Natural genetic recombination between co-circulating heterotypic enteroviruses

Gabriela Oprisan,1 Mariana Combiescu,1 Sophie Guillot,2 Valerie Caro,2 Andrei Combiescu,1 Francis Delpeyroux2 and Radu Crainic2

1 Cantacuzino Institute, Splaiul Independentei 103, Bucharest, R-70.100 Romania
2 Molecular Epidemiology of Enteroviruses, Pasteur Institute, 25 rue du Dr Roux, 75724 Paris Cedex 15, France

Natural recombination in poliovirus is a frequent phenomenon. In practice, whenever different genotypes have the opportunity to infect the same individual, a high proportion of viruses with recombinant genomes are excreted. To determine whether enteroviruses other than poliovirus can naturally produce viable virions with recombinant genomes, we studied the molecular features of two distant regions of the viral genomes – the VP1 coding region and the 3D polymerase coding region – of the echovirus serotypes associated with a large outbreak of aseptic meningitis. Nucleotide sequences of nine epidemic strains [belonging to echovirus serotypes 4 (E4), 7 (E7) and 30 (E30)] in the two genomic regions (300 nt of VP1 and 520 nt of 3D polymerase) were compared to prototype and field strains, and phylogenetic trees were generated from alignments. In the VP1 region, each of the three epidemic serotypes clustered with the homotypic prototype strain, whereas in the 3D polymerase region, E7 and E30 grouped as a single cluster, distant from the two corresponding prototype strains. This suggests that one of these two E7 and E30 strains has evolved through recombination with the other or that both have acquired the 3D polymerase coding region from a common ancestor. Our results suggest that such genetic recombinations between different echovirus serotypes are possible when multiple epidemic strains are circulating simultaneously.

Introduction

The Enterovirus genus from the family Picornaviridae comprises 64 human serotypes. These serotypes are now classified into five groups: (i) poliovirus (PV) types 1–3; (ii) human enterovirus A (HEV-A), including 11 coxsackie A viruses (CA) and enterovirus (EV) 71; (iii) HEV-B with 36 serotypes including all coxsackie B viruses (CB), all echoviruses (E), EV 69 and CA9; (iv) HEV-C, including 11 coxsackie A viruses; and (v) HEV-D, with only two serotypes, EV68 and EV70 (Mayo & Pringle, 1998; Pringle, 1999). Phylogenetic analysis of either the capsid protein coding regions or the 3D polymerase coding region has confirmed this classification (Hyypia et al., 1997; Pöyry et al., 1996).

Echoviruses are the largest subgroup of HEV-B with 28 serotypes. It is known that echoviruses cause a wide variety of human diseases ranging from subclinical infections and common cold-like illness to fatal encephalitis and meningitis (Melnick, 1996). Most enterovirus serotypes have been associated with aseptic meningitis, although some serotypes are more frequently implicated than others, particularly certain echovirus serotypes (Melnick, 1996; Muir et al., 1998; Nairn & Clements, 1999; Rotbart & Romero, 1995).

The enterovirus genome is a single-stranded RNA molecule of positive polarity, approximately 7500 nt long. The 5′ and 3′ non-coding regions (NCRs) are generally highly conserved, as are several regions encoding the non-structural proteins. The most variable regions of the genome are within the genes encoding the capsid proteins, VP1, VP2 and VP3, which are partially exposed at the virus surface. Since the VP1 gene contains major antigenic sites as well as receptor recognition

Author for correspondence: Gabriela Oprisan.
Fax +40 1 411 56 72 e-mail goprisan@cantacuzino.ro

The EMBL accession numbers of the sequences reported in this paper are AJ304433–AJ304442 for the VP1 region and AJ298084–AJ298097 for the 3D polymerase region.
sequences, the VP1 sequence is supposed to represent most optimally an enterovirus serotype. Sequence comparison and phylogenetic reconstructions have indicated that VP1 contains serotype-specific information that can be used for virus identification. Moreover, sequence analysis of the VP1 region has been shown to be useful in molecular epidemiological studies of enterovirus disease outbreaks (Caro et al., 2001; Künkel & Schreier, 2000; Oberste et al., 1999a, b, c).

The enteroviral RNA genome is replicated by the virus-encoded replicase, an RNA-dependent RNA polymerase (3D polymerase). The sequences encoding the non-structural proteins, including the 3D polymerase, show less variation than the coding regions for the structural proteins (Huttunen et al., 1996; Muir et al., 1998). Due to the absence of proofreading activity, the misinsertion rate by the 3D polymerase is high, and mutations accumulate during replication (Drake, 1993; Holland et al., 1982). Furthermore, recombination has been seen to occur frequently between polioviruses of vaccine and wild-type origin (Cammack et al., 1988; Furione et al., 1993; Georgescu et al., 1994) and perhaps even with non-polio enteroviruses (Guillot et al., 2000). It was recently demonstrated that recombination is a significant and relatively frequent mechanism in the evolution of enterovirus genomes. Bootstrap and genetic similarity analyses have revealed that genetic exchanges could occur within a given serotype (intratypic recombination) and between different serotypes (intertypic recombination) (Sanetti et al., 1999).

A large outbreak of aseptic meningitis with 5000 non-fatal cases occurred between July and September 1999 in Romania. The aetiological agents identified in this outbreak belonged to three different echovirus serotypes: 4, 7 and 30. In several cases, two of the three serotypes were co-isolated from the same patient, either from stools or from cerebrospinal fluid (CSF) samples. It is noteworthy that, in contrast to the E7 strains, the E4 and E30 strains were very rarely isolated in Romania before this outbreak.

In order to study the molecular features of the echovirus serotypes associated with this outbreak, two distant regions of the virus genome, the VP1 coding region and the 3D polymerase coding region, were analysed.

**Methods**

**Viruses.** Viral isolation was carried out from clinical specimens, including CSF, stools and throat swabs, at the National Reference Center for Enteroviruses (Cantacuzino Institute, Romania). Prototype strains used in this study were obtained from the World Health Organization (Copenhagen) and from the Pasteur Institute (Paris).

Serotyping of viruses isolated on rhabdomyosarcoma (RD) cells was carried out by neutralization with in-house pools of serotype-specific rabbit antisera (LBM pools). In order to ensure that viruses were separated from the mixtures, the results were verified with anti-echovirus reference sera obtained from the ATCC, with serum pools provided by RIVM (The Netherlands) and by plaque purification, for some viruses.

Nine epidemic strains (E4, E7 and E30) from Iasi, Suceava and Bacau were chosen for this analysis. Two non-epidemic E7 strains that had previously been isolated from Bucharest were also used for comparison. Table 1 shows the list of the strains used.

**RNA extraction, RT–PCR and sequencing.** RNA extractions were carried out with the Total Quick RNA (Talent) using 100 µl of infected RD cell-culture supernatant, according to the manufacturer’s instructions. The RNA pellet was dissolved in 30–70 µl RNase-free water and stored at −70 °C until analysis. For cDNA synthesis, 1 µl viral RNA, 20 U Rnasin (Promega) and 10 pmol antisense primer were heated at 80 °C for 5 min, then annealed at 42 °C for 5 min. To this reaction mixture were added 4 µl 5 × transcriptase buffer, 2 µl 0.1 M DTT, 1 µl of dNTP mix (10 mM each; Gibco BRL) and 200 U of SuperScript RNase H− reverse transcriptase (Life Technologies) in a final volume of 20 µl. The RT reaction mix was incubated at 42 °C for 30 min, heat-inactivated at 95 °C for 5 min and then chilled on ice.

Two PCR reactions were designed for the capsid and 3D polymerase coding regions: the first RT–PCR encompassed approximately 1500 bp from VP1 and spanned the non-structural coding region to 2C (nt 2941–4428, with reference to the E30 Bastianni strain, GenBank accession no. AF081340). The primer sequences and the PCR conditions have been previously described (Caro et al., 2001). The second amplicon included approximately 890 bp of the 5′ extremity of the 3D polymerase coding region (nt 6046–6938, with reference to E30 Bastianni). The 3D polymerase primers represented degenerate positions and were designed to amplify, when used in combination, a wide variety of EV serotypes, including polioviruses (Table 2). The amplification was carried out for 30 cycles consisting of 20 s at 95 °C, 1 min at 50 °C, 1 min at 72 °C, followed by an additional 10 min incubation at 72 °C. To minimize the risk of PCR contamination, filter tips, different PCR reagent aliquots, and positive and negative controls were systematically used for each experiment.

The two sets of primers were tested first by PCR amplification of the HEV prototype strains. PCR products were analysed in ethidium bromide-stained 1% agarose gels. Depending on the presence of a single or multiple bands in the gel, PCR products were either directly purified using the QIAquick PCR purification kit (Qiagen), or excised from the gel and purified for sequencing with the QIAquick Gel Extraction kit (Qiagen). Amplicons were eluted in 30–50 µl 10 mM Tris–HCl (pH 8.5) and stored at −20 °C.

The amplified DNA fragments were directly sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer Applied Biosystems) on the ABI Prism DNA 377 Sequencer (Perkin-Elmer Applied Biosystems) according to the protocol of the kit. The sequencing reactions were performed with the same forward primers as used in the PCR step. For some E7 and E30 isolates, the entire VP1–2C region was sequenced.

**Phylogenetic analysis.** Nucleotide sequences of epidemic strains in the two genomic regions (300 nt of VP1 and 520 nt of the 3D polymerase) were compared with those of the prototype and field strains isolated elsewhere in the world. Nucleotide sequences were aligned with clustal w (version 1.81) (Thompson et al., 1994). The GenBank DNA sequence library was screened for similar sequences using the FASTA 3.0 program (Pearson & Lipman, 1988). The phylogenetic analysis was performed using the programs included in the PHYLIP package version 3.5 (Felsenstein, 1993) and Puzzle version 4.0 (Strimmer & von Haeseler, 1996). Puzzle was executed by the use of the distance method of Kishino & Hasegawa (1989). The distance matrix was calculated by the Kimura two-parameter method using DNADIST. Tree reconstruction was performed with the kitchin program of the phylip package. The reliability of the phylogenetic reconstructions was estimated by bootstrap analysis with 100 pseudoreplicate data sets (kitchin) or by using 1000 puzzling
Table 1. List of enterovirus strains used

<table>
<thead>
<tr>
<th>Isolate name*</th>
<th>Location</th>
<th>Year of isolation</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>E7 RO-201/4/99</td>
<td>Iasi</td>
<td>1999</td>
<td>CSF</td>
</tr>
<tr>
<td>E7 RO-205/4/99</td>
<td>Iasi</td>
<td>1999</td>
<td>CSF</td>
</tr>
<tr>
<td>E30 RO-205/1/99</td>
<td></td>
<td>1999</td>
<td>Stool</td>
</tr>
<tr>
<td>E7 RO-208/2/99</td>
<td>Iasi</td>
<td>1999</td>
<td>Stool</td>
</tr>
<tr>
<td>E4 RO-208/4/99</td>
<td>Iasi</td>
<td>1999</td>
<td>CSF</td>
</tr>
<tr>
<td>E30 RO-263/2/99</td>
<td></td>
<td>1999</td>
<td>Stool</td>
</tr>
<tr>
<td>E4 RO-210/2/99</td>
<td>Suceava</td>
<td>1999</td>
<td>Stool</td>
</tr>
<tr>
<td>E4 RO-234/1/99</td>
<td>Bacau</td>
<td>1999</td>
<td>Stool</td>
</tr>
<tr>
<td>E30 RO-229/4/99</td>
<td></td>
<td>1999</td>
<td>CSF</td>
</tr>
<tr>
<td>E7 RO-434/2/81†</td>
<td>Bacau</td>
<td>1981</td>
<td>Stool</td>
</tr>
<tr>
<td>E7 RO-141/2/95†</td>
<td>Bucharest</td>
<td>1995</td>
<td>Stool</td>
</tr>
</tbody>
</table>

* Isolates are described by country of origin, laboratory identifier and year of isolation.
† Non-epidemic isolates.

Table 2. Oligonucleotide primers for RT–PCR and sequencing of the enterovirus 3D polymerase region

<table>
<thead>
<tr>
<th>Primer</th>
<th>Position*</th>
<th>Sense</th>
<th>5′ → 3′ Sequence†</th>
<th>Use†</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUC12-a</td>
<td>6938–6916</td>
<td>R</td>
<td>ACATGTCMCCCATATGCRATCAT</td>
<td>PCR/sequencing</td>
</tr>
<tr>
<td>EUC12-b</td>
<td>6938–6916</td>
<td>R</td>
<td>ACATCACTCCATAGGCACCAT</td>
<td>PCR</td>
</tr>
<tr>
<td>EUC12-c</td>
<td>6938–6916</td>
<td>R</td>
<td>ACGTCACTACCACATGCATCAT</td>
<td>PCR/sequencing</td>
</tr>
<tr>
<td>UG7#</td>
<td>6046–6068</td>
<td>F</td>
<td>TTTGAGGGGTGAGGAGCCACGC</td>
<td>PCR/sequencing</td>
</tr>
<tr>
<td>UG7-A</td>
<td>6046–6068</td>
<td>F</td>
<td>TTTGAGGGSAACAARGAACCAGC</td>
<td>PCR/sequencing</td>
</tr>
<tr>
<td>UG7-B</td>
<td>6046–6068</td>
<td>F</td>
<td>TTTGAGGGMAGTAAGAGCCWGC</td>
<td>PCR/sequencing</td>
</tr>
<tr>
<td>UG7-P</td>
<td>6046–6068</td>
<td>F</td>
<td>TTTGAGGGSGHCAAAGAACGC</td>
<td>PCR/sequencing</td>
</tr>
</tbody>
</table>

* Primer positions are given according to the orientation of the primer, either forward (F) or reverse (R) sense; positions refer to E30 Bastianni (GenBank accession no. AF081340).
† M = A/T; R = A/G; H = A/T/C; B = G/T/C; S = G/C; W = A/T (IUPAC–IUB/GCG code for degenerated bases).
‡ Poliovirus-specific primer previously described (Furione et al., 1993).

Results

In order to evaluate the phylogenetic relationships between the three echovirus serotypes isolated during the aseptic meningitis outbreak from Romania and their relationships with other enterovirus strains, two genomic regions were chosen for analysis by PCR and sequencing: the VP1 and 3D polymerase regions.

All three epidemic serotypes, E4, E7 and E30, and corresponding prototype strains were successfully amplified and sequenced by both sets of primers. These degenerate primers were designed to be applicable to all enterovirus sequences available in the database. Phylogenetic trees were constructed based on distance matrices.

Precautions to minimize the risk of PCR contamination were taken (see Methods). Amplifications in the VP1 and 3D polymerase regions were not carried out in the same experiment and reamplifications confirmed the previous results. In addition, as some patients were infected with multiple echoviruses, different methods of virus typing were used to ensure that virus strains were actually separated from the mixtures.

VP1 coding region

A segment of approximately 300 nt from the C-terminal third of the VP1 region was sequenced. Sequencing of the epidemic E4, E7 and E30 strains showed that the partial VP1 sequence fully correlated with the serotype determined by the conventional neutralization test. A dendrogram depicting the relationships of the epidemic strains in the VP1 and 3D polymerase regions is shown in Fig. 1. The analysis also included two non-epidemic E7 strains isolated before the outbreak (E7 RO-434/2/81 and E7 RO-141/2/95) in 1981 and 1995, respectively, and the sequences for the prototype
E4, E7 and E30 strains and other enteroviruses from the group HEV-B, which were obtained from the EMBL/GenBank database.

The E4 strains were very similar to each other and fell into the same cluster with the prototype E4 Pesacek. When the GenBank DNA sequence library was screened for similar sequences, the E4 epidemic strains were found to be genetically related to an American field isolate, E4 WA93 (numbers correspond to the year of isolation) with a 91% nucleotide identity and to a German epidemic isolate, enterovirus Stuttgart 97, with a 90% identity. The E7 strains were related to the E7 prototype strain Wallace and were similar to an American E7 isolate GA93 with 85% identity. As with the epidemic E4 isolates, the epidemic E7 isolates analysed were very homogeneous in the capsid region. This was not the case with the epidemic E30 strains, which appeared to belong to two slightly different lineages, one of them being very similar (91–98% identity) to some French isolates (E30 FR-1829/96, E30 FR-2689/91 and E30 FR-2090, isolated in 1989). Nevertheless, in the phylogenetic tree, the epidemic E30 strains clustered together with the prototype strain Bastianni.

**3D polymerase coding region**

An RT–PCR product of approximately 890 nt was obtained from the prototype strains tested and from the epidemic E4, E7 and E30 strains. Since there were no sequence data available for the reference strains E4 and E7 in the polymerase region chosen for analysis, the strains Pesacek and Wallace, respectively, were amplified and sequenced. A sequence of 520 nt located in the 3D polymerase region (nt 6127–6646, with reference to E30 Bastianni) was used in this analysis.

Virus sequences grouped quite differently in the 3D polymerase region when compared with the phylogenetic tree representing the VP1 region. The three serotypes of the epidemic isolates did not group with the corresponding prototype strains. The E30 epidemic isolates fell into the same cluster as the epidemic E7 strains. The robustness of the trees was supported by high bootstrap values (Fig. 1). Different phylogenetic methods (maximum-likelihood and maximum-parsimony) gave the same tree topology (not shown).

Sequence comparisons revealed that although epidemic E30 strains are distinct from other serotypes in terms of their
Table 3. Percentage nucleotide identity in the capsid (VP1) and 3D polymerase regions among some analysed enteroviruses

Sequences were compared pairwise using the CLUSTAL W program, and the percentage nucleotide identity values for the pairs of strains are shown. The prototype enterovirus strains are indicated in **bold** type and the two non-epidemic E7 strains in *italics*. The values corresponding to the recombinant viruses in the 3D polymerase region are highlighted.

<table>
<thead>
<tr>
<th>Isolates</th>
<th>E4</th>
<th>E4</th>
<th>E7</th>
<th>E7</th>
<th>E7</th>
<th>E7</th>
<th>E7</th>
<th>E30</th>
<th>E30</th>
<th>E30</th>
<th>E30</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>208</td>
<td>201</td>
<td>205</td>
<td>208</td>
<td>434</td>
<td>141</td>
<td>Wal</td>
<td>205</td>
<td>229</td>
<td>263</td>
<td>Bas</td>
</tr>
<tr>
<td>3D polymerase region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E4 RO-208/4/99*</td>
<td>-</td>
<td>79</td>
<td>79</td>
<td>79</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>81</td>
<td>79</td>
<td>78</td>
<td>83</td>
</tr>
<tr>
<td>E4 Pesacek</td>
<td>79</td>
<td>-</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>79</td>
<td>80</td>
<td>83</td>
<td>78</td>
<td>78</td>
<td>76</td>
</tr>
<tr>
<td>E7 RO-201/4/99</td>
<td>67</td>
<td>99</td>
<td>99</td>
<td>99</td>
<td>82</td>
<td>83</td>
<td>82</td>
<td><strong>95</strong></td>
<td><strong>95</strong></td>
<td><strong>95</strong></td>
<td>80</td>
</tr>
<tr>
<td>E7 RO-205/4/99</td>
<td>67</td>
<td>63</td>
<td>100</td>
<td>99</td>
<td>99</td>
<td>99</td>
<td>82</td>
<td>82</td>
<td>81</td>
<td><strong>95</strong></td>
<td><strong>95</strong></td>
</tr>
<tr>
<td>E7 RO-208/2/99</td>
<td>67</td>
<td>63</td>
<td>100</td>
<td>100</td>
<td>-</td>
<td>92</td>
<td>82</td>
<td>82</td>
<td>95</td>
<td><strong>96</strong></td>
<td><strong>95</strong></td>
</tr>
<tr>
<td>E7 RO-434/2/81</td>
<td>66</td>
<td>66</td>
<td>82</td>
<td>82</td>
<td>82</td>
<td>-</td>
<td>90</td>
<td>91</td>
<td>82</td>
<td>82</td>
<td>83</td>
</tr>
<tr>
<td>E7 RO-141/2/95</td>
<td>66</td>
<td>66</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>-</td>
<td>80</td>
<td>82</td>
<td>83</td>
<td>83</td>
</tr>
<tr>
<td>E7 Wallace</td>
<td>62</td>
<td>62</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>78</td>
<td>77</td>
<td>-</td>
<td>81</td>
<td>80</td>
<td>82</td>
</tr>
<tr>
<td>E30 RO-205/1/99</td>
<td>60</td>
<td>58</td>
<td>62</td>
<td>62</td>
<td>62</td>
<td>57</td>
<td>58</td>
<td>57</td>
<td>-</td>
<td><strong>95</strong></td>
<td>98</td>
</tr>
<tr>
<td>E30 RO-229/4/99</td>
<td>62</td>
<td>66</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>56</td>
<td>58</td>
<td>60</td>
<td>91</td>
<td>-</td>
<td><strong>95</strong></td>
</tr>
<tr>
<td>E30 RO-263/2/99</td>
<td>60</td>
<td>58</td>
<td>62</td>
<td>62</td>
<td>62</td>
<td>57</td>
<td>58</td>
<td>58</td>
<td>99</td>
<td>91</td>
<td>-</td>
</tr>
<tr>
<td>E30 Bastianni</td>
<td>59</td>
<td>62</td>
<td>59</td>
<td>59</td>
<td>59</td>
<td>57</td>
<td>58</td>
<td>56</td>
<td>84</td>
<td>87</td>
<td>85</td>
</tr>
</tbody>
</table>

*The three epidemic E4 isolates gave similar values.

Discussion

We analysed the partial nucleotide sequence of nine epidemic isolates of E4, E7 and E30 serotypes in two distant genomic regions encoding VP1 and the 3D polymerase. In order to investigate the relationships between the three epidemic strains, unrooted distance trees were constructed.

The results indicated that the phylogenetic relationships observed between epidemic strains in the VP1 region corresponded to the antigenic classification. Each of the three different echoviruses clustered together with the corresponding reference strain. Analysis of the VP1 sequence data was in agreement with previous studies with VP1 as a region of choice for molecular typing (Caro *et al*., 2001; Mulders *et al*., 2000; Oberste *et al*, 1999a). In the capsid coding region, the three epidemic serotypes clustered together with the corresponding reference strains, although they were rather distantly related to them in terms of nucleotide differences. The epidemic E30 strains were genetically distinct in the VP1 region from the reference strain Bastianni with 13–16% nucleotide differences. The E7 strains were different from the E7 Wallace prototype with 23% nucleotide differences, and the capsid protein VP1, the 3D polymerase region shares considerable identity with those of the epidemic E7 strains. The epidemic E7 strains were only distantly related to the E7 prototype strain Wallace and to the two Romanian non-epidemic E7 strains.

The relationships between the strains analysed as indicated by percentage nucleotide identity are presented in Table 3. The nucleotide identity between the epidemic E7 isolates in the VP1 coding region and 3D polymerase coding region was very high (100% and 99%, respectively). More heterogeneous, the epidemic E30 strains appeared as two different lineages with 91% versus 99% nucleotide identity in the capsid region and 95% versus 98% in the polymerase region. The two non-epidemic E7 strains (434/2/81 and 141/2/95), isolated at an interval of 14 years, presented 93% nucleotide identity in the capsid region and 90% in the 3D polymerase region. Both of them presented only 82–83% nucleotide identity with the epidemic E30 strains in the 3D polymerase coding region. In contrast, in the same region, the epidemic E7 strains analysed were very closely related to the epidemic E30 strains, with 95–96% nucleotide identity (Fig. 1).
E4 strains were different from the prototype strain Pesacek with 21–24% nucleotide differences. These results are consistent with other reports describing the differences between the echovirus strains currently circulating and the prototype strains isolated 40 years ago, as in the case of the E30 isolates (Kunkel & Schreier, 2000; Oberste et al., 1999c). Otherwise, little VP1 sequence variation was observed among homotypic E4 and E7 epidemic isolates. The E30 strains isolated in the outbreak were more heterogeneous, showing up to 8% nucleotide difference in their capsid region (Table 3).

While in the VP1 region each of the three epidemic serotypes clustered independently with the homotypic prototype strain, in the 3D polymerase region, epidemic E7 and E30 strains grouped in a single cluster. In contrast, the non-epidemic E7 strains (141/2/95 and 434/2/81) formed a separate cluster from the epidemic E7 and E30 strains (Fig. 1).

Mutations and recombination are the two mechanisms playing roles in picornavirus evolution. While recombinant polioviruses have been observed in isolates from vaccine recipients and in naturally circulating wild viruses (Furione et al., 1993; Georgescu et al., 1994; Kew & Nottay, 1984), very few other examples of a recombinant human enterovirus have been described (Hughes et al., 1989; Santti et al., 1999). According to our results, the same echoviruses were found in different genetic clusters when sequences of regions encoding the capsid protein and RNA polymerase were compared. This suggests that a genetic rearrangement between epidemic E7 and E30 strains may have occurred. Such a phenomenon of recombination involving different serotypes would be favored when multiple epidemic strains are circulating simultaneously, as happened in the outbreak of aseptic meningitis occurring in Romania in 1999. Strikingly, in some patients combinations of two different serotypes (E4 and E7, E7 and E30 or E4 and E30) were isolated in stool and CSF samples. It is noteworthy that the putative recombinants E7 and E30 (even when co-isolated, for instance E7 205/4/99 and E30 205/1/99) presented 95% nucleotide identity in the polymerase region but only 62% identity in the capsid coding region. The non-epidemic E7 strains and the E7 prototype strain Wallace were not closely related to the epidemic E30 strains (82–83%) in the 3D polymerase region. Such a substantial transition from 82% to 95% in terms of nucleotide identity is difficult to explain by natural mutation rate. However, precautions were taken to ensure that this high level of relatedness between sequences from different serotypes was not artefactual (see Methods and Results).

In the 3D polymerase region, all field isolates of each serotype, including the E4 isolates, appeared to be closely related. This pattern of subgrouping could be explained by the fact that these viruses were isolated in the same geographical space and time. In contrast, the field isolates were only distantly related in this genomic region to the corresponding prototype strains, which were actually isolated many decades ago. Similar grouping for field isolates and prototype strains have been recorded in the 5′ NCR (Kopecka et al., 1995) and in non-structural regions (Santti et al., 1999).

Genetic recombination during the evolution of enteroviruses could explain why the polymerase regions (and possibly all non-structural coding regions) are inappropriate regions to find correlations between distantly related strains of the same serotype and thus to identify virus serotype. However, very frequent recombination events would render the polymerase regions inappropriate to find relationships between any field isolates. This is clearly not the case since a close relationship was found between strains of the same serotype in this study. Intratype recombination of polioviruses appears to have a higher frequency than the intertype type (King, 1988; Kirkegaard & Baltimore, 1986). Similarly, the intertype recombination between non-polio enteroviruses could be restricted by the selective forces based on the functionality of the viral replicase. These constraints would maintain the frequency of enterovirus recombination at a moderate level.

Santti et al. (1999, 2000) demonstrated that intraspecies exchanges have occurred in the evolution of enterovirus genomes. High similarity values (90 %) were obtained by bootstrapping analysis in the regions encoding the non-structural proteins of the cluster B representatives. The authors interpreted these results as strong evidence that multiple recombination events, both within and between serotypes, had taken place in the evolution of this cluster. These findings evaluated the recombination at a macroevolutionary level, mainly on prototype enteroviruses. Sequencing data from several studies have suggested that there are genetic exchanges between more recent isolates when multiple genotypes are circulating simultaneously in the same outbreak (Kopecka et al., 1995).

If the epidemic E7 and E30 strains analysed arose through recombination, then the donor of sequences in the 3′ half of the genome could have been either an E7 or E30 strain. Unfortunately, we did not find the progenitors in the 3D polymerase region sequences to be different from those of the recombinant viruses. To confirm these assumptions and to find putative parents of the recombinant viruses, further investigations on other E7 and E30 isolates are needed. Moreover, location of the recombinant sites remains to be determined. Sequencing of genomic fragments in the 2C region of some E7 and E30 isolates (nt 3320–4400, with reference to E30 Bastianii) suggested that the recombination sites are located between this 2C fragment and the 3D polymerase fragment analysed (not shown).

Previous studies have shown that recombination can occur between polioviruses, either among vaccine strains or between vaccine strains and wild-type viruses. Moreover, the possibility that such genetic exchanges occur with non-polio enteroviruses has not been excluded (Georgescu et al., 1995; Guillot et al., 2000; Stanway, 1990). For many recombinant viruses, the genetic exchanges occurred in the 3′ moiety of
their genomes. Consequently, the combined analysis of two distant segments of the enterovirus genome represents an effective system of finding recombinant viruses (Crainic & Kew, 1993; Furione et al., 1993; Guillot et al., 2000).

In this study, we have presented an approach for the characterization of field isolates in two genomic regions using RT–PCR and subsequent sequence analysis. We interpreted the pattern of relatedness of sequences in the 3D polymerase region as evidence for recombination between epidemic E7 and E30 strains. Our findings suggest that intertypic recombination is possible when multiple enterovirus serotypes are circulating at the same time and in the same geographical area.

In the context of poliomyelitis eradication, we can expect an increasing circulation of non-polio enteroviruses. When vaccination with oral poliovirus vaccine is stopped, the ecological niche specific to the polioviruses will most probably be occupied by the non-polio enteroviruses.

The 3D polymerase sequence database available for field enteroviruses is very limited. Further studies concerning not only the 5’ part of the genome but also the 3’ part, including the 3D polymerase coding region, may reveal some interesting features concerning the actual evolution of enteroviruses.

We would like to thank Dr. Eckart Schreier, from Robert Koch Institut, Berlin, Germany, for the helpful sequencing data regarding the 5’ NCR and VP1 region of the Romanian epidemic echoviruses. We would also like to thank Adrian Onu, Dan Otelea and Ana Persu (Cantacuzino Institute, Bucharest) for helpful comments and Sandrine Rosario (Pasteur Institute, Paris) for skilful technical assistance. This work was partly supported by grants to G.O. and F.D. from the Deutscher Akademischer Austauschdienst, Hamburg, Germany, and from the European Community (Copernicus-CIPA NCR 107) and from the European Community (Copernicus-CIPA V26 170–179). For the helpful sequencing data regarding the 5’ NCR we thank Combescu, M., Combiescu, A. A., Guillot, S. & Crainic, R. (2000).

References


Received 28 November 2001; Accepted 18 April 2002