Molecular identification of new picornaviruses and characterization of a proposed enterovirus 73 serotype

M. Steven Oberste,1 David Schnurr,2 Kaija Maher,1 Suleiman al-Busaidy3 and Mark A. Pallansch1

1 Division of Viral and Rickettsial Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, 1600 Clifton Road NE, Mailstop G-17, Atlanta, GA 30333, USA
2 Viral and Rickettsial Diseases Laboratory, California Department of Health Services, Berkeley, CA 94704, USA
3 Department of Laboratories, Directorate General of Health Affairs, Ministry of Health, 113 Muscat, Oman

Enteroviruses (EV) have traditionally been identified by using serotype-specific antisera in a virus-neutralization test. Three EV strains isolated in California, USA, in 1955, 1964 and 1978, and a 1995 Oman isolate, were found to be antigenically related to one another; however, the strains were not neutralized by standard EV typing antisera, suggesting that they may represent a new EV serotype. The isolates were characterized genetically by RT–PCR coupled with amplicon sequencing and comparison to a database of enterovirus nucleotide sequences. The strains were 75-3 to 87-2% identical to one another in complete VP1 nucleotide sequence, but no more than 68% identical in sequence to the prototype strain of any EV serotype. Their complete capsid sequences were closely related to one another, but only distantly related to those of any EV prototype strain. The California and Oman isolates were most closely related to members of EV cluster B, suggesting that they are unclassified members (i.e. a new serotype) of cluster B. The complete genome sequence was determined for one isolate, CA55-1988, and the predicted polyprotein sequence was 86-5 to 89-2% identical to those of other cluster B EV and 56-7 to 61-9% identical to the polyprotein sequences of EV belonging to other clusters. Isolation of this new EV serotype from samples obtained on two continents and over a period of 40 years suggests continued circulation over a wide geographical area. In keeping with standard picornavirus nomenclature, we propose that this new serotype be named ‘enterovirus 73’ (EV73).

Introduction

Viruses in the family Picornaviridae were initially classified on the basis of physical characteristics and antigenic relationships measured by neutralization test (Committee on Enteroviruses, 1962; Melnick, 1996; Minor et al., 1995). Over 200 serotypes have been identified and nine genera have been defined (King et al., 2000; Mayo & Pringle, 1998): Enterovirus (89 serotypes), Rhinovirus (103 serotypes), Cardiovirus (3 serotypes), Aphthovirus (8 serotypes), Hepatovirus (2 serotypes), Parechovirus (2 serotypes), Tospovirus (10 serotypes), Kobuvirus (1 serotype) and Erbovirus (1 serotype). Of the 89 Enterovirus serotypes, 64 are known to infect humans (Melnick, 1996). Most of the human enterovirus serotypes were discovered and described between 1948 and 1963 as a result of the application of cell culture and suckling mouse inoculation to investigations of cases of infantile paralysis (paralytic poliomyelitis) and other central nervous system diseases (Committee on Enteroviruses, 1962; Panel for Picornaviruses, 1963).

The accepted approach to classification requires the investigator to generate antisera against each potentially new serotype and to perform reciprocal cross-neutralization testing, using a complete panel of prototype strains and antisera (Committee on Enteroviruses, 1962). The effect of this labour-intensive approach on the pace of enterovirus discovery was recognized as early as 1962, when a leader in the field remarked, ‘new human enteroviruses are still being discovered. The rate of making new discoveries has slowed, probably only because the labours involved in establishing “new” serotypes...’
are now so very great!’ (Wenner, 1962). Since that time, only nine new human enterovirus serotypes have been identified (Melnick et al., 1974; Panel for Picornaviruses, 1963; Rosen & Kern, 1965; Rosen et al., 1973; Schieble et al., 1967) and none have been identified in the last 25 years.

Recognizing the technical difficulties and limitations inherent in the classic approach to enterovirus identification, we have developed rapid molecular techniques for enterovirus typing (Oberste et al., 2000, 1999a, b). In the present study, we describe a comparison of immunological and molecular methods for the characterization and classification of a group of related human enteroviruses isolated over a period of 40 years and propose the classification of these isolates as members of a new human enterovirus serotype.

Methods

Virus isolation and antigenic characterization. The viruses analysed in this study are described in Table 1. CA55-1988, CA64-4454 and CA78-1480 were isolated in the Viral and Rickettsial Diseases Laboratory (VRDL), California Department of Health Services, Berkeley, CA, USA. The VRDL serves as a reference laboratory for the State of California, performing primary isolation and identification of viruses, as well as identification of viral cultures referred from other laboratories. CA55-1988 was isolated in primary rhesus monkey kidney cells (RMK). CA64-4454 was isolated in RMK and in human foetal diploid kidney fibroblast cells (HFDK). CA78-1480 was isolated in RMK, HFDK and diploid human foetal lung cells (IMR-90). None of the strains were lethal in intracerebrally inoculated newborn mice. Typing by neutralization test was attempted using the Lim and Benyesh-Melnick (LBM) or in-house immune serum pools and standard methods (Grandien et al., 1989; Lim & Benyesh-Melnick, 1960; Schmidt et al., 1961).

OMA95-6498 and OMA95-6499 were isolated in the National Poliovirus Laboratory, Department of Laboratory Services, Ministry of Health, Muscat, Oman, during the course of poliovirus surveillance activities in support of the Global Poliomyelitis Eradication Initiative. Specimens were inoculated into RD and HEp2(C) cells, observed for the development of enterovirus-like cytopathic effect, and tentatively identified as untypeable non-polio enteroviruses on the basis of the absence of neutralization using antisera against the three poliovirus serotypes and with pooled antisera specific for 27 non-polio enteroviruses (World Health Organization, 1997).

Immune sera were prepared against CA55-1988 and CA78-1480 in hamsters and against CA64–4454 in monkeys. Antisera were standardized with all other enterovirus serotypes, as previously described (Schnurr et al., 1996), and used to assess the antigenic relationships of CA55-1988, CA64-4454, CA78-1480 and OMA95-6498 by neutralization, using standard methods (Grandien et al., 1989).

Physical characterization. The California isolates were tested for stability to acid by incubation of the virus in pH 3.0 buffer for 1 h at 4 °C and inoculation of cell cultures after adjustment to pH 7 (Gwaltney et al., 1989). The isolates were purified by three terminal dilutions and further characterized by determination of size and morphology using electron microscopy, stabilization to heat by divalent cations, determination of genome type (DNA or RNA) and measurement of either stability. The Oman isolates were not physically characterized.

Molecular characterization of viruses. Viral RNA extraction, RT–PCR, nucleotide sequencing and sequence analysis were performed as described previously (Oberste et al., 2000, 1999a). For initial molecular characterization, viral RNA from each isolate was amplified by RT–PCR, using VP1 primers 012-011, 040-011, 187-222, 188-222 and 189-222, as described previously (Oberste et al., 2000, 1999a). The partial VP1 sequences were compared with a database of all complete enterovirus VP1 sequences (Oberste et al., 1999b) and other picornavirus VP1 sequences that are available in the GenBank database, as described previously (Oberste et al., 2000, 1999a), to determine whether the isolates were genetically related to any known picornavirus serotype. Sequences spanning the 3′ portion of the 5′ NTR to the 5′ end of VP2 were determined as described previously (Oberste et al., 1998). Complete capsid sequences were determined using the primer-walking method to amplify and sequence the region between the 5′-NTR-VP4-VP2 and VP1 sequences. The complete genome sequence of CA55-1988 was also determined by primer walking.

Results

Physical and antigenic properties

As part of routine enterovirus typing procedures, neutralization of the California isolates was attempted using standard pools of enterovirus typing antisera (Lim & Benyesh-Melnick,
Identification and characterization of EV73

Table 2. Antigenic relationships assayed by neutralization test

Titres are shown as reciprocal of 50% endpoint, determined by the method of Reed & Muench (1938). Numbers in parentheses are the virus dose (50% cell culture infectious dose), calculated from a back-titration. NT, Not tested.

<table>
<thead>
<tr>
<th>Virus</th>
<th>Antiserum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA55-1988</td>
</tr>
<tr>
<td>EV prototypes (48 strains)*</td>
<td></td>
</tr>
<tr>
<td>CA55-1988</td>
<td>1024 (320)</td>
</tr>
<tr>
<td>CA64-4454</td>
<td>&lt; 8 (320)</td>
</tr>
<tr>
<td>CA78-1480</td>
<td>256 (32)</td>
</tr>
<tr>
<td>OMA95-6498</td>
<td>512 (320)</td>
</tr>
<tr>
<td>OMA95-6499</td>
<td>1024 (32)</td>
</tr>
</tbody>
</table>

* Strains tested included the recognized prototype strains of poliovirus types 1–3, coxsackieviruses A7, A9, A16 and B1–B6; echoviruses 1–9, 12–29 and 31–33; and EV69–71, as well as non-prototype reference strains of echovirus types 11 and 30.

Table 3. VP1 nucleotide sequence relationships (percent identity) of candidate EV73 strains to other picornaviruses

<table>
<thead>
<tr>
<th></th>
<th>CA55-1988</th>
<th>CA64-4454</th>
<th>CA78-1480</th>
<th>Oman95*</th>
</tr>
</thead>
<tbody>
<tr>
<td>EV prototypes (64 strains)</td>
<td>47–0–67·0</td>
<td>45–9–67·8</td>
<td>46–0–67·4</td>
<td>45·4–66·7</td>
</tr>
<tr>
<td>EV Cluster A</td>
<td>47–0–50·5</td>
<td>45–9–50·2</td>
<td>46–0–50·4</td>
<td>45·4–51·2</td>
</tr>
<tr>
<td>EV Cluster B</td>
<td>59–1–67·0</td>
<td>58·5–67·8</td>
<td>58·7–67·4</td>
<td>58·8–66·7</td>
</tr>
<tr>
<td>EV Cluster C</td>
<td>49·3–54·7</td>
<td>49·8–53·9</td>
<td>49·8–55·5</td>
<td>50·1–55·1</td>
</tr>
<tr>
<td>EV Cluster D</td>
<td>48·1–48·9</td>
<td>49·5–51·2</td>
<td>48·1–49·2</td>
<td>47·2–49·4</td>
</tr>
<tr>
<td>Rhinoviruses†</td>
<td>47·6–51·5</td>
<td>47·5–50·8</td>
<td>48·5–51·6</td>
<td>48·0–50·7</td>
</tr>
<tr>
<td>Other picornaviruses</td>
<td>≤ 41·1</td>
<td>≤ 41·7</td>
<td>≤ 43·1</td>
<td>≤ 42·0</td>
</tr>
</tbody>
</table>

* ‘Oman95’ includes both OMA95-6498 and OMA95-6499, whose VP1 sequences were identical to one another. Because the VP1 sequences were identical, OMA95-6499 was excluded from further analyses.
† Human rhinovirus VP1 sequences available from the GenBank database: types 1B, 2, 3, 14, 16 and 89.
Molecular serotyping and genetic relationships

The complete VP1 sequence (867 nucleotides) of each of the isolates was determined and compared with those of the human enterovirus prototype strains and of other picornaviruses. In all cases, the nearest taxa were viruses in the Enterovirus genus, and the highest pair-wise nucleotide sequence identity score was between 67 and 68% (Table 3). The highest scoring serotypes were members of enterovirus cluster B, which contains the coxsackie B viruses, the echoviruses, coxsackievirus A9 and EV69. Scores for comparisons with available human rhinovirus sequences were between 47 and 52%, similar to the scores for comparisons with cluster A enteroviruses (Table 3). Scores for comparisons with picornaviruses of all other genera were below 44%. For the complete VP1 sequence comparisons, a highest identity score below 70% indicates that the query sequence is heterologous in serotype to all sequences in the database (Oberste et al., 1999a, b), suggesting that the California and Oman isolates are unclassified members (a new serotype) of cluster B. The complete VP1 sequence of OMA95-6499 was identical to that of OMA95-6498, and therefore OMA95-6499 was not further characterized genetically.

Phylogenetic relationships were inferred from complete VP1 coding sequences by the quartet puzzling method, using maximum likelihood distances. CA55-1988, CA64-4454, CA78-1480 and OMA95-6498 were monophyletic with respect to the prototype strains of all known enterovirus serotypes, with a 99% support value (Fig. 1 and data not shown). The nearest sibling taxa were E25, E29 and E6, in general agreement with the nucleotide identity scores. Within the new group, the phylogenetic relationships mirrored the antigenic relationships, as CA55-1988 and CA78-1480 clustered together (100% support) and CA64-4454 was more distant.

The nucleotide sequence for the region encoding the complete capsid protein was determined for CA55-1988, CA64-4454, CA78-1480 and OMA95-6498. Overall, the capsid-coding sequences were 75-9 to 87-0% identical to one another and the amino acid sequences of the predicted capsid polyproteins were 90-3 to 98-1% identical to one another (Table 4). Of the 854 predicted amino acids in the capsid polyprotein, 758 residues were invariant among all four isolates. Only five of 96 variant sites occurred in regions predicted to participate in conserved β-barrel structures (data not shown), and four of the differences in these sites were conservative changes (A–S, F–Y, I–V and K–R). VP1 was the most divergent capsid protein, varying by up to 12.5% (CA64-4454 versus CA78-1480). Nucleotide identity scores
Table 4. Nucleotide and amino acid sequence relationships (percent identity) among complete capsid sequences of CA55-1988, CA64-4454, CA78-1480 and OMA95-6498

Numbers above the diagonal are percent nucleotide sequence identity and numbers below the diagonal are percent amino acid sequence identity.

<table>
<thead>
<tr>
<th></th>
<th>CA55-1988</th>
<th>CA64-4454</th>
<th>CA78-1480</th>
<th>OMA95-6498</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA55-1988</td>
<td>100</td>
<td></td>
<td>76.0</td>
<td>87.0</td>
</tr>
<tr>
<td>CA64-4454</td>
<td>90.3</td>
<td>100</td>
<td></td>
<td>76.8</td>
</tr>
<tr>
<td>CA78-1480</td>
<td>98.1</td>
<td>90.4</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>OMA95-6498</td>
<td>96.1</td>
<td>91.4</td>
<td>96.0</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 5. Comparison of coding and noncoding sequences of CA55-1988 with those of representative enteroviruses

<table>
<thead>
<tr>
<th></th>
<th>Cluster A</th>
<th>Cluster B</th>
<th>Cluster C</th>
<th>Cluster D</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’NTR (nt 1–744)*</td>
<td>84.1–84.4†</td>
<td>79.9–91.0</td>
<td>68.2–71.3</td>
<td>73.4</td>
</tr>
<tr>
<td>Complete polyprotein</td>
<td>56.7–58.1</td>
<td>86.5–89.2</td>
<td>60.7–61.9</td>
<td>60.4</td>
</tr>
<tr>
<td>P1 (nt 745–3306)</td>
<td>46.0–48.9</td>
<td>71.4–77.7</td>
<td>54.5–55.6</td>
<td>47.8</td>
</tr>
<tr>
<td>P2 (nt 3307–5040)</td>
<td>65.2–65.7</td>
<td>93.8–96.4</td>
<td>57.6–60.8</td>
<td>63.7</td>
</tr>
<tr>
<td>P3 (nt 5041–7311)</td>
<td>61.0–62.7</td>
<td>96.0–98.7</td>
<td>68.3–68.8</td>
<td>72.0</td>
</tr>
<tr>
<td>3’NTR (nt 7312–7411)</td>
<td>43.0</td>
<td>73.0–99.0</td>
<td>41.0–49.0</td>
<td>44.0</td>
</tr>
</tbody>
</table>

* Nucleotide position relative to the complete genome sequence of CA55-1988.
† Numbers are percent nucleotide identity for 5’ and 3’ non-translated regions (NTR) and percent amino acid identity for protein coding regions.

for comparisons of the complete capsid sequences to available enterovirus prototype sequences were similar to those for comparisons of VP1 alone (data not shown).

Complete sequence of CA55–1988

The complete nucleotide sequence of CA55-1988 was 7411 residues long, exclusive of the poly(A) tract (Table 5). The genome organization was typical of the human enteroviruses, with a predicted 744 nucleotide non-translated region at the 5’ end and a 100 nucleotide non-translated region at the 3’ end of the genome. Eight AUG codons occur in the first 750 nucleotides of the genome, but comparison with other cluster B enteroviruses suggested that translation probably begins at nucleotide 745 (eighth AUG) and continues to a termination codon at residues 7309 to 7311, resulting in a primary translation product of 2188 amino acids. The complete polyprotein sequence was compared, by genome region (P1, P2 and P3), with each of the complete enterovirus sequences available in the GenBank database (21 different serotypes). Overall, the predicted polyprotein sequence of CA55-1988 was 86.5 to 89.2% identical to those of other cluster B enteroviruses (13 different serotypes) and 56.7 to 61.9% identical to the polyprotein sequences of enteroviruses belonging to other clusters (Table 5). The CA55-1988 P1 (capsid) region was the most divergent, differing from other cluster B capsid sequences by 22.3 to 28.6%. By contrast, the P2 and P3 (nonstructural) regions were much more highly conserved, with a maximum divergence from other cluster B viruses of 6.2% in P2 and 4.0% in P3.

Discussion

In the past, antigenic tests, including neutralization, haemagglutination and complement fixation, have been used to dissect the relationships among enterovirus isolates (Duncan, 1968; Wenner et al., 1967), to suggest modifications to the taxonomy of enteroviruses (Harris et al., 1973) and to establish new serotypes (Committee on the Enteroviruses, 1957; Melnick et al., 1974; Panel for Picornaviruses, 1963). Molecular sequence data and comparative genomics have recently been applied to the classification of newly discovered picornaviruses (Marvil et al., 1999; Niklasson et al., 1999; Yamashita et al., 1998) and to the reclassification of existing picornaviruses.
(Doherty et al., 1999; Hyypiä et al., 1992). Our previous studies have shown that human enterovirus VP1 sequences correlate with antigenically defined serotypes (Oberste et al., 1999b), and we have successfully applied this observation to the typing of clinical isolates, including those that could not be typed by traditional antigenic methods (Oberste et al., 2000, 1999a). Comparisons of VP1 sequences supported the suggestion, based upon antigenic relatedness (Melnick, 1996), that CA15 and CA18 should be reclassified as variants of CA11 and CA13, respectively (Oberste et al., 1999b). A number of other studies have also demonstrated the utility of using picornavirus capsid region sequences, and VP1 sequences in particular, as a molecular surrogate for antigenic type (Brocchi et al., 1997; Callens & De Clercq, 1997; Hyypiä et al., 1997; Santti et al., 1999). In the present study, sequences encoding the VP1 protein, as well as the complete capsid of the Oman and California isolates, correlated with the antigenic relationships of the viruses to one another. Sequence comparisons also fully supported the antigenic data that indicated CA55–1988 and related strains are distinct from the known human enterovirus serotypes.

Presumably, enteroviral structural diversity, and hence the number of possible serotypes, is constrained by the ability of different primary capsid sequences to form a functional viral particle capable of interacting with and infecting a susceptible host cell. However, the large number of enterovirus serotypes suggests that the viable conformation space may be very large, so that many more serotypes may remain to be discovered, either as historical, previously untyped enteroviruses, or as newly emerging serotypes. In either case, the molecular methods described here will be invaluable in rapidly identifying and characterizing these agents. For example, national enterovirus surveillance figures indicate that 3% of all enterovirus isolates for the years 1970 to 1983 (Strikas et al., 1986), and 3·8% of isolates for 1993 to 1996 (Centers for Disease Control and Prevention, 1997), were reported by state public health laboratories as ‘untyped enterovirus’. In our own collection, more than 12% of isolates from the period 1962 to 1997 were originally reported as untypeable (Centers for Disease Control and Prevention, 1999). This higher figure is partially explained by the fact that the Centers for Disease Control and Prevention is a reference centre and may be expected to receive those isolates that were most difficult to type in the primary clinical or public health laboratory. Our recent pilot study suggests that VP1 sequencing and database comparison will easily identify the vast majority of untypeable isolates as variants of known serotypes (Oberste et al., 2000). In that study, we identified one potentially new serotype, represented by four isolates from three states, isolated between 1985 and 1987, and those isolates could be analysed by methods similar to those outlined in the present study. Broadly reactive VP1-specific primers are available to rapidly screen untypeable isolate collections for potentially novel serotypes (Oberste et al., 2000, 1999a).

Practical criteria must be established before molecular sequence information can be applied routinely to picornavirus identification. A partial or complete VP1 nucleotide sequence identity of at least 75% (minimum 85% amino acid sequence identity) between a clinical enterovirus isolate and serotype prototype strain may be used to establish the serotype of the isolate (Oberste et al., 2000, 1999a, b). These criteria also appear to apply to comparisons among isolates of foot-and-mouth-disease virus (Vosloo et al., 1992), but a study directly comparable to the enterovirus studies has not yet been performed. A best-match nucleotide sequence identity of between 70 and 75% or a second-highest score of greater than 70% may provide a tentative identification, pending confirmation by other means, such as neutralization with monospecific antisera (Oberste et al., 2000) or more extensive sequencing. A best-match nucleotide sequence identity below 70% (less than 85% amino acid sequence identity) may indicate that the isolate represents an unknown serotype. Sequencing of the complete capsid-coding region may be useful in confirming this result, but complete capsid sequences are available for less than half of the known enterovirus serotypes, limiting the utility of complete capsid sequence comparisons. More extensive characterization, possibly including complete genome sequences, may be required for viruses that appear to represent previously unknown genera (Hyypiä et al., 1992; Marvil et al., 1999; Niklasson et al., 1999; Yamashita et al., 1998). Due to the high frequency of recombination among picornaviruses (King, 1988; Kopecka et al., 1995; Santti et al., 1999), sequence information from non-capsid regions is of little value in characterizing new serotypes within known genera.

On the basis of the antigenic and molecular comparisons presented here, we propose that CA55-1988, CA64-4454, CA78-1480, OMA95-6498 and OMA95-6499 be recognized as isolates of a new human enterovirus serotype, ‘enterovirus 73’ (EV73), subject to approval by the appropriate taxonomic authority, and that CA55-1988 be designated as the prototype strain. RT–PCR coupled with amplicon sequencing is a simple and rapid method for the typing and classification of picornaviruses and may lead to the identification of many new picornavirus serotypes. We propose this method as a general approach for the molecular classification of newly discovered picornaviruses and for the reclassification of known serotypes. In light of the increased use of molecular sequence data in virus classification (Calisher et al., 1995; Mayo & Pringle, 1998; Van Regenmortel et al., 1997) and the correlation of picornavirus capsid sequence data with classic antigenic typing, we further recommend that molecular data be given increased weight in the classification of picornaviruses.

The authors thank Marge Dondero of the Viral and Rickettsial Diseases Laboratory, California Department of Health Services, for excellent technical assistance. We also thank Esther de Gourville, Eastern Mediterranean Regional Office of the World Health Organization, for
encouraging the characterization of non-polio enteroviruses isolated from AFP patients and contacts.

References


Received 17 July 2000; Accepted 27 October 2000