

Emergence of enterovirus 71 “double-recombinant” strains belonging to a novel genotype D originating from southern China: first evidence for combination of intratypic and intertypic recombination events in EV71

Cyril C. Y. Yip · Susanna K. P. Lau · Boping Zhou · Ming-Xia Zhang · Hoi-Wah Tsoi · Kwok-Hung Chan · Xin-Chun Chen · Patrick C. Y. Woo · Kwok-Yung Yuen

Received: 23 November 2009 / Accepted: 3 June 2010 / Published online: 13 June 2010
© Springer-Verlag 2010

Abstract Hand–foot–mouth disease due to enterovirus 71 (EV71) and coxsackievirus A16 (CA16) has recently caused large outbreaks in mainland China in 2008. We performed complete genome sequencing on two EV71 (SZ/HK08-5 and SZ/HK08-6) and two CA16 (SZ/HK08-3 and SZ/HK08-7) strains from patients in Shenzhen, China. Phylogenetic, similarity plot and bootscan analyses revealed recombination between EV71 genotypes B and C

at the 2A–2B junction, and between EV71 genotype B and CA16 strain G-10 in the 3C region for EV71 strains. A similar phenomenon was also found upon further gene sequencing with other EV71 strains. Recombination between CA16 strain G-10 and EV71 genotype A at the 2A–2B junction was also observed for CA16 strains. The present “double-recombinant” EV71 strains circulating in China and other EV71 subgenotype “C4” strains represent an additional genotype, D. CA16 strains should also be classified into two genotypes. This represents the first evidence for a combination of intratypic and intertypic recombination in EV71 strains.

C. C. Y. Yip, S. K. P. Lau and B. Zhou contributed equally to the manuscript.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-010-0722-0) contains supplementary material, which is available to authorized users.

S. K. P. Lau · P. C. Y. Woo (✉) · K.-Y. Yuen (✉)
State Key Laboratory of Emerging Infectious Diseases,
Queen Mary Hospital, The University of Hong Kong,
University Pathology Building, Hong Kong, China
e-mail: pcywoo@hkucc.hku.hk

K.-Y. Yuen
e-mail: kyyuen@hkucc.hku.hk

S. K. P. Lau · P. C. Y. Woo · K.-Y. Yuen
Research Centre of Infection and Immunology,
The University of Hong Kong, Hong Kong, China

S. K. P. Lau · P. C. Y. Woo · K.-Y. Yuen
Carol Yu Centre for Infection, The University of Hong Kong,
Hong Kong, China

C. C. Y. Yip · S. K. P. Lau · H.-W. Tsoi · K.-H. Chan ·
P. C. Y. Woo · K.-Y. Yuen
Department of Microbiology, The University of Hong Kong,
Hong Kong, China

B. Zhou · M.-X. Zhang · X.-C. Chen
Shenzhen East Lake Hospital, Shenzhen,
People's Republic of China

Introduction

Human enteroviruses belong to the family *Picornaviridae* and are divided into four species: *Human enterovirus A* (HEV-A), HEV-B, HEV-C, and HEV-D, based on their molecular and biological properties [14]. The species HEV-A includes coxsackievirus A2–A8, A10, A12, A14, A16, enterovirus 71 (EV71), EV76 and EV89–92. Among these, EV71 and coxsackievirus A16 (CA16), which commonly infect young children, are two major causative agents of hand, foot, and mouth disease (HFMD) and herpangina. However, EV71 infection is more frequently associated with severe diseases such as acute flaccid paralysis, myocarditis, aseptic meningitis, and fatal encephalitis than CA16 [11].

Traditionally, human enteroviruses have been typed by neutralization assay, which is labour-intensive and time-consuming. In recent years, rapid and sensitive molecular diagnostic tests such as reverse transcription polymerase chain reaction (RT-PCR) have been increasingly used for enterovirus identification and classification [3, 4, 14, 25].

In addition, a molecular approach involving VP1 gene analysis has been developed for typing of human enteroviruses [5], which was in line with a previous study showing a high correlation between molecular typing and serotyping for enteroviruses [27]. EV71 is further classified into three genotypes, A, B and C, based on VP1 gene sequence analysis [3]. The prototype strain, BrCr, is the only member of genotype A. Most EV71 strains belong to genotype B (containing subgenotypes B1-B5) or C (containing subgenotypes C1-C5) [4, 13]. Other regions, including the 5' non-coding region (5' NCR) [1, 33] and VP4 [4, 8] of the EV71 genome have also been used for phylogenetic analysis.

Mutation and recombination are well-known phenomena for enterovirus evolution. The infidelity of the 3D polymerase of enteroviruses makes their mutation rates as high as one mutation per newly synthesized genome [9]. Several studies have shown that recombination occurs between vaccine and wild-type poliovirus strains as a result of template switching during negative strand synthesis, which is thought to be mediated by a "copy-choice" mechanism [15, 17]. Recombination among circulating HEV-B strains has also been demonstrated [23]. A molecular epidemiology study of coxsackievirus A9 suggested that recombination among circulating strains may have occurred in the non-structural region of the genome [29]. Recently, intratypic recombination between EV71 genotypes B and C was shown to occur in one EV71 isolate (N3340-TW-02) from Taiwan [12]. In addition to recombination among enteroviruses of the same type, intertypic recombination between EV71 subgenotype C2 and CA16 strain G-10 has also occurred in two EV71 isolates (SHZH98 and SHZH03) from China, and between CA16 strain G-10 and EV71 genotype A in one CA16 isolate (Tainan/5079/98) from Taiwan [35]. In China, 83,344 cases were reported in HFMD outbreaks in 2007, while 488,955 cases with 126 deaths were confirmed in HFMD outbreaks in 2008 [26]. Since EV71 and CA16 are major etiological agents of HFMD, the dramatic increase in HFMD cases in China in 2008 suggested that these viruses might have undergone mutation or recombination. To test this hypothesis, the complete genomes of two EV71 strains (SZ/HK08-5 and SZ/HK08-6) and two CA16 strains (SZ/HK08-3 and SZ/HK08-7) from four different patients from Shenzhen were amplified and sequenced.

Materials and methods

Patients and microbiological methods

In this study, six stool specimens and eight nasal swabs were collected from nine patients with HFMD at

Shenzhen East Lake Hospital, Shenzhen, China, in May 2008. All specimens were tested by RT-PCR for HEV-A. The clinical features, laboratory results and outcome of illness of patients who were positive for enterovirus were analyzed.

RT-PCR for enterovirus and sequencing

Viral RNA was extracted from stool specimens and nasal swabs using a QIAamp Viral RNA Mini Kit (QIAGEN, Hilden, Germany). RT was performed using random hexamers and a SuperScript III Kit (Invitrogen, San Diego, CA, USA) as described previously [18, 20]. PCR for HEV-A was performed using forward primer LPW8417 5'-ATAGCTATTGGATTGGCCAT-3' and reverse primer LPW8418 5'-TTCCARTACCA YCCCTTNGA-3', covering the VP4 region. The PCR mixture (25 µl) contained cDNA, PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2 mM MgCl₂ and 0.01% gelatin), 200 µM of each dNTP and 1.0 U *Taq* polymerase (Applied Biosystems, Foster City, CA, USA). The mixtures were amplified by 60 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min, with a final extension at 72°C for 10 min. The amplified products were detected by agarose gel electrophoresis. Both strands of PCR products were sequenced twice using an ABI Prism 3700 DNA Analyzer (Applied Biosystems, Foster City, CA, USA) with the PCR primers. The nucleotide sequences of the PCR products were compared with the sequences of enterovirus strains available in GenBank.

Complete genome sequencing and genome analysis of EV71 and CA16

Since PCR and sequence analysis suggested the presence of EV71 in clinical specimens from four patients and CA16 in those from five patients, the genomes of two EV71 strains and two CA16 strains from four patients were amplified and sequenced using the strategy described in our previous publications [19, 20, 34]. The RNA was converted to cDNA by a combined random-priming and oligo (dT) priming strategy. The cDNA was amplified using degenerate primers designed by multiple alignment of EV71 and CA16 genomes available in GenBank and additional primers designed from the results of the first and subsequent rounds of sequencing. These primer sequences are available on request. The terminal sequences were confirmed by rapid amplification of cDNA ends using a 5'/3' rapid amplification of cDNA ends kit (Roche, Mannheim, Germany). The nucleotide and the deduced amino acid sequences of the single open reading frame (ORF) in each genome were compared to those of EV71 and CA16 strains with available sequences in GenBank.

Phylogenetic analysis of complete VP4, partial 2B and partial 3D regions

Four EV71 and five CA16 strains were included in the phylogenetic tree of the VP4 region. Sequences for 207 nucleotide positions in each VP4 region were analyzed. Four EV71 and two CA16 strains were included in the phylogenetic trees of the partial 2B and partial 3D regions. Sequences for 297 nucleotide positions in each partial 2B region and 401 nucleotide positions in each partial 3D region were analyzed. Sequence alignment was performed using ClustalX version 1.81 [32]. The best evolutionary model for each dataset (TN93 + I for VP4 and partial 2B, and GTR + I for partial 3D) was determined using Modelgenerator [16], based on the Hierarchical Likelihood Ratio test and Akaike Information Criteria. Maximum-likelihood phylogenetic trees were constructed using the PhyML program [10], with bootstrap values calculated from 1,000 trees.

Phylogenetic analysis of the 5' NCR, P1, P2 and P3 regions

Two EV71 and two CA16 strains with complete genome sequences were included in the analysis. Sequences for 759 nucleotide positions in each 5' NCR region, 2,586 nucleotide positions in each P1 region, 1,734 nucleotide positions in each P2 region and 2,259 nucleotide positions in each P3 region were analyzed. The best evolutionary model (GTR + I) for each dataset was determined using Modelgenerator [16]. Maximum-likelihood phylogenetic trees were constructed using the PhyML program [10], with bootstrap values calculated from 1,000 trees.

Recombination analysis

A nucleotide alignment of the genome sequences of our EV71 and CA16 strains, EV71 genotype A prototype strain BrCr (GenBank accession no. U22521), EV71 genotype B strain MS/7423/87 (GenBank accession no. U22522), EV71 genotype C strain Tainan/4643/98 (GenBank accession no. AF304458), and CA16 prototype strain G-10 (GenBank accession no. U05876) was generated using ClustalX version 1.81 and edited manually. Once aligned, similarity plot and bootscan analyses were performed using SimPlot version 3.5.1 (window size, 1,000 bp; step, 20 bp).

Phylogenetic analysis of regions before and after the recombination sites

Two EV71 strains (SZ/HK08-5 and SZ/HK08-6) were included in the analysis. Sequences before nucleotide position 3600, those between nucleotide positions 3600 and

5430, and those after nucleotide position 5430 were analyzed. The best evolutionary model for each dataset (GTR + I for the region before position 3600 and the region between position 3600 and 5430, and GTR + I + G for the region after position 5430) was determined using Modelgenerator [16]. Maximum-likelihood phylogenetic trees were constructed using the PhyML program [10], with bootstrap values calculated from 1,000 trees.

Partial 2B and 3D gene sequencing of additional EV71 strains

Since our similarity plot and bootscan analyses indicated that recombination events have taken place in the genomes of the two EV71 strains, the partial 2B and 3D genes of the EV71 strains from the other two patients were amplified and sequenced. PCR for partial 2B was performed using forward primer LPW8419 (5'-GARGCTAGYGAGTAY TAYCC-3') and reverse primer LPW8420 (5'-TGCTC TTGAAGTGCATGTA-3'), and PCR for partial 3D was performed using forward primer LPW8453 (5'-CCTTACT CCACTTATGTCAA-3') and reverse primer LPW8456 (5'-TTGAGTTGAAAATGGATGTT-3'), using the same PCR conditions described above. Phylogenetic analysis was performed, and the results were compared with those obtained using the VP4 gene.

Nucleotide sequence accession numbers

The complete genome sequences of the two EV71 strains (SZ/HK08-5 and SZ/HK08-6) and the two CA16 strains (SZ/HK08-3 and SZ/HK08-7) have been deposited in the GenBank database under accession numbers GQ279368 to GQ279371.

Results

Detection of HEV-A in clinical specimens

Among the 6 stool specimens and 8 nasal swabs, 11 specimens from 9 patients were positive for HEV-A by RT-PCR. Most patients were young children (median age 20 months; range 8 months to 8 years). Six were males and two were females, while clinical information for the only patient who was not hospitalized was not available. The clinical characteristics of these nine patients are summarized in Table 1. All patients presented with HFMD, manifested by fever and vesicular lesions in the oral cavity, palms and in some cases, buttocks and knees. One case was complicated by acute encephalitis (patient 2) and two by acute bronchiolitis (patients 4 and 6). Leukocytosis with neutrophilia was observed in two patients (patients 1 and

Table 1 Results of enterovirus RT-PCR in the nine patients with hand-foot-mouth disease included in this study

Patient no.	Sex/age	Underlying disease	Diagnosis	Duration of fever (days)	Duration of hospitalization (days)	Enterovirus PCR and sequencing result		Strain no.
						Fecal sample	Nasal swab	
1	M/8m	None	HFMD	2	6	CA16	Not done	SZ/HK08-1
2	M/18m	Burn injury 1 month ago	HFMD, acute encephalitis	4	16	EV71	–ve	SZ/HK08-2
3	M/20m	None	HFMD	2	6	CA16	CA16	SZ/HK08-3
4	M/2	None	HFMD, acute bronchiolitis	1	8	CA16	–ve	SZ/HK08-4
5	F/4	None	HFMD	2	3	EV71	EV71	SZ/HK08-5
6	M/8	Kawasaki disease	HFMD, acute bronchiolitis	2	6	EV71	–ve	SZ/HK08-6
7	NA	NA	NA	NA	0	Not done	CA16	SZ/HK08-7
8	M/20m	History of pneumonia	HFMD	3	9	Not done	CA16	SZ/HK08-8
9	F/12m	None	HFMD	2	7	Not done	EV71	SZ/HK08-9

HFMD Hand-foot-mouth disease, NA not available because the patient only visited an outpatient clinic

8). In addition to the two patients with bronchiolitis, chest radiographs showed mild diffuse interstitial infiltrates in four other patients, suggesting that lung involvement may be common. All of the patients survived.

Phylogenetic analysis of the VP4 gene sequences of the HEV-A strains from these nine patients showed that four (SZ/HK08-2, SZ/HK08-5, SZ/HK08-6 and SZ/HK08-9) were EV71, and five (SZ/HK08-1, SZ/HK08-3, SZ/HK08-4, SZ/HK08-7 and SZ/HK08-8) were CA16 (Fig. 2a). Apart from the patient who had encephalitis due to EV71 infection, the clinical features of HFMD appeared to be similar in other patients infected with EV71 and CA16.

Complete genome analysis of EV71 and CA16

The genomes of the two EV71 strains was similar to other reported genomes of EV71, with 7,405 nucleotides and a G + C content of 48.1–48.2% (excluding the 3' polyA tract), and including a 5' NCR of 742 nucleotides, a single ORF of 6,579 nucleotides encoding a single polyprotein of 2,193 amino acids, and a 3' non-coding region (3' NCR) of 84 nucleotides preceding the polyA tract (Fig. 1). The predicted amino acid sequences of the polyprotein of the two EV71 strains (SZ/HK08-5 and SZ/HK08-6) showed 99% identity to that of EV71 strain 984-TW (GenBank accession no. DQ133458).

The genomes of the two CA16 strains were similar to other reported genomes of CA16, with 7,409 nucleotides and a G + C content of 47.13–47.33% (excluding the 3' polyA tract). The genome organization was similar to that of other CA16 strains, comprising a 5' NCR of 745 nucleotides, a single ORF of 6,579 nucleotides encoding a single polyprotein of 2,193 amino acids, and a 3' NCR of

85 nucleotides preceding the polyA tract (Fig. 1). The predicted amino acid sequences of the polyprotein of the two CA16 strains (SZ/HK08-3 and SZ/HK08-7) showed 98% identity to that of CA16 strain shzh00-1 (GenBank accession no. AY790926).

Phylogenetic and recombination analysis of EV71 strains

Phylogenetic trees using the nucleotide sequences of the 5' NCR, P1 (VP4–VP1), P2 (2A–2C) and P3 (3A–3D) regions of the two EV71 strains (SZ/HK08-5 and SZ/HK08-6) and other EV71 strains with complete genome sequences available in GenBank were constructed (Fig. 3). Phylogenetic analysis showed that our two EV71 strains were clustered with EV71 genotype C strains for the 5' NCR and P1 regions, but with EV71 genotype B strains for the P2 region. Intriguingly, our two EV71 strains and EV71 subgenotype B3 strains were most closely related to CA16 prototype strain G-10 in the P3 region.

Since the incongruent phylogenetic relationships observed among different regions of the genomes suggested that recombination events may have taken place in the two EV71 strains (SZ/HK08-5 and SZ/HK08-6), similarity plot and bootscan analyses were performed to identify potential recombination sites (Fig. 4a). In the similarity plot analysis, our two EV71 strains showed high sequence similarity ($\geq 80\%$) to EV71 genotype C strain Tainan/4643/98 before position 3600. However, a higher similarity ($\geq 78\%$) to EV71 genotype B strain MS/7423/87 between positions 3600 and 5430 was noted in EV71 strain SZ/HK08-6 but not SZ/HK08-5. After position 5430, our two EV71 strains had higher similarities ($\geq 80\%$) to CA16

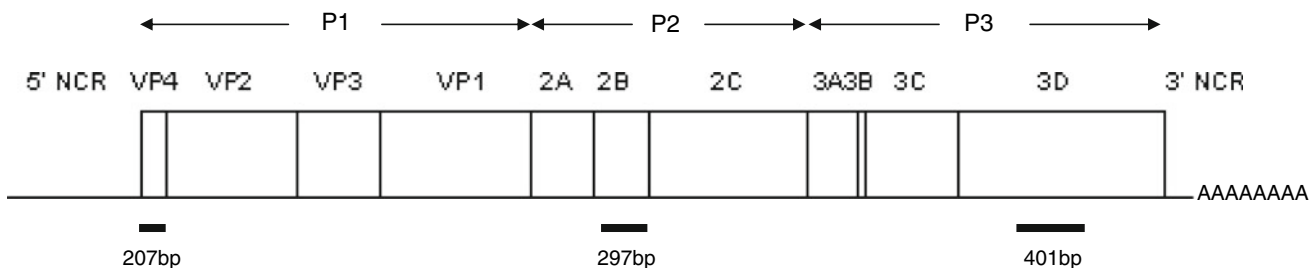


Fig. 1 Schematic representation of the human enterovirus A genome. VP4, partial 2B and partial 3D regions used for phylogenetic analysis are indicated by *black bars*

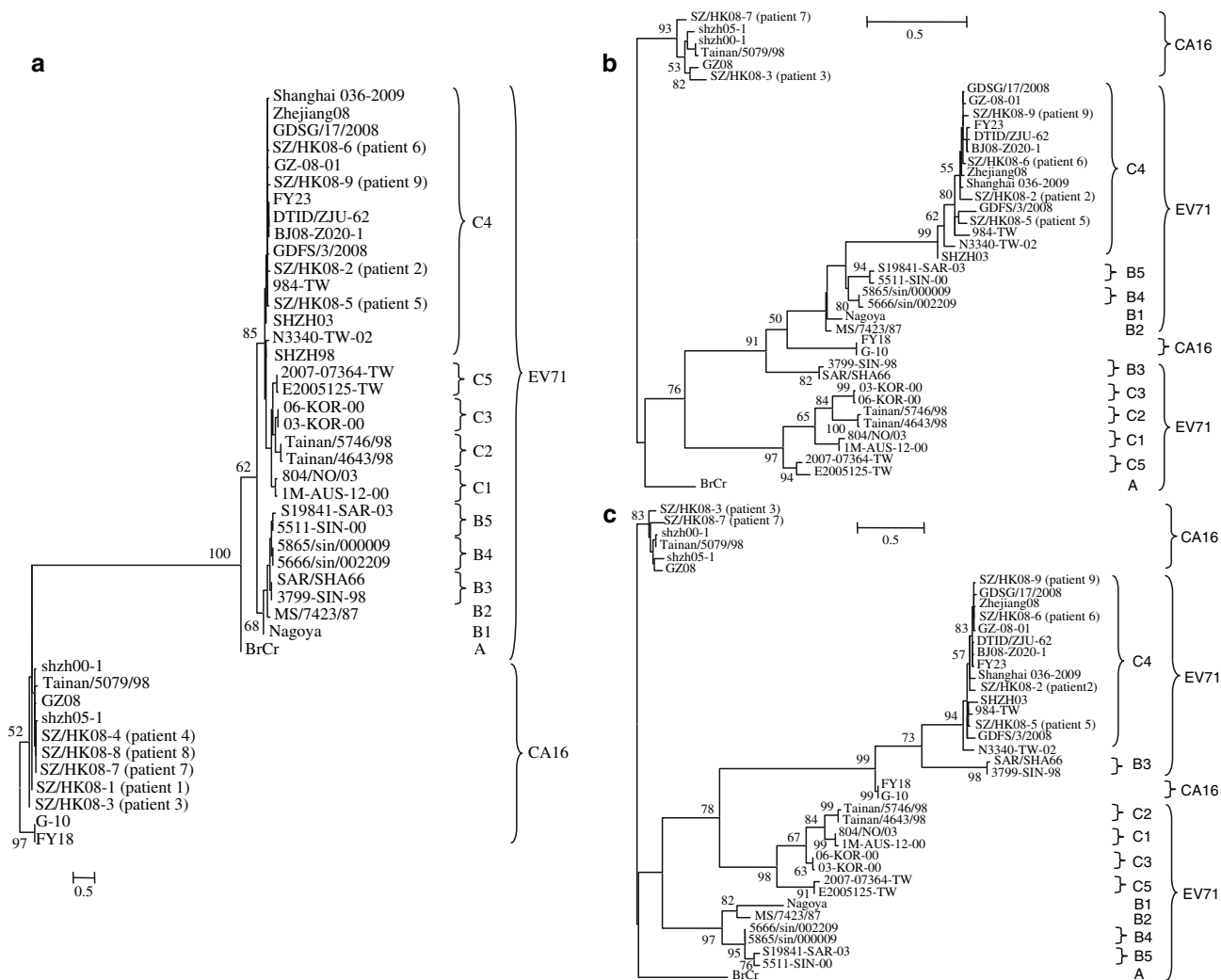


Fig. 2 Phylogenetic analysis of **a** complete VP4, **b** partial 2B and **c** partial 3D regions of EV71 and CA16. Bootstrap values expressed as percentage are shown at the *nodes*, and the scale reflects the number of nucleotide substitutions per site along the branches

strain G-10. In the bootscan analysis, the results showed that from the 5' end of the genome to position 3200, high bootstrap support for clustering between EV71 genotype C strain Tainan/4643/98 and our two EV71 strains was observed. From position 4000 to 5200, high bootstrap support for clustering between EV71 genotype B strain

MS/7423/87 and our two EV71 strains was observed. From position 5800 to the 3' end of the genome, high bootstrap support for clustering between CA16 strain G-10 and our two EV71 strains was observed. These findings indicated that recombination events have possibly occurred between nucleotide positions 3200 and 4000, corresponding to the

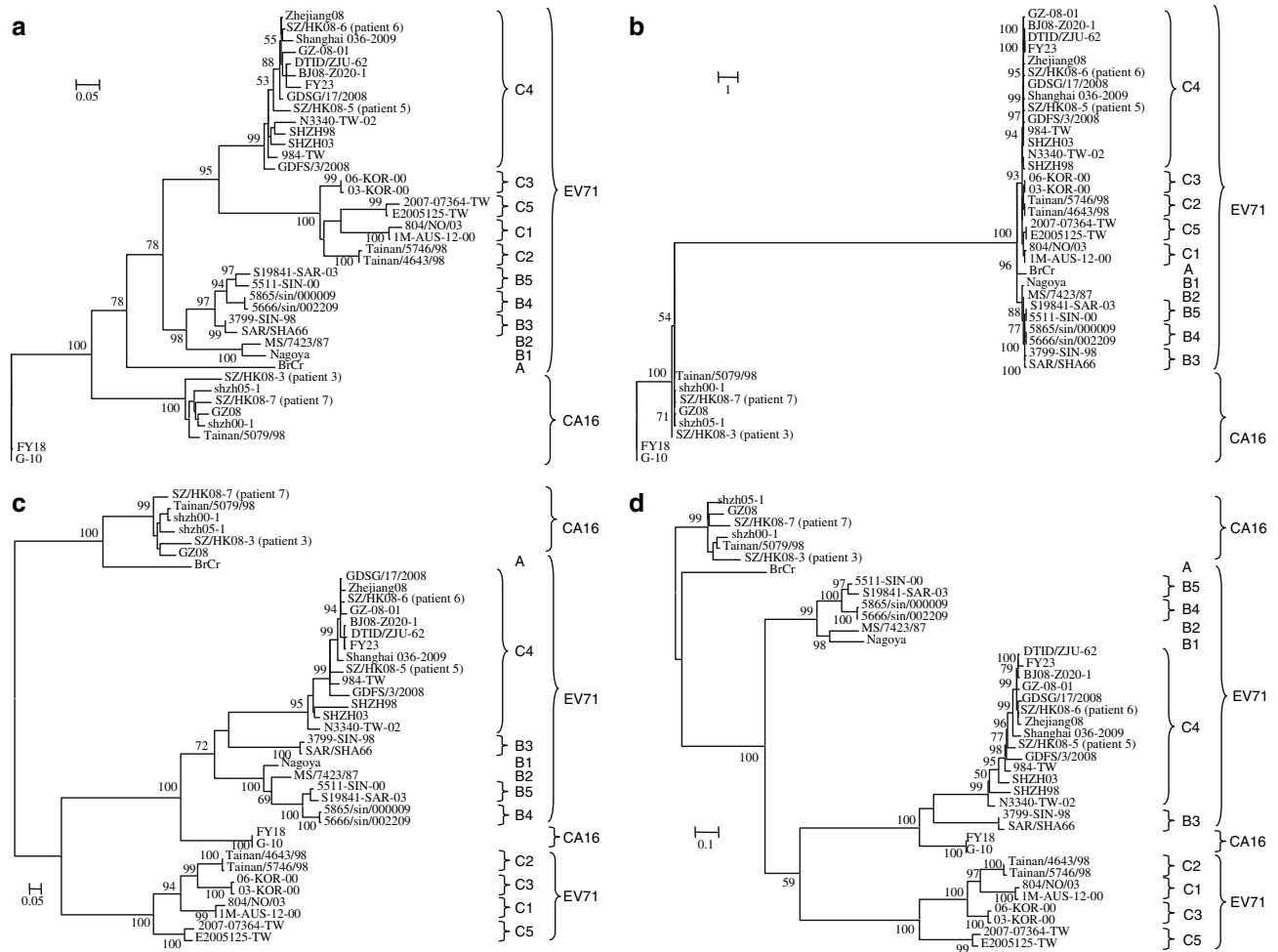


Fig. 3 Phylogenetic analysis of **a** 5' NCR, **b** P1, **c** P2 and **d** P3 regions of the two EV71 and two CA16 strains with complete genome sequences. Bootstrap values expressed as percentage are shown at the

nodes, and the *scale* reflects the number of nucleotide substitutions per site along the branches

2A–2B junction, and between nucleotide positions 5200 and 5800, corresponding to the 3A–3C region.

Based on the two recombination breakpoints revealed by SimPlot analysis, three phylogenetic trees were constructed using sequences before and after the recombination sites (Fig. 5). The results showed that our two EV71 strains clustered with EV71 genotype C strains for the region upstream of nucleotide position 3600, but with EV71 genotype B strains for the region between nucleotide positions 3600 and 5430. Interestingly, our two EV71 strains and the EV71 subgenotype B3 strains were most closely related to CA16 strain G-10 for the region downstream of nucleotide position 5430. This is in line with the results of the recombination analysis, indicating that both intratypic and intertypic recombination events may have occurred in our two EV71 strains.

Since both phylogenetic and recombination analysis showed that there was a possible recombination breakpoint at the 2A–2B junction and in the 3A–3C region, multiple

alignments among the nucleotide sequences of one EV71 strain SZ/HK08-5 and putative parental strains were also performed to identify the recombination sites. Upstream of nucleotide position 3590 (the first predicted breakpoint by bootscan analysis) of EV71 strain SZ/HK08-5, there was a high degree of nucleotide identity between the sequences of EV71 strain SZ/HK08-5 and the EV71 genotype C strain (Tainan/4643/98), whereas downstream of nucleotide position 3589 of EV71 strain SZ/HK08-5, there was a high degree of nucleotide identity between the sequences of EV71 strain SZ/HK08-5 and the EV71 genotype B strain (MS/7423/87) (see Fig. S1a in Supplemental Material). In addition, upstream of nucleotide position 5420 (the second predicted breakpoint by bootscan analysis) of EV71 strain SZ/HK08-5, there was a high degree of nucleotide identity between the sequences of EV71 strain SZ/HK08-5 and the EV71 genotype B strain (MS/7423/87), whereas downstream of nucleotide position 5419 of EV71 strain SZ/HK08-5, there was a high degree of nucleotide identity

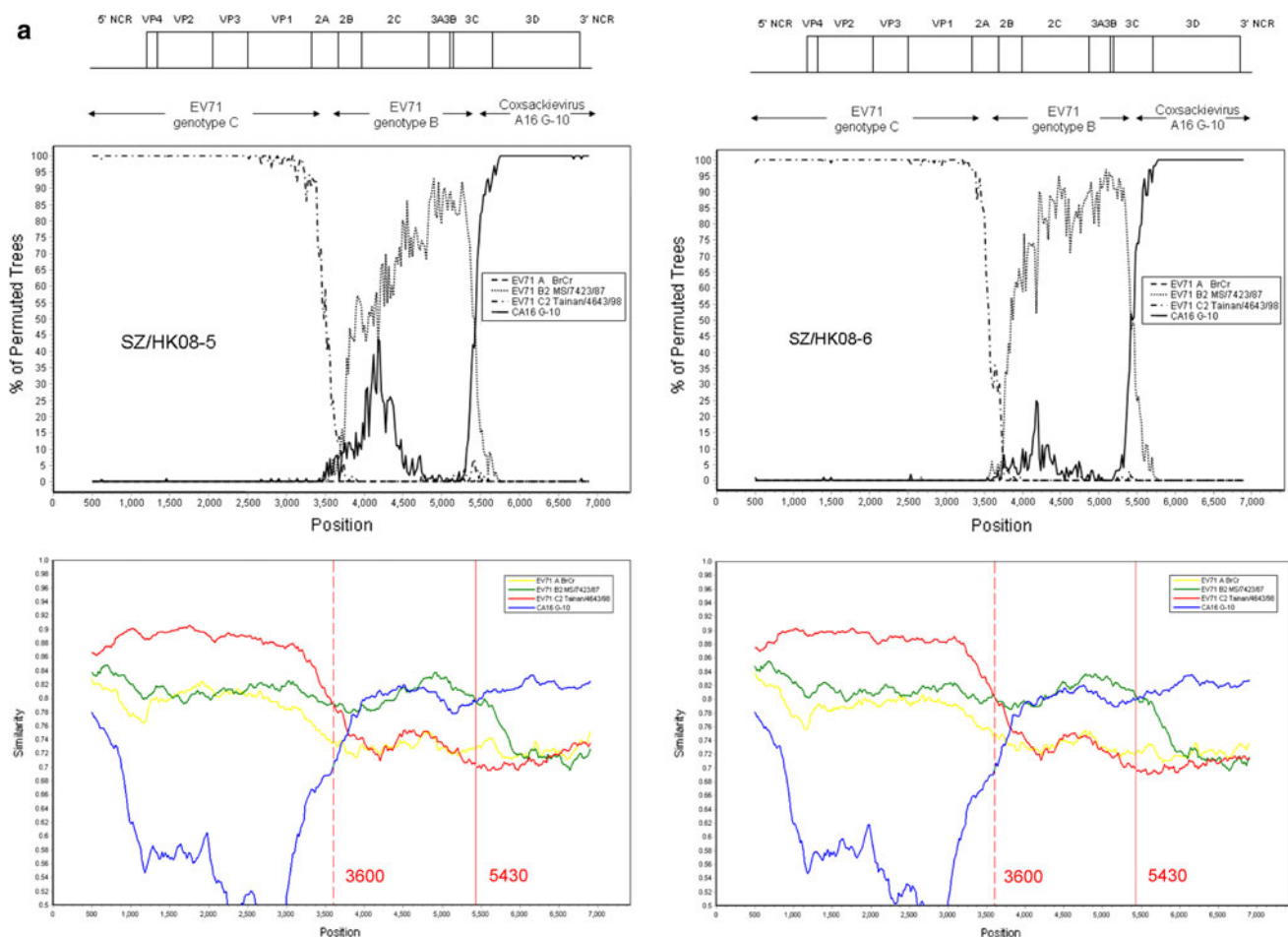


Fig. 4 Recombination analysis of the two EV71 and two CA16 complete genomes. Bootscanning (*upper panel*) and similarity plot analysis (*lower panel*) were conducted with SimPlot version 3.5.1 (Kimura distance model; window size, 1,000 bp; step, 20 bp) on a

gapless nucleotide alignment, generated with ClustalX, with the genome sequences of the **a** two EV71 strains (SZ/HK08-5 and SZ/HK08-6) and **b** two CA16 strains (SZ/HK08-3 and SZ/HK08-7) as the query sequences

between the sequences of EV71 strain SZ/HK08-5 and CA16 strain G-10 (see Fig. S1b in Supplemental Material).

Phylogenetic and recombination analysis of CA16 strains

Phylogenetic trees using the nucleotide sequences of the 5' NCR, P1, P2 and P3 regions of our two CA16 strains (SZ/HK08-3 and SZ/HK08-7) and other CA16 strains with complete genome sequences available in GenBank were constructed (Fig. 3). Phylogenetic analysis showed that our two CA16 strains clustered with CA16 strain G-10 in the 5' NCR and P1 regions, but with EV71 genotype A strain BrCr in the P2 and P3 regions. As this finding suggested that recombination may have occurred in our two CA16 strains, similarity plot and bootscan analyses were performed to identify potential recombination sites (Fig. 4b). In the similarity plot analysis, our two CA16 strains showed high sequence similarity ($\geq 69\%$) to CA16 strain

G-10 before position 3770 but shared higher similarity ($\geq 76\%$) with EV71 genotype A strain BrCr after position 3770. In the bootscan analysis, the results showed that from the 5' end of the genome to position 3600, high bootstrap support for clustering between CA16 strain G-10 and our CA16 strains was observed. From position 4000 to the 3' end of the genome, high bootstrap support for clustering between the EV71 genotype A strain BrCr and our CA16 strains was observed. These findings indicated that recombination had possibly occurred between nucleotide positions 3600 and 4000, corresponding to the 2A–2B junction.

Partial 2B and 3D gene analysis of additional EV71 strains

To determine whether recombination events have taken place in the other two EV71 strains in this study, their partial 2B and 3D genes, in addition to their VP4 genes,

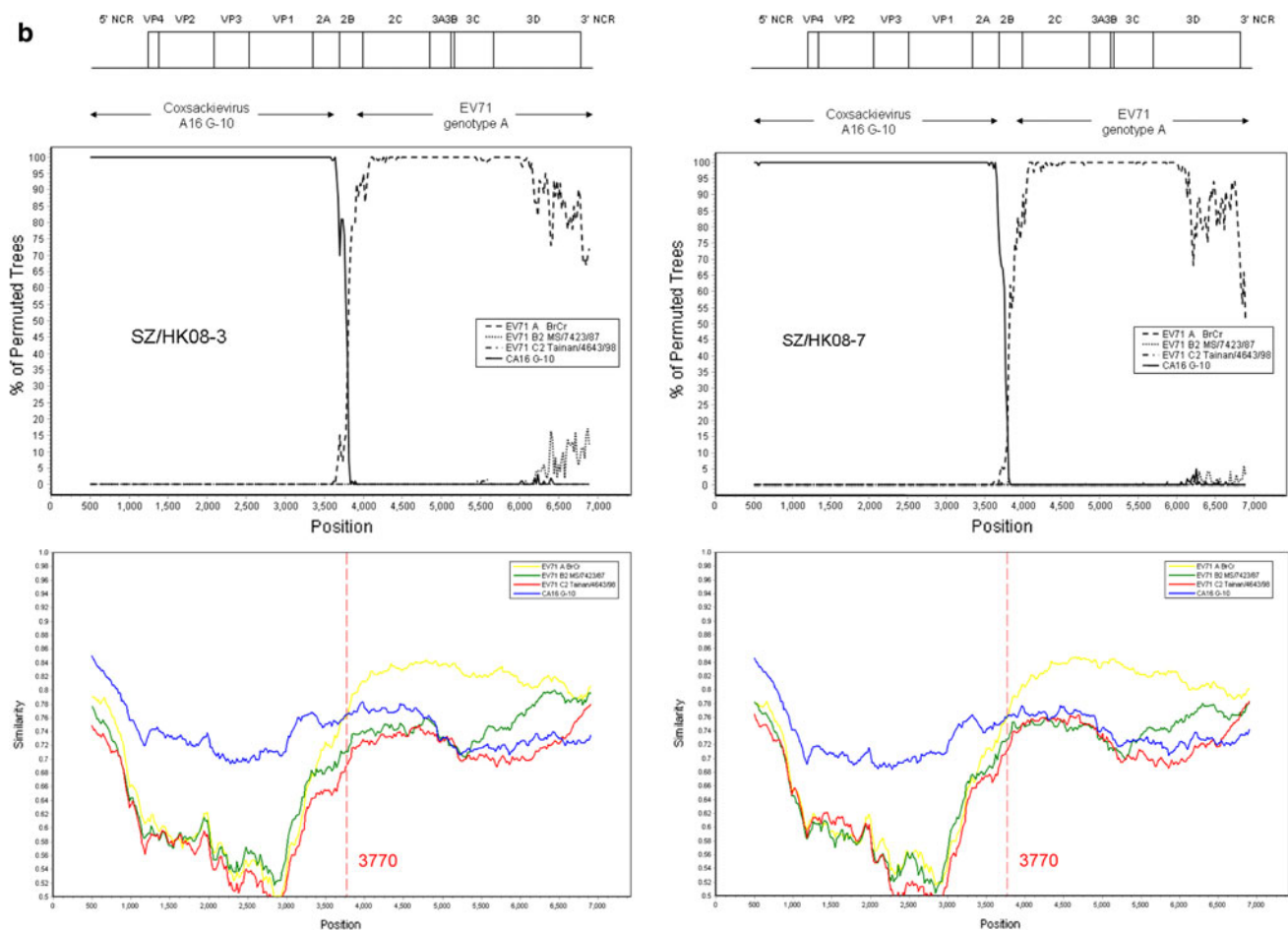


Fig. 4 continued

which correspond to the three different regions with incongruent phylogenetic positions resulting from recombination, were amplified and sequenced. Phylogenetic analysis supported similar recombination events in these additional EV71 strains, which clustered with EV71 genotype C strains in the VP4 gene (Fig. 2a) but with EV71 genotype B strains in the partial 2B gene (Fig. 2b) and with CA16 strain G-10 in the partial 3D gene (Fig. 2c).

Discussion

The present study demonstrates for the first time a combination of intratypic and intertypic recombination events occurring in EV71 strains associated with HFMD in children from Shenzhen in 2008. Both phylogenetic and recombination analysis of the complete genome sequences of two strains (SZ/HK08-5 and SZ/HK08-6) showed that their nucleotide sequences were most closely related to those of EV71 genotype C strains in the 5' NCR and P1 regions (Fig. 3a, b) and the region upstream of nucleotide

position 3600 (Fig. 5a). However, a notable change was observed in the topology of the phylogenetic tree for the P2 region (Fig. 3c) and the region between nucleotide positions 3600 and 5430 (Fig. 5b), where their nucleotide sequences were most closely related to those of EV71 genotype B strains. This is supported by high bootstrap support for EV71 genotype B from 2B to 3C upon bootscan analysis, suggesting that intratypic recombination had occurred between EV71 genotypes B and C in the two EV71 strains, and the recombination site was possibly located at the 2A–2B junction. Results from multiple alignments of nucleotide sequences of the 2A–2B region of EV71 strain SZ/HK08-5 and EV71 genotypes B and C (Fig. S1a) also demonstrated a recombination site located in the 2A–2B region, further supporting the findings of the phylogenetic and recombination analyses. Similar intratypic recombination between EV71 genotypes B and C has previously been shown to occur in one EV71 isolate, N3340-TW-02, from Taiwan in 2002 [12]. One of the recombination sites was located at the 3' terminus of the 2A region, similar to our two EV71 strains. In a previous

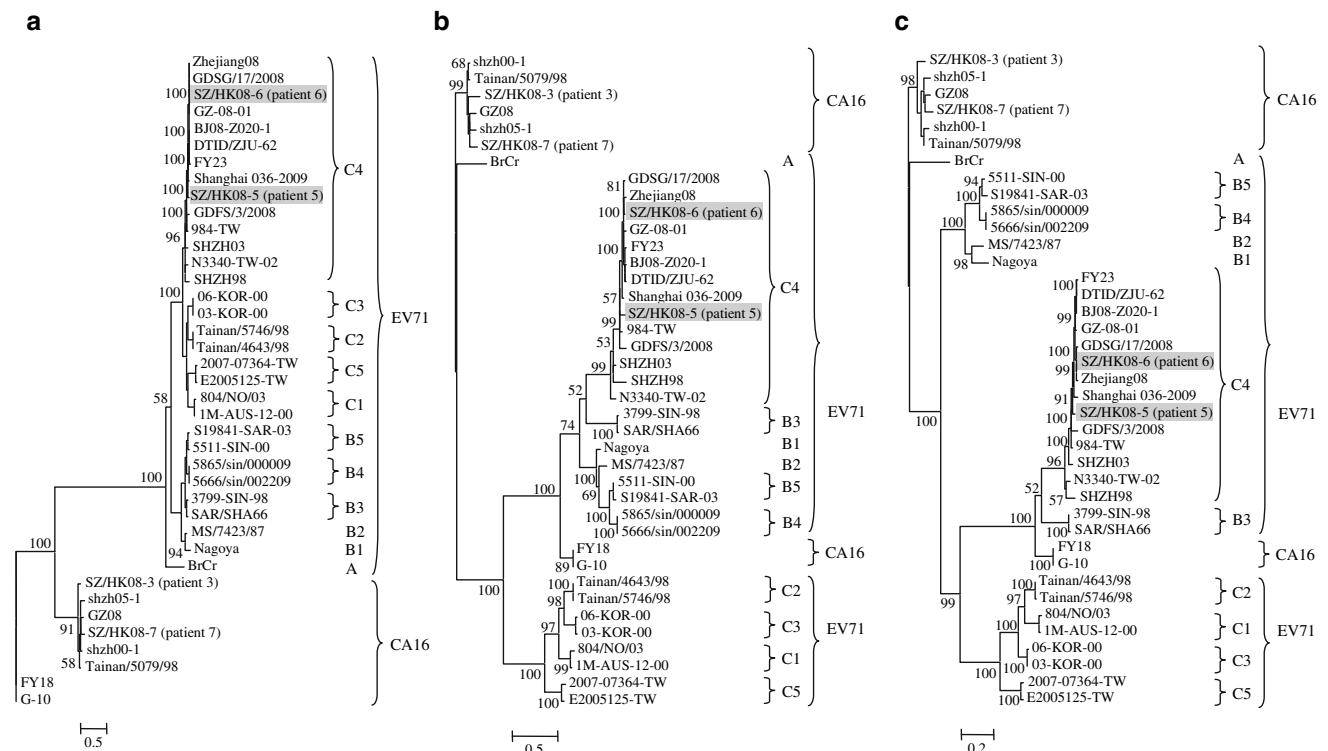


Fig. 5 Phylogenetic analysis of sequences before and after recombination breakpoints revealed by SimPlot analysis. Three phylogenetic trees were constructed using **a** the region upstream of nucleotide position 3600, **b** the region between nucleotide positions 3600 and

5430, and **c** the region downstream of nucleotide position 5430. EV71 strains with evidence of recombination are shaded. Bootstrap values expressed as percentage are shown at the *nodes*, and the scale reflects the number of nucleotide substitutions per site along the branches

study on recombination in enteroviruses, the 2A region was also found to be a hot spot for such events [23]. In addition to this intratypic recombination, intertypic recombination was also found in our two EV71 strains by complete genome analysis. The apparent close relationships between the P3 regions of the two strains and those of EV71 subgenotype B3 strains (Fig. 3d) implied that the two strains were actually closely related to those of CA16 strain G-10, as EV71 subgenotype B3 strains were previously shown to be recombinants of an EV71 genotype B strain and the CA16 strain G-10 [6]. This finding is in line with the result shown in Fig. 5c, indicating that intertypic recombination has occurred between EV71 genotype B and CA16 strain G-10 at the 3C region. Results from multiple alignments of nucleotide sequences of the 3A–3C regions of EV71 strain SZ/HK08-5, EV71 genotype B and CA16 strain G-10 (Fig. S1b) also revealed a recombination site located in the 3C region, further supporting the findings of phylogenetic and recombination analyses (Figs. 3, 4a, 5). A similar phenomenon has also been observed in a previous study, in which intertypic recombination between EV71 genotype C and CA16 strain G-10 was identified at the 3C region in two EV71 isolates from Shenzhen in 1998 and 2003 (SHZH98 and SHZH03) [35]. However, none of the previous strains were reported to be the result of both

intratypic and intertypic recombination events. In fact, when we performed similar phylogenetic and recombination analyses on previously reported strains with complete genome sequence available, strains N3340-TW-02, SHZH98 and SHZH03 were found to possess “double recombination” similar to that of our present strains. We speculated that the recent epidemic of EV71 infection in China may be due to the emergence of this “double recombinant” strain resulting from recombination between EV71 genotypes B and C, and CA16 strain G-10, which were present as early as 1998 in Shenzhen and spread to Taiwan no later than 2002. Partial 2B and 3D gene analysis of the other two EV71 strains in the present study supported our speculation. Further studies on complete genome sequence analysis of more EV71 strains from China are required to determine the prevalence of this “double recombinant” strain.

The present recombinant EV71 strains should be considered a novel genotype, genotype D, which requires sequence analysis of multiple gene regions for accurate genotyping. In the present study, we demonstrated recombination events between different CA16 and EV71 strains or between different genotypes of EV71 in our enterovirus strains, resulting in incongruent phylogenetic positions in different regions of their genomes. The

sequences of the P1, P2 and P3 regions of our EV71 strains were most closely related to those of EV71 genotype C, EV71 genotype B and CA16 strain G-10, respectively. Hence, amplification and sequencing of only one gene region may lead to false conclusion about the true genotype identity. This is similar to the phenomenon reported for coronaviruses, which are also positive-sense single-stranded RNA viruses, in which we have previously identified recombination between different genotypes of human coronavirus HKU1 in the generation of a new genotype [34]. Sequencing of at least three gene regions (e.g. VP4, 2B and 3D) is required to determine the actual genotype of these recombinant EV71 strains. In previous studies, EV71 isolates SHZH98 and SHZH03 were classified as subgenotype C4 based on phylogenetic analysis using the VP1 gene sequence only [21]. The existence of EV71 “subgenotype C4” was first reported in China (SHZH98) and Taiwan (3254-TAI-98) in 1998 [21]. This subgenotype was then identified in Japan, Vietnam and Thailand and was further divided into two clusters, C4a (2003–2008) and C4b (1998–2004) [36]. Since 2007, a large number of EV71 strains with VP1 sequences corresponding to subgenotype C4a have been identified in China, indicating that this subgenotype was predominant in China, causing HFMD outbreaks [36]. Sequencing of different gene regions or complete genome sequencing will reveal if these strains also belong to the novel EV71 genotype D. In fact, the reclassification of subgenotype C4 as a new genotype was also supported in a recent study by phylogenetic analysis using EV71 complete genome sequences [7].

In addition to the recombinant EV71 strains, intertypic recombination was also detected in two CA16 strains (SZ/HK08-3 and SZ/HK08-7) from Shenzhen in 2008. Both phylogenetic and recombination analysis showed that the nucleotide sequences of the two CA16 strains were most closely related to those of CA16 strain G-10 in the 5' NCR and P1 regions, but to those of EV71 genotype A strain BrCr in the P2 and P3 regions, suggesting that intertypic recombination had occurred between CA16 strain G-10 and EV71 genotype A. This is in line with a previous report of a similar recombination event between CA16 strain G-10 and EV71 genotype A in one CA16 isolate (Tainan/5079/98), where the site of crossover was located within a 166-bp region between nucleotide positions 3616 and 3781 in the 2A–2B region [35]. Based on the present phylogenetic analysis, CA16 strains should be divided into two genotypes, A and B. For the 5' NCR and P1 region, our two CA16 recombinant strains, together with shzh00-1, shzh05-1, GZ08 and Tainan/5079/98, were most closely related to the CA16 strains G-10 and FY18. However, for the P2 and P3 regions, the former CA16 strains were more closely related to EV71 genotype A strain BrCr than CA16 strains G-10 and FY18. Therefore,

we propose that CA16 strains G-10 and FY18 should be classified as genotype A, and the remaining CA16 strains should be classified as genotype B. Sequencing of at least two gene regions (5' NCR or P1 and P2 or P3) is needed to differentiate between the two genotypes of CA16.

The non-structural regions P2 and P3 are likely the hot spot for recombination in enteroviruses. Recombination is a well-known phenomenon in enterovirus evolution. Our present findings are in line with a previous study demonstrating that several recombination breakpoints were located within the P2 and P3 regions of enteroviruses of same species, while no breakpoint was found within the P1 region by genomic analysis [30]. Other studies have also shown recombination sites located within non-structural regions, such as 2A in EV11/EV19 [24], 2B in coxsackievirus B4 [22], and 2C in echovirus 9/18 [2]. One possible reason for frequent detection of recombination sites within non-structural regions of the enterovirus genome is that non-structural regions are more homologous than capsid regions among enteroviruses of the same species. Instead, mutation rather than recombination frequently occurred in the capsid region of enteroviruses, particularly in VP1, which bears most of the motifs essential for interaction with antibodies and the host-cell receptor, and these result from evasion of host immune attack. The combination of mutations in the P1 region and recombination in the P2 and P3 regions of the enterovirus genome is probably important for the generation of the high diversity of enteroviruses circulating in the human population. This unique mechanism for generation of novel enterovirus genotypes may allow for the emergence of new strains and new epidemics.

The reason for the dramatic increase in HFMD cases in 2008 in China remains to be determined. Although the present study identifies a new combination of recombination events in the present EV71 strains from children with HFMD in China, their sequences in the VP1 region, bearing major antigenic epitopes, did not show dramatic changes when compared to those of EV71 strains of subgenotype C4 in China in 2007. In addition, the amino acid residues ‘RAGLVGEIDLPLEGTTNP’, which have been predicted to comprise the immunogenic BC loop in the VP1 sequences of EV71 strains [28, 31], were also present in our two EV71 strains. The only difference observed was an amino acid substitution of alanine₇₅₃ by serine in strain SZ/HK08-6 (data not shown). Further studies are required to find out if this substitution and/or the present double recombinant events contribute to altered antigenicity that might explain the recent rapid emergence of EV71 in China.

Acknowledgments We are grateful for the generous support of Mr. Hui Hoy and Mr. Hui Ming in the genomic sequencing platform, and the support of Professor York Chow. This work is partly

supported by the Committee on Research and Conference Grants, University Development Fund, HKU Special Research Achievement Award and Outstanding Young Researcher Award, The University of Hong Kong; Research Grant Council Grant; the Croucher Senior Medical Research Fellowship 2006–2007; the HKSAR Research Fund for the Control of Infectious Diseases of the Health, Welfare and Food Bureau, and Consultancy Service for Enhancing Laboratory Surveillance of Emerging Infectious Disease for Department of Health, the Government of Hong Kong Special Administrative Region of China.

References

1. AbuBakar S, Chee HY, Al-Kobaisi MF, Xiaoshan J, Chua KB, Lam SK (1999) Identification of enterovirus 71 isolates from an outbreak of hand, foot and mouth disease (HFMD) with fatal cases of encephalomyelitis in Malaysia. *Virus Res* 61:1–9
2. Andersson P, Edman K, Lindberg AM (2002) Molecular analysis of the echovirus 18 prototype: evidence of interserotypic recombination with echovirus 9. *Virus Res* 85:71–83
3. Brown BA, Oberste MS, Alexander JP Jr, Kennett ML, Pallansch MA (1999) Molecular epidemiology and evolution of enterovirus 71 strains isolated from 1970 to 1998. *J Virol* 73:9969–9975
4. Cardosa MJ, Perera D, Brown BA, Cheon D, Chan HM, Chan KP, Cho H, McMinn P (2003) Molecular epidemiology of human enterovirus 71 strains and recent outbreaks in the Asia-Pacific region: comparative analysis of the VP1 and VP4 genes. *Emerg Infect Dis* 9:461–468
5. Caro V, Guillot S, Delpyroux F, Crainic R (2001) Molecular strategy for ‘serotyping’ of human enteroviruses. *J Gen Virol* 82:79–91
6. Chan YF, AbuBakar S (2004) Recombinant human enterovirus 71 in hand, foot and mouth disease patients. *Emerg Infect Dis* 10:1468–1470
7. Chan YF, Sam IC, Abubakar S (2010) Phylogenetic designation of enterovirus 71 genotypes and subgenotypes using complete genome sequences. *Infect Genet Evol* 10(3):404–412
8. Chu PY, Lin KH, Hwang KP, Chou LC, Wang CF, Shih SR, Wang JR, Shimada Y, Ishiko H (2001) Molecular epidemiology of enterovirus 71 in Taiwan. *Arch Virol* 146:589–600
9. Drake JW (1993) Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci USA* 90:4171–4175
10. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 5:696–704
11. Ho M (2000) Enterovirus 71: the virus, its infections and outbreaks. *J Microbiol Immunol Infect* 33:205–216
12. Huang SC, Hsu YW, Wang HC, Huang SW, Kiang D, Tsai HP, Wang SM, Liu CC, Lin KH, Su IJ, Wang JR (2008) Appearance of intratypic recombination of enterovirus 71 in Taiwan from 2002 to 2005. *Virus Res* 131:250–259
13. Huang YP, Lin TL, Kuo CY, Lin MW, Yao CY, Liao HW, Hsu LC, Yang CF, Yang JY, Chen PJ, Wu HS (2008) The circulation of subgenogroups B5 and C5 of enterovirus 71 in Taiwan from 2006 to 2007. *Virus Res* 137:206–212
14. Hyypiä T, Hovi T, Knowles NJ, Stanway G (1997) Classification of enteroviruses based on molecular and biological properties. *J Gen Virol* 78:1–11
15. Jarvis TC, Kirkegaard K (1992) Poliovirus RNA recombination: mechanistic studies in the absence of selection. *EMBO J* 11:3135–3145
16. Keane TM, Creevey CJ, Pentony MM, Naughton TJ, McInerney JO (2003) Assessment of methods of amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol Biol* 6:29
17. Kirkegaard K, Baltimore D (1986) The mechanism of RNA recombination in poliovirus. *Cell* 47:433–443
18. Lau SK, Woo PC, Yip CC, Tse H, Tsoi HW, Cheng VC, Lee P, Tang BS, Cheung CH, Lee RA, So LY, Lau YL, Chan KH, Yuen KY (2006) Coronavirus HKU1 and other coronavirus infections in Hong Kong. *J Clin Microbiol* 44:2063–2071
19. Lau SK, Yip CC, Que TL, Lee RA, Au-Yeung RK, Zhou B, So LY, Lau YL, Chan KH, Woo PC, Yuen KY (2007) Clinical and molecular epidemiology of human bocavirus in respiratory and fecal samples from children in Hong Kong. *J Infect Dis* 196:986–993
20. Lau SK, Yip CC, Tsoi HW, Lee RA, So LY, Lau YL, Chan KH, Woo PC, Yuen KY (2007) Clinical features and complete genome characterization of a distinct human rhinovirus (HRV) genetic cluster, probably representing a previously undetected HRV species, HRV-C, associated with acute respiratory illness in children. *J Clin Microbiol* 45:3655–3664
21. Li L, He Y, Yang H, Zhu J, Xu X, Dong J, Zhu Y, Jin Q (2005) Genetic characteristics of human enterovirus 71 and coxsackievirus A16 circulating from 1999 to 2004 in Shenzhen, People’s Republic of China. *J Clin Microbiol* 43:3835–3839
22. Lindberg AM, Andersson P, Savolainen C, Mulders MN, Hovi T (2003) Evolution of the genome of Human enterovirus B: incongruence between phylogenies of the VP1 and 3CD regions indicates frequent recombination within the species. *J Gen Virol* 84:1223–1235
23. Lukashev AN, Lashkevich VA, Ivanova OE, Koroleva GA, Hinkkanen AE, Ilonen J (2003) Recombination in circulating enteroviruses. *J Virol* 77:10423–10431
24. Lukashev AN, Lashkevich VA, Koroleva GA, Ilonen J, Hinkkanen AE (2004) Recombination in uveitis-causing enterovirus strains. *J Gen Virol* 85:463–470
25. McMinn P, Lindsay K, Perera D, Chan HM, Chan KP, Cardosa MJ (2001) Phylogenetic analysis of enterovirus 71 strains isolated during linked epidemics in Malaysia, Singapore, and Western Australia. *J Virol* 75:7732–7738
26. Ministry of Health, People’s Republic of China. <http://www.moh.gov.cn/publicfiles/business/htmlfiles/mohbgt/s3582/200902/39079.htm>. Accessed 10 Aug 2009
27. Oberste MS, Maher K, Kilpatrick DR, Pallansch MA (1999) Molecular evolution of the human enteroviruses: correlation of serotype with VP1 sequence and application to picornavirus classification. *J Virol* 73:1941–1948
28. Ranganathan S, Singh S, Poh CL, Chow VT (2002) The hand, foot and mouth disease virus capsid: sequence analysis and prediction of antigenic sites from homology modelling. *Appl Bioinforma* 1:43–52
29. Santti J, Harvala H, Kinnunen L, Hyypiä T (2000) Molecular epidemiology and evolution of coxsackievirus A9. *J Gen Virol* 81:1361–1372
30. Santti J, Hyypiä T, Kinnunen L, Salminen M (1999) Evidence of recombination among enteroviruses. *J Virol* 73:8741–8749
31. Singh S, Poh CL, Chow VT (2002) Complete sequence analyses of enterovirus 71 strains from fatal and non-fatal cases of the hand, foot and mouth disease outbreak in Singapore (2000). *Microbiol Immunol* 46:801–808
32. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882
33. Wang JR, Tsai HP, Chen PF, Lai YJ, Yan JJ, Kiang D, Lin KH, Liu CC, Su IJ (2000) An outbreak of enterovirus 71 infection in Taiwan, 1998. II. Laboratory diagnosis and genetic analysis. *J Clin Virol* 17:91–99

34. Woo PC, Lau SK, Yip CC, Huang Y, Tsoi HW, Chan KH, Yuen KY (2006) Comparative analysis of 22 coronavirus HKU1 genomes reveals a novel genotype and evidence of natural recombination in coronavirus HKU1. *J Virol* 80:7136–7145
35. Yoke-Fun C, AbuBakar S (2006) Phylogenetic evidence for intertypic recombination in the emergence of human enterovirus 71 subgenotypes. *BMC Microbiol* 6:74
36. Zhang Y, Tan XJ, Wang HY, Yan DM, Zhu SL, Wang DY, Ji F, Wang XJ, Gao YJ, Chen L, An HQ, Li DX, Wang SW, Xu AQ, Wang ZJ, Xu WB (2009) An outbreak of hand, foot, and mouth disease associated with subgenotype C4 of human enterovirus 71 in Shandong, China. *J Clin Virol* 44:262–267