Comprehensive full-length sequence analyses of human parechoviruses: diversity and recombination

K. S. M. Benschop,1† M. de Vries,2† R. P. Minnaar,1 G. Stanway,3 L. van der Hoek,2 K. C. Wolthers1 and P. Simmonds4

1Laboratory of Clinical Virology, Department of Medical Microbiology, Academic Medical Center, Amsterdam, The Netherlands
2Laboratory of Experimental Virology, Department of Medical Microbiology, Academic Medical Center, Amsterdam, The Netherlands
3Department of Biological Sciences, University of Essex, Colchester CO4 3SQ, UK
4Virus Evolution Group, Centre for Infectious Diseases, University of Edinburgh, Edinburgh EH9 1AJ, UK

Received 29 June 2009
Accepted 15 September 2009

Human parechoviruses (HPEVs) are highly prevalent pathogens among very young children. Although originally classified into two serologically distinct types, HPEV1 and -2, recent analyses of variants collected worldwide have revealed the existence of 12 further types classified genetically by sequence comparisons of complete genome sequences or the capsid (VP1) gene. To investigate the nature of HPEV evolution, its population dynamics and recombination breakpoints, this study generated 18 full-length genomic sequences of the most commonly circulating genotypes, HPEV1 and -3, collected over a time span of 14 years from The Netherlands. By inclusion of previously published full-length sequences, 35 sequences were analysed in total. Analysis of contemporary strains of HPEV1 and those most similar to the prototype strain (Harris) showed that HPEV1 variants fall into two genetically distinct clusters that are much more divergent from each other than those observed within other HPEV types. Future classification criteria for HPEVs may require modification to accommodate the occurrence of variants with intermediate degrees of diversity within types. Recombination was frequently observed among HPEV1, -4, -5 and -6, but was much more restricted among HPEV3 strains. Favoured sites for recombination were found to flank the capsid region, and further sites were found within the non-structural region, P2. In contrast to other HPEV types, the majority of the HPEV3 sequences remained monophyletic across the genome, a possible reflection of its lower diversity and potentially more recent emergence than other HPEV types, or biological and/or epidemiological constraints that limit opportunities for co-infections with potential recombination partners.

INTRODUCTION

Human parechoviruses (HPEVs) are members of the family Picornaviridae and are classified within the genus Parechovirus (Stanway et al., 1994). HPEVs contain a single-stranded RNA of positive polarity and are approximately 7300 nt in length. The RNA encodes a large polyprotein within a single open reading frame flanked by 5′ and 3′ untranslated regions (UTRs). The polyprotein is post-translationally cleaved into three structural proteins (VP0, VP3 and VP1 encompassing P1) and seven non-structural (NS) proteins in the P2 (2A–2C) and P3 (3A–3D) domains.

Recent advances in molecular screening and sequencing have led to the rapid identification and characterization of several new HPEV types (Al-Sunaidi et al., 2007; Benschop et al., 2006b, 2008a; Drexler et al., 2009; Ito et al., 2004; Li et al., 2009; Watanabe et al., 2007). Currently, a total of 14 genotypes (HPEV1–14) has been identified and classified based on a minimum 25–30 % nucleotide divergence threshold (approx. 15 % amino acid divergence) in the VP1 gene (Benschop et al., 2006b; Oberste et al., 2004b). Of these, HPEV1 and -3 have been identified as the most commonly circulating strains in Europe and elsewhere (Benschop et al., 2008a; Tapia et al., 2008; van der Sanden et al., 2008; Watanabe et al., 2007). They have been isolated...
predominantly from very young children and are associated with a range of diseases ranging from mild gastrointestinal and respiratory symptoms to severe diseases such as neonatal sepsis and meningitis (Abed & Boivin, 2006; Benschop et al., 2006a; Stanway et al., 2000; Verboon-Maciele et al., 2008; Wolthers et al., 2008). Clinical studies have indicated HPeV3 as the predominant type associated with severe symptoms in neonates and as the main type infecting the central nervous system (Benschop et al., 2008a; Verboon-Maciele et al., 2008; Wolthers et al., 2008).

Phylogenetic analyses of full-length genomes have shown the genomes of many HPeV types to be highly mosaic, consistent with the occurrence of frequent recombination events (Al-Sunaidi et al., 2007; Benschop et al., 2006b; Li et al., 2009; Zoll et al., 2009). The exception is HPeV3, where all sequences available to date (three complete genome sequences and 43 paired VP1/3Dpol region sequences) cluster together throughout the genome (Benschop et al., 2008b; Williams et al., 2009).

In the current study, we generated 18 additional full-length sequences of the predominant HPeV genotypes, HPeV1 and HPeV3, and their variants identified over a time span of 14 years. Our work doubles the size of the HPeV complete sequence database and this extended dataset allowed a more robust analysis of the relationship between HPeV divergence and recombination frequency, and the positions where breakpoints have occurred.

RESULTS

Full-length sequence determination

Each of the 18 genomes sequenced was approximately 7200 nt in length. Eight of the ten HPeV1 strains clustered with other contemporary HPeV1 variants on phylogenetic comparison of capsid gene sequences and were collected over two periods from 1993 to 1994 and from 2002 to 2005. Together with three previously published full-length sequences from 2000 and 2003 [HPeV1 BNI-789St (de Souza Luna et al., 2008) and 7555312 (Zoll et al., 2009), and PicoBank/HPeV1/a (Williams et al., 2009)], the dataset comprised 11 full-length contemporary strains of HPeV1 (designated clade 1B; Fig. 1 and Table 1) collected over a period of 12 years. These variants differed substantially from the HPeV1 Harris strain from 1956. Two other HPeV1 variants collected in 2004 and 2007 clustered loosely with the Harris strain, a grouping we labelled clade 1A (Fig. 1 and Table 1).

Among the eight HPeV3 strains collected over the same 12-year time span as the HPeV1 variants, seven were closely similar to each other and the two prototype sequences [HPeV3 A308-99 (Ito et al., 2004); Can82852-01 (Abed & Boivin, 2005)]. However, we also characterized an HPeV3 strain (651689) that was more divergent in structural gene sequences from other HPeV3 variants (Table 1). The ten HPeV1 sequences, eight HPeV3 sequences and 17 previously published full-length sequences from GenBank used for phylogenetic analysis are shown in Table 1.

Sequence relationships in the capsid region

Phylogenetic analyses of the capsid-encoding regions (Fig. 1, VP0, VP3 and VP1) of the 35 sequences identified five major clusters (HPeV1, -3, -4, -5 and -6) and three single sequence branchings (HPeV2, -7 and -8), corresponding to the first eight of the 14 HPeV types currently described. Phylogenetic grouping of the HPeV3, -4, -5 and -6 variants into type-specific clades was bootstrap supported in each structural gene. For HPeV1, there was a clear separation of the HPeV1 group into two clusters, with contemporary strains falling into the designated clade 1B and a looser grouping of Harris-like variants designated clade 1A (Table 1, Fig. 1). Bootstrap-supported grouping was only observed within VP0 and VP1 (≥70%). Despite their separate groupings, the two clades showed 9.6% amino acid divergence in the VP1 gene, lower than the amino acid divergence between previously designated HPeV types (mean 26.4%; Table 2), and greater than those observed within other HPeV types, although with a distinct distribution of pairwise distances (Fig. 2b). The range of pairwise distance values between the two HPeV1 clades did not overlap with the values calculated between and within other genotypic groups. Similarly, nucleotide sequence divergence between the two HPeV1 clades of 23.5% was substantially greater than between variants within each genotypic group (mean 10.8%; Table 2, Fig. 2a), yet lower than that observed between other types (31.7%; Table 2, Fig. 2a).

The degree of sequence divergence between capsid region sequences correlated closely with difference in years of isolation (Fig. 1). For example, HPeV1 contemporary strains isolated in the early 1990s were found to be distinct from the more recently isolated strains and clustered separately within the HPeV1 contemporary clade 1B. However, HPeV3 strains isolated in the early 1990s were found to be quite similar to the more recently isolated strains (95.6% nucleotide similarity), indicative of a more direct line of descent over the 12-year observation period (i.e. the 1994 variants were more closely related to the common ancestor of recently isolated HPeV3 strains), a hypothesis supported by differences in the shape of HPeV3 and HPeV1 subtrees.

Phylogenetic relationship of HPeV in different genome regions

By comparison of VP1 and 3Dpol sequences, we showed previously that sequence divergence in the structural gene, and the difference in years between the isolation dates of the HPeV1, -4, -5 and -6 strains, correlated with a greater frequency of recombination. This was demonstrated by grouping HPeV variants into a number of phylogenetically
distinct clades in the 3Dpol region (Benschop et al., 2008b). With the full-length sequences, we were able to extend this analysis with HPeV variants collected over a longer observation period and to determine positions between VP1 and 3Dpol where recombination occurred.

All HPeV strains that were closely related to each other (VP1 divergence < 2.5%) remained monophyletic throughout the NS region (Fig. 1, 2A–2C, 3A–3D), including nine out of ten of the HPeV3 isolates. The more divergent HPeV3 strain (651689, 14% VP1 divergence) lost its phylogenetic clustering with other HPeV3 variants in 2A and throughout the rest of the NS region (Fig. 1).

Analysis of the segregation of sequences by type across HPeV genomes showed a 100% or near to 100% segregation within the capsid (zero or low y-axis values; Fig. 3). The capsid region was sharply demarcated at its 5' and 3' ends, with the 5' UTR and NS region sequences showing high values, indicative of recombination. Small violations of the type-specific grouping could also be observed within the VP3 gene in the capsid-encoding region and were found to be caused solely by a lack of monophylogeny of the two HPeV1 clades (clades 1A and 1B). This in turn was consistent with the observed lack of bootstrap support for HPeV1 variants in the VP3 phylogenetic tree (Fig. 1). Reassignment of the Harris-like clade 1A into a different sequence group from clade 1B restored 100% segregation within the capsid-encoding region (data not shown).

The sharp demarcation of segregation scores at each end of the capsid region (Fig. 3) is consistent with the existence of frequent recombination breakpoints at the structural gene boundaries. However, identification of the precise positions of breakpoints was complicated by the absence of the sequences of both recombination parents of putative recombinant strains. For example, the sequence of the HPeV variant that recombined with the divergent 651689 HPeV3 isolate was unrepresented in the dataset. As a result, breakpoint assignment could only be made by recording positions where phylogenetic groupings change, although necessarily this method of analysis does not indicate which of the two separated lineages underwent the recombination event and which remained non-recombinant.

Changes in phylogenetic groupings could be identified by recording tree positions and boundaries of phylogenetically supported clades of the 35 complete genome sequences using output from the TreeOrder scan program. Using a window size of 252 nt and an increment of 9 nt, changes in phylogeny were located principally at the structural gene boundaries, consistent with the segregation analysis (Fig. 3). The recombination breakpoints were spread over the 5' UTR–P1 junction for HPeV1, -3, -4, -5 and -6. At the P1–P2 junction, recombination breakpoints were also dispersed, already starting at the 3' end of the VP1 gene. This included the break-up within the HPeV1 group into separate clusters, as well as those of HPeV4 and -3 (Fig. 3). Within our dataset, breakpoints in HPeV5 and -6 were found in the 2C protein, corresponding to the second rise in segregation values (Fig. 3).

The position where recombination events occurred could also be visualized by calculation of frequencies of phylogeny violations between trees constructed from different sequence fragments in the genome. Using the same fragment size (252 nt) and increment (9 nt), violation scores were consistently low on comparison of different fragments of the capsid gene (shaded dark blue in Fig. 4; i.e. trees constructed from VP0, VP3 and VP1 were largely congruent), whilst higher scores were scored across the capsid–NS gene boundary as well as within the NS region (e.g. between P2 and P3) and the 5' UTR. The latter observations are consistent with the existence of more recombination breakpoints.

**DISCUSSION**

In this study, we obtained 18 new complete genome sequences of HPeV1 and -3, allowing a comprehensive analysis of intra- and intertype diversity of HPeV and the positions of and constraints on recombination.

Phylogenetic comparison of HPeV1 strains based on the capsid region identified two distinct clusters. The diversity seen by the separate clustering of clade 1B with the original Harris strain has been observed in previous studies (Abed & Boivin, 2006; Baumgarte et al., 2008; Benschop et al., 2006a; Williams et al., 2009). With only one strain defining clade 1A, data were limited to warrant a separate classification of the two variants. Divergence within the VP1 gene between the two clades of 23% approached the threshold of 25% defining HPeV types (Benschop et al., 2006b; Oberste et al., 2004b). The diversity resembled that seen for enterovirus 71 (EV71), showing a 17–26% nucleotide divergence (approx. 8–13% amino acid divergence) based on the VP1 gene between three major genogroups (Wang et al., 2002). The same was found based on the VP4 gene of EV71 (15–22% nucleotide divergence; Mizuta et al., 2005; Munemura et al., 2003). Similarly, echovirus 11 (Kapoor et al., 2004) and echovirus 30 (Savolainen et al., 2001) were found to contain specific clusters within the genotype. However, the criteria for defining a subcluster are not specifically defined and are often based on clustering and divergence within different genes. Here, we showed a distinctive distribution of pairwise distances within the VP1 gene that extended through the entire capsid gene (data not shown), justifying a subclassification of the HPeV1 genotype rather than a separate classification into new genotypes. The distinction between the two clades was found specifically within the VP3 gene. The VP3 protein contains the majority of the proposed antigenic determinants found on the capsid surface (Stanway et al., 1994). It is possible that sequence changes in HPeV1 underlying this diversification have been driven by immune escape mechanisms, ultimately producing serologically distinct descendants able to circulate...
Fig. 1. Unrooted phylogenetic analyses of HPeV based on post-translationally cleaved protein and untranslated flanking regions (5' UTR, VP0, VP3, VP1, 2A–2C, 3A–3D, 3' UTR) based on the neighbour-joining method with JC-corrected distances. HPeV strains are colour-coded. One thousand replicates were used to generate the bootstrap values, and values over 70% are shown. Bars indicate nucleotide divergence.

Diversity and recombination among human parechoviruses
independently from each other, although this seems unlikely as strains identified in clade 1B can be neutralized by antisera directed against the Harris strain (Abed & Boivin, 2006, and unpublished data), consistent with the much lower amino acid sequence divergence (10 and 8.4% in VP1 and P1, respectively) than between types (26 and 24%, respectively). However, more research is needed to determine whether the variability within type 1 might influence the titre or affinity of heterologous antibody reactivity.

Recombination between HPeV types was frequently observed within the NS region among HPeV1, -4, -5 and -6 strains, similar to many picornaviruses (Lukashev, 2005; Lukashev et al., 2003; Oberste et al., 2004a, c; Santti et al., 1999; Simmonds, 2006). However, we found a loss of type-specific segregation in only one of the HPeV3 sequences (strain 651689), indicating that, whilst HPeV3 sequences are able to recombine, this occurs less frequently than in other genotypes. In this specific case, the high divergence in P1 and therefore greater temporal (evolutionary) separation between this isolate and other HPeV3 strains are probable contributory factors for a recombination event to have occurred in the past (Benschop et al., 2008b). However, we did not observe any recombination among other HPeV3 strains with a similar degree of temporal separation. In marked contrast, most HPeV1 variants in

### Table 1. Full-length sequences

<table>
<thead>
<tr>
<th>HPeV type</th>
<th>Strain</th>
<th>Year of isolation</th>
<th>Clade</th>
<th>GenBank accession no.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPeV1</td>
<td>Harris</td>
<td>1956</td>
<td>Clade 1A</td>
<td>S45208</td>
<td>Stanway et al. (1994)</td>
</tr>
<tr>
<td></td>
<td>BNI-788st</td>
<td>2003</td>
<td>Clade 1B</td>
<td>EF051629</td>
<td>de Souza Luna et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>7555312</td>
<td>2003</td>
<td>Clade 1B</td>
<td>FM178558</td>
<td>Zoll et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>PicoBank/HPeV1/a</td>
<td>2000</td>
<td>Clade 1B</td>
<td>FM242866</td>
<td>Williams et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>152478</td>
<td>2001</td>
<td>Clade 1B</td>
<td>CQ183018</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>252581</td>
<td>2002</td>
<td>Clade 1B</td>
<td>CQ183019</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>450343</td>
<td>2004</td>
<td>Clade 1B</td>
<td>CQ183020</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>550163</td>
<td>2005</td>
<td>Clade 1B</td>
<td>CQ183021</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>K54-94</td>
<td>1994</td>
<td>Clade 1B</td>
<td>CQ183024</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>K63-94</td>
<td>1994</td>
<td>Clade 1B</td>
<td>CQ183025</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>K129-93</td>
<td>1993</td>
<td>Clade 1B</td>
<td>CQ183022</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>K150-93</td>
<td>1993</td>
<td>Clade 1B</td>
<td>CQ183023</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>2007-863</td>
<td>2007</td>
<td>Clade 1A</td>
<td>CQ183034</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>452568</td>
<td>2004</td>
<td>Clade 1A</td>
<td>CQ183035</td>
<td>This study</td>
</tr>
<tr>
<td>HPeV2</td>
<td>Williamson</td>
<td>1956</td>
<td></td>
<td>AJ005695</td>
<td>Ghazi et al. (1998)</td>
</tr>
<tr>
<td></td>
<td>A308-99</td>
<td>1999</td>
<td></td>
<td>AB084913</td>
<td>Ito et al. (2004)</td>
</tr>
<tr>
<td></td>
<td>152037</td>
<td>2001</td>
<td></td>
<td>CQ183026</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>251360</td>
<td>2002</td>
<td></td>
<td>CQ183027</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>450936</td>
<td>2004</td>
<td></td>
<td>CQ183028</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>651689</td>
<td>2006</td>
<td></td>
<td>CQ183029</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>K8-94</td>
<td>1994</td>
<td></td>
<td>CQ183033</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>K11-94</td>
<td>1994</td>
<td></td>
<td>CQ183030</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>K12-94</td>
<td>1994</td>
<td></td>
<td>CQ183031</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>K20-94</td>
<td>1994</td>
<td></td>
<td>CQ183032</td>
<td>This study</td>
</tr>
<tr>
<td>HPeV4</td>
<td>T75-4077</td>
<td>1975</td>
<td></td>
<td>AM235750</td>
<td>Al-Sunaidi et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>K251176-02</td>
<td>2002</td>
<td></td>
<td>DQ315670</td>
<td>Benschop et al. (2006b)</td>
</tr>
<tr>
<td></td>
<td>FUK2005-123</td>
<td>2005</td>
<td></td>
<td>AB433629</td>
<td>Watanabe et al. (2007)</td>
</tr>
<tr>
<td>HPeV5</td>
<td>CT86-6760</td>
<td>1986</td>
<td></td>
<td>AF055846</td>
<td>Oberste et al. (1998)</td>
</tr>
<tr>
<td>HPeV6</td>
<td>NII561-2000</td>
<td>2000</td>
<td></td>
<td>AB252582</td>
<td>Watanabe et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>BNI-67/03</td>
<td>2003</td>
<td></td>
<td>EU024629</td>
<td>Baumgarte et al. (2008)</td>
</tr>
<tr>
<td>HPeV7</td>
<td>Pak5045</td>
<td>2007</td>
<td></td>
<td>EU556224</td>
<td>Li et al. (2009)</td>
</tr>
<tr>
<td>HPeV8</td>
<td>Br/217/2006</td>
<td>2006</td>
<td></td>
<td>EU716175</td>
<td>Drexler et al. (2009)</td>
</tr>
</tbody>
</table>

### Table 2. Nucleotide and amino acid divergence between and within types based on the VP1 gene

<table>
<thead>
<tr>
<th></th>
<th>Nucleotide divergence mean (%) (range)</th>
<th>Amino acid divergence mean (%) (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>31.7 (28.9–35.9)</td>
<td>26.4 (22.3–33.4)</td>
</tr>
<tr>
<td>Within groups</td>
<td>10.8 (4.4–16.1)</td>
<td>3.5 (1.6–5.3)</td>
</tr>
</tbody>
</table>
clade 1B collected in the same time span of 12 years showed evidence for recombination; i.e. those isolated in 1993 and 1994 did not remain segregated with strains isolated 7 years later, in 2001 (152478) and 2002 (252581). The lower observed frequency of recombination observed in HPeV3 sequences may reflect biological constraints such as differences in cellular tropism and spread to different tissues, limiting the opportunities for recombination to occur.

The most frequent sites for recombination flanked the capsid region and could readily be identified by the segregation scans and bootscan analysis (de Souza Luna et al., 2008; Zoll et al., 2009). These sites mirror those observed for other picornaviruses such as enteroviruses (Lukashev et al., 2003; Oberste et al., 2004a; Simmonds & Welch, 2006) and aphthoviruses (Heath et al., 2006; Tosh et al., 2002). Within this study, we were able to map these sites more precisely for HPeV. The recombination breakpoints identified in the current study occurred at several positions around the P1–P2 junction, already starting at the C-terminal end of the VP1 gene, confirming the findings by Williams et al. (2009). This region contains the RGD receptor-binding domain, and recombination at this position can thus have major consequences for receptor usage and tropism of different HPeV types and variants (Benschop et al., 2008a; Williams et al., 2009). Two breakpoints were found near the P2–P3 junction, within the 2C protein, for HPeV5 and HPeV6, an observation that is not uncommon among picornaviruses (Heath et al., 2006; Simmonds & Welch, 2006). Thus, the segregation of HPeV5 and -6 was found to carry over to a large portion of the P2 NS region, although we cannot rule out that this is due to a sampling effect of only two HPeV5 strains and three HPeV6 strains. These obvious crossover points flanking the capsid gene and near the P2–P3 boundary are relatively conserved among picornaviruses, facilitating recombination through template switching during replication (Kirkegaard & Baltimore, 1986; Lukashev, 2005). Alternatively, recombination at other sites (i.e. within functional protein domains) may be more likely to produce non-viable or less fit viruses, preventing their fixation in the population.

The data presented here address the need for clear classification criteria that accommodate the additional tier of variability within some types, especially in a time when new types are rapidly being identified. Recombination in

![Fig. 2. Distributions of nucleotide (a) and amino acid (b) sequence pairwise p distances measured in the VP1 region within HPeV types (open columns), between types (shaded columns) and the subset of comparisons between clades 1A and 1B within HPeV (filled columns). The distributions were calculated from a total of 2878 pairwise distances.](image)

![Fig. 3. Segregation scans across the HPeV genome using a window size of 252 nt with an increment of 9 nt. Segregation scores are plotted according to their position in the genome and are represented as a value between 0 (perfect segregation) and 1 (no segregation). A schematic of the genome is given for visualization purposes. The arrows are colour-coded (see key) and indicate the position at which specific clades broke up.](image)
polymorphism) combined with primer walking (de Vries van der Hoek et al., 2004) (Table 1). The 5’ UTR was partially amplified excluding the first 100 nt (loop A) (Nateri et al., 2000). The 3’ UTR was amplified with a tagged oligo(dT) primer. Partially overlapping fragments of 600–1500 nt were amplified and purified from an agarose gel, before sequencing using a Big Dye Terminator Cycle Sequencing kit (Applied Biosystems). Partial fragments were assembled using Vector NTI Advance 10 software (Invitrogen) and edited manually.

**Nucleotide sequence alignment, and diversity and distance measurements.** Sequences were aligned by nucleotide- and codon-based identity implemented in CLUSTAL W within the Simmonics Sequence Editor software package (version 1.6, http://www.virus-evolution.org; Simmonds & Smith, 1999). All full-length sequences available from GenBank were included in the alignment (Table 1), along with the following P1 sequences (type): HPeV4: AB433630, AB434673; HPeV1: AB112485, AB112486, AB112487, EU024630, EU024632, EU024633, EU024634, EU024635, EU024636; HPeV3: AB112484, AB112483, AB112482; and VP1 sequences: HPeV5: AB443848, EU077511; HPeV1: AB443802, AB443830, AB443809, AB443817, AB443814; HPeV6: AB300942. Nucleotide sequence numbering was based on the reference sequence from HPeV1 Harris isolate (S45208). Sequence distances within and between types were calculated using the program Sequence Distance within the Simmonic Editor package with sequence groups assigned to types 1–8. Sequence diversity was computed as mean pairwise Jukes and Cantor (JC)-corrected distances between nucleotide sequences or p distances for amino acid sequences.

**Phylogenetic and recombination analysis.** Phylogenetic trees were constructed with MEGA3.1 (Kumar et al., 2004) based on JC distances (Jukes & Cantor, 1969). The use of other evolutionary models for distance estimation (i.e. Kimura two-parameter, maximum composite likelihood distances, maximum parsimony and maximum likelihood) did not influence the tree topology or calculations of genetic distances. One thousand bootstrap replicates were analysed. Recombination analysis was automated within the Simmonics Sequence Editor software and phylogenetic incongruencies were recorded across the genome using the TreeOrder scan program (Simmonds, 2006). Sequence scans were run using a sliding window of 252 nt (84 codons) and an increment of 9 nt (three codons).

**REFERENCES**


Benschop, K. S. M., Schinkel, J., Luken, M. E., van den Broek, P. J. M., Beersma, M. F. C., Menelik, N., van Eijk, H. W. M., Zaaier, H. L.,
Diversity and recombination among human parechoviruses


Wang, J. R., Tuan, Y. C., Tsai, H. P., Yan, J. J., Liu, C. C. & Su, I. J. (2002). Change of major genotype of enterovirus 71 in outbreaks of...


