human bocavirus (HBoV) 3 and HBoV4 have recombinant genomes. Pair-wise Jukes-Cantor corrected distance scans of HBoV3 and HBoV4 sequences against representative sequences of HBoV1 and 2A or individual sequences, as indicated in legend.

ciation was observed between AFP and HBoV2 excretion. The prevalence of HBoV2 in adults with diarrhea from the United States and Nepal were also compared with those in healthy matched subjects. No associations were observed between HBoV2 shedding and diarrhea. Although the numbers of HBoV3 and HBoV4 detected were relatively small, it was noticed that the 8 HBoV3- and 4 HBoV4-positive samples were found among 192 samples from patients with AFP, whereas these 2 species were not detected among 96 healthy Tunisian matched contacts. Comparing the combined Nigerian and Tunisian patients with AFP infected with HBoV3 and HBoV4 to Tunisian controls yielded a 2-tailed Fisher's exact *P* value of .01, whereas comparing only the Tunisian patients with AFP with Tunisian controls gave a *P* value of .059. HBoV3 was also found in 1 patient with diarrhea from the United States and from 1 healthy person each from the United States and Nepal. HBoV4 was also found in a patient with diarrhea and 1 healthy adult from the United States. Further testing will be needed to

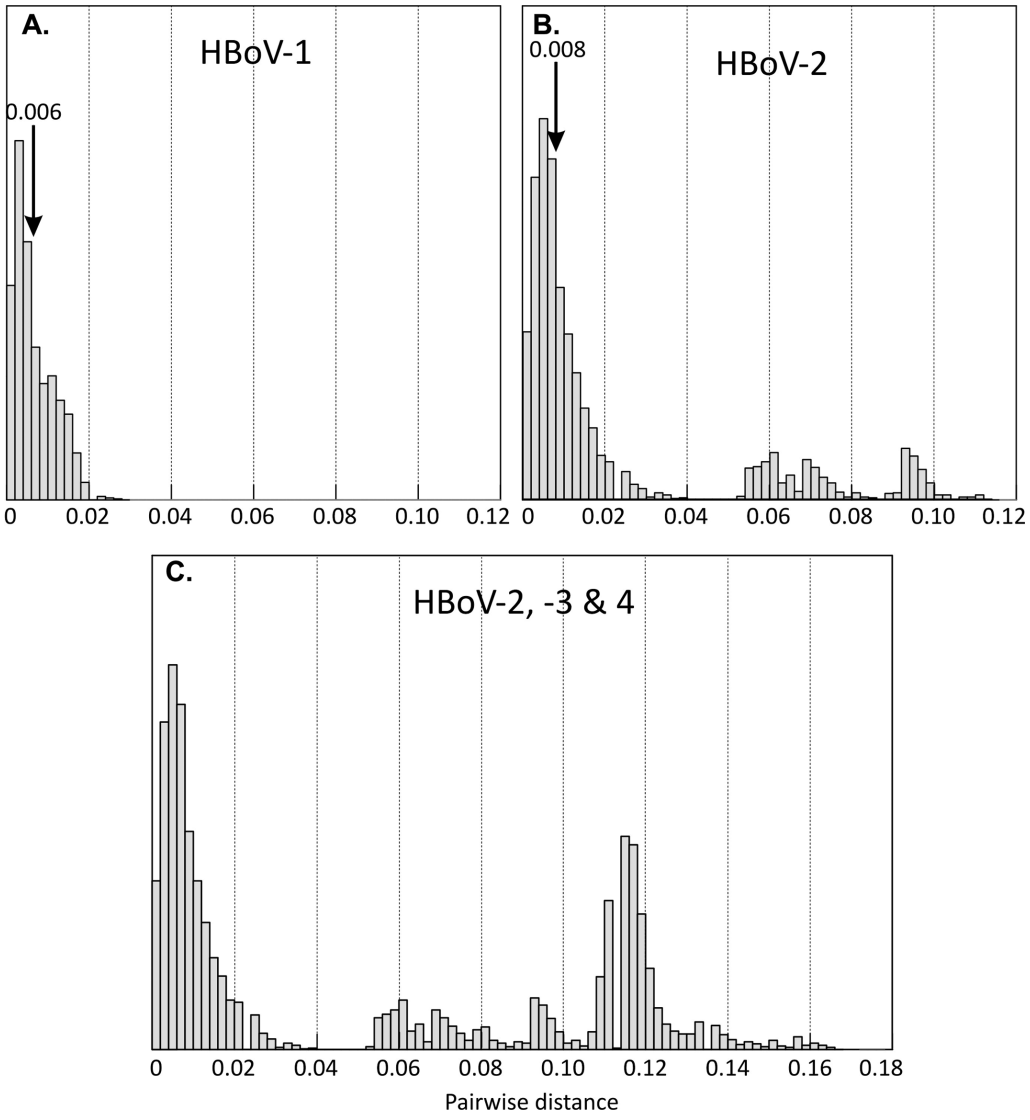

Figure 4. Representation of genetic diversity within and between different bocavirus species. Distribution of pair-wise uncorrected p-distances in the partial VP1 region amplified using pan-bocavirus polymerase chain reaction among human bocavirus (HBoV) 1 (*A*), HBoV2 (*B*), and HBoV2–4 combined (*C*).

confirm this trend of an association of HBoV3/4 with AFP in children or of the association of HBoV2 with diarrhea [8].

DISCUSSION

We report on a previously uncharacterized species of human bocavirus which we tentatively named HBoV4. HBoV1, HBoV2, and HBoV3 were also detected using a pan-PCR approach. HBoV2 has been detected 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) is closely related to a Tunisian strain (TA-210–07) (Figure 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 for the disease association studies that are now required for the 3 recently characterized enteric human bocavirus species (HBoV2, HBoV3, and HBoV4). Based on the phylogenetic clustering observed for a large number of partial VP1 sequences (Figure 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 the complete VP1 gene should be considered different species, whereas those showing >1.5% protein and >5% nucleotide difference should be considered different genotypes. This

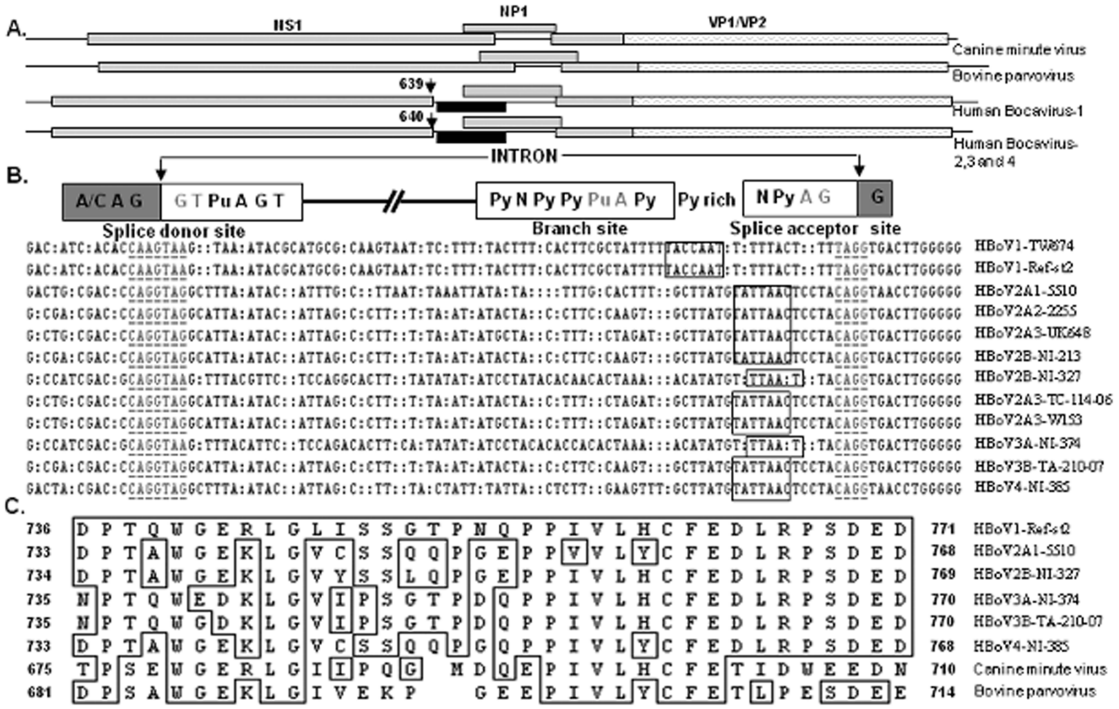

Figure 5. Evidence for RNA splicing in human bocavirus (HBoV) genomes to generate longer NS1 proteins. *A,* Proposed genomic organization of HBoVs relative to that of animal bocaviruses. *Black boxes* represent exon 2 of the NS1 protein. *B,* Canonical sequence of the splice donor, branch site, and splice acceptor (Pu = A/G, Py = T/C). All HBoV species were aligned to show the 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 bocavirus NS1. The amino acid positions used in the alignment are shown.

VP1-based classification retains the 4 proposed human bocavirus species. The VP1 locus was selected because it is likely to strongly influence tissue tropism and potentially pathogenesis [42].

HBoV1 is primarily, although not exclusively [12–31], a respiratory virus. We show here a higher prevalence of HBoV2 (particularly HBoV2B) than HBoV1 in stool samples from children and adults of different countries. A recent study testing for HBoV1 and HBoV2 DNA with use of species-specific nested PCRs failed to detect any HBoV2 in >6500 respiratory secretion samples from Edinburgh and Bangkok, whereas HBoV1 was found in 3% and 14% of these respiratory samples, respectively [43]. Another study found 5 (2%) of 212 nasopharyngeal samples from Korean children with acute lower respiratory track infections to be positive for HBoV2 DNA, but unexpectedly, no HBoV1 DNA was found [44]. 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 intraspecies recombination between HBoV2 variants [7, 8]. The high prevalence of bocavirus infection does provide the opportunity for coinfections, the first step in generating recombinant viruses. Indeed, a HBoV3 and HBoV4 coinfection was detected based on the pattern of mixed bases in 1 directly sequenced PCR product (data not shown).

The frequent detection of HBoVs in stool samples from both healthy children and adults supports the likelihood of long periods of viral shedding and/or frequent reinfections. Determination of whether symptoms such as diarrhea are associated with acute infection in subsequently healthy viral shedders will require quantitative viral load measurements, analysis of longitudinally collected samples, and serological assays. Whether prior infection provides any protection against reinfection with the same or different genotype or species is also unknown.

In both diarrhea sample sets analyzed here, consisting of samples from 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 without protective immune responses from prior infections. Coinfections with other enteric viruses may also exacerbate symptoms. Given the very large number of childhood infections (viral shedding prevalence

of >20% in some developing countries), even low virulence would 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 [11].

The totality of HBoV1 sequences collected worldwide by multiple groups show very low protein and nucleotide sequence diversity (Table 2 and Figure 4*A*) [35, 36]. In contrast, this single study found substantial diversity among HBoV2 and HBoV3, a fourth species (HBoV4), and extensive viral recombination (Figures 3, 4*B,* and 4*C*). 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 has also been reported [45–47]. Parvoviruses have also demonstrated the capacity to rapidly expand their host species tropism, resulting in a recent pandemic in dogs [42, 48, 49]. A recent study showed that HBoV1 could replicate in differentiated human airway epithelial cells [41]. 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 study.

Acknowledgments

We thank Dr Michael P. Busch and Blood Systems Research Institute for sustained support and Hope Biswas for statistical assistance.

References

1. Bloom, Marshall E, Young, Neal S. Knippe DM, Howley PM, eds. Parvoviruses. 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.
3. Patel JR, Heldens JG. Review of companion animal viral diseases and immunoprophylaxis. Vaccine **2009**; 27:491–504.
4. Appel MJ. Forty years of canine vaccination. Adv Vet Med **1999**; 41: 309–324.
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.
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.
7. Kapoor A, Slikas E, Simmonds P, et al. A newly identified bocavirus species in human stool. J Infect Dis **2009**; 199:196–200.
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.
9. Allander T. Human bocavirus. J Clin Virol **2008**; 41:29–33.
10. 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.
11. Manteufel J, Truyen U. Animal bocaviruses: a brief review. Intervirology **2008**; 51:328–334.
12. Choi EH, Lee HJ, Kim SJ, 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.
13. Bastien N, Brandt K, Dust K, Ward D, Li Y. Human bocavirus infection, Canada. Emerg Infect Dis **2006**; 12:848–850.
14. Foulongne V, Rodiere M, Segondy M. Human Bocavirus in children. Emerg Infect Dis **2006**; 12:862–863.
15. Kupfer B, Vehreschild J, Cornely O, et al. Severe pneumonia and human bocavirus in adult. Emerg Infect Dis **2006**; 12:1614–1616.
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.
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.
18. McIntosh K. Human bocavirus: developing evidence for pathogenicity. J Infect Dis **2006**; 194:1197–1199.
19. Kesebir D, Vazquez M, Weibel C, 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.
20. Manning A, Russell V, Eastick K, et al. Epidemiological profile and clinical associations of human bocavirus and other human parvoviruses. J Infect Dis **2006**; 194:1283–1290.
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.
22. Kahn J. Human bocavirus: clinical significance and implications. Curr Opin Pediatr **2008**; 20:62–66.
23. Kantola K, Hedman L, Allander T, et al. Serodiagnosis of human bocavirus infection. Clin Infect Dis **2008**; 46:540–546.
24. Schildgen O, Muller A, Allander T, et al. Human bocavirus: passenger or pathogen in acute respiratory tract infections? Clin Microbiol Rev **2008**; 21:291–304.
25. Sloots TP, Whiley DM, Lambert SB, Nissen MD. Emerging respiratory agents: new viruses for old diseases? J Clin Virol **2008**; 42(3):233–243.
26. Albuquerque MC, Rocha LN, Benati FJ, et al. Human bocavirus infection in children with gastroenteritis, Brazil. Emerg Infect Dis **2007**; 13:1756–1758.
27. Cheng WX, Jin Y, Duan ZJ, et al. Human bocavirus in children hospitalized for acute gastroenteritis: a case-control study. Clin Infect Dis **2008**; 47:161–167.
28. Lau SK, Yip CC, Que TL, 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.
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.
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.
31. Yu JM, Li DD, Xu ZQ, et al. Human bocavirus infection in children hospitalized with acute gastroenteritis in China. J Clin Virol **2008**; 42(3): 280–285.
32. Chieochansin T, Chutinimitkul S, Payungporn S, et al. Complete coding sequences and phylogenetic analysis of human bocavirus (HBoV). Virus Res **2007**; 129:54–57.
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.
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.
35. Simmonds P, Smith DB. Structural constraints on RNA virus evolution. J Virol **1999**; 73:5787–5794.
36. Blinkova O, Rosario K, Li L, et al. Frequent detection of highly diverse variants of cardiovirus, cosavirus, bocavirus, and circovirus in sewage samples collected in the United States. J Clin Microbiol **2009**; 47(11): 3507–3513.
37. Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA. Virus taxonomy. Eighth report of the International Committee on Taxonomy of Viruses. Academic Press, **2005**.
38. Lederman M, Patton JT, Stout ER, Bates RC. Virally coded noncapsid protein associated with bovine parvovirus infection. J Virol **1984**; 49: 315–318.
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.
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**; 81:12080–1**2085**.
41. Dijkman R, Koekkoek SM, Molenkamp R, Schildgen O, van der Hoek L. Human bocavirus can be cultured in differentiated human airway epithelial cells. J 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.
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.
44. Han T-H , Chung J-Y, Hwang E-S. Human bocavirus 2 in children, South Korea. Emerg Infect Dis **2009**; 15:1698–1699.
45. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet **2008**; 9:267–276.
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**; 82(13):6427–6433.
47. Shackelton LA, Hoelzer K, Parrish CR, Holmes EC. Comparative analysis reveals frequent recombination in the parvoviruses. J Gen Virol **2007**; 88:3294–3301.
48. 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.
49. 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.