Complete genome sequences for nine simian enteroviruses

M. Steven Oberste, Kaija Maher and Mark A. Pallansch

Polio and Picornavirus Laboratory Branch, Division of Viral Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA

Analysis of the VP1 capsid-coding sequences of the simian picornaviruses has suggested that baboon enterovirus (BaEV), SV19, SV43 and SV46 belong to the species Human enterovirus A (HEV-A) and SA5 belongs to HEV-B, whereas SV4/A2 plaque virus (two isolates of a single serotype), SV6 and N125/N203 (two isolates of a single serotype) appear to represent new species in the genus. We have further characterized by complete genomic sequencing the genetic relationships among the simian enteroviruses serotypes (BaEV, N125/N203, SA5, SV4/A2 plaque virus, SV6, SV19, SV43 and SV46) and to other enteroviruses. Phylogenetic and pairwise sequence relationships for the P1 region paralleled those of VP1 alone, and confirmed that SV4/A-2 plaque virus, SV6 and N125/N203 represent unique genetic clusters that probably correspond to three new species. However, sequence relationships in the P2 and P3 regions were quite different. In 2C, SV19, SV43 and SV46 remain clustered with the human viruses of HEV-A, but BaEV, SV6 and N125/N203 cluster together; in 3CD, SA5 (HEV-B) also joined this cluster. The 3' non-translated region (NTR) sequences are highly conserved within each of the four human enterovirus species, but the 3' NTRs of the simian enteroviruses are distinct from those of all human enteroviruses and generally distinct from one another. These results suggest that host species may have a significant influence on the evolution of enterovirus non-capsid sequences.

INTRODUCTION

The simian picornaviruses (family Picornaviridae) were isolated from various primate tissues during the development of general tissue culture methods in the 1950s and 1960s or from specimens derived from primates used in biomedical research (Fuentes-Marin et al., 1963; Hoffert et al., 1958; Kalter, 1982; Kalter et al., 1967; Rodriguez et al., 1977). Most were shown to be members of the genus Enterovirus (Oberste et al., 2002; Pöyry et al., 1999), but SV2, SV16, SV18, SV42, SV44, SV45 and SV49 were found to be related to one another (Oberste et al., 2003), to porcine enterovirus 8 (Krumbholz et al., 2002) and to a duck picornavirus, but distinct from all other picornavirus genera (Oberste et al., 2003). Analysis of the complete genome sequences of SV2 (Oberste et al., 2003) and PEV8 (Krumbholz et al., 2002) has resulted in a proposal to place them in a new genus, Sapelovirus (Knowles, 2006). Analysis of VP1 capsid-coding sequences has shown that SV19, SV26 and SV35 comprise a single virus type and that SV19 (SV19/SV26/SV35), SV43, SV46 and baboon enterovirus (BaEV) are members of the species Human enterovirus A (HEV-A), while SA5 is a member of HEV-B (Oberste et al., 2002). SA4, SV4, SV28 and A-2 plaque virus (A2plaque) comprise a single type, but they are distinct from other enteroviruses; as a result, A2 plaque has been assigned as the type strain of the new species Simian Enterovirus A (SEV-A) and the type has been named SEV-A1 (Knowles, 2006; Stanway et al., 2005). N125 and N203 comprise a single type that is somewhat related to SV6; together, N125/N203 and SV6 appear to comprise at least one additional species within the genus.

The genus Enterovirus is composed of more than 100 serotypes, most of which are known human pathogens (Pallansch & Roos, 2001). Most enterovirus infections are asymptomatic or result in only mild illness, such as non-specific febrile illness or mild upper respiratory symptoms (common cold). However, enteroviruses can also cause a wide variety of clinical illnesses, including acute haemorrhagic conjunctivitis, aseptic meningitis, undifferentiated rash, acute flaccid paralysis, myocarditis and neonatal sepsis-like disease (Pallansch & Roos, 2001). A number of simian enterovirus strains have been isolated from...
monkeys with diarrhoea, but there is little compelling evidence for a positive association with disease in primates (Kalter, 1960).

We have determined the complete genome sequences for BaEV, SA5, SV4, SV6, SV19, SV43, SV46, N125 and N203 and compared the sequences to one another and to those of enterovirus reference strains. The analysis illustrates the complex relationships between the simian and human enteroviruses.

METHODS

Viruses. The prototype strains of seven simian enteroviruses were obtained from American Type Culture Collection (Table 1). Two additional strains, N125 and N203, were kindly provided by Drs S. Kalter and R. Heberling, Esoterix (Table 1). Viral RNA was extracted from cell culture supernatant using a QIAamp Viral RNA kit (Qiagen) either directly from the original material or following one passage in LLC-MK2 cells (ATCC CCL 7), as described previously (Oberste et al., 2002).

Molecular characterization of isolates. To facilitate complete genome sequencing, RT-PCR primers were designed to anneal to sites encoding amino acid motifs that are highly conserved among enteroviruses (Brown et al., 2003; Oberste et al., 2004a, b, c, d). Specific, non-degenerate primers were designed from preliminary sequences to close gaps between the original PCR products. Sequences of the genome termini were determined by random amplification of cDNA ends (RACE), as described previously (Oberste et al., 2004a). For all sequence determinations, the PCR products were purified for sequencing by using a High-Pure PCR product purification kit (Roche Molecular Biochemicals), and both strands were sequenced by sequencing by using a High-Pure PCR product purification kit (Roche Molecular Biochemicals), and both strands were sequenced by automated methods, using fluorescent dye-terminator chain terminators (Applied Biosystems).

Sequence analysis. Analyses included comparison to the complete genome sequences of reference strains of each of the known human, bovine and porcine enterovirus serotypes, as well as to the six available human rhinovirus complete genome sequences (total of 107 sequences) (Supplementary Table S1, available with the online version of this paper). The nucleotide and deduced amino acid sequences of the simian enteroviruses were compared to one another and to those of other enteroviruses by using the programs Gap and Distances (Wisconsin Sequence Analysis Package, version 10.3, Accelrys). Alignments of nucleotide sequences and deduced amino acid sequences were generated using the PILEUP program (Wisconsin Package). Phylogenetic relationships in the 5'-non-translated region (NTR) were inferred from the aligned nucleotide sequences by the neighbour-joining method implemented in MEGA, version 3.1 (Kumar et al., 2001), using the Kimura two-parameter substitution model (Kimura, 1980), with a transition-transversion ratio of 10. Phylogenetic relationships in the coding region were inferred from the aligned amino acid sequences by the neighbour-joining method implemented in MEGA, using the JTT and Poisson distribution substitution models (Jones et al., 1992; Kumar et al., 2001). For all phylogenetic reconstructions, regions containing alignment gaps were excluded and support for specific tree topologies was estimated by bootstrap analysis with 1000 pseudo-replicate datasets.

Similarity plots depicting the relationships among the aligned polyprotein sequences were generated using SimPlot, version 3.2 beta (Lole et al., 1999). Similarity was calculated in each window of 200 amino acids by the Jukes and Cantor distance method (Jukes & Cantor, 1969). The window was successively advanced along the genome alignment in 20-residue increments.

RESULTS

5'-NTR

The simian enterovirus 5'-NTRs vary in length from 666 (BaEV) to 744 nt (SV4), with most in the 722–744 nt range; those of the human enteroviruses are 711–755 nt long (Brown et al., 2003; Oberste et al., 2004a, b, d, 2006), while the BEV 5'-NTRs are 812–820 nt (Earle et al., 1988; Zell & Stelzner, 1997) and those of the porcine enteroviruses are 809–814 nt (Krumbohlolz et al., 2002).

The 5'-NTR sequences of the simian enteroviruses were aligned with those of other enteroviruses. Phylogenetic trees were constructed by the neighbour-joining method, with 1000 bootstrap replicates (Fig. 1). The simian enterovirus 5'-NTR sequences formed several phylogenetic clusters, all of which were distinct from the previously known clusters of human, bovine and porcine enterovirus sequences. SV19, SV43 and SV46 cluster together in a

Table 1. Simian enteroviruses analysed in this study

The published sequence of A-2 plaque virus (GenBank AF201894), a strain of SV4, was also included in analyses.

<table>
<thead>
<tr>
<th>Virus</th>
<th>Strain</th>
<th>Species of origin*</th>
<th>Notes</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>BaEV</td>
<td>A13</td>
<td>Papio anubis</td>
<td></td>
<td>AF326750</td>
</tr>
<tr>
<td>SA5</td>
<td>B165</td>
<td>Cercoipithecus aethiops</td>
<td></td>
<td>AF326751</td>
</tr>
<tr>
<td>SV4</td>
<td>1715 UWB</td>
<td>Macaca mulatta</td>
<td>Same serotype as SA4 and SV28</td>
<td>AF326759</td>
</tr>
<tr>
<td>SV6</td>
<td>1631</td>
<td>Macaca mulatta</td>
<td></td>
<td>AF326766</td>
</tr>
<tr>
<td>SV19</td>
<td>M19s(P2)</td>
<td>Macaca fascicularis</td>
<td>Same serotype as SV26 and SV35</td>
<td>AF326754</td>
</tr>
<tr>
<td>SV43</td>
<td>OM112t(P12)</td>
<td>Macaca fascicularis</td>
<td></td>
<td>AF326761</td>
</tr>
<tr>
<td>SV46</td>
<td>OM22(P15)</td>
<td>Macaca sp.</td>
<td></td>
<td>AF326764</td>
</tr>
<tr>
<td>N125</td>
<td>N125</td>
<td>Papio cynocephalus</td>
<td>Same serotype as N203</td>
<td>AF414372</td>
</tr>
<tr>
<td>N203</td>
<td>N203</td>
<td>Papio cynocephalus</td>
<td>Same serotype as N125</td>
<td>AF414373</td>
</tr>
</tbody>
</table>

*Macaca mulatta, rhesus monkey; Macaca fascicularis, cynomolgus monkey; Cercoipithecus aethiops, African green monkey (Vervet); Papio anubis (or doguera), baboon; Papio cynocephalus, baboon.
Fig. 1. Phylogenetic relationships among enteroviruses based on alignment of 5'-NTR sequences. An unrooted tree was reconstructed using the neighbour-joining algorithm implemented in MEGA 3.1, as described in Methods. The left and right panels represent the same tree, in radial and rectangular configuration, respectively. The scale bars indicate genetic distance (nucleotide differences per residue). *, sequence available at http://www.virology.wisc.edu/acp/Aligns/aligns/rhino.rna.
position between human enterovirus 5’-NTR cluster II and the PEV-B cluster (Fig. 1). N125, N203 and SV6 also cluster together, between the PEV-B cluster and HEV cluster I. SA5 and BaEV each branch independently, while SV4 and A2plaque cluster together between SA5 and the rhinoviruses. In all of the simian enterovirus 5’-NTR sequences, the cloverleaf and internal ribosome entry site structures are maintained as expected (data not shown).

### 3’-NTR

Enterovirus 3’-NTR sequences are well conserved within a species (70–99% identity), but highly divergent between species (<62% identity) (Brown et al., 2003; Oberste et al., 2004a, b, d, 2006). Predicted secondary and tertiary structures are also conserved within species, and the major structural features are largely conserved between species (Mirmomeni et al., 1997). The 3’-NTRs of members of HEV-A and HEV-B are predicted to form three stem–loops, termed X, Y and Z, with base-pairing between the loops of stem–loops X and Y, termed the ‘kissing interaction’ (Supplementary Figure S1, available with the online version of this paper). The 3’-NTRs of viruses in HEV-C, HEV-D, BEV and PEV-B have only the X and Y stem–loops and maintain the X–Y kissing interaction (Supplementary Figure S1). The rhinovirus 3’-NTRs form only a single stem–loop, with no kissing interaction (Supplementary Figure S1). The 3’-NTR sequences of the simian enteroviruses are less than 67% identical to one another, with the exception of SA5, SV6, N125 and N203, which are 74–83% identical; the simian virus 3’-NTRs are less than 68% identical to those of other enteroviruses. However, all of the simian 3’-NTRs are predicted to form structures that are similar to those of other enteroviruses (Supplementary Figure S1). The 3’-NTR structures for all of the simian enteroviruses, except SEV-A1 (SV4 and A2plaque), contain stem–loops X, Y and Z, whereas the SEV-A1 3’-NTRs contain only stem–loops X and Y (Supplementary Figure S1).

### Capsid (P1) region

Analysis of VP1 sequences had suggested that BaEV, SV19, SV43 and SV46 are members of HEV-A, and SA5 is a member of HEV-B, while N125/N203, SV4 and SV6 probably represent three new species in the genus Enterovirus (Oberste et al., 2002). To further characterize the capsid sequence relationships of the simian enteroviruses to other members of the genus, we compared the complete capsid sequences of the simian enteroviruses to those of reference strains of all other known enterovirus serotypes. For the 10 simian enterovirus strains analyzed, the length of the capsid-coding region varied from 2535 (BaEV) to 2580 nt (SV4) (845–860 aa); the range for other enteroviruses and rhinoviruses is 832–887 aa (Brown et al., 2003; Oberste et al., 2004a, b, d, 2006). Phylogenetic and pairwise sequence comparisons revealed relationships that paralleled those of VP1 alone, and confirmed that SV4/A2plaque, SV6 and N125/N203 each represent unique genetic clusters that probably correspond to new species (Fig. 2 and Supplementary Table S2). As described above, A2plaque has already been established as the type strain for the species SEV-A.

The SA5 P1 amino acid sequence is 68–76% identical to those of the human viruses in HEV-B (which are 68–87% identical to one another), but less than 54% identical to those of other simian enteroviruses and less than 60% identical to those of viruses in all other enterovirus species (Supplementary Table S2, available with the online version of this paper); the human viruses in HEV-B are 44–59% identical in P1 amino acid sequence to human strains of other species. SA5 clusters with HEV-B (100% bootstrap support), branching from near the HEV-B root node. BaEV, SV19, SV43 and SV46 are 76–77% identical to one another in P1 amino acid sequence, except that SV19 and SV43 are somewhat more closely related to one another (~87%), and 66–79% identical to members of HEV-A, but less than 59% identical to members of all other species (Supplementary Table S2). EV76, EV89, EV90 and EV91 are the most closely related human viruses to BaEV, SV19, SV43 and SV46 (75–79% identity) (Supplementary Table S2), and EV76, EV89, EV90 and EV91 are 56–68% identical to the other HEV-A strains (Oberste et al., 2005). In the P1 phylogenetic tree, BaEV, SV19, SV43, SV46, EV76, EV89, EV90 and EV91 cluster together within HEV-A (91% bootstrap value), with 100% bootstrap support for the overall species cluster (Fig. 2).

In contrast to BaEV, SV19, SV43, SV46 and SA5, the P1 amino acid sequences of N125, N203, SV4 and SV6 are less than 64% identical to those of all other enteroviruses (Supplementary Table S2). The N125 and N203 capsid sequences are 92% identical to one another, confirming that they represent a single type, while the SV4 and A2plaque sequences are 95% identical to each other (Supplementary Table S2). SV6 is 60–64% identical to N125, N203 and the simian members of HEV-A, 57–63% identical to the human viruses in HEV-A, and less than 58% identical to all other enteroviruses (Supplementary Table S2). SV4 and A2plaque cluster together tightly in the P1 phylogenetic tree, as expected, and are as distinct from members of other enterovirus species as those species are from one another (Fig. 2). Similarly, N125 and N203 cluster together and are also distinct from other species (Fig. 2). SV6 is approximately equidistant from N125/N203 and from the simian viruses in HEV-A, branching from a point near the root of the N125/N203 branch of the P1 tree (Fig. 2).

To better visualize the distribution of pairwise P1 sequence identities, the individual P1 amino acid identities summarized in Supplementary Table S2 were plotted versus the P1 nucleotide sequence identities for the same comparisons (Fig. 3). The plot can be divided into four major zones. From upper right to lower left, the first zone is represented by viruses of the same type, i.e. SV4 compared to A2plaque,
Fig. 2. Phylogenetic relationships among enteroviruses based on alignment of P1 amino acid sequences. An unrooted tree was reconstructed using the neighbour-joining algorithm implemented in MEGA 3.1, as described in Methods. The left and right panels represent the same tree, in radial and rectangular configuration, respectively. The scale bars indicate genetic distance (amino acid differences per residue). *, sequence available at http://www.virology.wisc.edu/acp/Aligns/aligns/rhino.rna.
and N125 compared to N203. Pairwise comparisons within other serotypes would also be in this zone (Brown et al., 2003; Oberste et al., 2005, 2004c, 2001). The next zone, with a single point for the comparison of SV19 and SV43, represents viruses of closely related serotypes (Fig. 3). For the human enteroviruses, the comparison of echoviruses 11 and 19 or coxsackieviruses A3 and A8 would also be classified in this zone (Oberste et al., 2004a, d). The third zone represents comparisons of different serotypes within the same species and can be resolved into two subregions, with the higher identity region (toward the upper right) composed of comparisons of BaEV, SV19, SV43 and SV46 to one another and to EV76, EV89, EV90 and EV91, as well as comparisons of SA5 to the six coxsackie B viruses (Fig. 3). The lower portion of the intra-species comparison zone is made up of comparisons of SA5 to HEV-B serotypes other than the coxsackie B viruses and comparisons of BaEV, SV19, SV43 and SV46 to HEV-A viruses other than EV76, EV89, EV90 and EV91. The fourth zone comprises comparisons to viruses of heterologous species and is comparable to other comparisons between enteroviruses of different species (Brown et al., 2003; Oberste et al., 2004a, d).

Non-capsid (P2 and P3) region

To further characterize the relationships among the simian enteroviruses and between the simian and other enteroviruses, the deduced protein sequences for the 2C and 3CD proteins were compared. These two regions, in particular, were chosen because they have been used, along with P1, in the molecular classification of picornaviruses (Stanway et al., 2005). As in P1 comparisons, the 2C analysis shows that SA5 is most closely related to members of HEV-B (85–88 % amino acid identity) and that SV19, SV43 and SV46 are closely related to one another (84–94 % identity) and to the human viruses in HEV-A (82–91 % identity, Supplementary Table S3, available with the online version of this paper). N125 and N203 are also closely related to one another in 2C, as are SV4 and A2plaque (both pairs about 98 % identical) (Supplementary Table S3). In contrast to the P1 results, in 2C comparisons BaEV is more closely related to viruses in HEV-B (72–75 %) than to those in HEV-A (66–67 %), and most closely related to SV6 (85 %) and N125/N203 (89–90 %), SV6, N125, N203 and SV4/A2plaque are also 71–76 % identical in 2C sequence to the human viruses of HEV-B, as well as to SA5 (Supplementary Table S3). In the 2C phylogeny, SV19, SV43 and SV46 remain clustered with the human viruses of HEV-A (99 % bootstrap support) and most closely related to EV76, EV89, EV90 and EV91, but BaEV clusters with SV6, N125 and N203, with 99 % bootstrap support (Fig. 4). As in the P1 tree, SA5 clusters with members of HEV-B, with 99 % bootstrap support, and SV4/A2plaque form a distinct cluster.

Most of the relationships deduced from analysis of 2C were mirrored in the 3CD comparisons. SV19, SV43 and SV46 are most closely related to one another (84–91 % identity) and to members of HEV-A (81–86 % identity) (Supplementary Table S4, available with the online version of the paper). As in 2C, BaEV, N125, N203 and SV6 are most closely related to one another in 3CD (88–94 % identity). However, SA5 3CD is most closely related to BaEV, N125, N203 and SV6 (83–84 %), rather than to members of HEV-B (80–82 %) (Supplementary Table S4). In the 3CD phylogenetic tree, SV19, SV43 and SV46 remain clustered with EV76, EV89, EV90 and EV91 within HEV-A (Fig. 5). Whereas SA5 clusters with HEV-B viruses in the 2C tree, with 99 % bootstrap support, in 3CD SA5 is part of the BaEV/SV6/N125/N203 cluster, with 89 % bootstrap support (Fig. 5). In the 3CD tree, SV4 and A2plaque remain in an independent cluster.

To better visualize the sequence relationships that are discordant in different genome regions, we compared simian enterovirus polyprotein sequences to those of other enteroviruses by plotting amino acid sequence identities in a sliding window using the program SimPlot (Fig. 6). As expected, the SV4 polyprotein sequence is highly similar to that of A2plaque throughout its length, but distinct from those of all other enteroviruses (Fig. 6a). The polyprotein of SV19 is most closely related to those of SV43, SV46 and the human viruses in HEV-A (Fig. 6b). This pattern is typical of comparisons among human enteroviruses of the same species, such as coxsackievirus A2 compared to other viruses in HEV-A (Oberste et al., 2004d). The SV19 and SV43 polyproteins are most closely related to one another in all regions of the polyprotein; interestingly, these two viruses are the only strains that were isolated from Macaca fascicularis (Hoffert et al., 1958). The SimPlot patterns for
Fig. 4. Phylogenetic relationships among enteroviruses based on alignment of 2C amino acid sequences. An unrooted tree was reconstructed using the neighbour-joining algorithm implemented in MEGA 3.1, as described in Methods. The left and right panels represent the same tree, in radial and rectangular configuration, respectively. The scale bars indicate genetic distance (amino acid differences per residue). *, sequence available at http://www.virology.wisc.edu/acp/Aligns/aligns/rhino.rna.
Fig. 5. Phylogenetic relationships among enteroviruses based on alignment of 3CD amino acid sequences. An unrooted tree was reconstructed using the neighbour-joining algorithm implemented in MEGA 3.1, as described in Methods. The left and right panels represent the same tree, in radial and rectangular configuration, respectively. The scale bars indicate genetic distance (amino acid differences per residue). *, sequence available at http://www.virology.wisc.edu/acp/Aligns/aligns/rhino.rna.
SV43 (Fig. 6c) and SV46 (Fig. 6d) were similar to that of SV19 (Fig. 6b). In contrast to the patterns for SV19, SV43 and SV46, the SimPlot analysis for BaEV showed a very different pattern. As predicted by the pairwise comparisons and phylogenetic trees, the BaEV polyprotein is related to those of HEV-A viruses and SV19, SV43 and SV46 in the capsid region, but unrelated to these viruses in P2 and P3 (Fig. 6e). Instead, BaEV is related to N125 and N203 and, to a lesser degree, to SV6, in 2A-2B, with the similarity to SV6 increasing to the level of N125 and N203 in 2C through 3D. SA5 is related to other members of HEV-B throughout the length of the polyprotein (Fig. 6f), but its similarity to BaEV, SV6, N125 and N203 increases beginning in 2A. In 3D, its similarity to BaEV, SV6, N125 and N203 exceeds its similarity to members of HEV-B. SV6 is not closely related to any other enterovirus in P1, but in P2 and P3, its SimPlot pattern resembles that of BaEV (Fig. 6g). As with SV4-A2plaque, N125 is closely related to N203 throughout the polyprotein (Fig. 6h) and, as described above, it is also related to BaEV and SV6 in P2 and P3, and to SA5 and HEV-B viruses in 3D. As expected, the SimPlot pattern for N203 is similar to that of N125 (data not shown).

**Cis-acting replication element (cre)**

In picornaviruses, a *cis*-acting RNA structure, termed *cre*, is necessary for uridylylation of Vpg to form Vpg-pU-pU that serves as the primer for viral RNA synthesis (Paul, 2002). In the enteroviruses, *cre* is located within the 2C coding region, but it can be found in other genome regions in other members of *Picornaviridae*. In all of the simian enteroviruses analysed here, a *cre* sequence is located in 2C, as expected, with the characteristic loop sequence RX₃AAAX₆R atop a three-part, 15-residue stacked stem (data not shown).

**DISCUSSION**

The enterovirus 3’-NTR, the site of initiation of negative-strand RNA synthesis, is required for efficient genome replication (Brown *et al.*, 2004; Mirmomeni *et al.*, 1997; Rohll *et al.*, 1995) and the 3’-NTR is thought to mediate the binding to viral RNA of viral and cellular proteins that are essential for replication (Harris *et al.*, 1994; Mellits *et al.*, 1998). While 3’-NTR sequences vary widely among the various enterovirus species (Brown *et al.*, 2003; Oberste *et al.*, 2004a, b, d, 2006), the existence of highly conserved secondary structures suggests that these structures are the functional unit(s) involved in replication (Mirmomeni *et al.*, 1997; Pilipenko *et al.*, 1992, 1996). The predicted structures consist of three stem–loops termed X, Y and Z in HEV-A and HEV-B, and two stem–loops (X and Y) in HEV-C and HEV-D (Mirmomeni *et al.*, 1997). This grouping of species with regard to the number of 3’-NTR structural domains parallels the species groupings observed in the sequences of the 5’-NTR (Pöyry *et al.*, 1996), a property which may derive from the interaction of the 5’- and 3’-NTRs, via a bridge of viral and cellular proteins, to form a circular structure during RNA synthesis (Barton *et al.*, 2001; Lyons *et al.*, 2001; Teterina *et al.*, 2001). However, the 5’-NTRs of SA5, BaEV, SV6 and N125/N203, all of which have X-Y-Z 3’-NTRs, cannot be readily classified into either 5’-NTR group I or group II based on their phylogenetic relationships to the human enteroviruses. Clearly, other factors, such as interaction with specific host cell proteins, may influence NTR evolution (Mellits *et al.*, 1998; Todd *et al.*, 1995; Waggoner & Sarnow, 1998). The viral NTRs interact with numerous host proteins that are part of the viral replication complex, including PABP, PCBP2, ÆEF-1-¢, nucleolin and others. In addition, the internal ribosome entry site in the 5’-NTR interacts with the ribosome and various translation factors to direct translation of viral proteins (Jackson, 2002). Variation in these proteins among different hosts (humans, baboons, macaques, etc.) may induce variation in the viral sequences with which they interact, allowing the viruses to better adapt to the host. The unique phylogenetic positions of the simian enterovirus 5’-NTRs and the P2 and P3 regions of SA5, BaEV, SV6 and N125/N203 implies host species may have a significant influence on the evolution of enterovirus non-capsid sequences.

Human enterovirus 5’-NTR sequences form two phylogenetic clusters, termed cluster I (HEV-C and HEV-D) and cluster II (HEV-A and HEV-B) (Pöyry *et al.*, 1996). The bovine and porcine enteroviruses each form distinct clusters, with BEV differentiated from all other enteroviruses by the apparent duplication of sequences at the 5’ end of the NTR, resulting in a second predicted ‘cloverleaf’ structure (Zell & Stelzer, 1997). Enteroviruses 76, 89, 90 and 91 (members of HEV-A) are exceptions to the correlation of species with 5’-NTR group, as EV90 and EV91 cluster with group I, while EV76 and EV89 cluster with group II (Oberste *et al.*, 2005). Limited sequence analyses focusing on 5’-NTR, VP1 and 3D sequences have...
previously shown that the phylogenetic trees for these regions are non-congruent (Oberste et al., 2002; Pöyry et al., 1999). SV19/SV26/SV35, SV43, SV46 and BaEV VP1 sequences formed a single genetic cluster within HEV-A, whereas BaEV was distinct from all other enteroviruses in the 3D tree. The BaEV 5′-NTR sequence was also distinct from, but related to, enterovirus 5′-NTR group I, while SV19, SV43 and SV46 clustered together and were related to enterovirus 5′-NTR group II (Pöyry et al., 1999). These complex relationships have been confirmed with the analysis of the complete genome sequences presented here. The changing pairwise and phylogenetic relationships across the genome suggest that the simian enterovirus genomes have been shaped by either inter-species recombination or extensive evolution independent of the human enteroviruses, and call into question the utility of using sequence relationships from multiple genome regions for the molecular taxonomy of the enteroviruses. These results suggest that host species has a significant influence on the evolution of non-capsid sequences, probably due to the specific interactions between host and viral proteins or host proteins and viral RNA. The specific functions of the simian enterovirus non-structural proteins can be only inferred by homology to enteroviral proteins whose functions have been studied explicitly, but numerous picornavirus proteins are known to interact with one or more host cell proteins. For example, the 2A proteins of picornavirus proteins are known to interact with one or more host cell proteins to disrupt Golgi assembly, from the P2 and P3 regions bind, cleave or otherwise interact with host proteins to disrupt Golgi assembly, endoplasmic reticulum-Golgi traffic and secretory pathways, all of which help to shift the cellular machinery to viral replication and protein synthesis (Belov et al., 2007). Presumably, sequence difference in any of these host proteins, as well as differences in receptor homologues between humans and non-human primates may be linked to host range and drive virus divergence and adaptation to host species.

The phylogenetic and pairwise sequence relationships between SV19, SV43, SV46 and the human enteroviruses EV76, EV89, EV90 and EV91, in multiple genome regions, and the unique phylogenetic clustering of EV90 and EV91 in the 5′-NTR, suggests a heretofore unexplored relationship between the human and simian enteroviruses. Some EV76, EV89, EV90 and EV91 strains were isolated from human acute flaccid paralysis cases in Bangladesh (Oberste et al., 2005, 2006), a country where interaction between the human and primate populations is frequent and where conditions are favourable for transmission of viruses between simian species and humans (Jones-Engel et al., 2006). Studies are under way to determine whether primates in Bangladesh currently harbour enteroviruses that closely resemble EV76, EV89, EV90 and EV91.

The sequence comparisons presented here allow us to make specific recommendations for the taxonomic classification of the simian enteroviruses: (i) the SV4 and A-2 plaque virus sequences are distinct from other enteroviruses in all regions of the genome, thereby supporting the recent establishment of SEV-A as a distinct enterovirus species; (ii) similarly, the sequences of SV6 and N125/N203 are consistent with the existence of at least one, and possibly two, additional new species; and (iii) SV19, SV43 and SV46 should be classified within HEV-A, on the basis of their close relationship with members of HEV-A in P1, P2 and P3. However, the current enterovirus molecular classification scheme, based on relationships of the P1, 2C and 3CD amino acid sequences, may be inadequate to classify all known enteroviruses. For example, the taxonomic classification of BaEV and SA5 is inconsistent among the three sequence regions. We suggest that P1 sequence alone can be used for species classification (placing BaEV in HEV-A and SA5 in HEV-B), pending further studies to define the role of the divergent P2 and P3 region sequences in the biology of the simian enteroviruses. Finally, the use of host species names in the taxonomic classification of enteroviruses (e.g. ‘Human enterovirus A’) should be re-evaluated in light of the close relationship between the human and simian viruses in HEV-A.

REFERENCES


