Journal of General Virology (2011), **92**, 564–571 **DOI** 10.1099/vir.0.022525-0

Full genome sequence analysis of parechoviruses from Brazil reveals geographical patterns in the evolution of non-structural genes and intratypic recombination in the capsid region

Jan Felix Drexler,[1,2]† Klaus Grywna,[2]† Alexander Lukashev,[3]† Andreas Stöcker, Patrícia Silva Almeida,[4] Janett Wieseler,[2] Tereza Cristina Medrado Ribeiro,[4] Nadine Petersen,[5] Hugo da Costa Ribeiro, Jr,[4] Ilya Belalov,[3] Beate Mareike Kümmerer[2] and Christian Drosten[2]

Correspondence
Christian Drosten
drosten@virology-bonn.de

[1]Research Laboratory for Infectious Diseases, Hospital Professor Edgar Santos, Federal University of Bahia, Salvador, Bahia, Brazil

[2]Institute of Virology, University of Bonn Medical Centre, Bonn, Germany

[3]Chumakov Institute of Poliomyelitis and Viral Encephalitides, Moscow, Russia

[4]Department of Paediatrics, Hospital Professor Edgar Santos, Federal University of Bahia, Salvador, Bahia, Brazil

[5]Clinical Virology Group, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany

Due to high genome plasticity, the evolutionary fate and geographical history of picornaviruses is hard to follow. Here, we determined the complete coding sequences of eight human parechoviruses (HPeV) of types 1, 5 and 6 directly from clinical samples from Brazil. The capsid genes of these strains were not remarkably different from European, North American and Japanese HPeV. Full genome analysis revealed frequent intertypic recombination in the non-structural genome region. In addition, evidence of recombination between viruses of the same type in the capsid-encoding genome region among HPeV1 and HPeV4 was obtained. Bayesian phylogenetic analysis indicated that strains without evidence of recombination with each other in any genome region were separated by no more than 35 years of circulation. Interestingly, in the 3C gene, all Brazilian parechoviruses grouped together regardless of serotype. The most recent common ancestor of these strains dated back 108 years, suggesting long-term endemicity of this particular P3 genome lineage in South America. Our results support the idea that picornavirus replicative genes acquire capsid proteins introduced by new strains. Under certain epidemiological conditions, replicative genes may be maintained in circumscript geographical regions.

Received 8 April 2010
Accepted 29 November 2010

INTRODUCTION

Human parechoviruses (HPeV) are small non-enveloped viruses that define a genus of the family *Picornaviridae*. Their positive-sense RNA genome encodes a single polyprotein that is cleaved into three structural proteins, VP0, VP3 and VP1 (defining the P1 genome region), and seven non-structural proteins, 2A–2C (P2) and 3A–3D (P3). The first two HPeV serotypes were isolated in the 1950s and designated echovirus 22 and 23. In the 1990s, parechoviruses were reclassified into a new genus based on their distinct genome organization (Hyypiä *et al.*, 1992). High prevalence of HPeV was not truly appreciated until the last few years. Recently, HPeVs were shown to be highly diverse, with up to 14 provisionally assigned types (Al-Sunaidi *et al.*, 2007; Benschop *et al.*, 2008a; Drexler *et al.*, 2009; Ito *et al.*, 2004; Kim Pham *et al.*, 2010; Li *et al.*, 2009; Watanabe *et al.*, 2007). HPeV also proved highly prevalent in the human population, with seropositivity rates reaching 88 % at the age of 2 years (Joki-Korpela & Hyypiä, 1998; Takao *et al.*, 2001), and an important cause of meningitis and sepsis-like disease of newborns (Baumgarte *et al.*, 2008;

†These authors contributed equally to this work.

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are HQ696570–HQ696577.

Benschop *et al.*, 2006; Wolthers *et al.*, 2008). In general, recent reports showed that epidemiological and clinical features of HPeV are similar to the most common picornavirus genus, *Enterovirus*. HPeV genetic and evolutionary properties also resemble those of enteroviruses. They feature both high mutation rate (Faria *et al.*, 2009) and frequent natural recombination (Benschop *et al.*, 2008b, 2010; Calvert *et al.*, 2010; Williams *et al.*, 2009; Zoll *et al.*, 2009). Unfortunately, analyses of recombination patterns have been limited by the low number of available full-length genome sequences and, most critically, limited geographical coverage of HPeV isolates. We analysed the complete coding sequence of eight HPeV sampled in Brazil, adding significantly to the geographical scope of the 34 sequences currently available in GenBank. These data were used to analyse recombination patterns with special respect to geographical implications.

RESULTS AND DISCUSSION

Whole genome analysis of HPeV

We determined the full coding sequence of eight Brazilian HPeVs representing types 1 (n=5), 5 (n=2) and 6 (n=1). These strains comprise the first near-complete genomic sequences from South America, while all other available strains were sampled in Europe, North America and Asia.

Recombination in HPeV has been described previously (Benschop *et al.*, 2008b, 2010; Calvert *et al.*, 2010; de Souza Luna *et al.*, 2008; Williams *et al.*, 2009; Zoll *et al.*, 2009). In agreement with these reports, a preliminary analysis of the novel dataset using phylogenetic compatibility matrices implemented in the SIMMONICS package (Simmonds & Smith, 1999) and bootscanning (Salminen *et al.*, 1995) implemented in SIMPLOT v.3.5.1 (Lole *et al.*, 1999) detected multiple recombination events, with most crossover points mapping to the regions immediately bordering the P1 genome region (data not shown).

Treeorder analysis

Treeorder analysis (Simmonds & Welch, 2006) was used to get a first overview of dynamics of genetic information exchange within the global pool of HPeV genetic information (Fig. 1). Groups corresponding to defined HPeV types were supported over the whole P1 region, suggesting absence of recombination in the capsid-encoding region between viruses of different types (Fig. 1a). Tree order was disrupted upon transition to the 2A region for all HPeV types with the exception of HPeV3. Only one Dutch HPeV3 isolate (termed 651689) showed signs of recombination in the P2 and P3 genome region with other parechoviruses, while all other HPeV3 grouped consistently over the whole genome, confirming previous reports (Benschop *et al.*, 2008b, 2010; Calvert *et al.*, 2010; de Souza Luna *et al.*, 2008; Williams *et al.*, 2009; Zoll *et al.*, 2009).

Capsid region

For a closer analysis of recombination within the capsid region, Bayesian phylogenetic trees were constructed in the VP0 and VP1 protein genes and opposed as summarized in Fig. 2. Interestingly, there were a few cases of phylogenetic incongruence between the VP0 and VP1 genome regions that involved viruses of the same type. For example, the Brazilian strain HPeV1_BR_114 grouped reliably with strain HPeV1_863 in VP0 and with strain HPeV1_452568 in VP1. HPeV4 _T75-4077 grouped with strain HPeV4_123 in VP1 with a posterior probability of 1, and with strain HPeV4_K251176 in VP0, albeit at a low posterior probability. Therefore, there was evidence of intratypic recombination in the capsid-encoding genome region of parechoviruses.

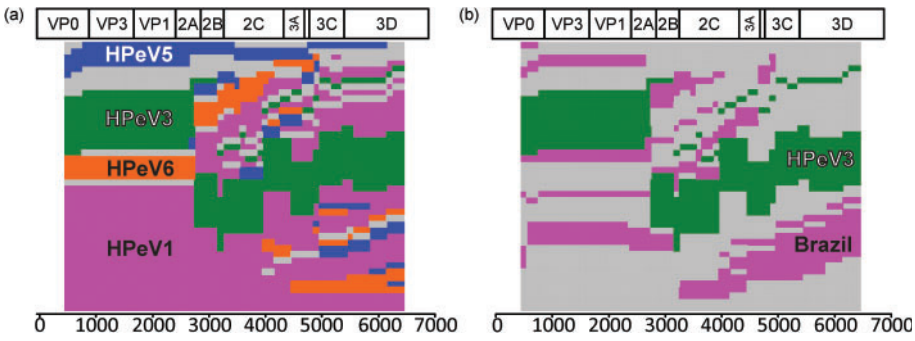

Fig. 1. Treeorder scan of parechovirus complete coding sequences performed with SIMMONICS 1.6 (Simmonds & Smith, 1999). Positions of individual sequences on phylogenetic trees (x-axis) are shown for a walking window of 1000 nt, step size 100 nt, bootstrap cut-off 70 %. y-Axis shows window midpoint position in the genome. (a) Colour-coded according to the VP1-type of analysed full genomes. (b) Coloured according to major groups observed in the 3D protein region. Viruses not relevant for the discussion (uncommon types or those not belonging to distinct 3D groups) are shown in grey.

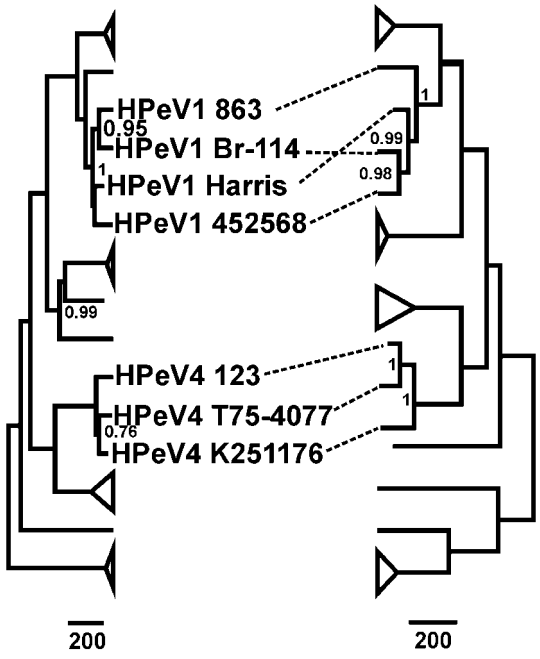

Fig. 2. Comparison of VP0 and VP1 in Bayesian phylogeny. Phylogenetic relation of parechoviruses in four genome regions. Only those clades containing strains with evidence of intratypic recombination are identified. Trees were created using Bayesian algorithm implemented in BEAST v.1.5.3 software (Drummond & Rambaut, 2007). Numbers at tree nodes are Bayesian posterior probabilities. Bars, indicate branch length in years.

To determine whether the addition of South American HPeV strains would change recent proposals on coalescent dating of HPeV capsid gene evolution, Bayesian molecular clock analysis was done on VP1 genes as summarized in Fig. 3 and Table 2. The most recent common ancestor (MRCA) of all extant HPeV capsid genes dated back 495 years, which was highly consistent with the recently published proposal (Faria *et al.*, 2009). It should be critically noted that the high mutation frequency in picornaviruses will lead to an underestimation of evolutionary distances at longer timescales, due to the fact the multiple nucleotide exchanges cannot appropriately be taken into account even by complex substitution models, such as the general time reversible (GTR) model chosen here. It should also be noted that the coalescent dating of picornavirus evolution is highly problematic and unreliable due to ubiquitous recombination that disrupts the molecular clock-like evolution. Nevertheless, the specific absence of recombination within VP1 genes of picornaviruses would qualify these genes as the least error-prone target regions for coalescent ancestor dating. Even if molecular clock-based dating cannot be taken as an absolute calibration of picornavirus evolution, at least the VP1-based results of Bayesian dating may be useful as reference estimates to compare the evolution of other genomic regions (see below).

Non-coding region

The results of a Treeorder scan sorted according to the most downstream portion of the non-structural genes (3D protein gene) are shown in Fig. 1(b). Except for HPeV3 for which absence of recombination has already been described (Benschop *et al.*, 2008b, 2010; Calvert *et al.*, 2010; de Souza Luna *et al.*, 2008; Williams *et al.*, 2009; Zoll *et al.*, 2009), different strains only grouped across rather short regions, often situated in or between the 2B and 3D portions of genomes. However, it was noted that all Brazilian strains grouped together in more than half of the genome, extending from the 3D gene into the 2C gene (Fig. 1b). This group included the formerly described eighth parechovirus type from Brazil (Drexler *et al.*, 2009), and did not include any strains from other geographical regions. Because the neighbour-joining algorithm used for Treeorder scanning could not consistently resolve phylogenetic positions for some viruses, more reliable phylogenetic analysis was done in order to support or rule out grouping of the Brazilian isolates in the non-structural part of the genome. Bayesian phylogenetic trees were created across the genome to obtain a reliable phylogeny for different virus protein genes. Phylograms in four key genome regions are shown in Fig. 3. Only in the 3C protease gene did all Brazilian strains form a clear monophyletic clade with very high Bayesian posterior probability. Already in the neighbouring upstream 3AB protein gene, this clade fell into two separate clades, compatible with a single recombination event between an ancestor of HPeV4_123 and a common ancestor of one group of Brazilian viruses forming a monophyletic subclade in the 3C protease gene (coalescent in node N in Fig. 3). Because of the monophyly of the Brazilian 3C gene sequences, we presumed for the Brazilian strains an absence of significant recombination in this gene and used the underlying data for Bayesian coalescent dating, allowing rate variation between lineages to reflect the fact that we could not systematically exclude recombination in other (non-Brazilian) branches of the same tree. With these restrictions in mind, it could be derived from the coalescent dating that the mentioned recombination event must have taken place between 97 and 43 years before present (compare nodes L and N in Fig. 3). The 95 % highest posterior density (HPD) intervals associated with these estimates are shown in Table 2. The fact that node G, identifying the same recombination event in the 3AB gene, is dated more recently (39 years), is not contradictory because is has to be assumed that further recombination has taken place after the described event, which lead to a shuffling of sequences and a levelling of divergence, resulting in a seemingly more recent recombination date in the 3AB gene. We thus prefer the older (3C-based) recombination date assumption. Interestingly, strain HPeV6_BR_104 has taken an independent evolutionary path as opposed to HPeV5_BR_53 and HPeV5_BR_77. Because this topology is congruent in the 3AB and 3C genes, the most parsimonious explanation is that no further recombination between 3AB and 3C has since taken place, but the lineage has split into two sublineages evolving

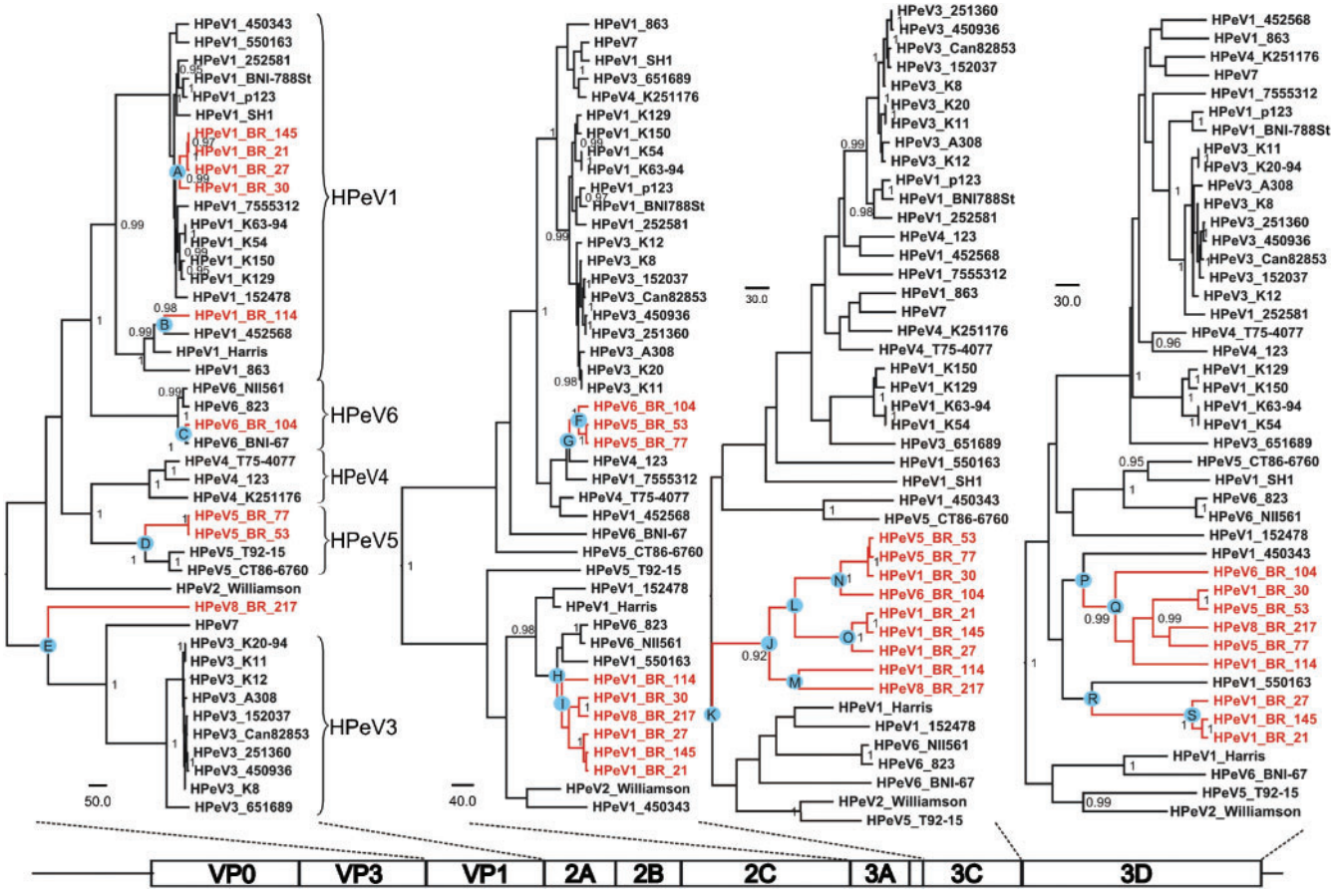

Fig. 3. Comparative Bayesian phylogenetic analysis in selected regions of the HPeV genome, as identified in the bottom of the figure. The VP1-based tree contains an identification of HPeV types based on distance in the VP1 fragment. Brazilian strains are shown in red. Numbers at tree nodes are Bayesian posterior probabilities. Selected root points are identified by letters in blue circles. The projected dates of coalescence are identified for these root points in Table 2. Trees were created using Bayesian algorithm implemented in BEAST v.1.5.3 software (Drummond & Rambaut, 2007). Bars, indicate branch length in years.

independently. This matches the two different VP1 types represented in this clade, probably generating evolutionary isolation. Thus, the type 6 capsid in HPeV6_BR_104 and/or the type 5 capsid in HPeV5_BR_53 and HPeV5_BR_77 must have been acquired after the split of both lineages, i.e. less than 43 years ago (node N in Fig. 3). From the coalescent dating of the VP1 genes it would appear more likely that the type 5 capsid gene would have been acquired, because it has split off from non-Brazilian ancestors earlier than the type 6 capsid gene (compare nodes C and D in Fig. 3). However, the topology of the 3C gene tree suggests a different most parsimonious explanation. The clade separating at node N would have acquired a type 6 capsid in one of the new sublineages (represented by HPeV6_BR_104). The remaining clade must have recently experienced another split by acquisition of either a type 5 or type 1 capsid. Acquisition of the type 5 capsid in a recent common ancestor of HPeV5_BR_53 and HPeV5_BR_77 appears more likely because basal and apical members of the Brazilian clade in the 3C gene are type 1 viruses, suggesting a type 1 virus as the MRCA of the whole clade identified by node J. Because strain HPeV1_BR_30 (the apical type 1 virus in the clade identified by node N) is co-segregating with the rest of the Brazilian viruses and not with the rest of the N-node clade in the 3AB gene, it appears likely that the recombination breakpoint that led to the acquisition of a type 6 capsid gene in the HPeV6_BR_104 lineage was between 3AB and 3C.

Remarkably, the monophyletic clade of Brazilian viruses in the 3C gene falls into two subclades in 3D. The most parsimonious explanation for this split is a single recombination event in the ancestral lineage preceding node S. The recombination event would have taken place between 22 and 145 years BP (nodes R and S in Fig. 3). However, these estimates have to be treated with caution because we have to consider further recombination within 3D in view of the change in apical topology as compared with the 3C gene tree. Nevertheless, the more reliable dating in the 3C gene suggests that the event has taken place between 29 and 97 years ago (nodes O and L), which is in good concordance. The basal relationship in the 3D gene between the main Brazilian 3D gene lineage (node Q)

Table 1. Full-length sequences used for analysis

HpeV type	Strain	Year of isolation	Country of isolation	GenBank accession no.	Reference
HPeV1	Harris	1956	USA	S45208	Stanway *et al.* (1994)
	BNI-788st	2003	Germany	EF051629	de Souza Luna *et al.* (2008)
	7555312	2003	Netherlands	FM178558	Zoll *et al.* (2009)
	PicoBank/HPeV1/a	2000	Finland	FM242866	Williams *et al.* (2009)
	152478	2001	Netherlands	GQ183018	Benschop *et al.* (2010)
	252581	2002	Netherlands	GQ183019	Benschop *et al.* (2010)
	450343	2004	Netherlands	GQ183020	Benschop *et al.* (2010)
	550163	2005	Netherlands	GQ183021	Benschop *et al.* (2010)
	K54-94	1994	Netherlands	GQ183024	Benschop *et al.* (2010)
	K63-94	1994	Netherlands	GQ183025	Benschop *et al.* (2010)
	K129-93	1993	Netherlands	GQ183022	Benschop *et al.* (2010)
	K150-93	1993	Netherlands	GQ183023	Benschop *et al.* (2010)
	2007-863	2007	Netherlands	GQ183034	Benschop *et al.* (2010)
	452568	2004	Netherlands	GQ183035	Benschop *et al.* (2010)
	Br/21	2006	Brazil	HQ696570	This study
	Br/27	2006	Brazil	HQ696571	This study
	Br/30	2006	Brazil	HQ696572	This study
	Br/114	2006	Brazil	HQ696573	This study
	Br/145	2006	Brazil	HQ696574	This study
HPeV2	Williamson	1956	USA	AJ005695	Ghazi *et al.* (1998)
HPeV3	Can82852-01	2002	Canada	AJ889918	Abed & Boivin (2005)
	A308-99	1999	Netherlands	AB084913	Ito *et al.* (2004)
	152037	2001	Netherlands	GQ183026	Benschop *et al.* (2010)
	251360	2002	Netherlands	GQ183027	Benschop *et al.* (2010)
	450936	2004	Netherlands	GQ183028	Benschop *et al.* (2010)
	651689	2006	Netherlands	GQ183029	Benschop *et al.* (2010)
	K8-94	1994	Netherlands	GQ183033	Benschop *et al.* (2010)
	K11-94	1994	Netherlands	GQ183030	Benschop *et al.* (2010)
	K12-94	1994	Netherlands	GQ183031	Benschop *et al.* (2010)
	K20-94	1994	Netherlands	GQ183032	Benschop *et al.* (2010)
HPeV4	T75-4077	1975	USA	AM235750	Al-Sunaidi *et al.* (2007)
	K251176-02	2002	Netherlands	DQ315670	Benschop *et al.* (2006)
	FUK2005-123	2005	Japan	AB433629	Wakatsuki *et al.* (2008)
HPeV5	CT86-6760	1986	USA	AF055846	Oberste *et al.* (1998)
	T92-15	1992	USA	AM235749	Al-Sunaidi *et al.* (2007)
	Br/53	2006	Brazil	HQ696575	This study
	Br/77	2006	Brazil	HQ696576	This study
HPeV6	NII561-2000	2000	Japan	AB252582	Watanabe *et al.* (2007)
	BNI-67/03	2003	Germany	EU024629	Baumgarte *et al.* (2008)
	2005-823	2005	Netherlands	EU077518	de Vries *et al.* (2008)
	Br/104	2006	Brazil	HQ696577	This study
HPeV7	Pak5045	2007	Pakistan	EU556224	Li *et al.* (2009)
HPeV8	Br/217/2006	2006	Brazil	EU716175	Drexler *et al.* (2009)

and strain HPeV_450343 can be explained by an ancient recombination event that would have taken place in the common ancestor of all Brazilian 3CD genes, at least 131 years ago (node J). Node P suggests that the earliest time point for this recombination was 156 years ago, but again this number has to be treated with great caution because of possible sequence shuffling by undetected recombination. In summary, we can assume that the oldest common element of extant Brazilian HPeV strains is the 3C gene. This gene has then accepted new upstream and downstream genomic elements without being displaced by new 3C gene elements that must have been encountered along with the arrival of upstream and downstream elements in this geographical region.

Validation of coalescent dating assumptions

The considerations above relied on comparisons of MRCA dates of clades of viruses that were preserved in neighbouring genomic regions (3AB vs 3C or 3C vs 3D). According to our Bayesian model allowing lineage-dependent rate variation, this approach assumed that the molecular clocks of such region-spanning clades were not disrupted in spite of underlying ubiquitous recombination

Table 2. Coalescent ancestor dating estimates for selected root point in Fig. 3

Node	Node height (years)	95 % HPD* (years)
A	29	15–46
B	70	38–105
C	13	7–19
D	122	35–84
E	384	192–608
F	20	5–45
G	37	19–62
H	59	19–104
I	50	12–96
J	131	44–226
K	202	112–322
L	97	31–178
M	95	22–174
N	43	12–78
O	29	9–52
P	156	97–224
Q	117	65–176
R	145	91–208
S	22	12–34

*Bayesian highest posterior density intervals.

in other branches of the same trees. To validate this assumption, we extended the comparison to clades without traces of inter- or intraclade recombination over the whole genome. Two HPeV1 clades, one HPeV6 clade and one HPeV3 clade fulfilled this criterion. Independently in all of these clades time to the most recent common ancestor (tMRCA) values were highly congruent across the genome (Table 3). In contrast, a comparison of tMRCA values for all HPeV revealed approximately 2.5 times younger MRCA dates in the non-structural genes than in the structural genes. We interpret this as evidence of the disturbing effect of recombination on the molecular clock in the non-structural genes. The conservation of ancestor dates in region-spanning clades supports our approach taken to reconstruct the evolutive fate of Brazilian non-structural gene elements. However, it has to be kept in mind that a global underestimation of evolutive distances will probably take effect in general because we cannot exclude the effects of intra-clade recombination on a smaller scale in any genomic region. With this limitation in mind, the observed range of 8–35 years limiting the time of existence of region-spanning clades suggests approximately 30–40 years as a maximum time span for which an isolate in our sample circulated without being involved in recombination.

In summary, our findings of frequent intertypic recombination between capsid and non-structural proteins and of rare intratypic, intra-capsid recombination confirm and extend the notion that parechoviruses are not different from most other human picornaviruses in their recombination behaviour (Benschop *et al.*, 2008b, 2010; Calvert *et al.*, 2010; Williams *et al.*, 2009; Zoll *et al.*, 2009). The major finding in this study is that the cohesion of non-structural protein genes of Brazilian viruses with total geographical restriction and exclusiveness of the 3C protein gene. Corresponding findings have been made for partial sequences from the Netherlands and the UK clustering with each other, as opposed to those from Thailand forming a separate entity (Calvert *et al.*, 2010). In summary, this suggests that HPeV non-structural protein genes are restricted to large geographical areas in an endemic (rather than epidemic) pattern. This parallels some enteroviruses, for which temporal and geographical grouping of the 3D genome region has been reported. Twenty-nine of 32 enteroviruses representing 11 serotypes isolated during 1998–2002 in the Former Soviet Union clustered uniformly in the 3D genome region and were strictly separated from other prototype human enterovirus B strains. In analogy to the genus *Enterovirus*, it has thus been proposed to analyse the fate of HPeVs in circulation by defining recombinant forms (RFs) that consist of stable combinations of at least capsid (VP1) and 3D polymerase genes (Calvert *et al.*, 2010; Leitch *et al.*, 2009). As an alternative proposal, one might also use the 3C gene (together with VP1) to define an HPeV RF, as it may indicate geographical and temporal clustering with better specificity than 3D in parechoviruses. It will be very interesting in the future to determine molecular mechanisms involved in the conservation of the 3C gene. The encoded protein functions as the main protease of HPeV with an extended range of cleavage sites as opposed to 3C proteins of enteroviruses including secondary cleavages in the P1 polyprotein portion (the 2A protein of parechoviruses has no proteolytic activity). Homologous interference as observed in other viruses, based on *trans*-cleavage of proteins of the superinfecting virus by activated proteases of the first virus would be conceivable (Johnston *et al.*, 1974). Whether homologous interference occurs in picornaviruses is inconclusive. It has been observed in foot-and-mouth disease virus (Polacino *et al.*, 1985), but could not be confirmed in rhinoviruses (Gercel *et al.*, 1985). Irrespective of the underlying mechanism, interference might render the genome of the second (overinfecting) virus incapable of replicating, while it still might serve as a template for recombination, explaining a *trans*-active excluding effect on superinfecting viruses. Our genetic data suggest that an excluding effect may be conferred by the 3C gene of the first virus, and might become epidemiologically apparent if a certain HPeV non-structural gene set predominates in a certain region. However, this idea is highly speculative and requires further investigation in epidemiological and molecular studies.

METHODS

Genome sequencing. HPeV strains sequenced here were identified previously in stool samples from patients with acute diarrhoea collected in 2006 in Salvador, Bahia, North-Eastern Brazil (Drexler *et al.*, 2008). Of the 54 samples that were positive in a consensus real-time RT-PCR assay targeting the genomic 5′-NTR (Baumgarte *et al.*, 2008), the full-length VP1 gene could be determined by consensus PCR as described for 11 samples (Drexler *et al.*, 2008). Published PCR

Table 3. tMRCA and 95 % HPD (years) of the non-recombinant parechovirus groups in different genome regions

Clade	VP0	VP1	3ABC	3D
HPeV1 Br/21	11 (6–18)	8 (5–13)	18 (6–35)	22 (12–34)
HPeV1 Br/27				
HPeV1 Br/145				
HPeV1 K54	30 (22–40)	29 (19–51)	27 (18–38)	34 (24–46)
HPeV1 K63/94				
HPeV1 K129				
HPeV1 K150				
HPeV6 NII561	13 (9–19)	20 (12–30)	17 (9–27)	17 (11–25)
HPeV6 823				
All HPeV3 except for HPeV3 651689*	19 (15–23)	20 (16–24)	21 (16–27)	22 (17–28)
All HPeV common ancestor	583 (315–905)	495 (287–751)	217 (109–353)	227 (140–328)

*HPeV3 651689 was excluded because it was recombinant in the non-structural part of the genome.

assays targeting the genomic VP3 and 3D regions (Calvert *et al.*, 2010; Harvala *et al.*, 2008) were used to further extend the available sequence information.

Since amplification of these PCR fragments was not equally successful for all strains, primers for amplification of several additional fragments of approximately 200 bp along the HPeV genome were then designed (sequences available upon request). Following nucleotide sequencing of all PCR products, specific primers were designed for each virus, cDNA was produced by using the Superscript III kit (Invitrogen), and the whole genome was amplified from clinical material in overlapping approximately 2–4 kb amplicons by using the Expand High Fidelity Plus kit (Roche). These PCR products were sequenced directly on both strands by primer walking. The 3′ terminus was amplified using the GeneRACEr kit (Invitrogen). The full genomic sequence could be identified for three strains. The 5′-terminal 450 nt of five other strains could not be amplified despite repeated trials and were abandoned due to limited amounts of clinical material.

Genetic analyses. All complete or nearly complete parechovirus sequences available in GenBank (Table 1) were aligned using CLUSTAL W (Thompson *et al.*, 1994). Alignments of coding genome regions based on translated sequences were joined to 5′-NTR sequences aligned separately. 3′-NTR sequences were omitted as this short and conserved region carries little phylogenetic information, and because 3′-NTR sequences were absent for a number of sequences from GenBank. Sequence handling was performed with BIOEDIT v.7.0.5.2 software (Hall, 1999). Phylogenies were calculated using a Bayesian likelihood-based algorithm implemented in BEAST 1.5.3 (Drummond & Rambaut, 2007). The SRD06 codon-based substitution model (Tu *et al.*, 1995) was used with a relaxed uncorrelated log-normal clock because it performed best in preliminary runs. A GTR nucleotide substitution model with a site heterogeneity model assuming gamma-distributed substitution rates across sites and allowing for invariable sites was used for 5′-NTR analysis. Each analysis was run over 100 000 000 generations and trees were sampled every 250 generations, resulting in 400 000 final trees. Trees were annotated with TREEANNOTATOR v.1.4.8 with a burn-in of 80 000 trees and visualized with FIGTREE v.1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/). Treeorder scans were done with SIMMONICS 1.6 (Simmonds & Smith, 1999).

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft grant no. DR772/2-1. Additional support was obtained from European Union projects RiViGene, EVA and EMPERIE. We are grateful to C. Pedroso, C. Brites and E. M. Netto for support. The study was partially funded by a research grant from the Foundation for research support of the State of Bahia (FAPESB), project code APR 125/2006.

REFERENCES

Abed, Y. & Boivin, G. (2005). Molecular characterization of a Canadian human parechovirus (HPeV)-3 isolate and its relationship to other HPeVs. *J Med Virol* **77**, 566–570.

Al-Sunaidi, M., Williams, C. H., Hughes, P. J., Schnurr, D. P. & Stanway, G. (2007). Analysis of a new human parechovirus allows the definition of parechovirus types and the identification of RNA structural domains. *J Virol* **81**, 1013–1021.

Baumgarte, S., de Souza Luna, L. K., Grywna, K., Panning, M., Drexler, J. F., Karsten, C., Huppertz, H. I. & Drosten, C. (2008). Prevalence, types, and RNA concentrations of human parechoviruses, including a sixth parechovirus type, in stool samples from patients with acute enteritis. *J Clin Microbiol* **46**, 242–248.

Benschop, K. S., Schinkel, J., Minnaar, R. P., Pajkrt, D., Spanjerberg, L., Kraakman, H. C., Berkhout, B., Zaaijer, H. L., Beld, M. G. & Wolthers, K. C. (2006). Human parechovirus infections in Dutch children and the association between serotype and disease severity. *Clin Infect Dis* **42**, 204–210.

Benschop, K., Thomas, X., Serpenti, C., Molenkamp, R. & Wolthers, K. (2008a). High prevalence of human parechovirus (HPeV) genotypes in the Amsterdam region and identification of specific HPeV variants by direct genotyping of stool samples. *J Clin Microbiol* **46**, 3965–3970.

Benschop, K. S., Williams, C. H., Wolthers, K. C., Stanway, G. & Simmonds, P. (2008b). Widespread recombination within human parechoviruses: analysis of temporal dynamics and constraints. *J Gen Virol* **89**, 1030–1035.

Benschop, K. S., de Vries, M., Minnaar, R. P., Stanway, G., van der Hoek, L., Wolthers, K. C. & Simmonds, P. (2010). Comprehensive full-length sequence analyses of human parechoviruses: diversity and recombination. *J Gen Virol* **91**, 145–154.

Calvert, J., Chieochansin, T., Benschop, K. S., McWilliam Leitch, E. C., Drexler, J. F., Grywna, K., da Costa Ribeiro, H., Jr, Drosten, C., Harvala, H. & other authors (2010). Recombination dynamics of human parechoviruses: investigation of type-specific differences in frequency and epidemiological correlates. *J Gen Virol* **91**, 1229–1238.

de Souza Luna, L. K., Baumgarte, S., Grywna, K., Panning, M., Drexler, J. F. & Drosten, C. (2008). Identification of a contemporary human parechovirus type 1 by VIDISCA and characterisation of its full genome. *Virol J* **5**, 26.

de Vries, M., Pyrc, K., Berkhout, R., Vermeulen-Oost, W., Dijkman, R., Jebbink, M. F., Bruisten, S., Berkhout, B. & van der Hoek, L. (2008). Human parechovirus type 1, 3, 4, 5, and 6 detection in picornavirus cultures. *J Clin Microbiol* **46**, 759–762.

Drexler, J. F., Luna, L. K., Stöcker, A., Almeida, P. S., Ribeiro, T. C., Petersen, N., Herzog, P., Pedroso, C., Huppertz, H. I. & other authors (2008). Circulation of 3 lineages of a novel Saffold cardiovirus in humans. *Emerg Infect Dis* **14**, 1398–1405.

Drexler, J. F., Grywna, K., Stöcker, A., Almeida, P. S., Medrado-Ribeiro, T. C., Eschbach-Bludau, M., Petersen, N., da Costa-Ribeiro-Jr, H. & Drosten, C. (2009). Novel human parechovirus from Brazil. *Emerg Infect Dis* **15**, 310–313.

Drummond, A. J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214.

Faria, N. R., de Vries, M., van Hemert, F. J., Benschop, K. & van der Hoek, L. (2009). Rooting human parechovirus evolution in time. *BMC Evol Biol* **9**, 164.

Gercel, C., Mahan, K. B. & Hamparian, V. V. (1985). Preliminary characterization of a persistent infection of HeLa cells with human rhinovirus type 2. *J Gen Virol* **66**, 131–139.

Ghazi, F., Hughes, P. J., Hyypiä, T. & Stanway, G. (1998). Molecular analysis of human parechovirus type 2 (formerly echovirus 23). *J Gen Virol* **79**, 2641–2650.

Hall, T. A. (1999). BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**, 95–98.

Harvala, H., Robertson, I., McWilliam Leitch, E. C., Benschop, K., Wolthers, K. C., Templeton, K. & Simmonds, P. (2008). Epidemiology and clinical associations of human parechovirus respiratory infections. *J Clin Microbiol* **46**, 3446–3453.

Hyypiä, T., Horsnell, C., Maaronen, M., Khan, M., Kalkkinen, N., Auvinen, P., Kinnunen, L. & Stanway, G. (1992). A distinct picornavirus group identified by sequence analysis. *Proc Natl Acad Sci U S A* **89**, 8847–8851.

Ito, M., Yamashita, T., Tsuzuki, H., Takeda, N. & Sakae, K. (2004). Isolation and identification of a novel human parechovirus. *J Gen Virol* **85**, 391–398.

Johnston, R. E., Wan, K. & Bose, H. R. (1974). Homologous interference induced by Sindbis virus. *J Virol* **14**, 1076–1082.

Joki-Korpela, P. & Hyypiä, T. (1998). Diagnosis and epidemiology of echovirus 22 infections. *Clin Infect Dis* **27**, 129–136.

Kim Pham, N. T., Trinh, Q. D., Takanashi, S., Abeysekera, C., Abeygunawardene, A., Shimizu, H., Khamrin, P., Okitsu, S., Mizuguchi, M. & Ushijima, H. (2010). Novel human parechovirus, Sri Lanka. *Emerg Infect Dis* **16**, 130–132.

Leitch, E. C., Harvala, H., Robertson, I., Ubillos, I., Templeton, K. & Simmonds, P. (2009). Direct identification of human enterovirus serotypes in cerebrospinal fluid by amplification and sequencing of the VP1 region. *J Clin Virol* **44**, 119–124.

Li, L., Victoria, J., Kapoor, A., Naeem, A., Shaukat, S., Sharif, S., Alam, M. M., Angez, M., Zaidi, S. Z. & Delwart, E. (2009). Genomic characterization of novel human parechovirus type. *Emerg Infect Dis* **15**, 288–291.

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W. & Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* **73**, 152–160.

Oberste, M. S., Maher, K. & Pallansch, M. A. (1998). Molecular phylogeny of all human enterovirus serotypes based on comparison of sequences at the 5′ end of the region encoding VP2. *Virus Res* **58**, 35–43.

Polacino, P., Kaplan, G. & Palma, E. L. (1985). Homologous interference by a foot-and-mouth disease virus strain attenuated for cattle. *Arch Virol* **86**, 291–301.

Salminen, M. O., Carr, J. K., Burke, D. S. & McCutchan, F. E. (1995). Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res Hum Retroviruses* **11**, 1423–1425.

Simmonds, P. & Smith, D. B. (1999). Structural constraints on RNA virus evolution. *J Virol* **73**, 5787–5794.

Simmonds, P. & Welch, J. (2006). Frequency and dynamics of recombination within different species of human enteroviruses. *J Virol* **80**, 483–493.

Stanway, G., Kalkkinen, N., Roivainen, M., Ghazi, F., Khan, M., Smyth, M., Meurman, O. & Hyypiä, T. (1994). Molecular and biological characteristics of echovirus 22, a representative of a new picornavirus group. *J Virol* **68**, 8232–8238.

Takao, S., Shimazu, Y., Fukuda, S., Noda, M. & Miyazaki, K. (2001). Seroepidemiological study of human parechovirus 1. *Jpn J Infect Dis* **54**, 85–87.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Tu, Z., Chapman, N. M., Hufnagel, G., Tracy, S., Romero, J. R., Barry, W. H., Zhao, L., Currey, K. & Shapiro, B. (1995). The cardiovirulent phenotype of coxsackievirus B3 is determined at a single site in the genomic 5′ nontranslated region. *J Virol* **69**, 4607–4618.

Wakatsuki, K., Kawamoto, D., Hiwaki, H., Watanabe, K. & Yoshida, H. (2008). Identification and characterization of two strains of human parechovirus 4 isolated from two clinical cases in Fukuoka City, Japan. *J Clin Microbiol* **46**, 3144–3146.

Watanabe, K., Oie, M., Higuchi, M., Nishikawa, M. & Fujii, M. (2007). Isolation and characterization of novel human parechovirus from clinical samples. *Emerg Infect Dis* **13**, 889–895.

Williams, C. H., Panayiotou, M., Girling, G. D., Peard, C. I., Oikarinen, S., Hyöty, H. & Stanway, G. (2009). Evolution and conservation in human parechovirus genomes. *J Gen Virol* **90**, 1702–1712.

Wolthers, K. C., Benschop, K. S., Schinkel, J., Molenkamp, R., Bergevoet, R. M., Spijkerman, I. J., Kraakman, H. C. & Pajkrt, D. (2008). Human parechoviruses as an important viral cause of sepsislike illness and meningitis in young children. *Clin Infect Dis* **47**, 358–363.

Zoll, J., Galama, J. M. & van Kuppeveld, F. J. (2009). Identification of potential recombination breakpoints in human parechoviruses. *J Virol* **83**, 3379–3383.