Sequence analysis of human rhinoviruses in the RNA-dependent RNA polymerase coding region reveals large within-species variation

Carita Savolainen, Pia Laine, Mick N. Mulders and Tapani Hovi

Enterovirus Laboratory, Department of Microbiology, National Public Health Institute (KTL), Mannerheimintie 166, FIN-00300 Helsinki, Finland

Human rhinoviruses (HRVs; family Picornaviridae), the most frequent causative agents of respiratory infections, comprise more than 100 distinct serotypes. According to previous phylogenetic analysis of the VP4/VP2-coding sequences, all but one of the HRV prototype strains distribute between the two established species, Human rhinovirus A (HRV-A) and Human rhinovirus B (HRV-B). Here, partial sequences of the RNA-dependent RNA polymerase (3D polymerase)-coding gene of 48 HRV prototype strains and 12 field isolates were analysed. The designated division of the HRV strains into the species HRV-A and HRV-B was also seen in the 3D-coding region. Phylogenetically, HRV-B clustered closer to human enterovirus (HEV) species HEV-B, HEV-C and poliovirus than to HRV-A. Intraspecies variation within both HRV-A and HRV-B was greater in the 3D-coding region than in the VP4/VP2-coding region, with the difference maxima reaching 48 % at the nucleotide level and 36 % at the amino acid level in HRV-A and 53 and 35 %, respectively, in HRV-B. Within both species, a few strains formed a separate cluster differing from the majority of strains as much as HEV-B from HEV-C. Furthermore, the tree topology within HRV-A differed from that for VP4/VP2, suggesting possible recombination events in the evolutionary history of the strains. However, all 12 field isolates clustered similarly, as in the capsid region. These results showed that the within-species variation in the 3D region is greater in HRV than in HEV. Furthermore, HRV variation in the 3D region exceeds that in the capsid-coding region.

INTRODUCTION

Human rhinoviruses (HRVs) are the most common agents in mild upper respiratory tract infections. Besides the common cold, they are associated with more severe illnesses, such as acute otitis media, pneumonia and exacerbations of asthma (Pitkäranta & Hayden, 1998). HRVs consist of more than 100 different serotypes. Taxonomically, they form the genus Rhinovirus in the family Picornaviridae and share morphological similarities with other members of the family. HRVs are small, spherical particles with a diameter of 25–30 nm. The virus capsid is non-enveloped and icosahedral in symmetry, consisting of 60 copies of each of the four capsid proteins (VP1–VP4). HRVs have a single-stranded positive-sense RNA genome of approximately 7.2 kb, encoding a single open reading frame. Within the family Picornaviridae, HRVs are genetically most closely related to human enteroviruses (HEVs).

Partial genomic sequencing has been used in phylogenetic studies of HEV strains. The capsid-coding region, and especially the VP1 gene, is known to correlate with serotype (Oberste et al., 1999a, b). Pulli et al. (1995) studied, some time ago, sequence variation among coxsackie A viruses (CV-As) in the genomic region encoding one of the non-structural proteins, the RNA-dependent RNA polymerase (3D). In their phylogenetic analysis, the 3D sequences studied clustered into three separate groups corresponding to those observed in the VP4/VP2 region. However, within the individual clusters, now defined as the HEV species (King et al., 2000), differences were found in the topology of the phylogenetic trees. Furthermore, in more recent studies, a number of field strains belonging to different HEV serotypes did not show corresponding clustering patterns between VP1 and 3D, suggesting past recombination between co-circulating heterotypic strains within a given HEV species (Oprisan et al., 2002; Lindberg et al., 2003; Brown et al., 2003).
We have recently reported genetic clustering of all known HRV serotypes in the VP4/VP2 region (Savolainen et al., 2002a). All HRV prototype strains were found to distribute between the two previously known major clades (Horsnell et al., 1995), except for one outlier, HRV87. This strain was further characterized in additional genomic regions and was found to be a member of the HEV-D species (Blomqvist et al., 2002). In addition, HRV field isolates were studied in the VP4/VP2 region, giving the same clustering into the major clades (Savolainen et al., 2002b). In order to know more about basic evolutionary relationships of the heterogeneous group of HRVs, we determined partial sequences of the 3D gene from a selection of HRV serotypes and compared the results of phylogenetic analysis with those for the VP4/VP2 region.

METHODS

HRV strains. HRV prototype strains were obtained from the ATCC, the Haartman Institute, Helsinki University, Janssen Pharmaceutica N.V. and the National Institute for Public Health and the Environment, The Netherlands. The field strains had been collected between 1994 and 1996 in the Tampere region, Finland, as part of the Finnish Otitis Media studies. The isolation procedure, epidemiological context and VP4/VP2 sequences of the strains have been described before (Blomqvist et al., 1999; Vesa et al., 2001; Savolainen et al., 2002b). All HRV strains were processed as described previously (Savolainen et al., 2002a).

RNA isolation and RT-PCR. Viral RNA was isolated from 100 µl tissue culture supernatants using the RNeasy Total RNA kit (Qiagen). Upon purification, RNA was eluted with 30 µl Rnase-free water and subsequently stored at −70 °C. RT-PCR using a forward primer (5′-GGIGGIRTNCMTCNGGNTG-3′; nt 6559–6578) and spanning the 3′ part of the 3D gene and the complete 3′ non-coding region. RT-PCR was carried out as described previously (Mulders et al., 2000) for 28 of the HRV prototype strains. The rest of the prototype strains and the field isolates were amplified under conditions described by Oberste et al. (1999a).

Sequence analysis. PCR products were either purified directly using a PCR purification kit (QiAquick; Qiagen) or excised from the gel and purified with a gel extraction kit (QiAquick) when multiple bands were obtained. The purified products were stored at −20 °C or used directly in cycle sequencing (BigDye Terminator cycle sequencing ready reaction kit versions 3.0, 1.0 and 1.1; Applied Biosystems) using the same RT-PCR primers. Sequencing was performed with an automated sequencer (model 377 or 310; ABI). Sequence data were analysed with Sequencing Analysis (version 3.1; ABI) and Sequence Navigator (version 1; ABI). Multiple sequence alignments were made using CLUSTAL X (Thompson et al., 1997). The 3D fragment (468 bp or 156 aa) of the obtained sequence was used for analysis. However, for prototype strains HRV55 and HRV100, as well as field isolates 7821oct95 and 7850sep96, the analysable fragment was shorter. Distance matrices were estimated with the DNADIST and PROTDIST programs (PHYLIP version 3.6) using the Kimura two-parameter method. For the truncated sequences, the relative gaps were filled with N. Dendrograms were drawn using the gap-striped neighbour-joining option in CLUSTAL X and were visualized with NJPLOT. Bootstrap analysis was performed with 1000 replicates. Sequences generated in this study were submitted to GenBank under accession numbers AY436646–AY436704 (details available in a Supplementary Table in JGV Online). Previously published sequences in the VP4/VP2 region of the same HRV strains are also presented in the supplementary table. HEV sequences used in the comparison had GenBank accession numbers V01150, X00595, X00925, AF499635–AF499643, D00538, D90457, AF311938, AF311939 and AF524866. In addition, GenBank sequences of Sinian picornavirus 1 (NC004445) and Porcine enterovirus B (NC004441) were included as outgroups.

RESULTS

A total of 48 of 101 known HRV prototype strains as well as 12 field isolates were partially sequenced in the 3D-coding region. The sequences of HRV1B, HRV14 and HRV89 were identical to those of the complete genome sequences in GenBank. The GenBank sequences of HRV2 (X02316) and HRV16 (NC001752) were added to the analysis.

Phylogenetic analysis

Based on the analysed 3D gene sequences, the HRV prototype strains could be divided in two distinct clusters representing the designated two HRV species, HRV-A and HRV-B (Fig. 1). The two clusters included the same members as the previously published VP4/VP2 region analysis (Savolainen et al., 2002a). The length of the partial 3D sequence obtained using the universal HRV primer pair was 468 nt for most of the studied strains. For the strains where only shorter sequences were able to be analysed, each of the missing nucleotides was replaced with N. Excluding the shorter sequences from the phylogenetic analysis did not change the observed clustering of the remaining strains (data not shown).

HRV serotypes 8, 45 and 95 and one field isolate (6155may96) formed a more distinct group in the 3D region (Fig. 1; designated HRV-A’). Similarly, HRV26 and HRV99 grouped together and differed strikingly from the rest of the HRV-B strains. In general, compared with enteroviruses, the intraspecies differences in the HRV 3D region were relatively closer to the interspecies differences. In addition, the tree topology and branching order within HRV-A was different from that in the VP4/VP2 region (Fig. 2). For instance, the locations of HRV1A, HRV1B, HRV16, HRV56, HRV59, HRV60, HRV66, HRV68 and HRV77 in the 3D region were different from those in the capsid region. HRV1A and HRV1B clustered together with HRV18 and HRV50 in the capsid region, while this clustering was not seen in the 3D region. HRV16, HRV56 and HRV60 formed a cluster in the capsid region, unlike in the 3D-coding region. In the VP4/VP2 region, HRV68 did not cluster closely with any of the other HRV-A but, in the 3D region, it was very close to HRV28. HRV59 and HRV63 clustered together in the capsid region, but not in the 3D region. HRV66 and HRV77 were grouped together in the capsid region, but were not the closest neighbours in the 3D-region analysis. However, there were no good bootstrapping values to support the apparent discrepancies for the two latter pairs of strains.
The maximum variation assessments were based on distance matrices. The maximum evolutionary distance between HRV-A and HRV-B at the amino acid level was 83.4 (the corresponding maximum direct difference in the alignment was 50%). The evolutionary distances between the two species were not determined at the nucleotide level because

**Fig. 1.** Neighbour-joining dendrogram showing clustering of HRV prototype strain and field isolate sequences in the 3D-coding region. The alignment was based on approximately 468 determined nucleotides. Numbers indicate bootstrap values for each branch calculated with 1000 replicates. In addition to the sequenced HRV, previously published HEV-C and poliovirus sequences and a selection of HEV-B sequences are included. As outliers, *Simian picornavirus 1* and *Porcine enterovirus B* were added to the analysis. For GenBank accession numbers, see Supplementary Table.
of the large variation. Within HRV-A, the distance maxima were 47.7 (35%) at the nucleotide level and 36.6 (31%) at the amino acid level. For HRV-B, the maxima were similar, with 53.4 (38%) at the nucleotide level and 35.0 (31%) at the amino acid level. The distinct group of HRV-A' was not included in the above 'within HRV-A' analysis. The maximum difference of this group compared with HRV-A was 55.5% and with HRV-B was 73.6% at the amino acid level.

**Fig. 2.** Clustering of the studied HRV prototype strains and field isolates for the VP4/VP2-coding sequences (data from Savolainen et al., 2002a). HEV-C, poliovirus and a selection of HEV-B sequences are included as in Fig. 1. For details, see Fig. 1 legend.
**Amino acid sequences**

There were no apparent insertions or deletions among the sequenced strains within the analysed part of the 3D region. Conserved amino acid motifs shared by all analysed strains were infrequent; 27.6% of the amino acids were identical in all studied strains. Yet, the YGDD motif thought to be a common nucleic acid recognition site and/or an active processing region in various polymerases (Kamer & Argos, 1984) was seen in all sequences. Likewise, the YGL, FLKR and SIRWT motifs downstream of YGDD were shared by all HRV strains as well as by the studied enteroviruses. The presence of the FLKR motif has been identified in enteroviruses previously (Brown et al., 2003). The remaining fully conserved amino acid positions occurred as doublets or solitary amino acids throughout the analysed part of the 3D region. Other individual positions were totally conserved in one group of viruses or in the other (HRV-A or HRV-B) but were not totally conserved within viruses of the other group. Yet another set of positions were specific for HRV-A or for HRV-B, and did not occur at all in the other group (Fig. 3). HRV-B contained more species-specific amino acids than did HRV-A. The overall amino acid patterns were group-specific, as visible in the alignment of all sequenced strains (see Supplementary Figure in JGV Online).

**DISCUSSION**

We found in this study that the 3D sequences of the 50 HRV prototype strains studied clustered according to the two established species, HRV-A and HRV-B, as in the clustering observed previously in the VP4/VP2 sequences (Savolainen et al., 2002a). The 3D sequences of HRV-B appeared to be closer to HEV-B and HEV-C than to HRV-A. This observation supported the recent suggestion of reclassifying enteroviruses and rhinoviruses as a single genus (‘Enterhinovirus’) in the family Picornaviridae, discussed within the Picornavirus Study Group of the International Committee for the Taxonomy of Viruses (T. Hovi, personal observation).

Based on the 3D region, the prototype strains of HRV serotypes 8, 45 and 95, a designated HRV-A’ group, clustered together and differed from the rest of HRV-A as much as HRV-B from enteroviruses; they may thus represent a phylogenetically distinct group of viruses. The same serotypes formed a cluster in the VP4/VP2 region, but here the difference from the rest of HRV-A strains was relatively small. A similar situation was evident for HRV26 and HRV99 of HRV-B. More extensive analysis, particularly of full-length genomes, may be necessary to establish the phylogenetic positions of these groups of viruses. While this manuscript was in preparation, two nearly identical sets of complete VP1 gene sequences of all HRV prototypes were published in GenBank with accession numbers AY355180–AY355281 and AY450462–AY450549 plus AY458604–AY458611, respectively. As expected, the overall clustering patterns based on these datasets (not shown) were very similar to the patterns based on our previously published VP4/VP2 sequences (Savolainen et al., 2002a).

For enteroviruses, it has previously been found that the genetic distribution of individual strains into different species based on the capsid-coding region and the 3D-coding region correlate with each other (Pulli et al., 1995). Our observations on HRV are in accordance with this. Sequence variation, especially at the amino acid level, within a given enterovirus species in the 3D-coding region is lower than that in the capsid-coding region (Huttunen et al., 1996; Pöyry et al., 1996; Pulli et al., 1995). This has been considered to be due to the role of the polymerase as a functionally important protein in the life cycle of viruses. Functional conservation would thus restrict sequence variation. Interestingly, our current results show that, in contrast to enteroviruses, among the sequenced HRV prototype strains, the within-species variation in the 3D region is greater than that found in the VP4/VP2 region (Savolainen et al., 2002a). The within-species variation of 3D sequences between different HRV serotypes is also greater than that within a given HEV species (Fig. 1). This was also evident at the amino acid level (data not shown). Apparently, a smaller number of amino acid motifs is needed to preserve the functionality of the polymerase in HRV, allowing a higher degree of variation for the rest of the polypeptide.

Picornavirus evolution is driven by point mutations, recombination and selection. Based on the current analysis, it seems that within-species recombination events are also possible for rhinoviruses, similar to those observed for enteroviruses (Santti et al., 1999; Oprisan et al., 2002; Kew et al., 2002; Lindberg et al., 2003). Clustering topology differed between the VP4/VP2 region and the 3D region for some prototype strains of HRV-A, which might be due to recombination in the evolutionary history of the strains.

**Fig. 3.** Conservation of amino acid sequences of the analysed part of the 3D polymerase-coding protein. Conserved amino acids within HRV-A are marked with light grey in HRV1B and those within HRV-B are marked with dark grey in the HRV14 sequence. Positions conserved in all HRV sequences are boxed. Black indicates species-specific amino acids.
Possibilities of recombination events in rhinovirus biology are obvious, since many different types of rhinovirus co-circulate simultaneously (Savolainen et al., 2002b) and could also appear in one person at a given time. However, all 12 field strains of HRV sequenced in this study clustered closest to the same prototype strain in both sequenced regions. Nor was any sequence distribution incongruence, suggesting recombination in the past, observed within the studied HRV-B prototype strains. This could be due to the small number of samples analysed; however, it is also possible that recombination events occur less frequently in HRV infections than in HEV infections because of differences in the natural course of infection and its duration. Enteroviral infection has been documented to be able to last for months, increasing the likelihood of a second enterovirus appearing in the same patient and thus providing possibilities for recombinatory events (as can be seen during oral poliovirus vaccinations, where a large proportion of progeny virus are recombinants; Alexander et al., 1997).

Rhinoviruses, on the other hand, usually have a much shorter duration of infection and often disappear within 2 weeks (Arruda et al., 1997). The short duration of rhinoviral infection may be a limiting factor for the recombinant progeny to be selected and become prevalent for transmission. In addition, the observed relatively greater within-species variation of the 3D sequences may mean that the interserotypic compatibility of different genomic domains is less than that within a given HEV species, decreasing the likelihood of viable interserotypic recombinant offspring.

In conclusion, phylogenetic analysis of partial 3D sequences of a wide range of HRV strains has suggested that, unlike the enterovirus species, within-species variation is relatively large. Furthermore, again unlike enteroviruses, the relative within-species differences of HRV 3D-coding sequences were greater than those in the VP4/VP2-coding region of the genome. Serotypes 8, 45 and 95 may represent a phylogenetically distinct subgroup of HRV-A, but further studies are needed to confirm this hypothesis. Similarly, the relative outlier positions of HRV26 and HRV99 within the HRV-B species call for a deeper analysis.

ACKNOWLEDGEMENTS

Päivi Hirttö is thanked for expert technical assistance and Solle Blomqvist for fruitful discussions. The authors wish to acknowledge Mika Salminen for valuable advice. This study was partially supported by grants from the Academy of Finland, Fääriviki and Sakari Sohlberg Foundation and Wyeth Vaccines and Pediatrics Ltd, Kansas, USA.

REFERENCES


