Characterization and manipulation of the human adenovirus 4 genome

Susan C. Jacobs, Andrew J. Davison, Sharon Carr, Alice M. Bennett, Robert Phillips and Gavin W. G. Wilkinson

Human adenovirus 4 (HAdV-4) and HAdV-7 were identified to cause epidemics of acute respiratory disease in US military recruits in the 1950s (Hilleman & Werner, 1954; Rubin & Rorke, 1994). In response, live HAdV-4 and HAdV-7 vaccines were developed that are administered orally in a tablet that consists of lyophilized virus encapsulated in an enteric coat. This coat allows the tablet to be handled safely, prevents infection of the pharynx during administration and also protects the virus during passage through the stomach. The vaccines are thus administered directly to the intestine, where the virus replicates, and elicit a comprehensive, protective immune response without inducing disease. The HAdV-4 and HAdV-7 strains that are utilized are tissue culture-adapted viruses that lack specific markers of attenuation (Chanock et al., 1966; Rubin & Rorke, 1994). In recent years, disruptions in vaccination due to logistical reasons have led to resurgences of disease (Barraza et al., 1999; McNeill et al., 1999). The live vaccines have been shown to be safe and efficacious over a period of up to 40 years and constitute the most successful application of an adenovirus in a therapeutic context that has been achieved thus far (Rubin & Rorke, 1994).

Vaccine strains have obvious potential as vehicles for delivery and expression of foreign antigens. Replication-competent HAdV-4 and HAdV-7 recombinants that encode human immunodeficiency virus or hepatitis B virus envelope proteins have been evaluated in animal models and limited human trials (Lubeck et al., 1989, 1994; Mason et al., 1990; Chengalvala et al., 1991, 1994, 1997; Tacket et al., 1992; Patterson et al., 2002). Encouraging results were obtained following oral administration in chimpanzees and humans. Safety indices were high, active virus replication was observed, immune responses were generated against the vector and transgene and there was evidence of enhanced responses following sequential delivery of different serotypes.

HAdV-4 and four simian viruses (SAdV-22–SAdV-25) comprise HAdV-E, one of the six species (HAdV-A–HAdV-F) that encompass primate adenoviruses (Benkö et al., 2000). HAdV-4 strains can be differentiated into two groups on the basis of minor differences in restriction endonuclease digest patterns; the prototype strain [RI-67, deposited in the American Type Culture Collection (ATCC) as VR-4] and vaccine strain (CL 68578) are members of the same group (Li & Wadell, 1988). DNA sequence data are available for several loci in a variety of HAdV-4 strains (Tokunaga et al., 1982, 1986; Hay, 1985; Kitchingman, 1985; Gruber et al., 1993; Kidd et al., 1995; Pring-Aåkerblom et al., 1995; Crawford-Miksza et al., 1999; Tarassishin et al., 1999; Li & Wold, 2000). Here, we report the complete DNA sequence of the vaccine strain, thus defining a therapeutic agent and enabling researchers to access a genome sequence for at least one human adenovirus from each of the six species. We also describe a system for construction of replication-competent HAdV-4 recombinants.

The GenBank/EMBL/DDBJ accession number for the HAdV-4 genome reported in this paper is AY487947.
HAdV-4 strain CL 68578 was cultured directly from the vaccine, strain RI-6 was kindly supplied by Professor R. T. Hay (University of St Andrews, UK) and strain VR-4 was obtained from the ATCC. These viruses were propagated and titrated in the cell line A549 (ECACC). Virus particles were purified from infected cells by extraction with Arkline P and CsCl gradient centrifugation (Scarpini et al., 1999) and then treated with RNase A. Virion DNA was purified by using a QIAamp DNA Blood kit (Qiagen) and sequenced directly by using customized primers that were based initially on published sequences. By using an ABI 377 instrument, both DNA strands were sequenced except for the terminal 300 bp sections, which were sequenced only in the direction towards the termini. Sequences were aligned and compiled by using the Lasergene programs (DNASTAR) and analysed by using the GCG programs (Accelrys), CLUSTALW (Thompson et al., 1994) and PTRANS (Taylor, 1986).

The genome of the HAdV-4 vaccine strain is 35 994 bp in length and has a G+C content of 57.7 mol%. It is thus marginally larger than the genomes of other characterized human adenoviruses and smaller than that of the chimpanzee adenoavirus SAdV-25 (36 519 bp), which is the closest relative of HAdV-4. Analysis of the HAdV-4 sequence revealed that its genetic content is similar to that of other human adenoviruses (Fig. 1a). Comparisons with published data for various HAdV-4 genes indicated very few differences, most of which can be attributed to strain variation and the remainder to previous errors. Differences between the published sequences of the inverted terminal repeat (ITR) for strains RI-6 and VR-4 were confirmed independently (Table 1). The corresponding vaccine strain sequence is identical to that of RI-6. The two sequences that are available for the SAdV-25 ITR also differ (Table 1).

As one of the two genome sequences for SAdV-25 (Farina et al., 2001; GenBank accession no. AF394196) is of poor quality (Davison et al., 2003b), the other (AR101859, reannotated as BK000413) was utilized for comparative purposes. The HAdV-4 and SAdV-25 genomes are highly conserved, particularly the portion to the left of the E3 region (Fig. 1b). However, two genes are characterized by strikingly divergent regions: hexon, which is known to

![Fig. 1.](image)

<table>
<thead>
<tr>
<th>HAdV-4 strain</th>
<th>Terminal sequence* (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vaccine</td>
<td>CATCATCAATAATACCTTA...</td>
</tr>
<tr>
<td>RI-6†</td>
<td>CATCATCAATAATACCTTA...</td>
</tr>
<tr>
<td>VR-4‡</td>
<td>CATCTATATACTACCTTA...</td>
</tr>
<tr>
<td>SAdV-25§</td>
<td>CACATCCTAAATATACCTCA...</td>
</tr>
<tr>
<td>SAdV-25</td>
<td></td>
</tr>
</tbody>
</table>

*The terminal 21–22 bp are shown; residues that differ from the vaccine strain are in bold. The remainder of the ITR is conserved between HAdV-4 strains, and also between SAdV-25 strains. The HAdV-4 and SAdV-25 ITRs exhibit several additional differences.

†From Hay (1985). Confirmed in this study.
‡From Tokunaga et al. (1986). Confirmed in this study.
§From GenBank accession no. AR101859.
||From GenBank accession no. AF394196 (Farina et al., 2001).
contain several hypervariable regions (Crawford-Miksza & Schnurr, 1996), and E3 CR1-β1.

Comparative aspects of the E3 region in human adenovirus species have been summarized by Davison et al. (2003a). The E3 region contains three genes (RID-x, RID-β and 14.7K) at its 3' end and, in all but HAdV-F, a single gene (12.5K) at its 5' end. Several glycoprotein genes are located in between; the numbers range from two in HAdV-A to five in HAdV-E. Many of the proteins that are encoded by these genes are related to each other via a loosely conserved immunoglobulin-like domain (CR1), although it appears that this domain has been lost in some instances. SAdV-25 contains four CR1 genes, whereas HAdV-4 contains three, i.e. it lacks a counterpart to CR1-β1. A small, unrelated gene was assigned previously to the 328 bp sequence between HAdV-4 CR1-β1 and CR1-δ1 (Li & Wold, 2000; Burgert et al., 2002). This assignment is unlikely to be correct, however, as it is clear that an alternative reading frame contains a 5'-truncated version of CR1-γ1 that represents a deletion relative to SAdV-25 (Fig. 1a). This vestigial element is unlikely to be expressed, as it lacks an initiating ATG and splice acceptor site. Like CR1-β1, it has diverged significantly from the corresponding region of its SAdV-25 counterpart. The deletion in CR1-γ1 is not restricted to the vaccine strain, as it is also a feature of other strains, including VR-4 (Li & Wadell, 1988; Li & Wold, 2000; Burgert et al., 2002). Rapid evolution of CR1 genes is evident in human adenoviruses from their variation in copy number and sequence; this evolution is seen here in the striking divergence of CR1-β1 (and probably CR1-γ1 prior to the deletion event) and the loss of CR1-γ1 function from HAdV-4. The CR1 family is likely to be involved in aspects of immune modulation, as is the case for other E3 genes.

Most genes in HAdV-4 and SAdV-25 exhibit highest similarity to those of HAdV-B (Kitchingman, 1985; Tokunaga et al., 1986; Gruber et al., 1993; Kidd et al., 1995; Pring-Åkerblom & Adrian, 1995). However, Gruber et al. (1993) observed that HAdV-4 was related most closely to HAdV-C in the fibre gene. This prompted the suggestion that HAdV-4 evolution may have involved recombination between ancestors of HAdV-B and HAdV-C. Our analysis (not shown) supported a closer relationship between HAdV-4 and HAdV-C for this gene. However, all other genes have a closer relationship to HAdV-B, although a few (including those that are immediately adjacent to the fibre gene) are too small to permit convincing discrimination. It is possible, therefore, that the proposed recombination event involved only the fibre gene as a major determinant of tropism.

Characterization of the HAdV-4 sequence facilitated the generation of replication-competent recombinants. A transfer plasmid (pAL605) was constructed to permit insertion of exogenous sequences at a site between the E4 promoter and the ITR, a region where disruptions were considered unlikely to be deleterious to virus growth. Briefly, to generate pAL605, a synthetic HAdV-4 DNA duplex, extending from the SalI site at nt 35 855 to the right genome terminus (nt 35 994), was subcloned into vector pSP70 (Promega) by using HindIII–EcoRI linkers. A HAdV-4 Bcl–SalI fragment (nt 31 459–35 855) was inserted at the SalI site to generate pAL422, which thus contains an insert that corresponds to nt 31 459–35 994. Next, the enhanced green fluorescent protein (EGFP) gene, under the control of an internal ribosomal entry site (IRES), was excised from pIRE2–EGFP (Clontech) and flanked with loxP sites; the resulting construct was placed under the control of the human cyto-megalovirus (HCMV) major immediate–early promoter (IIEP; residues −299 to +69; Jacobs et al., 1992; Wilkinson & Akrigg, 1992) and this was inserted into the unique SalI site in pAL422 to generate pAL605. The expression cassette thus comprised HAdV-4 (nt 31 459–35 855)–HCMV IEP–loxP–IRES–EGFP–loxP–HAdV-4 (nt 35 856–terminus). The unique BamHI site between the HCMV IEP and loxP site is available for transgene insertion (Fig. 2a).

The replication-competent HAdV-4 recombinant that encodes EGFP under the control of the HCMV IEP was generated by cotransfecting the transfer vector pAL605 with HAdV-4 DNA (Fig. 2a). Linearized pAL605 was cotransfected with HAdV-4 vaccine strain DNA into A549 cells by using the cationic transfection reagent DAC30 (Eurogentech). Virus was harvested at 6–10 days post-transfection, extracted with Arklone P and titrated in A549 cells under an agarose overlay. Recombinant virus (RAd605) was detected by fluorescent green plaques (Fig. 2b) and this property was used to plaque-purify the recombinant virus. Analysis of RAd605 genomic DNA by restriction endonuclease analysis and DNA sequencing demonstrated that the EGFP gene was inserted at the correct locus (not shown).

Generation of RAd605 demonstrated that the EGFP gene could be used as a selectable marker for the production of HAdV-4 recombinants. However, the EGFP gene was flanked by loxP sites in the transfer vector so that it could be excised specifically from HAdV-4 recombinants, leaving a transgene under the control of the HCMV IEP. The IRES and EGFP gene were removed by treating 1 μg RAd605 DNA with 1 U Cre recombinase (Novagen) at 37°C for 1 h. The enzyme was inactivated by heating to 70°C for 5 min and the reaction product was transfected into A549 cells. Non-fluorescent virus (RAd605ΔEGFP) occurred at a frequency of 0.5 and was subjected to three rounds of plaque purification. Viral DNA was analysed by restriction endonuclease analysis and DNA sequencing (not shown). RAd605ΔEGFP contains the HCMV IEP and a polyadenylation signal, but no EGFP gene.

To investigate whether transgene insertion affected the in vitro growth properties of the HAdV-4 recombinants, A549 cells grown in a 25 cm² tissue-culture flask were infected at an m.o.i. of 0.1 for 4 h with the Ad4 vaccine strain, RAd605 or RAd605ΔEGFP. Both cells and supernatant were harvested at a range of times post-infection and subjected to three rounds of freeze–thawing. The
supernatant was clarified by low-speed centrifugation, aliquotted and stored at −70°C; titres were determined subsequently by plaque assay. Growth kinetics of the HAdV-4 vaccine strain, RAd605 and RAd605ΔEGFP were similar (Fig. 2c), a result that is consistent with the theory that insertion at this locus is compatible with efficient HAdV-4 replication in vitro.

Successful experience with the HAdV-4 vaccine and generation of recombinants derived therefrom provide a solid basis for the development of replication-competent HAdV-4 vectors for use as vaccine carriers or in gene therapy. Although the EGFP gene is a useful marker, its inclusion would reduce the size capacity for additional transgenes. However, an alternative strategy is feasible, which would utilize the loss of EGFP expression from RAd605 by cotransfection with a transfer plasmid that contains an alternative transgene in place of the EGFP gene. This approach has been used to generate a replication-competent HAdV-4 recombinant that encodes the Venezuelan equine encephalitis virus glycoprotein complex E3–E2–6K (not shown). Larger transgenes could be accommodated by the deletion of sequences elsewhere in the genome, but this is likely to alter virus viability and antigen presentation in vivo (Tacket et al., 1992; Chengalvala et al., 1997; Patterson et al., 2002).

Fig. 2. Generation and characterization of the replication-competent HAdV-4 vector. (a) The HAdV-4 recombinant RAd605 was generated by homologous recombination. The linearized plasmid pAL605 was cotransfected into A549 cells with HAdV-4 DNA. Following recombination between the DNA molecules, replication-competent HAdV-4 with the HCMV IEP EGFP expression cassette inserted at the SalI site (nt 35855) was generated. (b) Fluorescence image of a RAd605 plaque in A549 cells showing EGFP expression at 5 days post-infection (p.i.). (c) Multiple-step growth-curve analysis of the HAdV-4 vaccine strain and recombinants derived therefrom. A549 cells (4·9 × 10⁵) were infected at an m.o.i. of 0·1 with the vaccine strain (○), RAd605 (□) or RAd605ΔEGFP (▲). Values at the 0 h time point correspond to total virus inocula that were retitrated from the diluents used. Samples harvested at the time points indicated were titrated on A549 cells; values correspond to total virus yield from the cultures.
Acknowledgements

We are grateful to William Wold for providing sequence data for the E3 region of HAdV-4 strain VR-4 prior to publication. Carole Rickards, Sian Llewellyn-Lacey and the MRC Cooperative Group Tissue Culture Facility (UWCM) provided invaluable advice and technical support. This work was facilitated by funding to G. W. G. W. from the Wellcome Trust and Cancer Research UK.

References


