Epidemiology and Infection

cambridge.org/hyg

Whole-genome sequencing analysis of human bocavirus detected in South Korea

L. H. Kang[1], Y. J. Won[1], A. R. Lee[1], S. G. Lee[2], H. G. Cho[3], Y. J. Park[4], J. I. Han[5] and S. Y. Paik[1]

[1]Department of Microbiology, College of Medicine, The Catholic University of Korea, 222, Banpo-daero, Seocho-gu, Seoul 06591, Republic of Korea; [2]Korea Zoonosis Research Institute, Chonbuk National University, Iksan, Republic of Korea; [3]Division of Public Health Research, Gyeonggi Province Institute of Health and Environment, Suwon, Republic of Korea; [4]Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea and [5]Laboratory of Wildlife Diseases, College of Veterinary Medicine, Chonbuk National University, Iksan 54596, Korea

Original Paper

Cite this article: Kang LH, Won YJ, Lee AR, Lee SG, Cho HG, Park YJ, Han JI, Paik SY (2018). Whole-genome sequencing analysis of human bocavirus detected in South Korea. *Epidemiology and Infection* **146**, 839–847. https://doi.org/10.1017/S0950268818000560

Received: 16 November 2017
Revised: 1 February 2018
Accepted: 14 February 2018
First published online: 14 March 2018

Key words:
Human bocarvirus; Korea; parvovirus; phylogeny; whole-genome sequencing

Author for correspondence:
S. Y. Paik, E-mail: paik@catholic.ac.kr

Abstract

Human bocaviruses (HBoVs) have been detected in human gastrointestinal infections worldwide. In 2005, HBoV was also discovered in infants and children with infections of the lower respiratory tract. Recently, several genotypes of this parvovirus, including HBoV genotype 2 (HBoV2), genotype 3 (HBoV3) and genotype 4 (HBoV4), were discovered and found to be closely related to HBoV. HBoV2 was first detected in stool samples from children in Pakistan, followed by detection in other countries. HBoV3 was detected in Australia and HBoV4 was identified in stool samples from Nigeria, Tunisia and the USA. Recently, HBoV infection has been on the rise throughout the world, particularly in countries neighbouring South Korea; however, there have been very few studies on Korean strains. In this study, we characterised the whole genome and determined the phylogenetic position of CUK-BC20, a new clinical HBoV strain isolated in South Korea. The CUK-BC20 genome of 5184 nucleotides (nt) contains three open-reading frames (ORFs). The genotype of CUK-BC20 is HBoV2, and 98.77% of its nt sequence is identical with those of other HBoVs, namely Rus-Nsc10-N386. Especially, the ORF3 amino acid sequences from positions 212–213 and 454 corresponding to a variable region (VR)1 and VR5, respectively, showed genotype-specific substitutions that distinguished the four HBoV genotypes. As the first whole-genome sequence analysis of HBoV in South Korea, this information will provide a valuable reference for the detection of recombination, tracking of epidemics and development of diagnosis methods for HBoV.

Introduction

Human bocaviruses (HBoVs) are members of the family Parvoviridae and are classified within the subfamily Parvoviridae. HBoVs are small non-enveloped viruses with an icosahedral symmetry and a single-stranded DNA genome of approximately 5 kb in length. The genome encodes two forms of non-structural proteins, NS1 and nuclear phosphoprotein 1 (NP1) and two major structural proteins, VP1 and VP2 [1–5]. Currently, HBoVs have been classified into four species, HBoV1, HBoV2, HBoV3 and HBoV4, based on the nucleotide (nt) divergence of the VP1 capsid region [6]. The HBoV genome contains three open-reading frames (ORFs); the first two sequential ORFs (ORF1 and ORF2) encode NS1 and NP1, respectively, and the third downstream ORF3 encodes VP1 and VP2 [1, 7]. NS1 is known to have regulatory functions, including transactivation or induction of apoptosis [8–10]. Although NP1 is a nucleoprotein, its specific functions are currently unknown, and VP1/VP2 are viral capsid proteins [10–12]. NS1 shows the greatest nt sequence conservation of the genome, with the lowest genetic diversity among all HBoV subtypes; thus, this gene has been preferentially used as a target for the detection of HBoV [13, 14]. By contrast, VP1/VP2 forms a variable region exhibiting high genetic diversity and has thus been mostly used for the phylogenetic analysis of HBoV [15, 16]. ORF1 and ORF3 are present in all parvovirus genomes, whereas the presence of ORF2 in the bocavirus genome is a specific feature distinguishing this group from the other parvoviruses [17]. HBoV NS1 is essential for replication of the viral single-stranded DNA genome, DNA packaging and may play several versatile roles in virus–host interactions [18]. HBoV is a parvovirus that was first identified in 2005 using a protocol based on DNase treatment, random polymerase chain reaction (PCR) amplification, high-throughput sequencing and bioinformatics analysis. When this virus-screening technique was initially applied to nasopharyngeal swabs and washings from children with unresolved respiratory tract infections, a positive result rate of 3.1% was obtained; hence, it was proposed that HBoV is a causative pathogen of respiratory tract diseases [19, 20]. Although HBoV1 and HBoV2 have been mainly detected in stool samples, HBoV3 and HBoV4 are only occasionally detected in these samples

CAMBRIDGE UNIVERSITY PRESS

[3, 21–23]. Since its discovery, HBoV has been found to be circulating globally, and is predominantly detected in urine, serum and stool specimens of children with respiratory infections [21, 24–26]. Conversely, recent research has raised concerns over its association with transfusion medicine [27]. HBoV is the main cause of respiratory tract infection symptoms in infants and toddlers, whose manifestation varies from no symptoms to symptoms such as fever, coughing and runny nose [10, 12, 28]. In many cases, patients with HBoV detected in their faecal specimens present with symptoms of viral gastroenteritis, including diarrhoea, vomiting and fever [13, 29]. HBoV1 has been most frequently detected in respiratory specimens, whereas HBoV2, HBoV3 and HBoV4 are most commonly found in faecal specimens, indicating a predisposition to causing gastrointestinal diseases [2, 30–33]. In 2009, Kapoor *et al.* [30] identified a new parvovirus, HBoV-2, in stool samples from 98 Pakistani children and 699 British individuals of a mixed-age population, which was suggested to be an aetiologic agent in acute gastroenteritis [22]. The high detection rate and a high degree of genetic diversity among these enteric viruses from stool specimens (especially for HBoV2) suggest that they may be pathogenic viruses in acute gastroenteritis, although current data show contradictory conclusions [23, 30, 33]. Thus, the aim of the present study was to analyse and present the first full-length genome sequence of an HBoV strain from South Korea. Phylogenetic analysis was performed for comparison of the strain with reported genotypes. We expect that these data will prove to be useful not only for advancing research in the molecular biology of HBoVs but also for basic epidemiologic analyses such as tracking of the international spread of the virus.

Materials and methods

Ethics statement

An HBoV-positive stool sample was obtained from a male infant who presented with fever and diarrhoea, provided by the Waterborne Virus Bank (WAVA). Most stool samples were obtained from infants (⩽3 years old) who were hospitalised, and the HBoV2 sample used in this study was also from an infant patient (2 years old). Because of difficulties in tracking the exact records of the patient from the donor hospital, informed consent from the parent of the patient could not be acquired. The Institutional Review Board reviewed and approved the use of this sample for the purpose of research as this study did not directly affect the patient. All of the experimental work and sample collections were supervised by the Catholic Medical Center Office of Human Research Protection Program (CMC OHRP) of South Korea (approval no. MC14SISI0096).

Table 1. Newly designed primers used in this study

Primer	Sequence (5′–3′)	Location[a]	Size (bp)
GSP1	CTGACCATGCTTCGCGTTCAA	770–790	
GSP2	AATTCAGCCATCACAAGGCCT	669–689	
Nested GSP	CGATCTCAGCCTGGCATATTA	563–583	
NS1-F1	GTGGTGAGTSAYACTATGGC	241–260	613
NS1-R1	GWGCATGAAGGTCWCCTCKS	834–853	
NS1-F2	GCA TGG KSA GGR SWY ACT GG	781–800	758
NS1-R2	CCTGTTGARGCAGGVCCRTAAAAGC	1514–1538	
NS1-F3	CACGCCMTSTGYTGTGTMCTKAAC	1462–1485	804
NS1-R3	GGWGTYYTCTCCGTCCGWATCC	2244–2265	
NS1/NP1-F1	CCT CAT AGT GCT GAC RAY TCT RTG	2098–2121	345
NS1/NP1-R1	KAKGMGCGATGCYTGTSYTTCATA	2419–2442	
NP1-F1	CCW MGW GAC GAA GAT GAG CTC	2392–2412	673
NP1-R1	CYAGGYTGKCKYTTAATTGGAGGC	3041–3064	
VP1-F1	GAA KCA GAS GAG RTA ACW GAC G	2981–3002	789
VP1-R1	GTWAGGCGCTGCCARTCYTGTG	3748–3769	
VP1-F2	GGWTCTGGTGTGGGKATWTCCAC	3522–3544	671
VP1-R2	CCAGGAGGAATGTATGCTCTTTCR	4169–4192	
VP1-F3	GAA CTG GWG ARA GCA CWG AAT TYA C	4114–4138	732
VP1-R3	CRCAGCTGACTTGTCCWGTRCAG	4823–4845	
VP1-F4	GCA ATG GAT CAT CCT CCW GGH AC	4731–4753	423
VP1-R4	CRTCGGRCTGTRGYCTYGAACYY	5178–5153	
3′-Oligo (dT)-anchor-R	CAA TGA GGT TAT GGC TTT GGA ACT TTT TTT TTT TTT TT	3′-end poly A tail	
3′-Anchor-R	CAA TGA GGT TAT GGC TTT GGA AC		

[a]According to GenBank accession number GU048663.

https://doi.org/10.1017/S0950268818000560 Published online by Cambridge University Press

Table 2. Nucleotide (nt) sequence identities of the full-length genome, and NS1, NP1, and VP1/VP2 genes between the HBoV CUK-BC20 strain and reference strains

Reference strain	Accession number	Genotype	Collection date	Country	Nt sequence identity (%)			
					Full-length	NS1	NP1	VP1/VP2
HK4	EF450720	1	2005	Hong Kong	75.39	74.15	75.76	78.21
HK10	EF450726	1	2005	Hong Kong	75.31	74.15	75.91	78.06
HK7	EF450723	1	2005	Hong Kong	75.39	74.21	76.06	78.16
TW27156	EU984232	1	2006	Taiwan	75.33	74.21	75.91	78.11
TW2717	EU984233	1	2006	Taiwan	75.39	74.1	76.21	78.16
TW141	EU984242	1	2007	Taiwan	75.33	74.26	75.76	78.01
JPOC	AB481080	1	2007	Japan	75.27	74.15	75.61	78.06
LWK	GU338055	1	2009	China	75.25	74.15	75.76	78.06
TUN2207	JF327786	1	2010	Tunisia	75.37	74.15	75.91	78.21
TUN4134	JF327788	1	2010	Tunisia	75.35	74.15	75.76	78.21
HBoV10-N1117	JQ964114	1	2010	Russia	75.33	74.1	75.91	78.16
CBJ030	KM464730	1	2008	China	75.25	74	75.76	78.21
ST2	NC007455	1	2004	Sweden	75.37	74.15	75.91	78.21
KU3	JQ411251	1	2005	Brazil	75.31	74.21	75.76	78.06
HZ1402	KP710212	1	2014	China	75.33	74.15	75.61	78.21
HZ1403	KP710213	1	2014	China	75.29	74	75.61	78.26
Irish	KC823115	1	2012	Ireland	75.29	74.21	75.76	78.06
Mty1117	KX373885	1	2004	Mexico	75.33	74.21	75.76	78.11
P214	KX373884	1	2010	Mexico	75.23	74.21	75.76	77.86
HBoV2A-TU-A-114-06	FJ973558	2a	2006	Tunisia	98.08	97.71	98.92	98.06
Rus-Nsc10-N386	JQ964116	2a	2010	Russia	98.77	97.82	99.07	99.5
UK648	FJ170280	2a	2008	United Kingdom	95.56	97.76	99.07	99
W153	EU082213	2a	2001	Australia	98.58	97.82	99.07	99
HBoV2B-NI-213	FJ973560	2b	2007	Nigeria	96.08	97.92	96.45	94.02
HBoV2B-NI-327	FJ973559	2b	2007	Nigeria	96.1	98.08	96.3	93.97
SH3	FJ375129	2b	2008	China	96.72	97.66	97.22	95.61
W298	FJ948860	2	2001	Australia	98.35	97.87	98.61	98.6
W208	EU082214	2	2001	Australia	98.6	97.87	99.07	99.5
Rus-Nsc10-N751	JQ964115	2	2010	Russia	98.75	97.82	99.07	99.4
CU1557UK	GU048664	2	2008	United Kingdom	98.52	97.82	98.77	98.95
BJQ435	JX257046	2	2011	China	97.08	98.02	97.07	95.86
LZ53819	GU301644	2	2007	China	97.16	98.08	99.07	95.37
PK-2255	FJ170279	2	2008	Pakistan	97.01	98.02	97.99	95.66
PK-5510	NC012042	2	2008	Pakistan	93.21	92.46	91.98	95.46
CU47TH	GU048662	2	2007	Thailand	97.14	98.13	99.38	95.47
CU54TH	GU048663	2	2007	Thailand	97.24	97.3	99.38	96.26
W471	NC_012564	3	2001	Australia	79.67	72.61	75.38	87.54
W855	FJ948861	3	2002	Australia	79.67	72.76	75.23	87.49
HBoV3	JN086998	3	2003	USA	79.69	72.61	75.23	87.59

(*Continued*)

https://doi.org/10.1017/S0950268818000560 Published online by Cambridge University Press

Table 2. (*Continued.*)

Reference strain	Accession number	Genotype	Collection date	Country	Nt sequence identity (%)			
					Full-length	NS1	NP1	VP1/VP2
MC8	GQ867666	3	2003	Brazil	79.69	72.76	75.53	87.49
IM10	GQ867667	3	2004	Brazil	79.67	72.76	75.68	87.39
HBoV3B-TU-A-210-07	FJ973562	3	2007	Tunisia	79.69	72.66	75.23	87.54
46-BJ07	HM132056	3	2007	China	79.61	72.56	74.92	87.54
CU2139UK	GU048665	3	2008	United Kingdom	79.65	72.61	75.38	87.44
HBoV4-NI-385	NC_012729	4	2007	Nigeria	88.35	90.64	87.65	87.18
CMH-S011-11	KC461233	4	2011	Thailand	89.25	91.68	92.28	86.79

Sample preparation and viral DNA extraction

The stool sample was stored at −20 °C until further analysis. Viral DNA was extracted into 50 µl of elution buffer manually from 140 µl of a 10% faecal suspension prepared in phosphate-buffered saline using the QIAamp DNA mini kit (Qiagen, Hilden, Germany), according to the manufacturer's protocol. Isolated DNA was stored at −20 °C until further use.

Polymerase chain reaction

For the detection of complete HBoV, PCR was performed with the 2× Emerald Amp PCR Master Mix (TaKaRa, Shiga, Japan) on an S1000 thermal cycler (Bio-Rad, Hercules, CA, USA), using primers designed based on the full genome sequence of the detected HBoV (Table 1). The PCR steps comprised initial activation (94 °C for 5 min), 25 cycles of three-step cycling (94 °C for 30 s, 53.8–60.1 °C for 30 s and 72 °C for 1 min) and final extension (72 °C for 7 min). All PCR products were examined by electrophoresis on ethidium bromide-stained 2% agarose gels.

Determination of the 5′ and 3′-ends of the HBoV genomic DNA

To determine the 5′-ends of HBoV genomic DNA, rapid amplification of cDNA ends (RACE) was performed with the 5′ RACE System for RACE Version 2.0 Kit according to the manufacturer's recommendations (Invitrogen, Carlsbad, CA, USA). Three primers (GSP1, GSP2 and nested GSP) were designed based on the NS1 sequence for 5′-end RACE PCR (Table 1). To obtain the exact sequence of the 3′-end of the HBoV genomic DNA, cDNA was synthesised using reverse transcription with 3′-oligo (dT)-anchor-R (Table 1). The second PCR was conducted using the VP1/VP2-F and 3′-anchor-R primers (Table 1) under the following conditions: 30 cycles of three-step cycling (94 °C for 30 s, 55 °C for 30 s and 72 °C for 1 min) and 72 °C for 10 min.

Cloning and sequencing of the complete genome

All PCR products obtained were extracted using the HiYield Gel/PCR DNA Fragments Extraction Kit (RBC, Taipei, Taiwan) and were cloned into pGEM-T Easy Vectors (Promega, Madison, WI, USA). The cloned vector was transformed into *Escherichia coli* DH5α competent cells (RBC) according to the manufacturer's instructions and was selected from Luria–Bertani agar plates (Duchefa, Haarlem, the Netherlands) containing 40 mg/ml X-gal, 0.1 mM isopropyl-β-D-thiogalactoside and 50 mg/ml ampicillin at 37 °C for 16–18 h. Selected clones were inoculated in Luria–Bertani broth (Duchefa) and incubated overnight in a shaking incubator (IS-971R, Jeiotech, Daejeon, South Korea) at 37 °C and 200 rpm. Plasmid DNA was purified using the HiYield Plasmid Mini Kit (RBC) and sequenced (Macrogen, Seoul, South Korea). The sequencing results were analysed using BLAST (National Center for Biotechnology Information, NCBI).

Phylogenetic analysis

Comparative sequence analysis, including sequence alignments and estimation of genetic distances, was performed with Clustal W using the Molecular Evolutionary Genetic Analysis (MEGA) software version 7.0 [34]. Phylogenetic trees were constructed using the neighbour-joining method with a Kimura two-parameter model in MEGA [35] and branch support was calculated based on 1000 bootstrap replicates. The complete genome sequences and partial genome sequences were obtained from the NCBI database.

Results

Nucleotide sequence identities

The complete coding sequences of nt and deduced amino acids (aa) of the two non-structural proteins, NS1 and NP1, as well as the two major structural proteins, VP1 and VP2, of the newly obtained South Korean strain CUK-BC20 (GenBank no. MF680549), were compared with those of established reference strains of HBoV1–4. The coding sequence starting from NS1, NP1 to VP1/VP2 (excluding the 5′- and 3′-untranslated sequences) is 4787 nt long. Analysis of the full-length sequence of CUK-BC20 revealed that this similar HBoV strain is most closely related to the Russian HBoV2 strain Rus-Nsc10-N386 (GenBank accession no. JQ964116) with 98.77% similarity (Table 2). Moreover, NS1 and NP1 of strain CUK-BC20 showed nt sequence identity with the HBoV4 strains KC461233 (CMH-S011-11) and FJ973561 (NI-385) at 90.64% and 92.28%, respectively. The NS1 gene of strain CUK-BC20 is 1923 nt long, encoding a polypeptide of 640 aa residues. Surprisingly, the NS1 nt sequence comparison indicated that CUK-BC20 is most closely related to the HBoV2 reference strains at 92.46–98.13%

https://doi.org/10.1017/S0950268818000560 Published online by Cambridge University Press

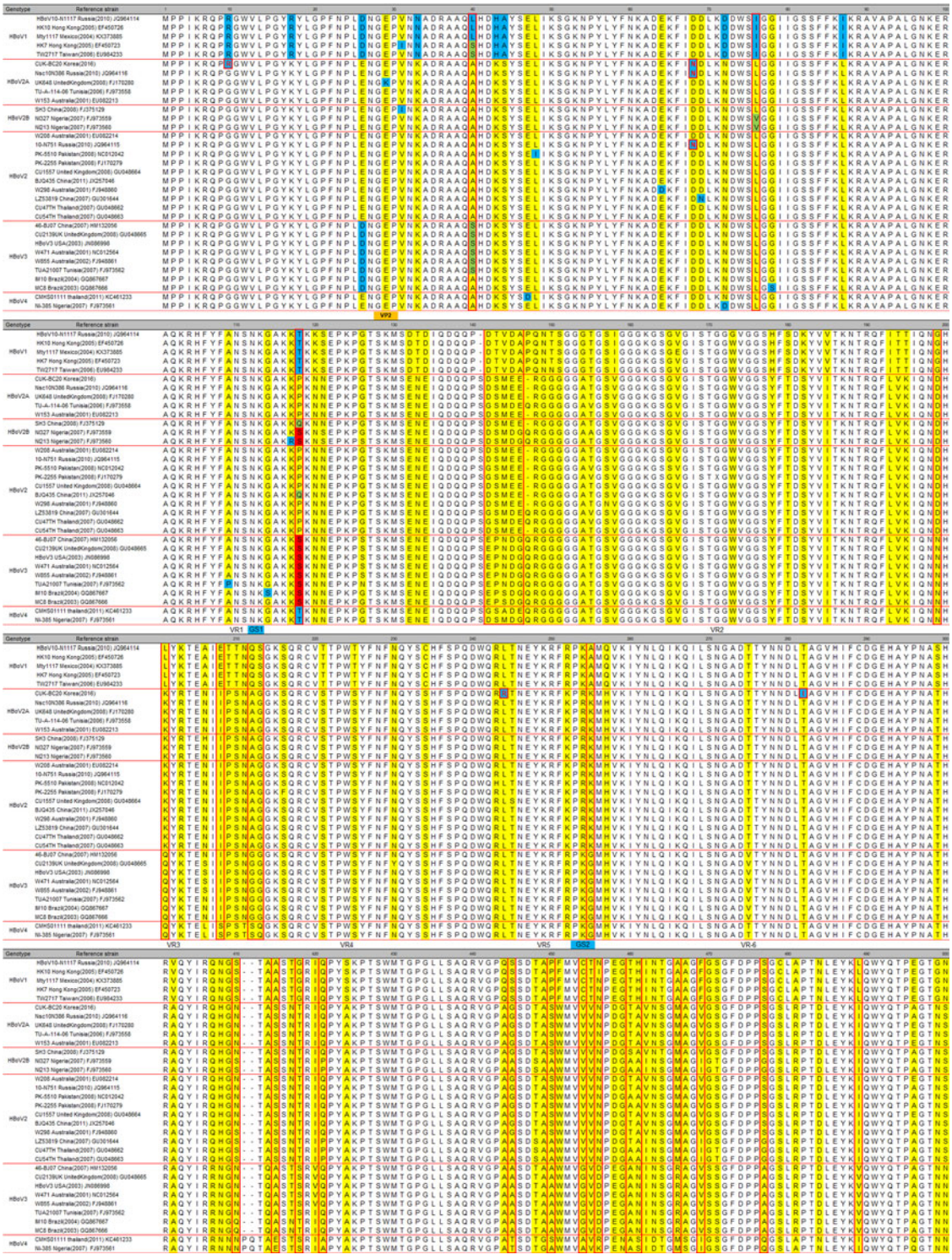

Fig. 1. Comparison of VP1/2 amino acid substitutions among HBoV genotypes. Representative strains were selected for each HBoV genotype. The red borders indicate the sites of amino acid alterations. The amino acid sequences that differ among the virus strains are highlighted in yellow. *GS, genotype-specific; VR, variable region.

(Table 2). The highest sequence identity (98.13%) was found with the reference strain GU048662 (CU47TH) detected in Thailand. Analysis of the complete NP1 gene sequence revealed a sequence of 648 nt long encoding a polypeptide of 215 aa residues, similar to the NP1 genes of the reference strains GU048662 (CU47TH) and GU048663 (CU54TH). The complete VP1 gene coding sequence of the strain CUK-BC20 is 2004 nt long encoding 667 aa residues. In addition, within this VP1 sequence, another start codon of VP2 (1617 nt long encoding 538 aa residues) was found at nt position 388 from the VP1 starting point. Comparative analysis of the nt sequences of the full-length VP1/VP2 gene with those of the HBoV1–4 reference sequences

https://doi.org/10.1017/S0950268818000560 Published online by Cambridge University Press

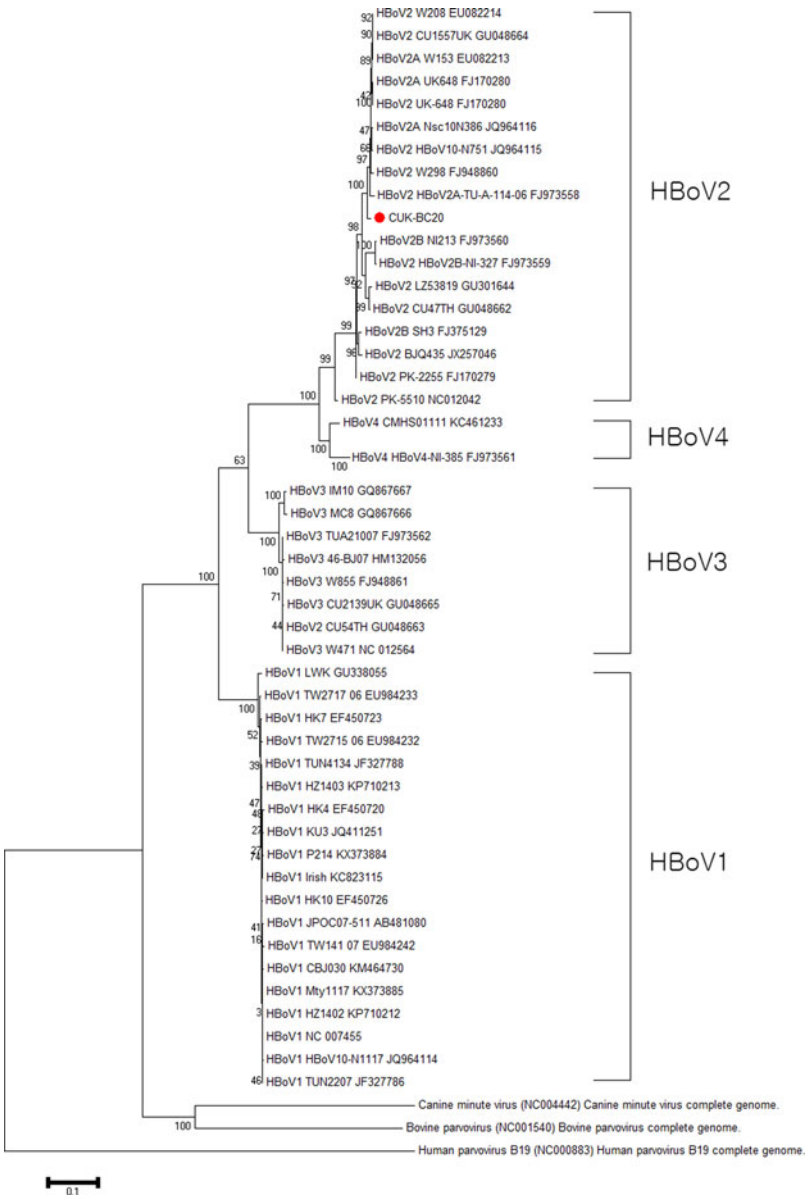

Fig. 2. Phylogenetic analysis of human bocavirus based on complete genome sequences. Phylogenetic trees were generated with 1000 bootstrap replicates using the neighbour-joining method with Mega 7 software. Parvovirus B19, bovine parvovirus, canine minute virus and representative strains of HBoV1–4 were used as reference strains for the genotype analysis of CUK-BC20 (red circle). The HBoV strain obtained in this study (CUK-BC20) was ultimately identified as HBoV2.

indicated that the VP1/VP2 gene of CUK-BC20 is most closely related to those of the reference strains JQ964116 (Rus-Nsc10-N386) and EU082214 (W208) detected in Russia and Australia, respectively, at 99.5% (Table 2).

Nucleotide and amino acid polymorphisms

Based on the three HBoV ORFs (NS1, NP1 and VP1/VP2) and the complete genomes of 17 HBoV2 strains, 10 of 1923 nts analysed for the NS1 gene were found to be variable, with five (50%) transitions and five (50%) transversions; four of 648 nts analysed for the NP1 gene were found to be variable, with three (75%) transitions and one (25%) transversion; and seven of 2004 nts analysed for the VP1 gene were found to be variable, with six (85%) transitions and one (15%) transversion (Fig. 1). VP1 amino acid sequences were compared among HBoV1, HBoV2A, HBoV2B, HBoV3 and HBoV4 genotypes using representative strains for each genotype (Fig. 3). In particular, the amino acid sequences from aa 212 to 213 and aa 454 corresponding to a variable region (VR)1 and VR5, respectively, showed genotype-specific substitutions that distinguished the four HBoV genotypes. Among 640 NS1 aa sequences, there were three aa substitutions detected: aa178 (Ser→Gly), aa180 (His→Val) and aa612 (Gly→Arg). Among 215 NP1 aa sequences, there was one aa substitution detected (data not shown): aa9 (Arg→Lys). Among 667 VP1/2 aa sequences, there were four aa substitutions detected (data not shown): aa9 (Gly→Arg), aa68 (Asp→Asn), aa244 (Leu→Arg) and aa282 (Thr→IlE).

Phylogenetic analysis

The HBoV strains used in the phylogenetic analysis included the target Korean strain CUK-BC20, as well as representative strains of HBoV1–4, human parvovirus B19, bovine parvovirus and canine minute virus. Based on the complete genome sequences, the phylogenetic analysis showed that CUK-BC20 is most genetically close to HBoV2 (Fig. 2), which was consistent with the sequence comparison analysis results. The phylogenetic tree

https://doi.org/10.1017/S0950268818000560 Published online by Cambridge University Press

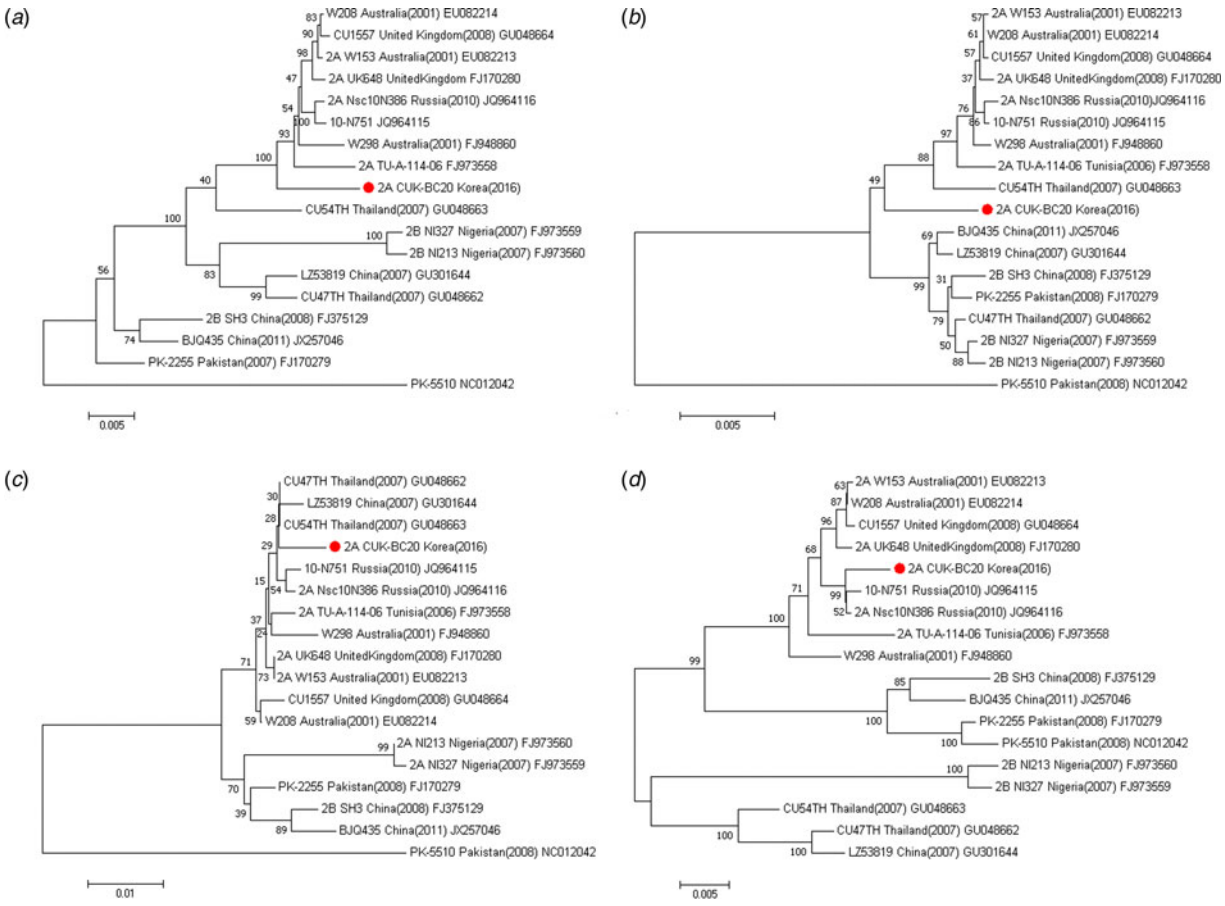

Fig. 3. Phylogenetic analysis of the Korean strain CUK-BC20 and other human bocavirus reference strains based on the full-length genome (a), and the complete NS1 (b), NP1 (c) and VP1/VP2 (d) gene sequences. Phylogenetic trees (1000 bootstrap replicates, Kimura two-parameter model) were based on CUK-BC20 (red circle) and the following HBoV2 reference strains: Australian strains W208, W153, W298; British strains CU1557UK and UK-648; Russian strains Nsc10N386 and Nsc10N751; Thai strains CU54TH and CU47TH; Nigerian strains 2BNI327 and NI213; Chinese strains SH3, LZ53819 and BJQ435; Pakistani strains PK-2255 and PK-5510; and Tunisian strain TU-A-114-06. GenBank accession numbers are given next to each strain name in the trees.

based on the complete genome further showed that the different strains of HBoV2 could be clearly divided into three groups (Fig. 3a). No apparent genotypic differences were evident among the phylogenetic trees based on the three HBoV ORFs (NS1, NP1 and VP1/VP2) and the complete genomes of 17 HBoV2 strains (Figs. 3b–d). As previously reported, NS1 appeared to be the most highly conserved gene, whereas VP1/VP2 had the greatest number of nt polymorphisms. Moreover, the phylogenetic trees based on the VP1/VP2 gene and complete genome were nearly identical, which indicated that VP1/VP2 can be used instead of the complete genome to analyse the genetic relationships among HBoVs.

Discussion

HBoV is a newly discovered Parvoviridae virus. However, the pathogenicity of HBoV is still uncertain because of its high co-infection rate with other pathogens; thus, it remains unclear whether HBoVs are the sole etiologic agent or simply a concomitant virus bystander in these cases. Therefore, to gain a better understanding of the prevailing status and pathogenicity of HBoV, more strains need to be simultaneously examined. HBoV is considered a major agent of several respiratory tract diseases, and according to several reports, HBoV is associated with gastrointestinal disorders, commonly by coinfection with Rotavirus, Norovirus and Adenovirus. The proportion of multiple viral infections in HBoV-infected patients has been reported to be between 35% and 91% [21, 36–39]. In this study, HBoV-infected patients were shown to have been coinfected with Human astrovirus 5. Human astroviruses are recognised as an important cause of infantile gastroenteritis around the world. The conflicting ideas of the pathogenic role of HBoVs are mainly due to the fact that Koch's revised postulates cannot be applied to HBoV, because neither an effective method for virus culture nor an animal model of infection is currently available in practice [20]. Moreover, little is known about the epidemiology and genetic characteristics of HBoV circulating in Korea. This is the first study to determine the whole genome sequence of HBoV in South Korea, based on a strain isolated from a patient with acute gastroenteritis. The identified HBoV strain, CUK-BC20, belongs to genotype 2 and showed no intra- or inter-genogroup recombination of the non-structural protein-encoding region and the VP1,2-encoding region. CUK-BC20 is very similar to the Russian HBoV isolate Rus-Nsc10-N386 (JQ964116), the prototype of which was reported from a patient with acute gastroenteritis at Novosibirsk Child Hospital, Russia in 2010–2011 [40]. Analysis of the full-length and VP1/2 sequences of the CUK-BC20 strain revealed high similarity (98.77–99.5%) to the Rus-Nsc10-N386 reference strain (HBoV2A), whereas the NS1

https://doi.org/10.1017/S0950268818000560 Published online by Cambridge University Press

sequence showed relatively lower similarity (97.82%) to the reference. Both the NS1 and NP1 sequences of the CUK-BC20 strain showed high similarity (97.3–99.38%) to the CU47TH and CU54TH reference strains (HBoV2). The CUK-BC20 strain showed the lowest similarity (91.98–95.46%) to the PK-5510 strain and showed high (98.2%) amino acid similarity with the CU47TH/Thailand strain isolated in 2007, although higher similarity levels were detected for the full-length and VP1/2 sequences (98.77% and 99.5%, respectively). This result showed that the common ancestor of HBoV2 may have been co-circulating in both Russia and Thailand in 2007–2016. Amino acid substitutions were also detected in NS1, NP1 and VP1/2 for the South Korean strain.

This is the first study reporting the full-length sequence of an HBoV2 strain isolated in South Korea from a clinical sample. This sequence will be useful for comparisons with the full-length HBoV2 sequences of other strains identified globally. Moreover, the information acquired from the whole-genome sequence of strain CUK-BC20 may prove useful for obtaining more accurate diagnoses of HBoV as well as for advancing basic research toward the elucidation of the genetic functions, the prediction of newly appearing pandemic variants via comparison with HBoVs in neighbouring countries and in vaccine development. Overall, broadening the information and genetic resources of HBoVs circulating globally will have important benefits for public health and help to identify new emerging strains of HBoV.

Acknowledgements. This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number, HI16C0443), as well as by the Korea Ministry of Environment (MOE) as a Public Technology Program based on Environmental Policy (grant number, 2016000210002).

Declaration of Interest. None.

References

1. **Chen AY, *et al.*** (2010) Characterization of the gene expression profile of human bocavirus. *Virology* **403**(2), 145–154.
2. **Kapoor A, *et al.*** (2010) Human bocaviruses are highly diverse, dispersed, recombination prone, and prevalent in enteric infections. *The Journal of Infectious Diseases* **201**(11), 1633–1643.
3. **Khamrin P, *et al.*** (2013) Complete genome sequence analysis of novel human bocavirus reveals genetic recombination between human bocavirus 2 and human bocavirus 4. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* **17**, 132–136.
4. **Xu ZQ, *et al.*** (2011) Development of a real-time PCR assay for detecting and quantifying human bocavirus 2. *Journal of Clinical Microbiology* **49** (4), 1537–1541.
5. **Dijkman R, *et al.*** (2009) Human bocavirus can be cultured in differentiated human airway epithelial cells. *Journal of Virology* **83**(15), 7739–7748.
6. **Zhao H, *et al.*** (2012) Detection of a bocavirus circular genome in fecal specimens from children with acute diarrhea in Beijing, China. *PLoS ONE* **7**(11), e48980.
7. **Babkin IV, *et al.*** (2015) A study of the human bocavirus replicative genome structures. *Virus Research* **195**, 196–202.
8. **Moffatt S, *et al.*** (1998) Human parvovirus B19 nonstructural (NS1) protein induces apoptosis in erythroid lineage cells. *Journal of Virology* **72**(4), 3018–3028.
9. **Raab U, *et al.*** (2002) NS1 protein of parvovirus B19 interacts directly with DNA sequences of the p6 promoter and with the cellular transcription factors Sp1/Sp3. *Virology* **293**(1), 86–93.
10. **Schildgen O, *et al.*** (2008) Human bocavirus: passenger or pathogen in acute respiratory tract infections? *Clinical Microbiology Reviews* **21**(2), 291–304, table of contents.
11. **McIntosh K**. (2006) Human bocavirus: developing evidence for pathogenicity. *The Journal of Infectious Diseases* **194**(9), 1197–1199.
12. **Lindner J and Modrow S**. (2008) Human bocavirus – a novel parvovirus to infect humans. *Intervirology* **51**(2), 116–122.
13. **Chieochansin T, *et al.*** (2007) Complete coding sequences and phylogenetic analysis of Human Bocavirus (HBoV). *Virus Research* **129**(1), 54–57.
14. **Maggi F, *et al.*** (2007) Human bocavirus in Italian patients with respiratory diseases. *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology* **38**(4), 321–325.
15. **Blinkova O, *et al.*** (2009) Frequent detection of highly diverse variants of cardiovirus, cosavirus, bocavirus, and circovirus in sewage samples collected in the United States. *Journal of Clinical Microbiology* **47**(11), 3507–3513.
16. **Alam MM, *et al.*** '(2015) Human bocavirus in Pakistani children with gastroenteritis'. *Journal of Medical Virology* **87**(4), 656–663.
17. **Fields BN, Knipe DM and Howley PM** (2007) *Fields Virology*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.
18. **Tewary SK, *et al.*** (2013) Structure of the NS1 protein N-terminal origin recognition/nickase domain from the emerging human bocavirus. *Journal of Virology* **87**(21), 11487–11493.
19. **Allander T, *et al.*** (2005) Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proceedings of the National Academy of Sciences of the United States of America* **102**(36), 12891–12896.
20. **Guido M, *et al.*** (2016) Human bocavirus: current knowledge and future challenges. *World Journal of Gastroenterology* **22**(39), 8684–8697.
21. **Allander T, *et al.*** (2007) Human bocavirus and acute wheezing in children. *Clinical Infectious Diseases: an Official Publication of the Infectious Diseases Society of America* **44**(7), 904–910.
22. **Han TH, *et al.*** (2009) Detection of human bocavirus-2 in children with acute gastroenteritis in South Korea. *Archives of Virology* **154**(12), 1923–1927.
23. **Jin Y, *et al.*** (2011) High prevalence of human bocavirus 2 and its role in childhood acute gastroenteritis in China. *Journal of Clinical Virology: the Official Publication of the Pan American Society for Clinical Virology* **52**(3), 251–253.
24. **Kesebir D, *et al.*** (2006) Human bocavirus infection in young children in the United States: molecular epidemiological profile and clinical characteristics of a newly emerging respiratory virus. *The Journal of Infectious Diseases* **194**(9), 1276–1282.
25. **Monavari SH, *et al.*** (2013) Human bocavirus in Iranian children with acute gastroenteritis. *Medical Journal of the Islamic Republic of Iran* **27** (3), 127–131.
26. **Xiang JY, *et al.*** (2014) [Etiological study of human bocavirus 1–4 in children with acute diarrhea in Lanzhou, China]. *Bing Du Xue Bao=Chinese Journal of Virology* **30**(4), 402–407.
27. **Li H, *et al.*** (2015) The genomic and seroprevalence of human bocavirus in healthy Chinese plasma donors and plasma derivatives. *Transfusion* **55**(1), 154–163.
28. **Arnold JC, *et al.*** (2006) Human bocavirus: prevalence and clinical spectrum at a children's hospital. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* **43**(3), 283–288.
29. **Kang LH, *et al.*** (2013) Simultaneous detection of waterborne viruses by multiplex real-time PCR. *Journal of Microbiology* (*Seoul, Korea*) **51**(5), 671–675.
30. **Kapoor A, *et al.*** (2009) A newly identified bocavirus species in human stool. *The Journal of Infectious Diseases* **199**(2), 196–200.
31. **Cashman O and O'Shea H**. (2012) Detection of human bocaviruses 1, 2 and 3 in Irish children presenting with gastroenteritis. *Archives of Virology* **157**(9), 1767–1773.
32. **Albuquerque MC, *et al.*** (2007) Human bocavirus infection in children with gastroenteritis, Brazil. *Emerging Infectious Diseases* **13**(11), 1756–1758.

33. **Arthur JL, *et al.*** (2009) A novel bocavirus associated with acute gastroenteritis in Australian children. *PLoS Pathogens* **5**(4), e1000391.
34. **Tamura K, *et al.*** (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**(12), 2725–2729.
35. **Saitou N and Nei M**. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4** (4), 406–425.
36. **Choi EH, *et al.*** (2006) The association of newly identified respiratory viruses with lower respiratory tract infections in Korean children, 2000–2005. *Clinical Infectious Diseases* **43**(5), 585–592.
37. **Manning A, *et al.*** (2006) Epidemiological profile and clinical associations of human bocavirus and other human parvoviruses. *The Journal of Infectious Diseases* **194**(9), 1283–1290.
38. **Fry AM, *et al.*** (2007) Human bocavirus: a novel parvovirus epidemiologically associated with pneumonia requiring hospitalization in Thailand. *The Journal of Infectious Diseases* **195**(7), 1038–1045.
39. **Foulongne V, *et al.*** (2006) Human bocavirus in French children. *Emerging Infectious Diseases* **12**(8), 1251–1253.
40. **Babkin IV, *et al.*** (2013) Evolutionary time-scale of primate bocaviruses. *Infection, Genetics and Evolution* **14**, 265–274.

https://doi.org/10.1017/S0950268818000560 Published online by Cambridge University Press