Open reading frame sequence of an Asian enterovirus 73 strain reveals that the prototype from California is recombinant

Helene Norder, Lotte Bjerregaard and Lars O. Magnius

Department of Virology, Swedish Institute for Infectious Disease Control, S-171 82 Solna, Sweden

Phylogenetic analysis within the VP1 region now enables molecular typing of enteroviruses consistent with neutralization results. Three untypable isolates, 2776/82, 57/99 and 22/00, from Korea, North India and Bangladesh, respectively, showed within this region 98·0–99·0% amino acid identities. These were less than 77% to the previous enterovirus prototypes, but 91·5–92·5% to CA55-1988, the recently identified enterovirus 73 (EV73) prototype from California. All three strains were, however, most similar to CA64-4454, an EV73 prime strain, to which they shared 96·5–98·5% identity. Seven compared EV73 strains formed two clusters in the VP1 dendrogram, one cluster with strains from South and East Asia and CA64-4454, and the other with strains from Oman and California including the prototype. When sequencing the complete open reading frame of 2776/82, its non-structural region was found to be divergent from all human enterovirus B (HEV-B) strains, including CA55-1988, indicating that one or other strain was recombinant. Boot scanning of the genomes showed a recombination point within the P2 region. Therefore, part of this was sequenced for 57/99 and 22/00 and was found similar to 2776/82, while CA55-1988 was similar to coxsackievirus B3, demonstrating that CA55-1988 was the recombinant. Since all strains of EV73 isolated so far outside California originate from Asia, where it has a broad geographical distribution, it seems that EV73 may have been introduced to California from Asia. Further analysis of EV73 strains will reveal if the recombination occurred in the USA or in Asia and will help to elucidate the origin of this virus.

Introduction

Enteroviruses form one genus of the *Picornaviridae* family of small non-enveloped plus-sense RNA viruses. The genome, 7500 nucleotides in length, has a single open reading frame (ORF), with three main genomic regions designated P1–P3 flanked by untranslated 5′- and 3′-extremities. The P1 region encodes the capsid proteins VP1–VP4, while the P2 and P3 regions encode seven nonstructural proteins with the same functions in different types (involvement in the viral RNA replication and processing of viral proteins). The 5′-noncoding region contains the initiation site for synthesis of the genomic RNA and the internal ribosomal entry site enabling cap-independent translation, while the 3′-noncoding region is

involved in the initiation and synthesis of the complementary RNA strand.

The junction between the P1 region, and the P2 and P3 regions, may be the site for recombination events, which do not change the serotype or affect important virus functions. Such recombinants between wild-type and vaccine strains of polioviruses, as well as between different enterovirus types belonging to the same species, have been described (Furione *et al.*, 1993; Santti *et al.*, 1999; Guillot *et al.*, 2000).

The *Enterovirus* genus comprises 65 types, which are assigned to four genomic groups or species designated human enterovirus (HEV-A through D), based on degree of similarity within the VP2 region (Pöyry *et al.*, 1994, 1996; Dahllund *et al.*, 1995; Pulli *et al.*, 1995; Huttunen *et al.*, 1996; Oberste *et al.*, 1999a). Coxsackie A virus 2–8, 10, 12, 14 and 16 and enterovirus (EV) 71 belong to HEV-A. All echo- and coxsackie B (CB) viruses, coxsackie A9 virus, EV69 and EV73 form HEV-B. Coxsackie A virus 1, 11, 13, 15, 17–22 and 24 form group HEV-C. Poliovirus 1–3 also cluster within HEV-C, although

Author for correspondence: Lars Magnius.

Fax +46 8 32 83 30. e-mail lars.magnius@smi.ki.se

The sequences of strains 2776/82, 57/99 and 22/00 have been deposited in GenBank, accession numbers AF504533–AF504537.

Fig. 1. The branch formed by 105 HEV-B strains in an UPGMA dendrogram based on the 199 amino-terminal codons of the VP1 region of 127 enterovirus strains.

they have distinct receptor usage and unique clinical features of infection. HEV-D is formed by EV68 and EV70 (Pöyry *et al.*, 1996; Hyypiä *et al.*, 1997; Oberste *et al.*, 1999b).

Classical typing of enteroviruses relies on virus neutralization in cell cultures with pools of type-specific polyclonal antisera followed up by neutralization with monospecific antisera. Recently, enterovirus strains belonging to the same type have been shown to share sequence similarity within the VP1 region (Oberste *et al.*, 1999a, b, 2000; Norder *et al.*, 2001). Although, the exact molecular counterpart, i.e. the type-specific amino acid residues of the capsid proteins that define the different enterovirus serotypes, remains to be elucidated, this sequence similarity enables molecular typing of enteroviruses (Oberste *et al.*, 1999a, b; Norder *et al.*, 2001). VP1 sequencing has also been applied to strains not neutralized by available antisera to investigate whether these are just divergent or represent putative new types (Oberste *et al.*, 2000).

A new enterovirus type within HEV-B designated enterovirus 73 (EV73) was recently proposed (Oberste *et al.*, 2001). We report here on three enterovirus isolates with Asian origins showing > 91% sequence identity to EV73. The complete polyprotein of an isolate originating from Korea was deduced by sequencing and compared with that of the EV73 prototype. Part of the P2 region for all three strains was also compared with that of the prototype.

Methods

■ **Virus strains.** Three enterovirus isolates, 2276/82, 57/99 and 22/00, were investigated. 2276/82 was isolated on human lung fibroblasts from stool of a 7-month-old Korean infant without symptoms. 57/99 was isolated on Green Monkey kidney cells from a 23-year-old man hospitalized with myocarditis after a travel in North India, where he fell ill with fever and diarrhoea. 22/00 was isolated on RD cells from a two and a half-year-old girl with a 3 day history of nausea and vomiting after returning from Bangladesh. Typing attempts had been unsuccessful for all three isolates using neutralizing antisera according to current laboratory procedures (Grandien *et al.*, 1995).

■ **Molecular methods.** RNA extraction and reverse transcription

were performed as described (Norder *et al.*, 2001). The VP1 region was amplified by PCR with primers 76C/62C and sequenced in both directions (Norder *et al.*, 2001). The complete ORF of the genome of 2776/82 was amplified with primers 6N (sense; 5' GTACCTTTGTGAGCCTGTT 3') and 3PR (antisense; 5' TAWMGRARAATTTACCCCYAC 3') with the Expand Long Template PCR System, according to the manufacturer (Roche Diagnostics). The amplicate was sequenced by primer walking in both directions using cycle sequencing with fluorescein-labelled dideoxy-chain terminators and the reagents in the BigDye sequencing kit (Applied Biosystems). Primers used in the sequencing reactions are available upon request.

The P2 region of isolates 57/99 and 22/00 was amplified with primers 296 (sense; 5' TGGGNGGHHGARGGNGTIGTIGGNTT 3') and 1471 (antisense; 5' TGACNGGRCARCACTCNCRTCCACA 3'). The amplicate was sequenced in both directions with the primers used in the PCR and primer 1295 (5' GCATYYTGTTYACBTCHCCRTTYGT 3') as sequencing primers.

■ **Sequence analysis.** The VP1 sequences of the three isolates were aligned with those of 127 enterovirus strains belonging to HEV-A through D. The sequence of the complete P1 region for 2776/82 was aligned with those of 54 enterovirus strains retrieved from GenBank. The sequence obtained of the P2 and P3 regions of strain 2776/82 as well as 1188 nucleotides of the P2 region of strains 57/99 and 22/00 were aligned with the corresponding sequence of 51 enterovirus strains retrieved from GenBank. The number of nucleotides and predicted amino acid differences were calculated with the Mega program package, version 1.02 (Kumar *et al.*, 1993). Genetic distances were calculated using the Kimura two-parameter method for the nucleotide sequences and the Dayhoff PAM matrix for the deduced amino acid sequences. The dendrograms were constructed using the UPGMA and neighbour-joining algorithms in the PHYLIP 3.57 program package (Felsenstein, 1993). Bootstrap analysis was performed on 1000 replicas using the programs Seqboot and Consense (Felsenstein, 1993). Similarity and boot-scanning analyses were performed using the SimPlot program (Lole *et al.*, 1999; Ray, 1997) and a boot-scanning program package using the PHYLIP package for phylogenetic analysis (Salminen & Cobb, 1998). The sequences were also analysed using the Recombinant Identification Program (RIP) (Siepel & Korber, 1995).

Results

The VP1 region of the three strains 2776/82, 57/99 and 22/00 was 92% identical to that of the recently described EV73 prototype strain, CA55-1988. The identity was lower,

Table 1. Percent divergence of 199 amino acids of the amino-terminal part of the VP1 region (upper triangle) and of the 854 amino acids forming the P1 region (lower triangle)

	2276/82	57/99	22/00	CA64-4454	CA55-1988	CA78-1480	OMA95-6498
2276/82	—	2·0	2·0	3·5	7·5	8·5	7·5
57/99	ND	—	1·0	1·5	8·5	8·5	7·0
22/00	ND	ND	—	2·5	8·5	8·5	7·5
CA64-4454	5·4	ND	ND	—	8·0	8·0	7·5
CA55-1988	10·2	ND	ND	9·7	—	3·0	4·5
CA78-1480	10·2	ND	ND	9·6	1·9	—	5·0
OMA95-6498	9·7	ND	ND	8·6	3·9	4·0	—

ND, Not determined.

Fig. 2. UPGMA dendrogram based on 854 amino acids of the P1 region in 55 HEV-A through HEV-D strains. The numbers at the branches indicate bootstrap values for 1000 replicas.

up to 77% for echovirus 29, as compared to all other 64 enterovirus prototypes. Our three strains, together with the four previously published EV73 isolates, formed a separate cluster in a dendrogram based on the 199 deduced amino-

terminal codons of the VP1 region for 105 enterovirus strains belonging to HEV-B (Fig. 1). These seven EV73 strains formed two subclusters. One subcluster was formed by the strains CA55-1988, CA78-1480 and OMAN95, and the other

Fig. 3. UPGMA dendrogram based on 4005 nucleotides of the P2 and P3 regions of 31 HEV-B and HEV-C strains. The numbers at the branches indicate bootstrap values for 1000 replicas.

subcluster was formed by 2776/82, 57/99, 22/00 and CA64-4454. The amino acid sequences of the strains within each subcluster varied by from 1.0 to 5.0%, while the strains in

different subclusters diverged by 7.5 to 8.5% (Table 1). All EV73 strains showed 22.1 to 42.2% divergence from other prototype strains within HEV-B.

Fig. 4. For legend see facing page.

In a dendrogram based on the deduced 854 amino-terminal codons of the P1 region, nucleotides 745–3306, strain 2776/82 formed a subcluster with CA64-4454 in a cluster including all four previously published EV73 strains (Fig. 2). The amino acid sequence of this region of 2776/82 diverged by 5.4 to 10.2% from that of other EV73 strains (Table 1), and by 22.1 to 40.7% from the other HEV-B strains.

Strains 2776/82 and CA55-1988 were on separate branches in the cluster formed by the HEV-B strains in a phylogenetic tree based on the P2 and P3 regions, corresponding to nucleotides 3307–7311 (Fig. 3). Strains 2776/82 and CA55-1988 diverged by 20% in this region, and by 17 to 22% and 13 to 21%, respectively, from other HEV-B strains, which diverged by 1 to 23% from each another in this region.

Similarity analysis of the complete P1–P3 regions of strains 2776/82, CA55-1988 and 24 other HEV-B strains showed a similarity between the two EV73 strains in the P1 region, consistent with that found when comparing homotypic viruses. However, in the P2 and P3 regions the similarity between 2776/82 and CA55-1988 was of the same magnitude as that found when comparing heterotypic HEV-B members. On the other hand, CA55-1988 showed 93% similarity with CB3 strain Woodruff in the P2 region, between nucleotides 3700–5300, and 90% similarity with CB3 strain 31-I-93 in the P3 region, between nucleotides 6400–6900. Bootscanning of the P1 through P3 regions showed 100% support for trees with close association of CA55-1988 and 2776/82 in the P1 region. The analysis supported a clustering of CA55-1988 with CB3 strain Woodruff in the P2 region from nucleotide position 3500. There was no support for an association between 2776/82 and any other HEV-B strain. The putative recombination point in CA55-1988 was further located to position 3497 by using RIP.

In a dendrogram based on part of the P2 region, nucleotides 3593–4780, isolates 2776/82, 57/99 and 22/00 formed a separate cluster, while CA55-1988 formed a cluster with CB3 strain Woodruff (Fig. 4). In the compared region isolates 2776/82, 57/99 and 22/00 diverged by 3.9 to 7.5% from each other, by 20.6 to 21.3% from CA55-1988, and 17.2 to 21.3% from other HEV-B members.

Discussion

Sequencing the amino-terminal part of the VP1 region within the enterovirus genome now enables, in addition to molecular typing of enteroviruses, the identification of genomes that represent putative new enterovirus types. The three strains characterized in this study could be classified as EV73 based on degree of sequence similarity with the prototype and phylogenetic analysis within the VP1 region. These three EV73 strains originate from Korea, Bangladesh

and India, while the four recently published strains originate from California and Oman (Oberste *et al.*, 2001), showing a broad geographical distribution of this enterovirus type within Asia. The seven strains formed two subclusters in the dendrogram based on partial VP1 sequences, one with three previously described strains, and another comprising CA64-4454 and the three Asian strains described in this study. An antiserum produced against CA64-4454 has been shown to neutralize the strains from California and Oman, while antisera produced against the other Californian strains did not neutralize CA64-4454, indicating that CA64-4454, as well as our Asian EV73 isolates, might be prime strains of EV73. The finding that all EV73 strains so far isolated outside California originate from Asia, implies that this enterovirus type may not occur worldwide. Since EV73 has not been isolated in Europe before, it is possible that the strains from California also represent imports from Asia. If so, the finding of Asian representatives of both the prototype and the putative prime strain may indicate that these are independent imports from Asia.

The P2 and P3 regions of 2776/82 showed a high divergence from that of CA55-1988. While 2776/82 was divergent from all HEV-B members sequenced so far in these regions, CA55-1988 was most similar to CB3. The difference between the EV73 strains was also evident in the compared part of the P2 region, where the three Asian strains formed a separate cluster, while the prototype strain CA55-1988 clustered with CB3. The P2 and P3 regions may vary significantly between homotypic strains due to recombinations involving different serotypes (Furione *et al.*, 1993; Santti *et al.*, 1999; Guillot *et al.*, 2000). Thus the CA55-1988 strain seems to be a recombinant between a parental EV73 strain and CB3. Since CA78-1480 and OMAN95 cluster with CA55-1988 in the P1 region, it would be of interest to characterize the P2 regions of the former strains, and in particular OMAN95, to elucidate if the recombination event occurred before or after CA55-1988 like strains were imported to California.

The finding that CA55-1988 is a recombinant virus has relevance with regard to CA64-4454 being a prime strain. CA64-4454 cannot have been formed by genetic drift from CA55-1988, since only the latter strain should be recombinant, considering that none of our three Asian isolates in the same genetic subcluster as CA64-4454 was recombinant. The two subclusters of EV73 may represent two different genotypes, the geographical distribution of which should be further investigated and include limited sequencing within VP1 and P2 regions. This will also elucidate if the recombinant strain has spread outside California. Further analysis of EV73 strains will reveal whether the recombination event occurred in the United States or in Asia, and will elucidate the evolutionary history of this virus. The future description of recombinants for other

Fig. 4. UPGMA dendrogram based on 1188 nucleotides of the P2 region in 45 HEV-A through HEV-D strains. The numbers at the branches indicate bootstrap values for 1000 replicas.

enteroviruses may therefore be important to understand the evolution and origins of these viruses in general.

References

- Dahllund, L., Nissien, L., Pulli, T., Hyttinen, V.-P., Stanway, G. & Hyypiä, T. (1995). The genome of echovirus 11. *Virus Research* **35**, 215–222.
- Felsenstein, J. (1993). PHYLIP: phylogeny inference package, version 3.52c. University of Washington, Seattle, Washington, USA.
- Furione, M., Guillot, S., Otelea, D., Balanant, J., Candrea, A. & Crainic, R. (1993). Polioviruses with natural recombinant genomes isolated from vaccine associated paralytic poliomyelitis. *Virology* **196**, 199–208.
- Grandien, M., Forsgren, M. & Ehrnst, A. (1989). Enterovirus and reovirus. In *Diagnostic Procedures for Viral, Rickettsial and Chlamydial Infections*, 6th edn, pp. 513–569. Edited by E. H. Lennette & N. J. Schmidt. Washington, DC: American Public Health Association.
- Guillot, S., Caro, V., Cuervo, N., Korotkova, E., Combiescu, M., Persu, A., Combiescu, A., Delpyroux, F. & Crainic, R. (2000). Natural genetic exchanges between vaccine and wild poliovirus strains in humans. *Journal of Virology* **74**, 8434–8443.
- Huttunen, P., Santti, J., Pulli, T. & Hyypiä, T. (1996). The major echovirus group is genetically coherent and related to coxsackie B viruses. *Journal of General Virology* **77**, 715–725.
- Hyypiä, T., Hovi, T., Knowles, N. J. & Stanway, G. (1997). Classification of enteroviruses based on molecular and biological properties. *Journal of General Virology* **78**, 1–11.
- Kumar, S., Tamura, K. & Nei, M. (1993). Molecular Evolutionary Genetics Analysis Version 1.02. Pennsylvania State University, Pennsylvania, USA.
- Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W. & Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of Virology* **73**, 152–160.
- Norder, H., Bjerregaard, L. & Magnius, L. O. (2001). Homotypic echoviruses share aminoterminal VP1 sequence homology applicable for typing. *Journal of Medical Virology* **63**, 35–44.
- Oberste, M. S., Maher, K. & Pallansch, M. A. (1999a). Molecular phylogeny of all human enterovirus serotypes based on comparison of sequences at the 5' end of the region encoding VP2. *Virus Research* **58**, 35–43.
- Oberste, M. S., Maher, K., Kilpatrick, D. R., Flemister, M. R., Brown, B. A. & Pallansch, M. A. (1999b). Molecular evolution of the human enteroviruses: correlation of serotype with VP1 sequence and application to picornavirus classification. *Journal of Virology* **73**, 1941–1948.
- Oberste, M. S., Maher, K., Flemister, M. R., Marchetti, G., Kilpatrick, D. R. & Pallansch, M. A. (2000). Comparison of classic and molecular approaches for the identification of 'untypable' enteroviruses. *Journal of Clinical Microbiology* **38**, 1170–1174.
- Oberste, M. S., Schnurr, D., Maher, K., al-Busaidy, S. & Pallansch, M. A. (2001). Molecular identification of a new picornavirus and characterization of a proposed enterovirus 73 serotype. *Journal of General Virology* **82**, 409–416.
- Pöyry, T., Hyypiä, T., Horsnell, C., Kinnunen, L., Hovi, T. & Stanway, G. (1994). Molecular analysis of coxsackie A16 reveals a new genetic group of enteroviruses. *Virology* **202**, 982–987.
- Pöyry, T., Kinnunen, L., Hyypiä, T., Brown, B., Horsnell, C., Hovi, T. & Stanway, G. (1996). Genetic and phylogenetic clustering of enteroviruses. *Journal of General Virology* **77**, 1699–1717.
- Pulli, T., Koskimes, P. & Hyypiä, T. (1995). Molecular comparison of coxsackie A virus serotypes. *Virology* **211**, 30–38.
- Ray, S. C. (1997). SimPlot for Windows 95/NT, version 1.2.2 (distributed by author). <http://sray.med.som.jhmi.edu/RaySoft>.
- Salminen, M. O. & Cobb, W. (1998). Boot scanning package for Unix/Linux version 1. National Public Health Institute, Helsinki, Finland.
- Santti, J., Hyypiä, T., Kinnunen, L. & Salminen, M. (1999). Evidence of recombination among enteroviruses. *Journal of Virology* **73**, 8741–8749.
- Siepel, A. C. & Korber, B. T. (1995). Scanning the database for recombinant HIV-1 genomes. *Human Retroviruses and AIDS Compendium III*, 35–60.

Received 12 November 2001; Accepted 14 March 2002