Analysis of the complete genome sequence of epidemic keratoconjunctivitis-related human adenovirus type 8, 19, 37 and a novel serotype

Hisatoshi Kaneko,1 Tomohiro Iida,2 Hiroaki Ishiko,3 Takeshi Ohguchi,4 Toshihide Ariga,4 Yoshitsugu Tagawa,4 Koki Aoki,4 Shigeaki Ohno4 and Tatsuo Suzutani1

Correspondence
Hisatoshi Kaneko
h-kane@chive.ocn.ne.jp

1Department of Microbiology, Fukushima Medical University School of Medicine, Fukushima, Japan
2Department of Ophthalmology, Fukushima Medical University School of Medicine, Fukushima, Japan
3Host Defense Laboratory, Mitsubishi Chemical Medience Corporation, Tokyo, Japan
4Department of Ophthalmology and Visual Sciences, Hokkaido University Graduate School of Medicine, Sapporo, Japan

Received 27 November 2008
Accepted 10 February 2009

We determined the complete genome sequence of epidemic keratoconjunctivitis (EKC)-related human adenoviruses (HAdVs). We analysed a total of 12 HAdV strains; three prototype strains and two HAdV-8, three HAdV-19 and three HAdV-37 clinical isolates from EKC patients in Japan, and one novel serotype of HAdV. Genome organization of these serotypes was identical to those of the recently determined HAdV-19 and HAdV-37. The identities of the whole genome were over 99% among strains from the same serotype, except for HAdV-19p, which is not associated with conjunctivitis, resulting in the formation of a distinct cluster in the phylogenetic analysis. The penton, loop 1 and loop 2 of hexon, early region 3 (E3) and fiber were hypervariable regions between serotypes. Results suggest that the HAdV-19 clinical strain is a recombinant of HAdV-19p-like and HAdV-37-like strains, and that the acquisition of the penton, E3 or fiber may be related to ocular tropism.

Adenoviruses are nonenveloped, double-stranded DNA viruses with icosahedral capsids (Swenson et al., 2003). Human adenoviruses (HAdVs) belong to the genus *Mastadenovirus* of the family *Adenoviridae* and are classified into six species, A to F (HAdV-A to HAdV-F) (Benkö et al., 2000; Wold & Horwitz, 2007). Adenoviral conjunctivitis is mainly caused by HAdV-3 (in HAdV-B), HAdV-4 (in HAdV-E), and HAdV-8, HAdV-19 and HAdV-37 (in HAdV-D), with the three HAdV-D serotypes being known to cause epidemic keratoconjunctivitis (EKC). HAdV-8 is the original causative agent of EKC and remains the predominant HAdV serotype isolated in association with EKC in many countries (Ishii et al., 1987; Chang et al., 2001; Vainio et al., 2001; Aoki & Tagawa, 2002; Jin et al., 2006). In Japan, although HAdV-8 and HAdV-19 have been described, a novel serotype of HAdV recently isolated from EKC patients and HAdV-37 are the predominant causative serotype of EKC in Japan (Higuchi et al., 1987; Aoki & Tagawa, 2002; Ishiko et al., 2008; Kaneko et al., 2008).

Nucleotide polymorphisms in HAdV strains isolated from EKC patients can be classified into discrete genotypes within a specific HAdV serotype on the basis of their restriction endonuclease cleavage pattern (Wadell et al., 1980; Adrian et al., 1986). It has been speculated that the appearance of new genotypes might contribute to the incidence of outbreaks of each serotype (Aoki & Tagawa, 2002; Ariga et al., 2005). DNA sequence analysis has allowed us to appreciate the molecular evolution of HAdV in greater detail, and revealed that the penton, hexon and fiber genes were the most variable among the different serotypes (Pring-Akerblom & Adrian, 1995; Arnberg et al., 1997; Ebner et al., 2005; Madisch et al., 2005, 2007; Miura-Ochiai et al., 2007). The complete genome sequences of 24 HAdV serotypes (HAdV-1, 2, 3, 4, 5, 7, 9, 11, 12, 14, 16, 17, 19, 21, 26, 34, 35, 37, 40, 41, 46, 48, 49 and 50) have now been determined.

In this study, we describe the complete genome sequences of the prototype strains HAdV-8p, HAdV-19p and HAdV-37p together with a novel HAdV serotype and eight clinical

The GenBank/EMBL/DDBJ accession numbers for the human adenovirus complete sequences determined in this study are AB448767–AB448778.
isolates associated with EKC: two HAdV-8, three HAdV-19 and three HAdV-37.

While the HAdV-19p strain was not associated with conjunctivitis (Wadell & De Jong, 1980), the clinical strains described were all isolated from EKC patients in Japan. The HAdV-8 clinical strains were HAdV-8b and HAdV-8e, and belonged to different genotypes (Adrian et al., 1990; De Jong et al., 1992). The clinical strains of HAdV-19 and HAdV-37 were HAdV-19a, HAdV-19/1997 and HAdV-19/2001, and HAdV-37/1991, HAdV-37/1996 and HAdV-37/2004, respectively. The numbers after the slash for each clinical strain indicate the year of isolation. The HAdV-19a strain has been classified into the most common HAdV-19 genotype, which infects the conjunctiva and frequently causes EKC (Wadell & De Jong, 1980; Tanaka-Yokogui et al., 2001). For analysis of the novel HAdV serotype, the Kobe-H strain, which was isolated in 2000 and might be serologically and phylogenetically characterized as a new serotype, was used (Ishiko et al., 2008). All viruses were cultured in A549 cells and the serotype of clinical isolates identified by neutralization test with type-specific antisera.

Sequencing was carried out using a PCR-directed sequencing method (Kaneko et al., 2005). The primers for PCR and sequencing were designed with reference to the complete genome sequences of HAdV-9 (GenBank accession no. AJ854486) and HAdV-17 (AF108105). To determine the sequences of both ends of the genome, we extracted whole viral DNA from HAdV-infected cells (Shinagawa et al., 1983), and performed a sequencing reaction using the extracted viral DNA as template. The genome size and G+C content of five HAdV strains are shown in Table 1. With the exception of HAdV-19p, the genome size and G+C content among strains from the same HAdV serotype were found to be almost identical.

Multiple alignments were performed and analysed using CLUSTAL W software. The phylogenetic analysis showed the HAdVs formed six major clusters, A to F, and the tree closely corresponded to those obtained from the analyses of alignments of partial regions, such as the hexon, penton and fiber regions (Madisch et al., 2005, 2007; Miura-Ochiai et al., 2007; Robinson et al., 2008). All 12 strains of HAdV-8, HAdV-19, HAdV-37 and the novel HAdV were segregated into the D cluster, where there are distinct subclusters containing strains belonging to the same serotype (Fig. 1). However, HAdV-19p strain did not form a monophyletic cluster with the other three HAdV-19 clinical strains. The novel HAdV showed the highest identity (95.53 %) with HAdV-8p over the complete genome sequence and was clustered with three HAdV-8 strains in the phylogenetic analysis; however, within this cluster, the lineage of the novel HAdV differed from that of the three HAdV-8 strains.

The genome organization of the 12 strains was determined by analyses of their open reading frames (ORFs) and alignment of their nucleotide sequences. The layout of the

Table 1. General properties of the genome of HAdV-8p, HAdV-19p, HAdV-19a, HAdV-37p and the novel HAdV, and the complete genome nucleotide sequence identities among these strains

<table>
<thead>
<tr>
<th>Virus</th>
<th>Genomic size (bp)</th>
<th>DNA G+C content (%)</th>
<th>Identities</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAdV-8p</td>
<td>34980</td>
<td>54.59</td>
<td>-</td>
</tr>
<tr>
<td>HAdV-19p</td>
<td>35125</td>
<td>56.91</td>
<td>-</td>
</tr>
<tr>
<td>HAdV-19a</td>
<td>35231</td>
<td>56.57</td>
<td>-</td>
</tr>
<tr>
<td>HAdV-37p</td>
<td>35215</td>
<td>56.57</td>
<td>-</td>
</tr>
<tr>
<td>Novel HAdV</td>
<td>34919</td>
<td>54.92</td>
<td>-</td>
</tr>
</tbody>
</table>

Downloaded from www.microbiologyresearch.org by
On: Wed, 12 Dec 2018 04:32:39
early and late regions in the complete sequences of these strains is shown in Fig. 2(a). The number of genes and their organization for each of the 12 strains were identical to those of the recently determined HAdV-19 and HAdV-37 sequences (Robinson et al., 2008, 2009), though the size of some ORFs varied slightly between strains or serotypes (data not shown).

We surveyed the complete nucleotide sequences using similarity plot analysis to find variable regions within the entire genome, as described previously (Ishiko et al., 2008; Aoki et al., 2008). The similarity between sequences was calculated for each window of 200 nucleotides. The window was successively advanced along the genome alignment in 20-nucleotide increments. The identities among strains of the same serotype were high with little variability over the entire genome (Fig. 2b), except for HAdV-19p. Most of the variations observed among the same serotype were nucleotide substitutions randomly distributed throughout the entire genome. In general, the identities of the penton, loop 1 and loop 2 of hexon, early region 3 (E3) and fiber genes between the different serotypes were noticeably low (Fig. 2c, d and e). However, the sequences of the E3 and fiber genes between the HAdV-19 clinical strains and HAdV-37 were almost identical (>99%) and identical (100%), respectively (Fig. 2g). The identities of the E3 and fiber genes between HAdV-8 and the novel HAdV were higher than those among the other serotypes (Fig. 2h). These results indicate that evolutional changes can be introduced over the whole genome, but that these four hypervariable regions might be hot spots for mutations or be targeted by evolutional pressures. The similarity plots between HAdV-37p and HAdV-4 (GenBank accession no. AY487947) in HAdV-E are shown in Fig. 2(i). The identity over the entire genome was much lower than that between members of the same species, and large differences were observed in the E3 regions in particular. A similar result was commonly observed between HAdVs of different species (Robinson et al., 2008). HAdV-52, the candidate strain of the new HAdV serotype (Jones et al., 2007), also showed a similar pattern (data not shown).

HAdV-19p, which is not associated with conjunctivitis, showed a slightly lower identity to and did not form a monophyletic cluster with the other three HAdV-19 clinical strains, though the three HAdV-19 clinical strains were almost identical (99.99%) over the entire genome (Table 1 and Fig. 1). These results are in agreement with those from partial sequence analyses (Miura-Ochiai et al., 2007; Aoki et al., 2008). The complete genome sequences of the HAdV-19 clinical strains were highly homologous with those of HAdV-37, except loop 1 and loop 2 of the hexon and penton genes (Fig. 2g). The loop 1 and loop 2 of the three HAdV-19 clinical strains were almost identical to those of HAdV-19p (Fig. 2f); therefore, a neutralization test was performed that confirmed their classification as HAdV-19 strains. However, the penton gene in HAdV-19 clinical strains showed a much lower identity to both HAdV-19p and HAdV-37. These results suggest that the
Fig. 2. The genome organization of the complete HAdV-8, -19, -37 and novel HAdV genomes (a). The grey horizontal bar in the centre shows the linear double-strand HAdV genome with each vertical line representing 5000 bp. Transcription units are shown by black arrows relative to their position and orientation in the HAdV genome. Grey arrows show the positions of the penton, hexon and fiber genes. The similarity plots of the complete genome sequences between HAdV-37p and HAdV-37/1996 (b), HAdV-8p and HAdV-19a (c), HAdV-8p and HAdV-37p (d); between HAdV-19p and HAdV-37p (e), HAdV-19p and HAdV-19a (f); between HAdV-19a and HAdV-37p (g), HAdV-8p and the novel HAdV (h); and between HAdV-37p and HAdV-4 (GenBank accession no. AY487947) (i) were generated using SimPlot 3.5.1. The vertical axis indicates the nucleotide identities, expressed as a decimal. The horizontal axis indicates the nucleotide positions of the complete sequences. Double-headed arrows in (g) show the crossover sites between HAdV-19a and HAdV-37p.
HAdV-19 clinical strains isolated from conjunctivitis patients might be a recombinant between HAdV-19p-like and HAdV-37-like strains. To date, three recombinant HAdV strains from conjunctivitis patients have been reported in Japan and Germany (Noda et al., 1991; Engelmann et al., 2006; Aoki et al., 2008). Moreover, the possibility that HAdV-19a might be a recombinant of HAdV-19p and HAdV-37 has been reported on the basis of partial sequences analysis and because of the crossover within the DNA-binding protein (DBP) (Blusch et al., 2002). Our results appear to support their hypothesis. On the basis of similarity plot analysis for the complete sequences, the identities from the 5′-end of the genome to within the 52/55K protein and from the DBP to the 3′-end of the genome were almost identical (the fiber gene, in particular, was completely identical) among the HAdV-19 clinical strains and HAdV-37 strains (Fig. 2g), and the crossover points were suggested to be located within the 52/55K protein and the DBP. HAdV-19p is not an ocular strain but its derivative, HAdV-19a, acquired ocular infectivity by acquiring the penton, E3 and fiber. This suggests that one or more of these three regions are related to ocular tropism.

Partial hexon sequence analysis of the novel HAdV strain showed that it closely resembled HAdV-8, first suggesting it to be a HAdV-8 variant (Ishiko et al., 2008). On the basis of the nucleotide sequence identities and phylogenetic analysis, our study also showed that this strain was much closer to the HAdV-8 strains over the entire genome. Recently, a number of novel HAdV strains were isolated from many EKC patients in Japan instead of HAdV-8. We speculate that the novel HAdV strain in this study might have evolved from HAdV-8.

Three of the four hypervariable regions on the HAdV genome encode antigenic determinants within the capsid proteins penton, hexon and fiber. Our analysis also found these three genes to be highly variable among the different HAdV serotypes, only E3 was as variable. Loop 1 and loop 2 of hexon reacted with type-specific antisera in neutralization assays (Gall et al., 1998), and were found to be hypervariable between different serotypes yet conserved among strains from the same serotype (Fig. 2). The fiber and penton base capsid proteins are the major players in adenovirus cell entry. The lengths of the loops in the penton base vary significantly between HAdV species and serotypes (Arnberg et al., 2000; Madisch et al., 2007). Our results showed that penton was a serotype-specific hypervariable region. On the other hand, the knob region of the fiber has haemagglutinating properties that are employed in haemagglutination inhibition (HI) tests. However, HI tests cannot be used to differentiate between all 51 HAdVs because of several cross-reactions (Swenson et al., 2003). This result implies that identical fiber knob sequences might exist among serotypes (Pringle & R. B. Wickner. San Diego: Academic Press. 1995; Arnberg et al., 1997), and our results confirm that this is indeed the case for the fiber gene among the HAdV-19 clinical strains and HAdV-37 strains.

Recent reports have suggested that the E3 regions are variable and informative components of the adenovirus genome due to unique differences in the E3 region between serotypes (Blusch et al., 2002; Ursu et al., 2004). The number or size of ORFs, particularly in the E3 region, varied according to species and the identity of the E3 regions was low. The E3 region of the members of species HAdV-D contains eight genes. By comparing sequences, E3 was also found to be a serotype-specific hypervariable region, although complete sequences of the E3 region in all HAdV serotypes have not yet been determined. In particular, three ORFs (23K, 49K and 31.6K proteins) in the E3 showed low identity by similarity plots (Fig. 2). E3 encodes the immunomodulatory functions; however, the detailed function of these ORFs with regard to ocular infection and the reason for the diversity among serotypes remain unknown.

In conclusion, the determination of serotype can be achieved by analysis of the penton, loop 1 and loop 2 of hexon, fiber and E3 sequences. In addition to the other three genes, the E3 region was identified as a hypervariable region between HAdV serotypes and further studies are recommended to clarify the functions of the genes in the E3 region in the pathogenesis and evolution of HAdV.

References


