Characterization of the haemorrhagic enteritis virus genome and the sequence of the putative penton base and core protein genes

M. T. Jucker, J. R. McQuiston, J. V. van den Hurk, S. M. Boyle and F. W. Pierson*

Center for Molecular Medicine and Infectious Diseases, Virginia-Maryland Regional College of Veterinary Medicine, VPI and SU, Blacksburg, VA 24061-0442, USA

Haemorrhagic enteritis virus (HEV) is a member of a genetically ill-defined group within the genus Aviadenovirus which causes significant clinical disease in gallinaceous fowl. Using DNA obtained from a low virulence isolate of HEV passed in turkeys, we developed a genomic restriction map and estimated an apparent genomic length of 25.5 kb. No evidence for extensive DNA hybridization was found between the HEV genome and either the hexon or penton base genes of human adenovirus 2 (HAdV-2) and fowl adenovirus 10 (FAdV-10). The HEV penton base gene was identified by PCR using primers based on conserved adenoviral DNA sequences. The penton base gene was expressed in Escherichia coli as a fusion protein and detected by anti-HEV serum in both colony and denaturing gel immunoblots. DNA sequencing revealed a putative penton base ORF with a predicted amino acid sequence showing approximately 39.0%, 53.0% and 44.2% similarity with the penton base of HAdV-2, human adenovirus 40 (HAdV-40) and FAdV-10, respectively. The penton base gene was located at 43.3-48.6 m.u. on the HEV genome and had a remarkably low G + C content (33.8%). DNA sequencing also revealed ORFs for putative core proteins resembling pVII, p-mu and a partial ORF similar to pVI (hexon-associated protein) of HAdV-2 and HAdV-40. The results support the claim that HEV represents a distinct group of viruses within the genus Aviadenovirus.

Introduction

Haemorrhagic enteritis virus (HEV) causes a disease of economic significance in turkeys, characterized by bloody diarrhoea, splenomegaly and death (Domermuth & Gross, 1991). HEV possesses a non-enveloped, icosahedral capsid, with characteristic hexon, penton base and penton fibre proteins (van den Hurk, 1992) and it has been classified as an adenovirus based mostly on morphological criteria (Tolin & Domermuth, 1975). The avian adenoviruses (family Adenoviridae) have been assigned to the genus Aviadenovirus on the basis of shared host range and lack of antigenic similarity with the mammalian adenoviruses, which constitute the genus Mastiadenovirus (Norrby et al., 1976). HEV, marble spleen disease virus and splenomegalgy virus are serologically indistinguishable isolates constituting what has been historically referred to as the group II avian adenoviruses (Domermuth & Gross, 1991). To date, these viruses have shown no significant serological similarity to other avian or mammalian adenoviruses (Domermuth & Gross, 1991). However, they do possess a single fibre at each vertex (van den Hurk, 1992), as is the case with the mastiadenoviruses. The only genetic characterization of the group has been a restriction fragment fingerprint study (Zahng & Nagaraja, 1989). The other major group within the genus Aviadenovirus has been historically referred to as the group I avian adenoviruses and is classically represented by fowl adenoviruses 1-12 (McFerran, 1991). These viruses are serologically cross-reactive, have a vertex penton consisting of two fibres and a trimeric base and show little similarity to the mastiadenoviruses at the DNA level (Aleström et al., 1984).

Live vaccines against haemorrhagic enteritis in turkeys have been developed which contain either naturally occurring, low virulence pathotypes (Domermuth et al., 1977) or cell culture adapted strains (Fadly et al., 1985). However, there is evidence to suggest that these vaccines...
are immunosuppressive and as such are capable of predisposing turkeys to secondary bacterial infections (Nagaraja et al., 1982, 1985; Larsen et al., 1985; Pierson, 1993). All isolates of HEV appear to be lymphotrophic and lymphocidal, providing a plausible explanation for their immunosuppressive effect (van den Hurk, 1990; Hussain et al., 1993; Saunders et al., 1993).

To initiate development of a recombinant vaccine against HEV and to shed light upon the genetic relationship of this poorly understood group of viruses with other adenoviruses, we set out to characterize the HEV genome. This study includes the first map of the HEV genome, a description of DNA hybridization experiments comparing the HEV genome with other adenovirus genomes, and the DNA sequence of a region encoding the penton base and two putative core protein genes.

Methods

Viral DNA purification. A low virulence isolate of HEV was propagated in the spleens of young turkeys (Domermuth et al., 1977) and purified by a combination of methods previously described in the literature. These included, in order: sonication of freeze-thawed splenic homogenates (van den Hurk, 1986), precipitation of virus with PEG (Ossa et al., 1982), a single sucrose/CCL ultra-centrifugation step (Green & Pina, 1963), DNAase digestion to remove residual host cell DNA, proteinase K treatment to release the viral DNA from purified intact virions and finally, extraction of viral DNA with phenol-chloroform (Sambrook et al., 1989).

Restriction mapping. A restriction map was constructed by standard techniques, using restriction digestion, cloning in plasmid vectors, and alkaline blotting and hybridization (Sambrook et al., 1989). Genomic size was calculated by averaging the totals of the apparent sizes of DNA fragments generated by each of several restriction enzymes.

DNA probes and hybridization. Restriction digests of HEV DNA were resolved on agarose gels and transferred to nylon membranes (Boehringer Mannheim) by alkaline blotting (Sambrook et al., 1989). Human adenovirus 2 (HAdV-2) DNA was purchased from Gibco BRL (Life Technologies). Plasmid pHPII containing the cloned HAdV-2 hexon gene (Scott-Taylor & Hammond, 1992) was kindly provided by T. Scott-Taylor of the University of Manitoba, Canada. The cloned fowl adenovirus 10 (FAdV-10) penton base gene (Sheppard & Trist, 1992) and the cloned FAdV-10 hexon gene (unpublished) were generously provided by T. Lehrbach of Arthur Webster's Pty Ltd, Parkville, Australia. The viral DNA inserts were excised from these plasmids by restriction digestion and purified on low melting-point agarose gels (Sambrook et al., 1989). The resulting DNA fragments were labelled with digoxigenin by the random primer method using a Genius kit (Boehringer Mannheim). Hybridizations were conducted according to the manufacturer's instructions under low stringency conditions, at temperatures between 40–65 °C ([Na+] = 825 mM).

Identification of the penton base gene. To identify the HEV penton base gene, degenerate PCR primers were synthesized corresponding to the sequences P231EGNY and D231TRNF (beginning at 14661 bp and 14778 bp in the HADV-2 genome, respectively), which were conserved in an alignment of the HADV-2, HADV-40 and FADV-10 penton base amino acid sequences (Roberts et al., 1984; Pieńazek et al., 1989; Sheppard & Trist, 1993). These primers yielded a PCR product which was used as a DNA probe and was found to hybridize with a 5.5 kb
by the dideoxynucleotide chain termination method using Sequenase version 2.0 T7 DNA polymerase (United States Biochemical). Both strands of double-stranded plasmid DNA templates were sequenced using a primer walking strategy. To confirm the sequence around the 5′ terminus of the putative penton base gene ORF, sequencing was repeated with a DNA template generated by PCR directly from HEV genomic DNA, using 18 amplification cycles and other conditions designed to minimize misincorporation (Gelfand & Sninsky, 1993). Searches on the NCBI database (version 9, April 1994) were done with the BLAST program (Altschul et al., 1990); sequence alignments were done with the CLUSTAL program (Higgins & Sharp, 1989) and protein analysis with algorithms contained in the DNASTAR program package. Conserved amino acids were defined as residues matching the consensus within three distance units on the PAM250 matrix (Altschul et al., 1990).

DNA hybridization experiments

Hybridization experiments were conducted in an attempt to identify genes on the HEV genome and to examine the relationship between HEV and other adenoviruses. DNA probes derived from the whole HAdV-2 genome, the HAdV-2 hexon gene, and the FAdV-10 penton base and hexon genes did not hybridize significantly with HEV DNA under low stringency conditions. However, positive signals were observed with HindIII digests of bacteriophage λ and salmon sperm DNA used as negative controls and with homologous DNA used as positive controls (not shown). These results indicate that overall DNA similarity between HEV and the above DNA probes is low, perhaps less than 50%, and/or that regions of DNA similarity are short. Again, these data support the conclusion that viruses in the HEV group are not closely related to the fowl adenoviruses or other adenoviruses.

These hybridization results are in agreement with our DNA sequence data. First, the HEV penton base ORF DNA sequence showed only weak overall similarity with the HAdV-2 and FAdV-10 penton base sequences (39.0% and 44.2% similarity index, respectively, data not shown; CLUSTAL method, DNASTAR), although conserved regions of twenty bases or less showed up to 70% identity. Secondly, the FAdV-10 penton base gene had a G+C content of 59.8%, while that of the HEV target DNA was 33.8%, a disparity that would be expected to lead to a high background and a lack of hybridization.

Zakharchuk et al. (1993) performed Southern hybridization experiments with egg drop syndrome 1976 virus (EDS-76), an avian adenovirus with a 34.2 kb genome and one fibre per penton base. They found DNA sequence similarity between EDS-76 and bovine adenoviruses, but found no similarity, under low stringency conditions, with either FAdV-1 (chick embryo lethal orphan virus, CELO) or HAdV-5. Our results parallel these, in that HEV DNA failed to hybridize with FAdV-10 DNA. Additional hybridization studies with HEV,
Fig. 2. Immunoblot showing E. coli extracts probed with anti-HEV serum. Lane 1, extract of the same E. coli strain transformed with pRSET C; lane 2, extract of E. coli transformed with pHPP4 encoding a portion of the putative penton base ORF cloned in expression vector pRSET C; lane 3, a preparation of HEV penton protein dissociated base and fibre.

EDS-76 and bovine adenovirus DNA could clarify their relationship, but it appears at this point that the division of avian adenoviruses into at least two groups is appropriate.

**Putative penton base ORF**

The penton base (polypeptide III) is a major structural component of the adenovirus virion. One penton base trimer is positioned at each of the 12 vertices of the icosahedral capsid, surrounded by five hexons, with one penton fibre in the case of HEV (van den Hurk, 1992) and the mastiadenoviruses and two in the case of the fowl adenoviruses, extending from the penton base (Van Oostrom & Burnett, 1985). To identify the HEV penton base gene, we sequenced DNA in a region of the genome which had been expressed in E. coli transformants that reacted positively with anti-HEV serum (Fig. 2). An ORF encoding a predicted 450 aa protein (Fig. 3) exhibited strong overall amino acid sequence similarity with published penton base amino acid sequences, e.g. 39.0%, 53.0% and 44.2% overall similarity with HAdV-2, HAdV-40 and FAdV-10 penton base ORFs, respectively (Fig. 4a). A more highly conserved region, from amino acids P158 to D267 of the HEV ORF1, showed similarities of 59.7%, 59.7% and 65.3% with HAdV-2, HAdV-40 and FAdV-10, respectively. The molecular mass of the predicted penton base protein is 51.1 kDa, in close agreement with the reported value of 52 kDa (van den Hurk, 1992) for HEV penton protein on Western blots. These results indicate that the putative HEV penton base ORF is homologous to other adenoviral penton base genes.

The putative HEV penton base ORF is unique in several respects. First, it encodes 450 aa, shorter than the HAdV-2 (572 aa), HAdV-40 (504 aa) and FAdV-10 (525 aa) penton base ORFs. The relative brevity of the penton base gene is not surprising in view of the shortness of the HEV genome as a whole. Secondly, the HEV penton base gene showed a G+C content of 33.8%, as compared to a typical average of 57.0% for the other adenoviruses. The entire 2643 bp region we have sequenced so far had a G+C content of 34.0%. Again, HEV differs dramatically from other adenoviruses, including the fowl adenoviruses, in this respect.

A curious feature of the HEV sequence is the implied amino acid sequence starting with the residues MRR, 21 aa upstream of the initial ATG codon of the proposed penton base ORF (Fig. 4a). This sequence shows similarity to the N-terminal regions of the penton base genes of HAdV-2 and HAdV-40, implying that this sequence would mark the start of the HEV gene also (Figs 3 and 4a). However, this region in HEV is followed by two stop codons, the second of which is followed immediately, in the same reading frame, by the ATG designated as the start of the putative penton base ORF. To confirm this result, we sequenced the DNA in this region repeatedly in both directions using four different primers and a cloned DNA template. To eliminate the possibility that an aberrant mutant genome had been cloned, we repeated sequencing using DNA template generated by PCR directly from a pool of HEV genomic DNA. N-terminal sequencing of penton base protein could validate our proposed penton base ORF, but repeated attempts to accomplish this have failed due to N-terminal blockage.

Possibly, the upstream penton-resembling region in HEV (Fig. 4a) is a vestige of an evolutionary predecessor ORF which encoded the more complete sequence. Alternatively, one can speculate that this arrangement somehow facilitates variable expression of the penton base protein, providing an explanation for the variable appearance of penton base protein bands on Western blots, e.g. 52 kDa and 51 kDa with low and high virulence isolates of HEV, respectively (van den Hurk, 1992). Van Oostrom & Burnett (1985) proposed a similar mechanism of alternative start codons, based on the observation of HAdV-2 penton base protein as a double or triple band on denaturing gels. The sequence TTTTAGAT at 342 bp (Fig. 3) is a good match for the HAdV-2 major late splice acceptor site consensus sequence PyPyNPyAG/GT (Jones, 1986). This sequence overlaps the initial ATG of the putative penton base
Fig. 3. For legend see page 475.
Fig. 3. For legend see facing page.
Characterization of the HEV genome

Fig. 3. Nucleotide sequence of a 2643 bp region spanning 49.2-54.8 m.u. (10.7 - 13.3 kb) of the HEV genome, showing predicted amino acid sequences of partial ORF5, putative penton base and core protein ORFs (core 1, core 2) and partial putative pVI (hexon-associated protein) ORF.

ORF so that a typical adenoviral RNA splicing event would place a hypothetical upstream leader region exactly 5' of the initial ATG. Since nothing is known about transcriptional or post-transcriptional regulation of gene expression in HEV, only further analysis can confirm whether this sequence is functional. However, the location of this possible splice acceptor site supports the designation of the putative penton base ORF shown in Fig. 3.

The sequence AATAAA at 200 bp, upstream of the penton base ORF (Fig. 3), is an exact match for the adenoviral polyadenylation signal. However, this sequence is 138 bp 5' of the penton base ORF, much further than the 10-30 bp spacing typical for other adenoviruses (Jones, 1986). Further analysis is needed to determine whether this is a coincidental sequence or a functional signal for the polyadenylation of a transcript analogous to the HAdV-2 L1 major transcript.

ORF2, putative major core protein 1

In human adenoviruses, the L2 major late transcripts encode the penton base gene, followed by the genes for pVII, pV and p-mu; these are basic core polypeptides associated with the viral DNA and bear recognition sites for the virus-encoded L3,-23K endopeptidase (Akusjarvi et al., 1986). ORF2, encoding 121 aa (Fig. 3), showed a significant degree of sequence similarity within the first 45 aa of the pVII major core protein sequences of HAdV-2, HAdV-40, FAdV-10 (41.3% similarity, Fig. 4b; Aleström et al., 1984; Sung et al., 1983) and canine adenovirus type 1 (CAV-1, 41.3% similarity, not shown; Cai & Weber, 1993). However, the overall similarity between the entire HEV ORF2, HAdV-2, HAdV-40 and FAdV-10 was only 21.7%, 23.8% and 51.9%, respectively. To the right of this region within ORF2, little significant similarity was found. The sequence LIGGAL at 26 aa in ORF2 closely matches the consensus sequence for the L3,-23K endopeptidase cleavage sites found in pVII, pV, p-mu and pVI (hexon-associated protein) (Freimuth & Anderson, 1993). With runs of positively charged amino acids (18% Arg), giving a predicted pI of 12.3, and extensive hydrophobic regions, the putative protein encoded by ORF2 shares characteristics typical of adenoviral major core proteins. These findings, along with the location of ORF2 immediately 3' of the penton base gene, imply that ORF2 encodes a major core protein similar in function to pVII of HAdV-2, HAdV-40 and the major core protein of FAdV-10. The predicted molecular mass of this putative protein is 13-3 kDa, implying a relationship with the 12.5 kDa HEV core protein band seen by immunoblot analysis (van den Hurk, 1992). Van den Hurk (1992) described a stronger 9.5 kDa band representing a second core protein. This band may correspond to the mature form of the ORF2 encoded protein, in which the 29 aa leader has been cleaved from the precursor protein to produce a mature core protein with a predicted molecular mass of 10.5 kDa (Fig. 4b).

The HEV ORF2 sequence shows a longer region of sequence similarity with HAdV-2 and HAdV-40 pVII than with the FAdV-10 major core protein ORF (ORF1) (Fig. 4b), which lacks an obvious protease cleavage site consensus motif (Sheppard & Trist, 1993) and validates the taxonomic distinction between the two avian adenovirus groups. The proposed HEV major core protein 1 ORF is shorter (121 aa) than the HAdV-2 (199 aa) and HAdV-40 (186 aa) pVII ORFs and longer than FAdV-10 major core protein 1 ORF (73 aa). The G+C content of 44.4% for HEV ORF2 also differs drastically from that...
Fig. 4. For legend see facing page.

of 70.0% for HAdV-2, 64.9% for HAdV-40 and 66.7% for FAdV-10 ORF2.

It is worth noting that the proposed ATG start codon of the core protein 1 gene overlaps by 1 bp the TAA stop codon of the putative penton base gene. If demonstrated experimentally, this would represent an efficiency of
Fig. 4. Alignments of HEV ORFs with amino acid sequences of adenoviral genes obtained from GenBank. Black indicates amino acids that match HEV exactly; shading indicates conserved amino acids; dots indicate stop codons. (a) Alignment of the putative HEV penton base ORF with penton base sequences of HAdV-2, HAdV-40 and FAdV-10. (b) Alignment of HAdV-2 and HAdV-40 pVII ORFs, HEV putative penton base ORF sequences. (c) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative p-mu ORF sequences. (d) Alignment of HAdV-2 and HAdV-40 p-20RF and the C-terminal portion of FAdV-10 putative core protein 10RF and FADV-10 core protein 1 ORF sequences with penton base ORF with penton base sequences of HAdV-2, HAdV-40 and FAdV-10 core protein 2 ORF sequences. (e) Alignment of HAdV-2 and HAdV-40 p-20RF and the C-terminal portion of FAdV-10 putative core protein 10RF and FADV-10 core protein 1 ORF sequences with penton base ORF with penton base sequences of HAdV-2, HAdV-40 and FAdV-10 core protein 2 ORF sequences. (f) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (g) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative p-mu ORF sequences. (h) Alignment of HAdV-2 and HAdV-40 p-20RF and the C-terminal portion of FAdV-10 putative core protein 10RF and FADV-10 core protein 1 ORF sequences. (i) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (j) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative p-mu ORF sequences. (k) Alignment of HAdV-2 and HAdV-40 p-20RF and the C-terminal portion of FAdV-10 putative core protein 10RF and FADV-10 core protein 1 ORF sequences. (l) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (m) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative p-mu ORF sequences. (n) Alignment of HAdV-2 and HAdV-40 p-20RF and the C-terminal portion of FAdV-10 putative core protein 10RF and FADV-10 core protein 1 ORF sequences. (o) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (p) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative p-mu ORF sequences. (q) Alignment of HAdV-2 and HAdV-40 p-20RF and the C-terminal portion of FAdV-10 putative core protein 10RF and FADV-10 core protein 1 ORF sequences. (q) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (r) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative p-mu ORF sequences. (s) Alignment of HAdV-2 and HAdV-40 p-20RF and the C-terminal portion of FAdV-10 putative core protein 10RF and FADV-10 core protein 1 ORF sequences. (t) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (u) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative p-mu ORF sequences. (v) Alignment of HAdV-2 and HAdV-40 p-20RF and the C-terminal portion of FAdV-10 putative core protein 10RF and FADV-10 core protein 1 ORF sequences. (w) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (x) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative p-mu ORF sequences. (y) Alignment of HAdV-2 and HAdV-40 p-20RF and the C-terminal portion of FAdV-10 putative core protein 10RF and FADV-10 core protein 1 ORF sequences. (z) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (a) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (b) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (c) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (d) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (e) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (f) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (g) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (h) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (i) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (j) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (k) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (l) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (m) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (n) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (o) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (p) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (q) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (r) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (s) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (t) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (u) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (v) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (w) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (x) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (y) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences. (z) Alignment of HAdV-2 and HAdV-40 p-mu ORFs, HEV putative penton base ORF sequences.
genetic coding space not observed in other adenoviruses at this location.

**ORF3, putative p-mu-like core protein**

Mu is a small, highly basic dsDNA-binding protein in HAdV-2 and HAdV-40 that is derived from the p-mu precursor by two cleavage reactions mediated by a virus-encoded endopeptidase (Anderson et al., 1989). ORF3, encoding a protein of 59 aa with a predicted PI of 11.0, showed 38.3% and 43.0% amino acid sequence similarity with HAdV-2 and HAdV-40 p-mu, respectively (Fig. 4d; Aleström et al., 1984), and 23.6% similarity with FAH-10 putative core protein 1 ORF (ORF2) (Sheppard & Trist, 1993). Located between the core protein 1 ORF (ORF2) and directly 5' of the putative pVI ORF (ORF4) (Fig. 1), ORF3 also showed short regions of similarity to core protein pV of HAdV-2 and HAdV-40, although introduced gaps are needed to align the related sequences. The adenoviral protease cleavage site consensus motif, LRGGF (beginning at L24), corresponding to the site in the C-terminal half of HAdV-2 and HAdV-40 p-mu, is well-conserved in ORF3 (Fig. 4c). Another possible protease recognition site in the N-terminal half of HEV ORF3, at G11, is not as well conserved. ORF3 may encode a core protein similar to HAdV-2 and HAdV-40 p-mu, but has undergone significant deletions and/or lacks insertions as compared to the related HAdV-2 (79 aa), HAdV-40 (71 aa) and FAH-10 (102 aa) ORFs. Although no HEV core proteins of this size were detected by immunoblotting (van den Hurk, 1992), the conditions used would not be expected to resolve a predicted 6-1 kDa protein or its cleavage products.

Although fewer gaps need to be introduced in order to align HEV ORF3 with HAdV-2 or HAdV-40 p-mu, the HEV sequence shows stronger similarity to the FAH-10 core protein 2 ORF; this is in contrast to the situation with HEV core protein 1 ORF (ORF2). Furthermore, like FAH-10, and unique to the avian adenoviruses, HEV appears to lack a pV core protein gene. It is possible that the major core protein 2 ORFs of these two avian adenoviruses serve the functions of pV and p-mu.

**Partial ORF4, putative hexon-associated protein**

HEV partial ORF4 showed some sequence similarity with the pVI (hexon-associated protein) ORFs of HAdV-2 and HAdV-40 (19.6% and 16.0%, respectively; Aleström et al., 1984) and the partial FAH-10 putative hexon-associated protein ORF (44.4% similarity with the first 40 residues, Fig. 4d; Sheppard & Trist, 1993). This putative HEV hexon-associated protein ORF also has a possible protease cleavage site consensus motif (LRGGKI), where it shows the best alignment with the FAH-10 sequence.

Partial ORF5 extended from the left end of the sequenced region to the start of the penton base ORF (ORF1) (Fig. 3), a location occupied by penton-associated protein IIIa genes in other adenoviruses (Akusjarvi et al., 1986). However, ORF5 showed only a short region of weak similarity to the IIIa genes of HAdV-2 and HAdV-40 (not shown) (Roberts et al., 1984; Davidson et al., 1993). Thus, this region of the HEV genome appears highly divergent and its function remains unclear.

Since the HEV ORFs presented here are based on DNA sequence data, much work remains to be done to demonstrate the function of these features on the transcriptional and translational levels. Our results confirm that HEV can indeed be considered a member of the Adenoviridae, but that it is an adenovirus with novel characteristics with regard to genomic size and organization. In many respects, but not all, HEV appeared to be as closely related to mammalian adenoviruses as it was to fowl adenoviruses. The HEV genome, having apparently undergone deletions in relation to other adenoviruses, presents an interesting model of viral genome evolution in the light of minimal functional sequences.

The authors wish to thank Wampler-Longacre Inc. and ROCCO Turkeys Inc. for grants in support of this work, Dr J.L. Johnson for helpful discussion and Dr F.A. Murphy of the ICTV for assistance with taxonomic nomenclature and classification.

**References**


Characterization of the HEV genome


(Received 1 August 1995; Accepted 18 October 1995)