

## Clinical epidemiology and molecular profiling of human bocavirus in faecal samples from children with diarrhoea in Guangzhou, China

D.-M. ZHANG<sup>1,2†</sup>, M.-M. MA<sup>1†</sup>, W.-T. WEN<sup>2,3†</sup>, X. ZHU<sup>2,3</sup>, L. XU<sup>2,3</sup>,  
Z.-J. HE<sup>1,2,3</sup>, X. HE<sup>4</sup>, J.-H. WU<sup>2,3</sup>, Y.-W. HU<sup>2,3</sup>, Y. ZHENG<sup>2</sup>, Y. DENG<sup>1,2</sup>,  
C.-J. LIN<sup>2,3</sup>, J.-H. LU<sup>1,2</sup>, M.-F. LI<sup>2,3\*</sup> AND K.-Y. CAO<sup>2,3\*</sup>

<sup>1</sup> Department of Medical Statistic and Epidemiology, School of Public Health, Sun Yat-sen University, Guangzhou, Guangdong, China

<sup>2</sup> Key Laboratory for Tropical Disease Control, Ministry of Education, Guangzhou, Guangdong, China

<sup>3</sup> Department of Microbiology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, Guangdong, China

<sup>4</sup> Department of Clinical Laboratory, Sun Yat-sen University Cancer Centre, Guangzhou, Guangdong, China

Received 28 April 2014; Final revision 28 August 2014; Accepted 4 November 2014;  
first published online 3 December 2014

### SUMMARY

To understand the clinical epidemiology and molecular characteristics of human bocavirus (HBoV) infection in children with diarrhoea in Guangzhou, South China, we collected 1128 faecal specimens from children with diarrhoea from July 2010 to December 2012. HBoV and five other major enteric viruses were examined using real-time polymerase chain reaction. Human rotavirus (HRV) was the most prevalent pathogen, detected in 250 (22·2%) cases, followed by enteric adenovirus (EADV) in 76 (6·7%) cases, human astrovirus (HAstV) in 38 (3·4%) cases, HBoV in 17 (1·5%) cases, sapovirus (SaV) in 14 (1·2%) cases, and norovirus (NoV) in 9 (0·8%) cases. Co-infections were identified in 3·7% of the study population and 23·5% of HBoV-positive specimens. Phylogenetic analysis revealed 14 HBoV strains to be clustered into species HBoV1 with only minor variations among them. Overall, the detection of HBoV appears to partially contribute to the overall detection gap for enteric infections, single HBoV infection rarely results in severe clinical outcomes, and HBoV sequencing data appears to support conserved genomes across strains identified in this study.

**Key words:** Children, diarrhoea, human bocavirus, phylogenetic analysis.

### INTRODUCTION

Diarrhoea is one of the most common clinical symptoms to occur among children, and in 2011 was the

fourth (9·3%) greatest cause of reported mortality in children aged <5 years, worldwide. Human rotavirus (HRV), norovirus (NoV), enteric adenovirus (EADV), human astrovirus (HAstV), as well as sapovirus (SaV) are recognized as five major viral aetiological agents of diarrhoea. HRV, particularly group A HRV (HRV-A), is identified as the most common cause of severe paediatric diarrhoea throughout the world, mostly impacting children in developing countries [1]. NoV is an emerging enteric infectious agent

\* Authors for correspondence: Dr Kaiyuan Cao or Dr Mengfeng Li, Zhongshan School of Medicine, Sun Yat-sen University, 74 Zhongshan Road II, Guangzhou, Guangdong 510080, China.

(Email: caoky@mail.sysu.edu.cn) [K.-Y. Cao]

(Email: limf@mail.sysu.edu.cn) [M.-F. Li]

† These authors contributed equally to this work.

with highly diverse genogroups, of which genogroups I (GI) and II (GII) are recognized as a primary cause of epidemic gastroenteritis in individuals of all ages. Precisely, NoV is responsible for nearly 50% of all-cause diarrhoeal outbreaks worldwide [2]. EADV (primarily serotypes 40 and 41) is also recognized as an important cause of both community-acquired and hospital-acquired childhood diarrhoea. It is estimated that 3.2–12.5% of acute diarrhoeal cases in children are related to EADV [3]. Evidence also demonstrates that HAstV and SaV are associated with acute gastroenteritis [4, 5]. Recently a newly identified human respiratory pathogen, namely, human bocavirus (HBoV), was also frequently detected in stool specimens in cases of acute gastroenteritis [6–12], suggesting that HBoV may be a causative agent for human enteric infections.

HBoV was first described as a human pathogen in 2005 when Allander *et al.* [13], using a specific molecular screening procedure, cloned a human parvovirus from nasopharyngeal aspirates collected from children with respiratory tract diseases in Sweden. This agent was ultimately listed under the genus *Bocavirus* (subfamily Parvoviridae, family Parvoviridae) according to phylogenetic analysis, showing a close relationship to bovine parvovirus (BPV) and canine minute virus (CnMV). Since then HBoV has been frequently detected in respiratory tract samples throughout various regions of the world [14–16]. Such evidence has demonstrated HBoV to be a causative agent for respiratory illnesses, which has been further supported by more recent serodiagnostic evidence [17].

During the subsequent 5 years since its discovery, numerous epidemiological and molecular detection studies have characterized HBoV into four species, namely, HBoV1 (known as Swedish prototype), HBoV2 [7], HBoV3 [8], and HBoV4 [10]. These classifications are based on a >5% variation in NS-gene nucleotide sequences according to the 8th report of the International Committee on Taxonomy of Viruses (ICTV). HBoV1 has been detected frequently in both respiratory and faecal samples, whereas human stool seems to be the main source of HBoV2–4 [8, 10, 18, 19].

In 2007, HBoV DNA was first documented in human faeces from children with gastroenteritis at a similar frequency as for children with respiratory illnesses [6]. Subsequently, detection of HBoV in faecal specimens in children has been frequently reported despite a lack of understanding of the causative

association between HBoV and gastrointestinal symptoms [20–25]. Our current study was performed to achieve a better understanding of the viral aetiological spectrum of diarrhoea in children, as well as to determine the prevalence and molecular characteristics of HBoV enteric infections in Guangzhou. Guangzhou is a city located in South China featuring prominent socio-natural factors, such as a subtropical climate and high-density residential and non-residential populations, allowing for greater susceptibility to food-borne and airborne viral infections. For this purpose, a cross-sectional study was performed to investigate faecal specimens from children with diarrhoea for HBoV and other common diarrhoeal viruses using real-time polymerase chain reaction (real-time PCR) methods under a viral surveillance programme conducted in Guangzhou. Furthermore, complete genome sequencing of isolated HBoV strains was conducted and used for phylogenetic analysis.

## METHODS

### Ethics statement

All research involving human participants was approved by the Institutional Review Board of Zhongshan School of Medicine, Sun Yat-sen University, in accordance with the guidelines for the protection of human subjects. Participants received written informed consent on the study's purpose and of their right to keep information confidential. Written consent was obtained from their guardians.

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

### Study population and sample collection

In this cross-sectional study, we collected 1128 stool specimens from children aged <14 years (66.5% male) with diarrhoea symptoms between July 2010 and December 2012. These children were inpatients or outpatients (17.0% and 83.0%, respectively) presenting with diarrhoea in four general hospitals selected as sentinel sites located in Guangzhou, China. Diarrhoea was defined as having at least three defecations per day with abnormal changes in faecal character (loose, watery, mucopurulent or bloody purulent). One stool specimen was collected

per patient within 24 h of the patient's arrival at the hospital. Upon collection, specimens were added to 2 ml of viral transport media (VTM) consisting of Earle's Balanced Salt Solution (Bio Source International, USA), 4.4% bicarbonate, 5% bovine serum albumin (BSA), 100 µg/ml vancomycin, 30 µg/ml amikacin, and 40 U/ml nystatin according to a standard protocol and transported within 8 h at 4 °C to a bio-safety level 2 laboratory at Sun Yat-sen University. Once there, samples were divided into equal aliquots, and stored at -80 °C until processed. Demographic, epidemiological, and clinical information including case history, symptoms, physical signs, and examination results, were collected using a standardized and validated questionnaire. All specimens were tested for five common viral aetiological agents of diarrhoea, including human rotavirus (HRV) groups A, B and C (HRV-A, HRV-B, HRV-C), EADV, NoV), SaV, and HAstV, as well as HBoV, using the real-time PCR method as described below. Data of children whose faecal specimens were found to be positive for HBoV were analysed.

#### Nucleic acid extraction

Total DNA and RNA were extracted from 200 µl of thawed 10% stool suspension with the QIAamp MinElute Virus Spin kit (Qiagen, Germany), and reverse transcription of virus RNA was conducted using SuperScript III Reverse Transcriptase and Random Hexamer primers (Invitrogen, Life Technology, USA), according to the manufacturer's instructions.

#### Viral agent screening

HRV (HRV-A, HRV-B, HRV-C), EADV, NoV, SaV and HAstV were detected by real-time PCR techniques as previously described [5, 26, 27]. For screening of HBoV, TaqMan real-time PCR primers (NP1-F: 5'-AGAGGCTCGGGCTCATATCA-3' and NP1-R: 5'-TCTTCATCACTTGGTCTGAGGTCTT-3') and probe: 5'-FAM-AGGAACACCCAATCARCCACC TATCGT-TAMRA-3' (synthesized by Invitrogen, Life Technology, USA), were designed to bind the highly conserved NP1 region of various HBoV strains and analysed using Primer Express software v. 3.0 (Applied Biosystems, USA), with regard to optimal G+C content (55%, 44% and 60% for NP1-F, NP1-R and probe, respectively), melting temperature (58.4 °C, 59.3 °C and 70.0 °C for NP1-F, NP1-R and probe, respectively), and amplicon length

(88 bp). The complete genome sequence of HBoV from the st2 strain (GenBank accession number DQ000496, or NC\_007455) was used as a reference. Each reaction mixture consisted of 10 µl 2×IQ Supermix reaction mixture (Bio-Rad, USA), 2 µl of viral DNA, 0.5 µM each of the forward and reverse primer, 0.3 µM of the probe, and nuclease-free water to reach a final volume of 20 µl. Real-time PCR was conducted at 95 °C for 15 min, followed by 45 cycles of 95 °C for 15 s, and 60 °C for 1 min on an ABI7500 Real-time PCR system (Applied Biosystems, USA).

#### Complete genome sequencing for HBoV

The complete genome sequences of HBoV strains were amplified using primers designed using Primer Premier 5.0 software, which targeted conserved regions of HBoV strains referenced in the GenBank database (primer sequences are shown in Table 1). PCR amplification was performed using the Platinum pfx *Taq* polymerase (Invitrogen, USA) in a prepared reaction mixture according to the following condition: 95 °C for 3 min, followed by 40 cycles of 95 °C for 30 s, 52 °C to 58 °C (see Table 1 for melting temperature of different primers) for 45 s, and 72 °C for 1 min, with a final extension at 72 °C for 10 min. PCR products for genome analysis was purified using an agarose gel DNA purification kit (Bio Teke, China), and the PCR products of terminal sequences were cloned into pMD18-T vectors using a T-A cloning kit (Takara, China). All PCR products used for cloning and sequencing were from three independent PCR reactions. Sequencing was performed by a commercial service of Invitrogen Co. (Guangzhou, China), according to previously published methods [28]. Sequences were determined and analysed using the DNASTAR software package, and were also submitted to the GenBank database (accession numbers KJ684065–KJ684078).

#### Phylogenetic analysis

The genomic sequences and open reading frames (ORFs) of HBoV were comparatively analysed against complete genome sequences of other HBoV strains found in the GenBank database, including HBoV prototype strains st1 and st2, representative strains for three additional HBoV species, and other strains in the bocavirus genera, including parvovirus B19, BPV, and CnMV. These sequences were

Table 1. *The primers used for human bocavirus (HBoV) complete coding sequence analysis*

Primer	Sequence (5'–3')	Position*	Amplicon length	Melting temp. (°C)
HBoV-1F	GCCGGCAGACATATTGGATT	1–20	722	57
HBoV-1R	GGACGTGTAGCCAGAAGAGATT	701–722		
HBoV-2F	GGGAGAAGGACTAAGCAAGAG	615–635	784	52
HBoV-2R	GTCAAGCGAAGTAACAGGATG	1378–1398		
HBoV-3F	GGCAGCCTTATAGAACAAGT	1293–1313	800	56
HBoV-3R	GGTAACCCACCGCAATCAGT	2074–2092		
HBoV-4F	GAGATTGCAGCTCTTCTACAG	1912–1932	729	58
HBoV-4R	CCTCGTCTTCATCACTTGGT	2621–2640		
HBoV-5F	AAGACAAGCATCGCTCCTAC	2429–2448	719	55
HBoV-5R	ACAGGTTACCGTTATCAAGT	3127–3147		
HBoV-6F	CACAGACAGAAGCAGACGAGAT	2990–3011	776	55
HBoV-6R	GGTGAGAAGTGACAGCTGTATTG	3743–3765		
HBoV-7F	TGGTCACCTTACAAAACAGA	3646–3666	742	55
HBoV-7R	GTCTGATGACTGTGGTCCTACT	4366–4387		
HBoV-8F	TCCAACAAGAAGAGTTCAGT	4242–4262	1058	56.2
HBoV-8R	TGTACAACAACAACACATTA AAAAG	5276–5299		

\* According to GenBank accession number NC\_007455.

aligned using the Clustal X2 program ([www.clustal.org/clustal2/](http://www.clustal.org/clustal2/)), and a neighbour-joining tree was constructed using MEGA 5.2 software ([www.megasoftware.net/](http://www.megasoftware.net/)).

### Statistical analysis

Analyses for demographic, epidemiological and clinical information of patients were performed using SPSS v. 21.0 statistical software (SPSS Inc., USA). Specifically, correlation between main clinical symptoms and causative viral agents in children from whom a single virus was confirmed was determined using a non-parametric test or Fisher's exact test.

## RESULTS

### Virological surveillance

Of the 1128 faecal samples collected from enrolled children with diarrhoea symptoms between July 2010 and December 2012, at least one virus (HRV-A, -B and -C, EADV, NoV, SaV, HAstV and HBoV) was detected in 370 (32.8%) stool samples. The detection rates of each virus are summarized in Table 2. Of the cases with a single or multiple viral infection, 20.3% (229) were positive for HRV-A (median age 1 year), 2.2% (25) for HRV-B (median age 1 year), 0.4% (5) for HRV-C (median age 9 months), 6.7% (76) for EADV (median age 1 year), 0.8% (9) for NoV (median age 1 year), 1.2% (14) for SaV (median

age 1 year), and 3.4% (38) for HAstV (median age 1 year). A single virus was detected in 329 specimens (29.2% of the whole study group), of which HRV-A was the most frequent. HBoV DNA was detected in 17 (1.5%) samples (median age 11 months) by real-time PCR, including 16 from outpatients and one from an inpatient. The mean age of the 1128 children was 23.5 months, ranging from 10 days to 14 years, with a gender ratio (male:female) of 2:1 (750:378). Overall, HRV and EADV predominated in almost all age groups, being most common in the <3 years age group (Fig. 1). This was also true for HAstV and NoV, while SaV seemed to have a greater prevalence in children aged <2 years. There was no clearly defined age distribution for HBoV despite its high prevalence in children aged <1 year, as 5/17 positive specimens were collected from children aged >6 years.

Assuming relative stability of seasonal distribution of targeted enteric viruses during continuous surveillance for the 2.5-year period (from July 2010 and December 2012), the cumulative monthly distribution of each viral aetiological agent assessed in the study was analysed. HRV was detected throughout the entire year with the exception of from March to May. HRV appeared to spike between August and December. EADV featured a similar distribution as HRV in spite of no hits in March and August and a relatively higher percentage in April and June. HAstV, SaV and NoV were all

Table 2. Distribution of co-infections and single-virus infections in children with diarrhoea

Viral agents detected	No. of specimens	% in whole study group
Positive	370	32.8
One virus detected	329	29.2
HRV-A	196	17.4
HRV-B	16	1.4
HRV-C	3	0.3
EADV	58	5.1
NoV	6	0.5
SaV	10	0.9
HAstV	27	2.4
HBoV	13	1.2
Two viruses detected	39	3.5
HRV-A, HRV-B	5	0.4
HRV-A, HRV-C	1	0.1
HRV-A, EADV	13	1.2
HRV-A, NoV	2	0.2
HRV-A, SaV	3	0.3
HRV-A, HAstV	6	0.5
HRV-A, HBoV	1	0.1
EADV, SaV	1	0.1
HRV-B, HRV-C	1	0.1
EADV, HAstV	3	0.3
EADV, HBoV	1	0.1
HAstV, HBoV	1	0.1
HRV-B, NoV	1	0.1
Three viruses detected	2	0.2
HRV-A, HRV-B, HAstV	1	0.1
HRV-A, HRV-B, HBoV	1	0.1
Negative	758	67.2
Total	1128	100.0

sporadically detected with the highest hits in October, October and November, respectively. While HBoV was detectable during winter and summer, the percentage of detection (7.1%) was greatest in June (Fig. 2).

### Co-infection

As shown in Table 2, there were 15 different combinations of co-infections identified in 41 stool specimens (3.7% of the study population). HRVs were the most common viruses to be identified in a mixed infection, with HRV-A and EADV co-infection being detected in 13/41 stool samples (1.2% of the study population). Only two triple mixed infections were identified, one being HRV-A/HRV-B/HAstV, and the other being HRV-A/HRV-B/HBoV. Co-infections were identified

in four of the 17 HBoV-positive patients (23.5%), one of which was the individual infected with both HRV-A and HRV-B in addition to HBoV.

### Clinical characteristics of targeted single-virus positive cases

Of the 329 cases from which a single virus was detected, 314 (95.4%) had one or more episodes of diarrhoea, and 51 (15.5%) reported vomiting, while 20 (6.1%) had symptoms of respiratory illness (Table 3). Other manifestations such as fever, dehydration and neurological symptoms were rare. There was a significant correlation between the causative viral agents and frequencies of diarrhoea ( $P < 0.01$ , Kruskal–Wallis test) and vomiting ( $P < 0.01$ ), while no difference was found with the occurrence of respiratory symptoms ( $P > 0.05$ ). In comparing HRV (featuring the most frequent bouts of diarrhoea and episodes of vomiting, suggesting greater disease severity) with HBoV, diarrhoea was more frequent in cases with HRV infection ( $P < 0.05$ ), although not so for those with vomiting ( $P > 0.05$ ). HBoV-positive patients were all in a mild disease condition with no vomiting, fever or respiratory symptoms.

### Complete genome sequences and phylogenetic analysis

The complete genomes of HBoV identified from positive faecal samples with  $C_t \leq 34.7$  (14/17) were amplified and sequenced. All the genomes were 5299 bp with an average G + C content of 42.31%, except for 2012GZ1185, 2012GZ8500, 2012GZ5497 (5298, 5298 and 5297 bp, respectively). Genomic organization of the 14 HBoV sequences was the same as that of the two Swedish prototype HBoV strains st1 and st2 and other bocaviruses detected thus far including HBoV2, HBoV3, HBoV4, human parvovirus B19, CnMV and BPV, having three ORFs with the order of NS1, NP1, and VP1/VP2. Representative strains from the GenBank database described above were used to conduct phylogenetic analysis based on complete genome, and results showed that the 14 Guangzhou strains were genetically close to HBoV1 (Fig. 3). The 14 strains were highly conserved with 99% similarity to one another in all three genes as well as to the prototype strains when conducting BLASTn in NCBI, with only limited variations found in predicted amino acid sequences as shown in tBLASTx. Further complete genome phylogenetic analyses of these 14 strains and the various strains



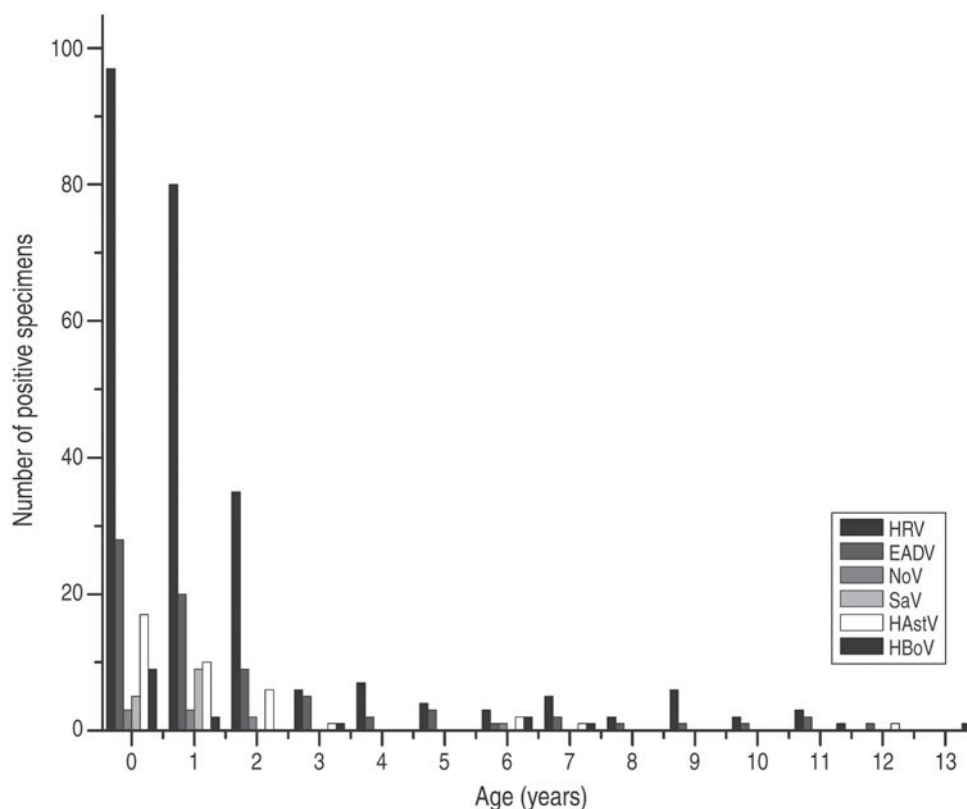
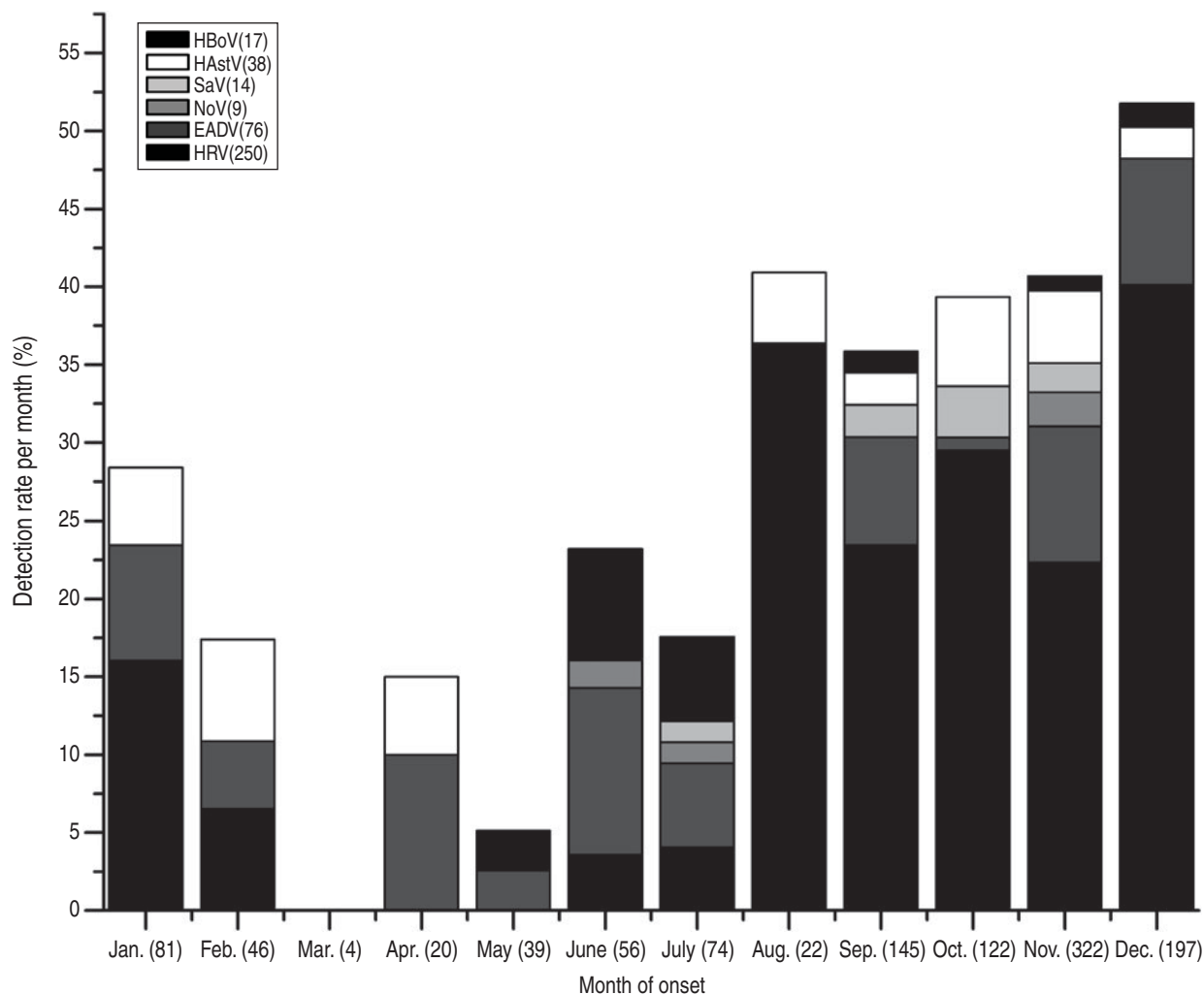


Fig. 1. Age distribution of viral agents in children aged <14 years with diarrhoea.

of HBoV1 from the GenBank database (45 strains in total) showed that there were three clearly divided groups (Fig. 4a). The strains detected in this study were all classified as group I, along with prototype st1 and st2 and most other HBoV1 strains, while Thailand strain CU74 and Taiwan strain TW2717\_06 form group II and group III along with the two Guangzhou strains, LWK and GZ9081. Based upon the phylogenetic analysis of the three genes, there were no apparent genotypic changes between the phylogenetic trees for the 14 strains (Fig. 4b–d). Consistent with previous studies [13, 29], NS1 appeared to qualify as the most conserved gene, whereas most nucleotide polymorphisms occurred in the VP1/VP2 gene, although there were only limited overall nucleotide and amino acid variations in different strains (<2% and 1%, respectively). It should be noted that the VP1/VP2 gene and the complete genome have almost identical phylogenetic trees. This suggests that it may be possible to sequence VP1/VP2 for routine laboratory surveillance instead of characterizing the entire genome for circulating HBoVs.

## DISCUSSION

Diarrhoea is a common clinical syndrome of various gastrointestinal diseases, such as acute gastroenteritis, which most notably affects children who might easily develop life-threatening complications if appropriate treatment is not provided in a timely manner. Viruses are recognized as a major cause of diarrhoea in children, which are largely preventable. These viruses include HRV, NoV, EADV, HAstV, SaV and others, with HRV being the most common cause of severe diarrhoea in children aged <5 years [1]. Our study identified that 31.3% of specimens collected from 1128 diarrhoea cases were attributed to the five enteric viruses tested for in this study, with HRV being the most frequently detected virus in 250 (22.2%) cases, followed by EADV in 76 (6.7%) cases, HAstV in 38 (3.4%) cases, SaV in 14 (1.2%) cases, and NoV in nine (0.8%) cases. These findings are consistent with other published work [30, 31]. However, NoV, a common cause of diarrhoeal outbreaks, is the least reported cause of diarrhoea in children in this study. This may be due to rare outbreaks



**Fig. 2.** Cumulative (2.5 years, from July 2010 and December 2012) monthly distribution of viral agents in children aged <14 years with diarrhoea. The number in parentheses after each month indicates the number of specimens tested per month.

in subjects enrolled into this programme. In addition, differently from HRV, EADV, HAstV and SaV which are usually endemic in children, NoV can cause epidemics in all age groups. [2]. Mixed infections were found in 3.7% of the study population, also consistent with previous publications [32, 33]. The high proportion (67.2%) of samples testing negative for any of the viruses suggests an impressive detection gap that researchers could further explore. This high negative rate is why numerous studies have been dedicated to tracing enteric infection for HBoV and investigating its causative role in gastrointestinal diseases.

Since first being identified in faecal samples in 2007, HBoV has been suspected as a possible cause of acute gastroenteritis in children [6, 32]. This has resulted in the implementation of many enteric aetiological

surveillance programmes throughout the world. In this study, HBoV was detected in 17 (1.5%) cases, which falls into the previously described detection range of 0.8–9.1% [20, 22, 34, 35]. Interestingly, 4/17 HBoV-positive patients (23.5%) were co-infected with other viruses, one of which was with two groups of HRV.

The age distribution of the major five enteric agents (HRV, EADV, HAstV, SaV, NoV) targeted in this study was similar to previous reports [1–5], while atypical distribution occurred in patients who were HBoV-positive, with 5/17 positive specimens being collected from children aged >6 years and 9/17 collected from children aged <1 year. These findings suggest that HBoV enteric infections should not be ruled out in older children, although additional studies

Table 3. Correlation between primary clinical symptoms and causative viral agents in children where a single viral infection was confirmed

Disease manifestation	HRV (n = 215)	EADV (n = 58)	NoV (n = 6)	SaV (n = 10)	HAstV (n = 27)	HBoV (n = 13)	P value	
Number of bouts of diarrhoea (stools/24 h)								
0	13	2	0	0	0	0	$P < 0.01^*$	$P < 0.05^\ddagger$
1–3	19	9	0	1	1	4		
4–6	110	43	4	7	22	9		
≥7	73	4	2	2	4	0		
Number of bouts of vomiting (episodes/24 h)								
0	166	57	6	9	27	13	$P < 0.01^*$	$P > 0.05^\ddagger$
1–2	2	0	0	0	0	0		
3–4	27	0	0	1	0	0		
≥5	20	1	0	0	0	0		
Respiratory symptoms								
Yes	11	5	0	0	4	0	$P > 0.05^\ddagger$	–
No	204	53	6	10	23	13		

\* Kruskal–Wallis test comparing all viruses.

† Mann–Whitney *U* test comparing human bocavirus and rotavirus.

‡ Fisher's exact test comparing all viruses.

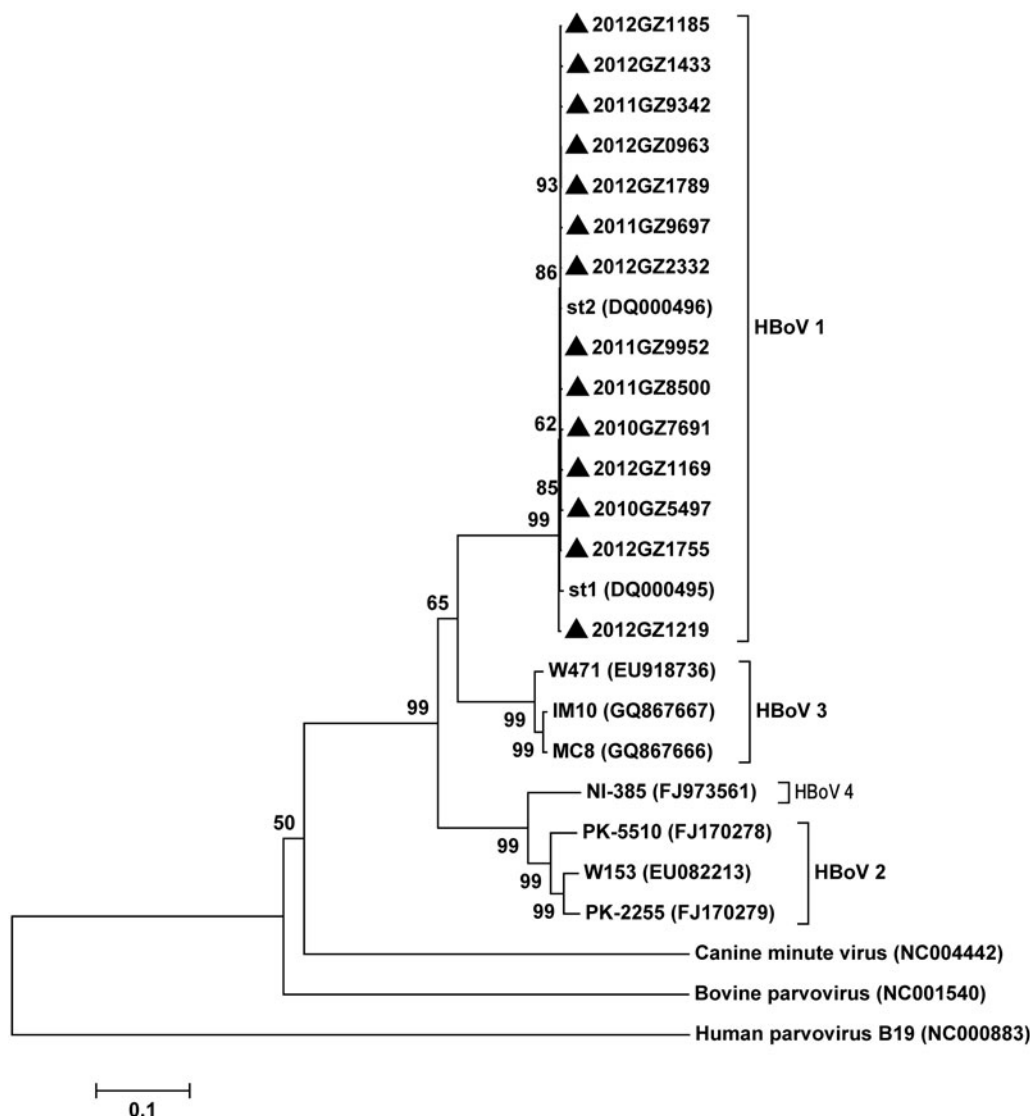
should be performed to explore these findings. Study results showed a slight variation in seasonal distribution of HBoV from previous studies [29, 32], with the highest rate (7.1%) of HBoV occurring in June. It is likely that this higher rate is due to the small sample size during the summer months. Seasonal patterns of other viruses such as HRV (peaks in summer and winter) and NoV (peaks in cold weather months) were consistent with previous findings [2, 36]. Patients positive for HBoV all had mild clinical symptoms with no vomiting, fever or respiratory symptoms, in contrast to those positive for HRV which were more likely to have severe symptoms. This suggests that HBoV, without concomitant infection with other pathogens, largely results in mild disease outcomes.

While the detection of HBoV in children in this study might explain part of the diagnostic gap, the real aetiological role of HBoV in enteric diseases still remains largely controversial. Most studies exploring this area are based on the cross-sectional design, which cannot establish causality. So far, only one case-control study conducted in Australia has demonstrated that acute gastroenteritis was related to higher prevalence of HBoV2 (odds ratio 2.6, 95% confidence interval 1.2–5.7) [8]. No similar findings have been found in other studies [25, 37]. HBoV3 and HBoV4 are also primarily identified in stool samples without clear correlation with enteric diseases [8, 10, 21, 37, 38]. Evidence from recent studies tend to support that the presence of HBoV1 in faeces may be

due to prolonged virus shedding from previous respiratory tract infections based on a persistence characteristic attributed to viruses in the family Parvoviridae. This hypothesis could also partly explain the high co-infection rate and frequent simultaneous detection in both respiratory and enteric secretions, as well as the indiscriminate distribution in diarrhoeal symptomatic and asymptomatic subjects [20, 21, 23, 37, 39, 40]. Our research is still too limited to reveal the real pathogenesis of HBoV in enteric infections although the fact of high co-infection rate and mild clinical outcomes in single HBoV infection cases were identified. More studies based on a prospective and controlled epidemiological design are needed to better establish the causality between HBoV and diarrhoeal diseases.

In this study, the complete genomes of HBoV in positive faecal samples with a  $C_t \leq 34.7$  (14/17) were amplified and sequenced. The 5.2-kb genome of HBoV contains three ORFs encoding NS1 (a non-structural protein), NP1 (a protein of unknown function), and VP1/VP2 (a viral capsid protein). Of the three genes, NS1 appears to be the most conserved, with only minor nucleotide polymorphisms identified. Most sequence variations were instead found to occur in the VP1/VP2 gene, although the overall nucleotide and amino acid differences in different strains were only 2% and 1% for VP1 and VP2, respectively [13]. All of the HBoV strains were found to be clustered into species HBoV1 in the constructed phylogenetic



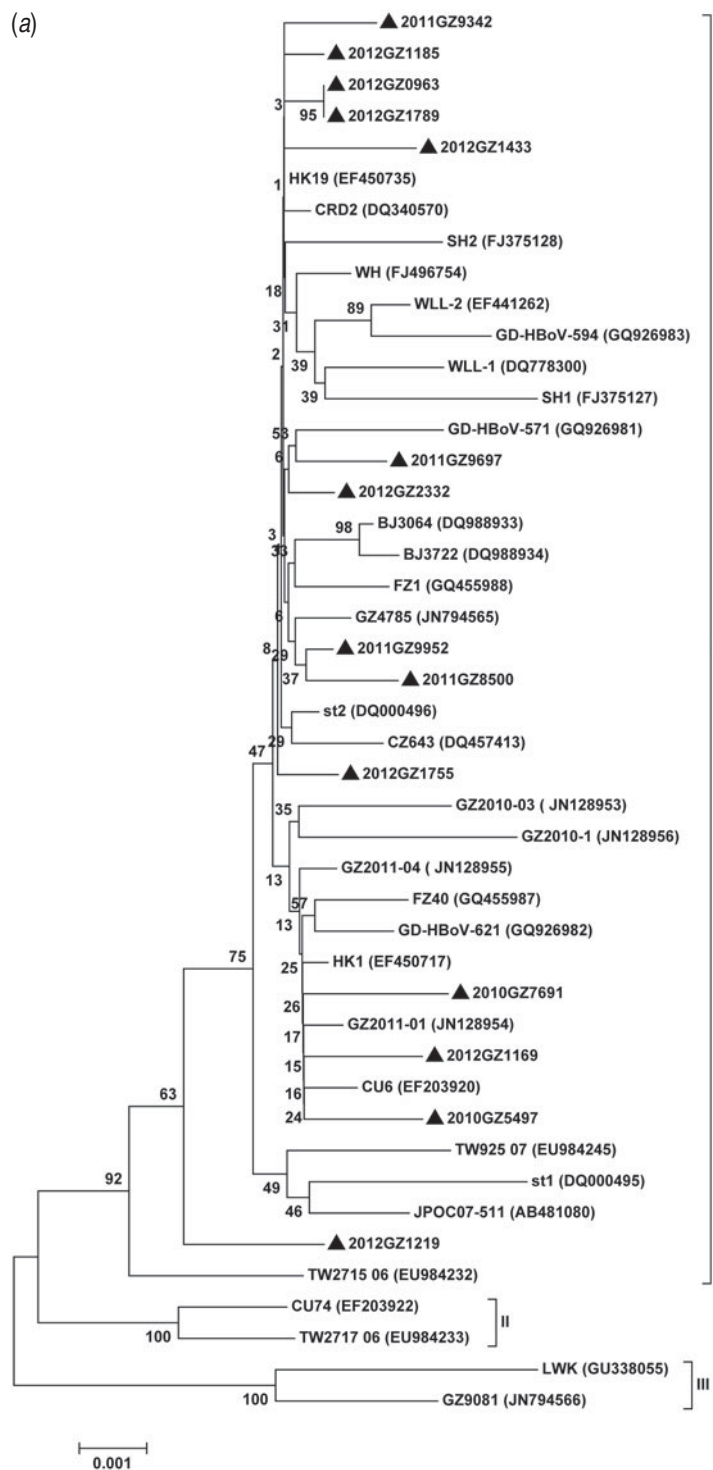


**Fig. 3.** Phylogenetic tree constructed from the complete genomes of 14 human bocavirus strains identified in faecal samples, along with other representative bocavirus genotypes available in the GenBank database. Phylogenetic trees were inferred from complete genome data with 1000 bootstrap replicates using the neighbour-joining method with Mega 5.2 software. The scale bar indicates the estimated number of substitutions per 500 bases by Kimura's two-parameter model. Parvovirus B19, bovine parvovirus, canine minute virus, and representative strains of HBoV1–4 were used as reference strains for genotype analysis of 14 human bocavirus strains identified in faecal samples (labelled with a black triangle). The 14 HBoV strains obtained in this study were identified as HBoV1.

tree, with no HBoV2 strains detected. They displayed limited sequence variation among themselves and in relation to the sequences of previously described HBoV strains in the GenBank database.

In conclusion, the present study investigated the aetiological viral spectrum of diarrhoea in children in Guangzhou, China and analysed the clinical characteristics and genomic phylogeny of HBoV infection. While HRV and EADV were most frequently

detected, HBoV was identified in 1.5% of all faecal specimens tested, helping to reduce the detection gap of enteric viruses. In addition, we found that 23.5% of the HBoV-positive samples were also positive for another enteric virus, and single HBoV infection rarely resulted in severe clinical outcome. Moreover, phylogenetic analysis of 14 HBoV complete genomes, all classified as HBoV1, revealed rather conserved sequences across analysed strains.



**Fig. 4.** Phylogenetic trees of complete genomes and NS1, NP1, and VP1/VP2 gene sequences of 14 human bocavirus (HBoV) strains identified in faecal samples, along with other representative HBoV1 strains available in the GenBank database. Phylogenetic trees were inferred from complete genome (a), NS1 (b), NP1 (c), and VP1/VP2 (d) gene data with 1000 bootstrap replicates using the neighbour-joining method with Mega 5.2 software. The scale bar indicates the estimated number of substitutions per 500 bases by Kimura's two-parameter model. Specifically, 1920 nucleotide positions in each NS1 gene, 660 nucleotide positions in each NP1 gene, and 2016 nucleotide positions in each VP1/VP2 gene were included in the analysis. The 14 HBoV1 strains studied are labelled with a black triangle. Swedish prototype strains st1 and st2, American strain CRD2, Japanese strain JPOC07-511, Thailand strains CU6 and CU74, Taiwanese strains TW925\_07, TW2715\_06 and TW2717\_06, Chinese strains HK1, HK19, WLL-1, WLL-2, CZ643, FZ1, FZ40, BJ3064, BJ3722, GD-HBoV-571, GD-HBoV-594, GD-HBoV-621 were included as HBoV1 reference strains.

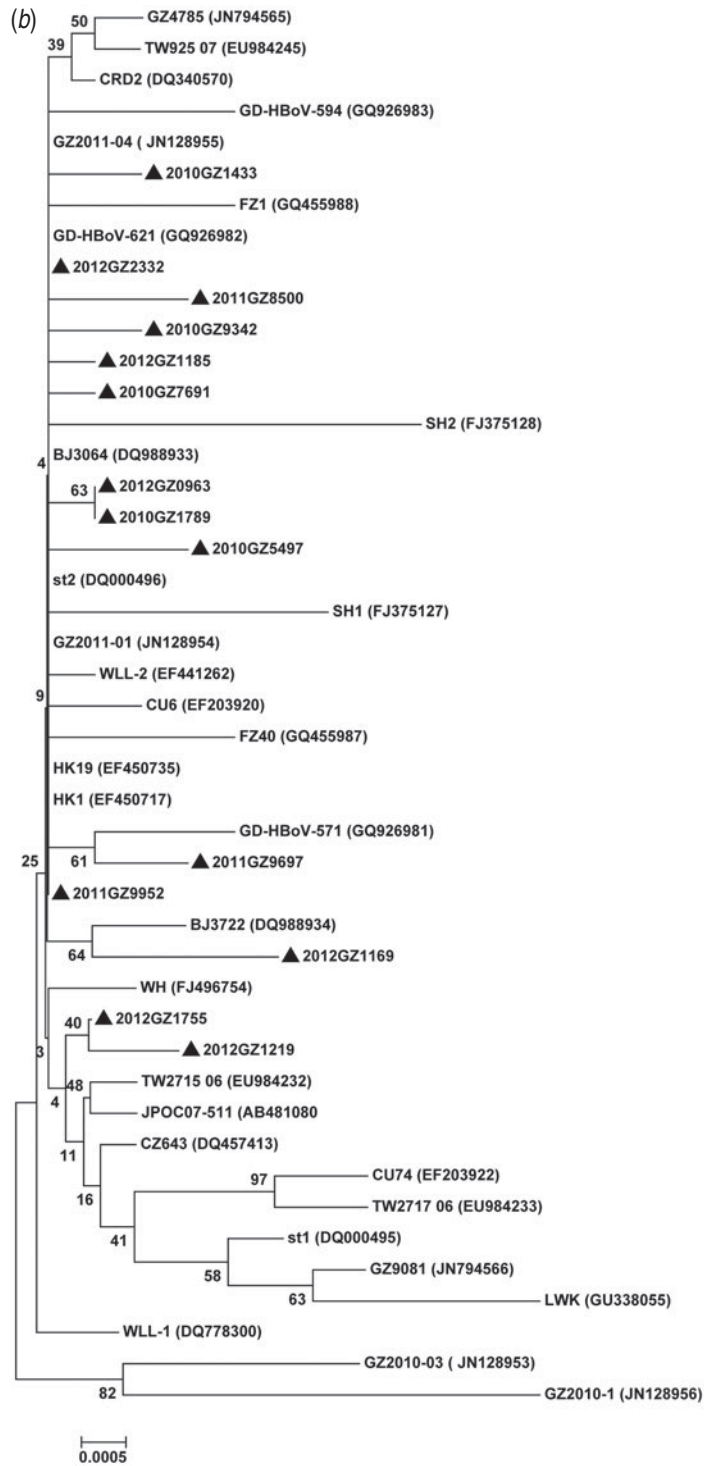


Fig. 4. (cont.)

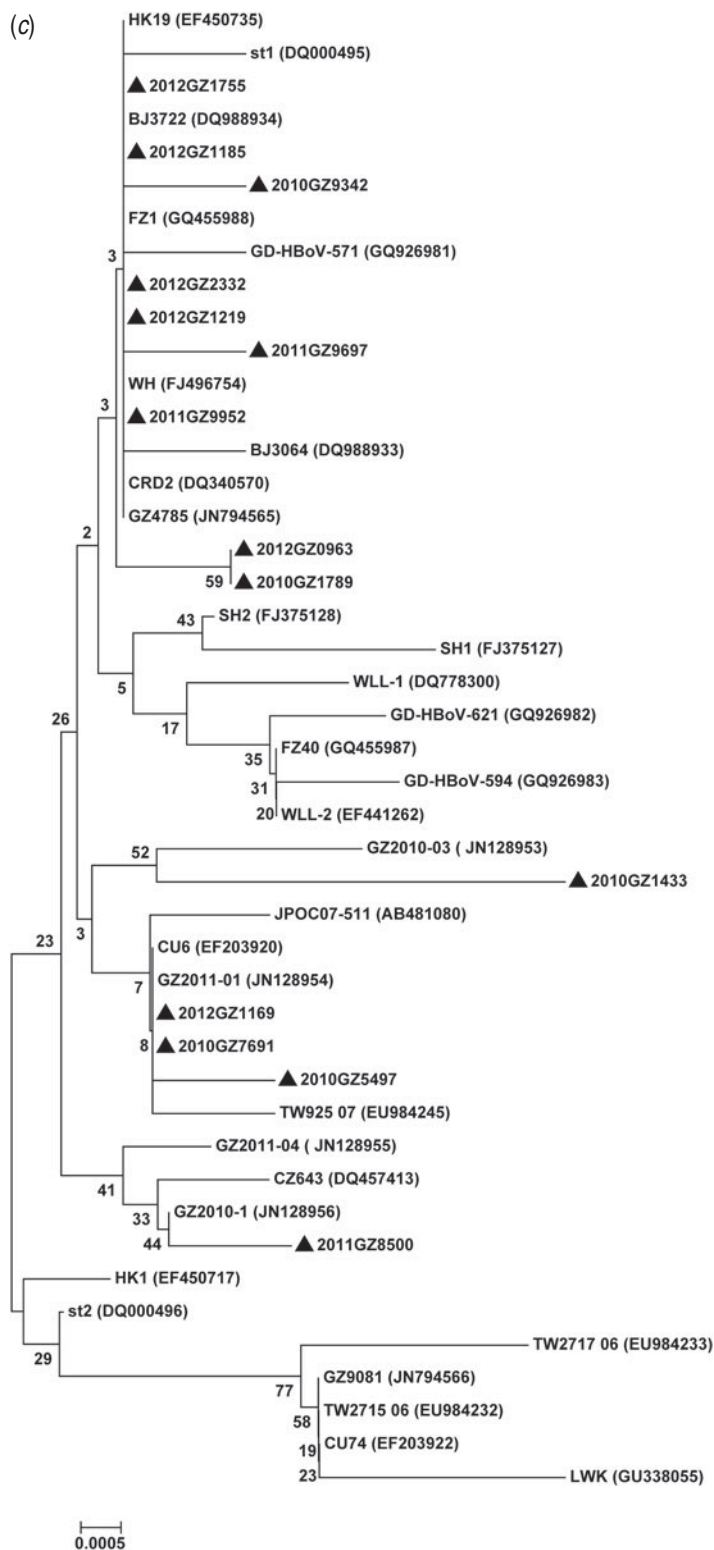


Fig. 4. (cont.)

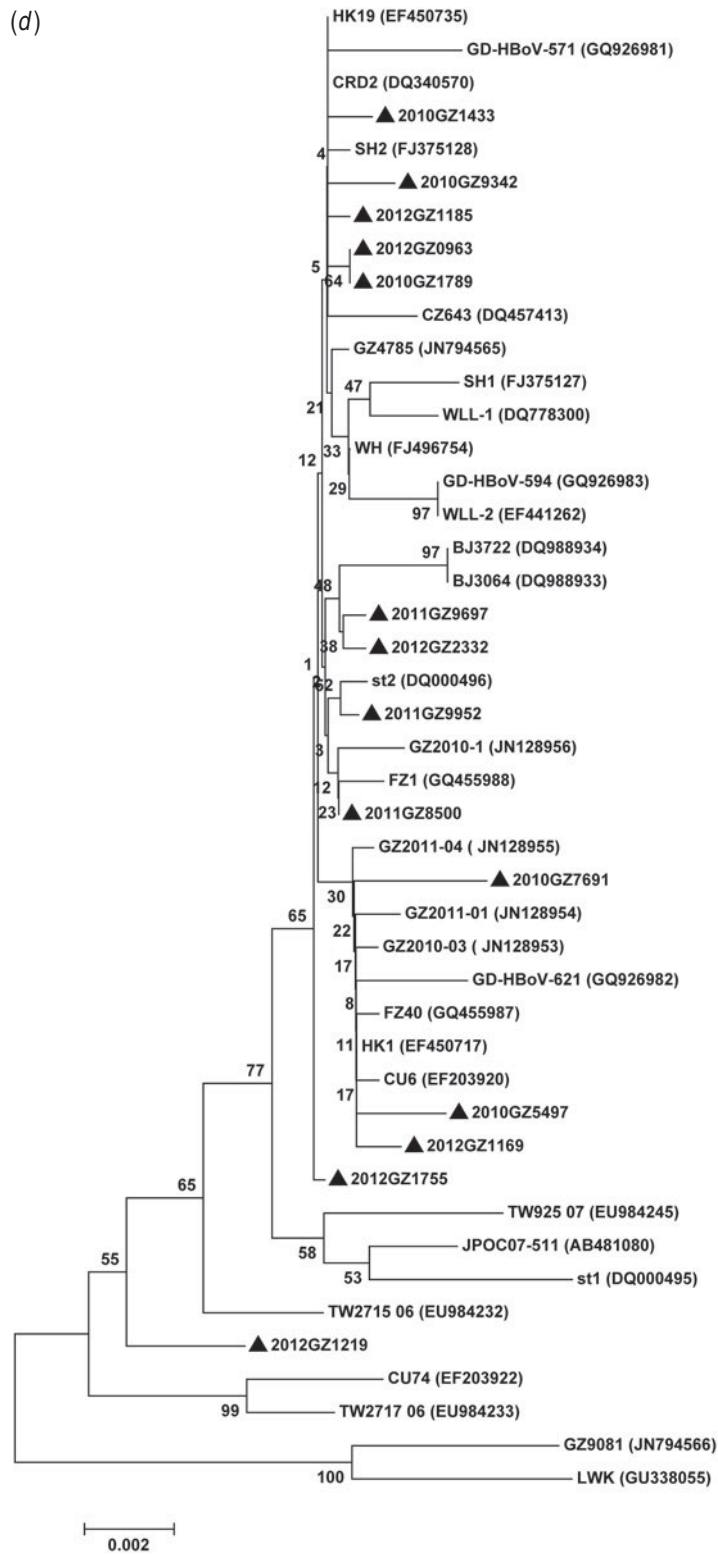


Fig. 4. (cont.)



## ACKNOWLEDGEMENTS

We owe special thanks to the doctors and nurses of the participating hospitals for their help in collecting samples. We thank PhD candidate Benjamin D. Anderson at Duke University, USA and PhD candidate Mary M. Merrill at the University of Florida, USA, for kindly reviewing and editing the manuscript.

This work was supported by the Natural Science Foundation of China (grant nos. 81201283, 81102370, 41176128); the Key (Key grant) Project of the Chinese Ministry of Education (grant no. 311030); the National Science and Technique Major Project (grant nos. 2012ZX10004-213, 201305017); Guangdong Provincial Natural Science Foundation (grant no. 8151008901000210); and Guangdong Recruitment Program of Creative Research Groups (grant no. 2009010058).

## DECLARATION OF INTEREST

None.

## REFERENCES

1. Parashar UD, *et al.* Global mortality associated with rotavirus disease among children in 2004. *Journal of Infectious Diseases* 2009; **200**: S9–S15.
2. Patel MM, *et al.* Noroviruses: a comprehensive review. *Journal of Clinical Virology* 2009; **44**: 1–8.
3. Liu L, *et al.* Adenoviruses associated with acute diarrhea in children in Beijing, China. *PLoS ONE* 2014; **9**.
4. Li C-y, *et al.* Outbreak of neonatal gastroenteritis associated with astrovirus serotype 1 at a hospital in Inner Mongolia, China. *Journal of Clinical Microbiology* 2010; **48**: 4306–4309.
5. Wu W, *et al.* Surveillance of pathogens causing gastroenteritis and characterization of norovirus and sapovirus strains in Shenzhen, China, during 2011. *Archives of Virology* 2014.
6. Vicente D, *et al.* Human bocavirus, a respiratory and enteric virus. *Emerging Infectious Diseases* 2007; **13**: 636–637.
7. Kapoor A, *et al.* A newly identified bocavirus species in human stool. *Journal of Infectious Diseases* 2009; **199**: 196–200.
8. Arthur JL, *et al.* A novel bocavirus associated with acute gastroenteritis in Australian children. *PLoS Pathogens* 2009; **5**: e1000391.
9. Han TH, *et al.* Detection of human bocavirus-2 in children with acute gastroenteritis in South Korea. *Archives of virology* 2009; **154**: 1923–1927.
10. Kapoor A, *et al.* Human bocaviruses are highly diverse, dispersed, recombination prone, and prevalent in enteric infections. *Journal of Infectious Diseases* 2010; **201**: 1633–1643.
11. Santos N, *et al.* Human bocavirus species 2 and 3 in Brazil. *Journal of Clinical Virology* 2010; **48**: 127–130.
12. Khamrin P, *et al.* Circulating of human bocavirus 1, 2, 3, and 4 in pediatric patients with acute gastroenteritis in Thailand. *Infection, Genetics and Evolution* 2012; **12**: 565–569.
13. Allander T, *et al.* Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proceedings of the National Academy of Sciences USA* 2005; **102**: 12891–12896.
14. Weissbrich B, *et al.* Frequent detection of bocavirus DNA in German children with respiratory tract infections. *BMC Infectious Diseases* 2006; **6**: 109.
15. Smuts H, Hardie D. Human bocavirus in hospitalized children, South Africa. *Emerging Infectious Diseases* 2006; **12**: 1457–1458.
16. I PM, *et al.* Pediatric hospitalization of acute respiratory tract infections with human bocavirus in Hong Kong. *Journal of Clinical Virology* 2008; **42**: 72–74.
17. Kantola K, *et al.* Serodiagnosis of human bocavirus infection. *Clinical Infectious Diseases* 2008; **46**: 540–546.
18. Chieochansin T, *et al.* Absence of detectable replication of human bocavirus species 2 in respiratory tract. *Emerging Infectious Diseases* 2009; **15**: 1503–1505.
19. Chow BD, Ou Z, Esper FP. Newly recognized bocaviruses (HBoV, HBoV2) in children and adults with gastrointestinal illness in the United States. *Journal of Clinical Virology* 2010; **47**: 143–147.
20. Cheng WX, *et al.* Human bocavirus in children hospitalized for acute gastroenteritis: a case-control study. *Clinical Infectious Diseases* 2008; **47**: 161–167.
21. Nawaz S, *et al.* Human bocaviruses are not significantly associated with gastroenteritis: results of retesting archive DNA from a case control study in the UK. *PLoS ONE* 2012; **7**: e41346.
22. Chieochansin T, *et al.* Human bocavirus infection in children with acute gastroenteritis and healthy controls. *Japanese Journal of Infectious Diseases* 2008; **61**: 479–481.
23. Proenca-Modena JL, *et al.* Viral load of human bocavirus-1 in stools from children with viral diarrhoea in Paraguay. *Epidemiology and Infection* 2013; **141**: 2576–2580.
24. Streiter M, *et al.* Does human bocavirus infection depend on helper viruses? A challenging case report. *Virology Journal* 2011; **8**: 417.
25. Jin Y, *et al.* High prevalence of human bocavirus 2 and its role in childhood acute gastroenteritis in China. *Journal of Clinical Virology* 2011; **52**: 251–253.
26. Trang NV, *et al.* Detection and molecular characterization of noroviruses and sapoviruses in children admitted to hospital with acute gastroenteritis in Vietnam. *Journal of Medical Virology* 2012; **84**: 290–297.
27. Garcia-Magan C, *et al.* Microbiological profile of acute viral gastroenteritis attended in a paediatric department

- of an area with high vaccine cover against rotavirus [in Spanish]. *Enfermedades Infecciosas y Microbiologia Clinica* 2014; **32**: 246–249.
28. **Sanger F, Nicklen S, Coulson AR.** DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* 1977; **74**: 5463–5467.
  29. **Lau SK, et al.** Clinical and molecular epidemiology of human bocavirus in respiratory and fecal samples from children in Hong Kong. *Journal of Infectious Diseases* 2007; **196**: 986–993.
  30. **Nakanishi K, et al.** Detection of enteric viruses in rectal swabs from children with acute gastroenteritis attending the pediatric outpatient clinics in Sapporo, Japan. *Journal of Clinical Virology* 2009; **46**: 94–97.
  31. **Glass RI, et al.** Gastroenteritis viruses: an overview. *Gastroenteritis Viruses* 2001; **238**: 5–25.
  32. **Lee JI, et al.** Detection of human bocavirus in children hospitalized because of acute gastroenteritis. *Journal of Infectious Diseases* 2007; **196**: 994–997.
  33. **Jartti T, et al.** Human bocavirus—the first 5 years. *Reviews in Medical Virology* 2012; **22**: 46–64.
  34. **Yu JM, et al.** Human bocavirus infection in children hospitalized with acute gastroenteritis in China. *Journal of Clinical Virology* 2008; **42**: 280–285.
  35. **Pham NT, et al.** Human bocavirus infection in children with acute gastroenteritis in Japan and Thailand. *Journal of Medical Virology* 2011; **83**: 286–290.
  36. **Sarkar R, Kang G, Naumova EN.** Rotavirus seasonality and age effects in a birth cohort study of Southern India. *PLoS ONE* 2013; **8**.
  37. **Risku M, et al.** Human bocavirus types 1, 2 and 3 in acute gastroenteritis of childhood. *Acta Paediatrica* 2012; **101**: e405–e410.
  38. **Kantola K, et al.** Real-time quantitative PCR detection of four human bocaviruses. *Journal of Clinical Microbiology* 2011; **49**: 4029–4029.
  39. **Kapusinszky B, Minor P, Delwart E.** Nearly constant shedding of diverse enteric viruses by two healthy infants. *Journal of Clinical Microbiology* 2012; **50**: 3427–3434.
  40. **Paloniemi M, et al.** Human bocaviruses are commonly found in stools of hospitalized children without causal association to acute gastroenteritis. *European Journal of Pediatrics* 2014; **173**: 1051–1057.