VIROLOGY 249, 307–315 (1998)
ARTICLE NO. VY989336

The Complete DNA Sequence and Genome Organization of the Avian Adenovirus, Hemorrhagic Enteritis Virus

J. Pitcovski,*[1] M. Mualem,* Z. Rei-Koren,* S. Krispel,* E. Shmueli,* Y. Peretz,* B. Gutter,† G. E. Gallili,† A. Michael,† and D. Goldberg*

MIGAL, South Industrial Zone, Kiryat Shmona 10200, Israel; and †ABIC, Jerusalem, Israel

Received February 11, 1998; returned to author for revision April 10, 1998; accepted July 7, 1998

Hemorrhagic enteritis virus (HEV) belongs to the Adenoviridae family, a subgroup of adenoviruses (Ads) that infect avian species. In this article, the complete DNA sequence and the genome organization of the virus are described. The full-length of the genome was found to be 26,263 bp, shorter than the DNA of any other Ad described so far. The G + C content of the genome is 34.93%. There are short terminal repeats (39 bp), as described for other Ads. Genes were identified by comparison of the DNA and predicted amino acid sequences with published sequences of other Ads. The organization of the genome in respect to late genes (52K, IIIa, penton base, core protein, hexon, endopeptidase, 100K, pVIII, and fiber), early region 2 genes (polymerase, terminal protein, and DNA binding protein), and intermediate gene IVa2 was found to be similar to that of other human and avian Ad genomes. No sequences similar to E1 and E4 regions were found. Very low similarity to ovine E3 region was found. Open reading frames were identified with no similarity to any published Ad sequence.

INTRODUCTION

Adenoviruses (Ads) are divided into two genera: mastadenoviruses, which infect mammals, and aviadenoviruses, which infect avian species (Shenk, 1996). Hemorrhagic enteritis virus (HEV) and marble spleen disease of pheasants and splenomegaly virus of chickens are classified as type II avian Ads (Domermuth and Gross, 1984). This group is serologically distinct from types I and III avian Ads, which are isolated from chickens [fowl Ad (FAV) 1–12 and egg drop syndrome (EDS) virus, respectively]. HEV infects turkeys and causes a disease characterized by depression, splenomegaly, intestinal hemorrhage, and immunosuppression (Domermuth and Gross, 1984). The virus replicates in B cells and macrophages (Suresh and Sharma, 1996) and is concentrated in large amounts in the spleen. HEV, like other Ads, is nonenveloped and has an icosahedral capsid with a diameter of 70 nm. The genome is linear, double-stranded DNA and was estimated to measure 25.5 kb (Jucker *et al.,* 1996). These authors also published a partial sequence (~4 kb) of HEV. The full sequences of several human Ads (types 2, 5, 12, and 40), avian Ads (CELO and EDS), ovine Ad (OVa), and others have been published and can be found in the EMBL and GenBank databases. All human Ads have the same general organization (Shenk, 1996). However, in some nonhuman Ad sequences (CELO, OVa, EDS), no similarity was found for various regions of human Ad (Chiocca *et al.*, 1996; Vrati *et al.*, 1996; Hess *et al.*, 1997, respectively). In the past decade, Ads became a subject for research as vectors in gene therapy (Kozarsky and Wilson, 1993). Sequencing of the complete viral genome is essential if the virus is to be used for this purpose. The complete sequence is also needed for the development of recombinant subunit vaccines (Pitcovski *et al.*, 1996) and if the virus is used as a carrier of foreign viral genes, in a similar manner to the way in which vaccinia is used in mammals and fowlpox virus is used in chickens (Yamanouchi *et al.*, 1993; Boursnell *et al.*, 1990).

The objective of this study was to obtain a detailed sequence of the virus, to understand HEV as a member of Aviadenoviridae, and to facilitate the isolation of genes for vaccination purposes and as a potential vector for gene delivery in recombinant vaccines or gene therapy.

RESULTS AND DISCUSSION

This article reports the complete sequence and genome organization of HEV. After sequence determination and translation of the nucleotide sequence of the open reading frames (ORFs), the deduced amino acid (aa) sequence was compared with that of several other published Ad sequences. Based on these comparisons, the putative genes of HEV were identified (Table 1 and Fig. 1). There are no functional data to support these determinations, but high similarity, consensus, and control sequences that are common to all Ads allow the prediction of the gene sites. Moreover, other human and animal Ads genes have been simi-

[1] To whom reprint requests should be addressed at MIGAL, P.O.B. 90000, Rosh Pinna 12100, Israel. Fax: 972–6-6944980. E-mail: jp@migal.co.il.

0042-6822/98 $25.00

TABLE 1

Location and Predicted Sizes of HEV Genes and the Size of the Putative Proteins for which They Code

Protein	ATG	Stop	No. of residues	HEV aa relative to CELO	MW
L1					
52K	8,569	9,471	300	−78	33,800
IIIa	9,461	10,978	505	−70	50,293
L2					
Penton base	11,000	12,346	448	−67	50,903
CP I	12,346	12,708	120	−68	11,134
CP II	12,711	12,887	58	−14	6,111
L3					
pVI	12,905	13,600	231	+8	24,947
Hexon	13,609	16,329	906	−36	101,089
EP	16,331	16,975	214	+8	25,008
L4					
100k	18,184	20,226	680	−304	78,283
pVIII	20,768	21,370	200	−45	21,769
E3?	21,213	22,115	300		
L5					
Fiber	22,518	23,882	452	−751	48,770
IVa2	3,436	2,333	367	−72	41,802
E2A					
DBP	18,012	16,972	346	−95	38,910
E2B					
POL	6,767	3,429	1,112	−9	129,160
pTP	8,557	6,764	597	+22	70,582

Note. The difference in size between similar HEV and CELO proteins is also indicated. CPI, core protein I; CPII, core protein II; EP, endoproteinase; DBP, DNA binding protein; POL, DNA polymerase; pTP, precursor terminal protein.

larly predicted (Davison *et al.*, 1993; Vrati *et al.*, 1996; Hess *et al.*, 1997). The genome map and location of genes are presented in Figure 1. For comparison, the map of another avian Ad (EDS) and a human Ad (Ad2) have been added to the figure. It can clearly be seen that the genome of HEV is organized very similarly to these Ads. The start and stop codon sites and size of the predicted proteins of HEV are given in Table 1. We have found many ORFs that bear no similarity to any known Ad genes. These ORFs are listed in Table 2 and discussed below. Most of the HEV genes were found to be smaller than those found in CELO, which has the largest reported Ad genome (Table 1). On the other hand, the capsid diameter of both viruses is similar: CELO is 70–80 nm (Laver *et al.*, 1971) and HEV 72 nm (van der Hurk, 1992). This similarity in the size of the capsid was expected because the hexon, which is the major capsid protein, is very similar in size (only 4% longer in aa in CELO).

Features of the nucleotide sequence of HEV

The genome length of HEV was found to be 26,263 bp. This is the smallest genome of all Ads isolated so far (Jucker *et al.*, 1996) but larger than estimates reported in earlier studies. This difference probably emanates from the inaccuracy of DNA size determination by agarose gel electrophoresis. The overall G + C content is 34.93%. This ratio is very low compared with that of other Ads. The G + C content of EDS is 42.5% (Hess *et al.*, 1997), and that of CELO is 54.3% (Chiocca *et al.*, 1996). In human Ads, the G + C content was found to be 48–59%. In ovine Ad, the G + C content is 33.6% (Vrati *et al.*, 1996). Inverted terminal repeats (ITRs) are 39 bp long. ITRs were found to be a parameter of the relationship between Ads. In human Ads, the ITRs of members belonging to the same subgenus are closely related to each other (Shinagawa *et al.*, 1987). The length of the viral genome, its G + C content, and its ITRs are used as parameters for the classification of Ads. In all these three parameters, HEV is very similar to ovine Ad, which is 29,544 bp in size and has a G + C content of 33.6% and an 46-bp ITR (Vrati *et al.*, 1996). This points to a close relationship between the two viruses.

Early regions (E)

E1 and E4. No similarity to any E1 and E4 regions were found in HEV genome. However, HEV has two clusters of ORFs: one that includes ORFs 1, 2, 3, and 4, and one that includes ORFs 7 and 8 (Fig. 1). The locations of the first and second clusters correspond to the locations of E1 and E4, respectively, in other Ad genomes. A similar situation exists in CELO, in which no region with simi-

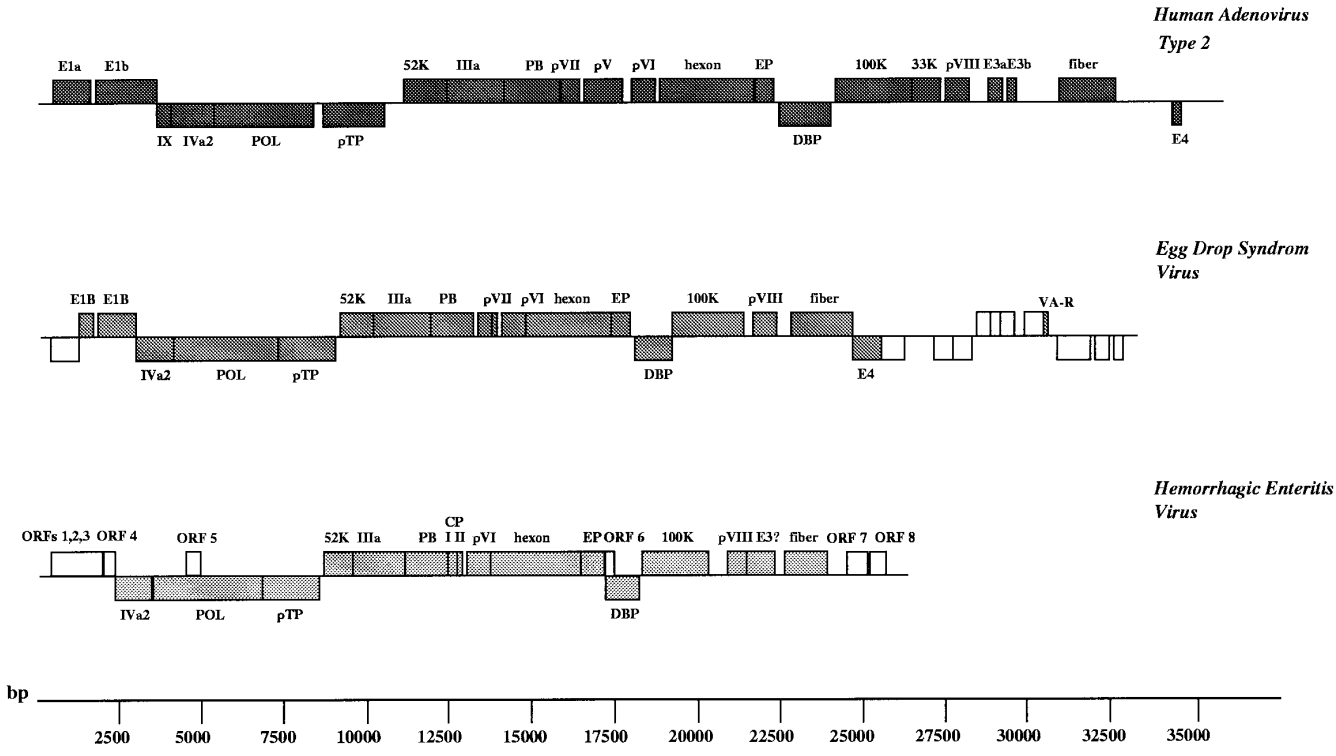

FIG. 1. A schematic representation of the genome organization of HEV and a comparison with the EDS and Ad2 genomes. Genes in HEV are named according to their similarity to known Ad genes and position along the genome. (Full boxes) The locations of the predicted genes. ORFs starting with methionine, >100 aa and with no similarity to any Ad gene, are designated "ORF1" to "ORF8" (empty boxes). EDS and Ad2 sequences were taken from GenBank (accession No. Y09598 and J01917), respectively. (Top boxes) Genes found on the upper strand. (Bottom boxes) Genes found on the bottom strand. POL, DNA polymerase; pTP, precursor terminal protein; DBP, DNA binding protein; EP, endoproteinase.

larity to E1 has been identified (Chiocca *et al.*, 1996). However, CELO has the ability to generate tumors in baby hamsters (Jones *et al.*, 1970), and it is therefore likely that at least some of the functions of E1 are retained by its unassigned ORFs. It is possible that HEV resembles CELO in this respect.

E2B. The E2B region of HEV consists of POL, preterminal protein (pTP), and DBP genes. The pTP primes human Ad DNA replication (Salas, 1991; Smart and Stillman, 1982; Stillman *et al.*, 1981). pTP is processed at two sites by viral protease to a mature terminal protein (TP).

TABLE 2

Sites of Open Reading Frames of >300 Nucleotides and Starting with Methionine for which No Similarity Was Found to Any Ad Gene

	Start	Stop
ORF1	399	1,952
ORF2	616	918
ORF3	1,498	1,872
ORF4	1,964	2,311
ORF5	4,477	4,974
ORF6	17,179	17,541
ORF7	24,509	25,165
ORF8	25,201	25,698

This processing takes place through two cleavages by viral endoprotease and was shown to be conserved in all human serotypes of Ad (Webster *et al.*, 1997). In HEV, pTP consist of 597 residues. Two cleavage sites were identified at residues 169 and 297, similarly to the EDS TP, which is 581 aa long and is cleaved at residues 156 and 265. A sequence similar to that reported as a nuclear localization signal in Ad2 (RLPVRRRRRRVP) and EDS (TLPARTRRTRRP) (residues 380–391 and 308–319, respectively) was identified in HEV (SLPLIRRIRRPP, residues 341–352).

The DNA polymerase (POL) of HEV is similar in size and aa sequence to other adenoviral polymerases (Table 3). Analysis of its predicted protein sequence by the GCG Motif program resulted in identification of a DNA polymerase family B signature in residues 925–933 (YGDTDSLFL).

The DNA-binding protein (DBP) is divided into two clusters in Ad2. The C-terminal cluster is involved in binding to the viral DNA and activating replication of the major late promotor. High identity levels were found in this cluster among the C termini of DBPs from various Ads. DBP is located between the endoproteinase (EP) and 100k genes, on the complementary strand. Its size in HEV (345 aa) is smaller than in other

TABLE 3

The Percentage of Amino Acid Sequence Identity and Similarity of HEV Proteins and Their Counterparts from Ad2, EDS, OVa, and CELO

	Ad2		EDS		OVINE		CELO	
Virus	% id	% s	% id	% s	% id	% s	% id	% s
52K	23.4	35.5	27.1	34.9	25.6	34.4	28.7	40.4
IIIa	31.2	40.9	34.2	43.7	32.3	44.3	30.8	41.2
Penton base	48.6	58.1	52.9	61.4	52.6	61.4	49.7	59.3
pVI	29.9	38.4	37.3	45.0	37.3	45.6	41.0	49.7
Hexon	52.8	59.4	53.9	62.3	55.9	63.3	54.7	62.0
EP	44.2	53.8	44.8	55.7	41	52.5	50.7	59.5
100K	36.1	45.3	38.4	47.5	36.6	47.4	38.6	48.8
pVIII	23.1	33.7	29.3	34.3	38.4	44.9	26.3	32.9
Fiber	25.7	33.6	30.1	37.6	29.3	35.4	24.8	33.4
IVa2	33.3	44.3	35.3	46.5	36.9	46.2	36.7	46.9
POL	41.1	52.9	43.7	54.9	44.8	54.5	42.0	53.7
pTP	34.3	44.8	37.4	45.8	34.5	46.6	35.9	47.7
DBP	28.4	40.5	28.4	40.5	36.0	43.7	34.4	43.9

Note. The comparisons were made by the Gap program for global comparison, which is part of the GCG software package. The gap opening penalty parameter of Gap was set to 6, and its gap extension penalty parameter was set to 2. The sequences for the comparison were taken from GenBank and EMBL databases. Accession Nos.: Ad2, J01917; CELO, U46933; EDS, Y09598; ovine adenovirus, U18755, U40837, U31557, U40839. POL, DNA polymerase; id, identity; s, similarity; pTP, precursor terminal protein; DBP, DNA binding protein; EP, endoproteinase.

Ads (382, 387, 441, and 539 in OVa, EDS, CELO, and Ad2, respectively). The sequence VFQCCNP of HEV DBP (aa 255–261) was found to be conserved in Ad2, EDS, and OVa. The OVa DBP has a charged motif KKRK (aa 11–14), which is used for nuclear localization. HEV DBP has a similar motif (KKNK; aa 35–38), which may have a similar function. The motifs of Leu-515, Pro-516, and Pro-526, which serve in the cooperative binding of the subunits, are partially conserved in HEV. Leu and Pro are at residues 316 and 317, but the next Pro is only 6 residues apart. This motif was completely conserved in EDS and is missing in CELO.

E3. An ORF of 900 bp lies between the regions coding for HEV pVIII and fiber. This ORF bears little similarity to one of the gene products of the predicted E3 of OVa (Vrati *et al.*, 1996). E3 has been mapped to this region in most Ads. The size of E3 varies from 2.5 kb in human Ad to 1.0–1.5 kb in canine Ad (Dragulev *et al.*, 1991) and 1 kb in mouse Ad (Beard *et al.*, 1990). These size variations show that a large diversity in size and sequence might be expected for this protein.

Delayed early genes

pIX. pIX of human Ad is known to strengthen hexon–hexon interactions (Boulanger *et al.*, 1979; van Oostrum and Burnet, 1985) and have transcriptional properties, such as stimulating the major late promoter (Lutz and Kedinger, 1996). No sequences similar to pIX were found in any of the avian Ads or in ovine Ad.

IVa2. IVa2 was recently found to be a component that contributes to the activation of the major late promoter (Lutz and Kedinger, 1996). This gene was found in to be shorter in HEV than in other Ads.

Late proteins

The genes encoded at this region were identified at a similar position to that in other Ads (Fig. 1).

Penton base. The penton base is bound to hexon protein at each vertex of the icosahedral structure. It is the base for the fiber, and together they constitute the penton. In most human Ads, a tripeptide sequence, Arg/Gly/Asp (RGD), was identified and reported as the binding site of the virus to cellular integrin, causing endocytosis of the virus by the cell (Mathias *et al.*, 1994). The RGD sequence was not identified in HEV, as in other avian and mammalian Ads (Kleiboeker, 1995; Meissner *et al.*, 1997). The fact that no RGD sequence was found on the penton base indicates that the penetration of this Ad differs from that described for human Ad. Probably an alternative way exists to allow the virus to penetrate cells. CELO virus, which has no RGD, has been shown to enter baby hamster cells and cause tumors (Sarma *et al.*, 1965).

Core proteins. The core of human Ads consists of four proteins, pV, pVII, mu, and terminal protein. pV, pVII, and mu are basic, and all contact the viral DNA. HEV counterparts of pV and pVII are core proteins I and II, respectively. Their genes are located between the regions coding for the penton base and pVI. A similar location for these genes was reported for other Ads (Jucker *et al.*, 1996). CPI was found to be rich in arginine, proline, and alanine (17%, 11%, and 13%, respectively). The molecular weights (11,000 and 6,000) are lower than those of Ad2 (48,000 and 18,000) but close to those of CELO (20,000, 12,000, and 9,500) and EDS

FIG. 2. Alignment of the last 11 aa of C termini of pVI of HEV and other Ads. The alignment was done by the Pileup program, which is part of the GCG software package. (Black boxes) The aa identical in all five viruses. (Gray boxes) The aa identical or similar in some of the viruses.

(18,000 and 7,000). pV is missing in all avian and ovine Ads. The fourth protein in the core is terminal protein, which is a product of the E2B region.

pVI. pVI is almost identical in size to pVI from other avian Ads. Two cleavage sites for the viral endoprotease were identified on both ends of the protein. Residues 24–28 (LRGGK) in the N terminus resemble a conserved cleavage site (Freimuth and Anderson, 1993). Cleavage of the second putative site in the C terminus of pVI can produce a peptide of 11 aa (Fig. 2). Sequence alignment of this C-terminal region shows partial similarity to several C termini of various pVIs. This small cleavage product was shown to be a cofactor of the viral protease (Mangel *et al.*, 1993). It forms a disulfide bond with the protease via a conserved cysteine residue that is also found in HEV pVI (Fig. 2). The consensus sequence that was proposed for this short peptide by Vrati *et al.* (1996) (G.XXXXX.R.R.y/r.C.f/y) may be different in avian Ads and is suggested to be GVXXXXXXXCY.

Hexon. The hexon protein is similar in size to that of other avian Ads (102 kDa). Because hexon is the major protein in the capsid, its size is important in determining the diameter of the capsid. Some of the other viral proteins are bound to the hexon, and its sequence is important in antigenic determination of the type and group of Ads.

Viral proteinase. HEV protease was found to have several more aa at the C-terminal end than proteases of other Ads. Alignment of proteases of different Ads shows some regions that are identical in all the other Ads compared (Fig. 3). Of the three aa that were determined in the active site of Ad2 (His54, Glu71, and Cys122) by Ding *et al.* (1996), the glutamic acid at position 71 was replaced by the similar aa, aspartate, in all the Ads compared (Fig. 3). The three aa at the active site are surrounded by regions that are conserved and probably form the active site. The protease is one of the most conserved proteins of Ads (Table 3). Cleavage of several structural proteins by EP is part of the maturation of the virion. The sites for protease cleavage, (M,L,I)XGXG and (M,L,I)XGGX that were determined for other Ads, were also identified in HEV proteins at similar locations.

Fiber. One fiber protein was found in HEV, projecting from the penton base (van den Hurk, 1992), with a relatively small predicted molecular weight compared with other fibers from avian or mammalian Ads. The small size is in agreement with the measured length of the HEV fiber, which is 17 nm, whereas that of Ad2 is 31 nm. All

FIG. 3. The aa sequence alignment of endoprotease of five Ads. The alignment was done with Clustal W software. The pairwise and multiple alignment parameters were 6 for gap opening penalty and 0.1 for gap extension penalty. The sequences are ordered according to their similarity. (Gray) Conserved residues of the active site. * Identical residues in all sequences. (Colon) Conserved substitutions. (Dot) Position in which >50% of the residues are identical.

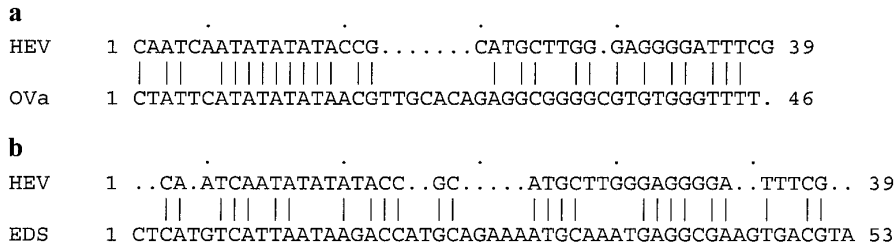

FIG. 4. The ITRs of HEV and their comparison with the ITRs of OVa (a) and EDS (b). The alignment was done by the Bestfit program, which is part of the GCG software package. The gap opening parameter of the program was set to 15 and the gap extension parameter to 1.

the elements described for fiber in other Ads can be found in HEV fiber using the organization rules described by Green *et al.* (1983) and Stouten *et al.* (1992). The attachment of the virus to the cell receptor is specific and at high affinity (Mei and Wadell, 1996). It is mediated by the fiber protein. The knob is the part that interacts with the host cell receptor (Mei and Wadell, 1996).

Undefined ORFs

Eight ORFs coding for putative polypeptides of >100 aa were found in the HEV genome. No similarity was found between them and any of the proteins of other Ads. Additional unidentified ORFs coding for smaller polypeptides, 50–100 aa in length, were found scattered throughout the HEV genome: 16 on the upper strand and 21 on the lower strand.

Sequence elements

Inverted terminal repeats. Although relatively short (39 bp), HEV inverted terminal repeats (ITRs) include two of the motifs that have been described for other Ads. The sequence between nucleotides 9 and 18 that is conserved in human Ads is fully conserved in HEV, except for a missing adenine residue at position 12. This adenine is part of the conserved TAAT region that is involved in the initiation of DNA replication in human Ads (Rawlins *et al.*, 1984; Stillman *et al.*, 1982). This residue is also missing in the OVa ITR (Fig. 4A). It has been suggested that a complex of two viral proteins (pTP and POL) binds to this sequence (Chen *et al.*, 1990; Mul and van der Vliet, 1992; Temperley and Hay, 1992). Nucleotides 19–39 are described in human Ads as domain B. The GGGAGG region (nucleotides 26–31 in HEV) may be the sequence reported as binding the SP1 transcription factor. Nucleotides 40–51 are described in human Ads as domain C. Cellular factors bind to domains B and C to enhance the efficiency of the initiation of DNA replication. It seems that HEV has domain B but that domain C is missing. It was found by Shinagawa *et al.* (1987) that members of the same Ad subgenus are closely related to each other with respect to their ITR. In the alignment of the HEV and EDS ITR (Fig. 4B), it seems that EDS does have the TAAT (nucleotides 10–13) conserved sequence. On the other hand, the SP1 binding site on domain B is missing in other members of the avian Ads, namely EDS (Fig. 4B) and CELO (data not shown). A comparison of the ITRs of HEV with those of EDS and OVa reveals that in both cases, 26 of the 39 nucleotides are identical. However, only the sequences of HEV and OVa LTRs are arranged in clusters that fit the two domains described for binding proteins (nucleotides 9–18 and 26–33 in HEV).

Stop codons. At the end of some genes, an additional stop codon appears within 20 nucleotides downstream from the first one (Fig. 5) in the same frame or in different frames. It is possible that this additional stop codon ensures the accurate termination of translation when such accuracy is crucial to the function of the protein or of other proteins.

Overlaps. Very short overlaps appear in the main genes. Only four of the unassigned ORFs overlap identified genes. This shows that the short length of the HEV genome cannot be accounted for by overlaps in the coding sequences.

General discussion

HEV was found to be similar in the features of its control regions, late genes, IVa, and E2 genes to other Ads. E1 genes were not identified in this Ad, as is the case in CELO and OVa. Some of these genes might be included in the unassigned ORFs, but the overall coding sequence of the ORFs cannot code all the proteins missing relative to the human Ads. We suggest that proteins that have control properties and are in tight interaction with cellular components are less similar among viruses that infect different species of animals (e.g., E1a, E1b, and E3 and E4 gene products). This may be a consequence of the need of the virus to introduce proteins that will be able to interact with specific host proteins. This hypothesis would explain why no sequences similar to E1A and E1B regions were found in HEV despite the fact that ORFs exist at similar sites of its genome.

When the HEV genome is compared with a larger Ad genome, it can be seen that its genes are shorter and that fewer ORFs can be identified. On the other hand, the size of the virion is very similar to that of other Ads. This leaves space to clone large genes into HEV for purposes

1. **52K**
9461 GTC TTC TAA GGA TGT GGC AGA GAT CTT ATC TGG AAA 9497

2. **100K**
20217 CCC AGG TAG TTA TTC TAA ATG GAA GCA ATA CTG CAC 20252

3. **PENTON BASE**
12336 TTG CAA TAA TGC ATT CTG TTG TTT ATT CTC CCA GGG 12360

4. **CPI**
12699 GTA GGT TAG AAA TGT TTG AAA ATT TAG CAC CCA GAA 12734

5. **CPII**
12878 TCT AAG TAG TGT TTT TTC TTA CAG ATA TGT TTT CAA 12913

6. **HEXON**
16323 GTA TAA AAT GGC TGG AAC TTC AAG TTC AGA ATT GAT 16358

7. **PVI**
13591 TGT TAT TAA TTT TTT AGA TGG ACA TAT CAA ATG CTA 13636

8. **ENDOPROTEASE**
16966 GCG CGT TAA CAA AAA GCG TCA TCA TCA CTT TCC ACC 17001

9. **PVIII**
21361 AAA GCA TAA ACA TGA GTG TCA ATC TAT TAA TTG TAT 21396

10. **PTP**
6770 CCA AGA TGA GCA AAT ACA TTT ATA CCA CTC AAA ACG 6734

11. **FIBER**
23873 ATA GGC TGA TGA AAA TAA ACT AGT GAT GCA ACT TTC 23907

12. **IIIa**
10968 CCA TAC TGA CCA ATT CCT GGT TCA TTT AGA TGG AAT 11004

13. **Iva2**
2339 GAA ATT TAA TAA AAT TTT TAT TCA AAC TTT TCA CAT 2303

14. **E2A DBP**
16979 TTT TGT TAA CGC GCA AAT AAA AAC AAT ACA ATA AAG 16943

15. **E3**
22106 TCT TAT TGA TAG ATT GTT TAC GAA GCA GAA GCA GAG 22141

16. **POL**
3435 GAT GAA TAA TGA AGT TGA AGA ATT CTA CGA TGT AGT 3399

FIG. 5. The predicted translational termination sites of HEV genes and the following 20 nucleotides. (Underlined) Stop codons.

of gene delivery. Because HEV infects chickens, as well as turkeys, it could be used as a vector in both species. More studies are required to identify unessential regions into which foreign genes could be inserted.

In the past few years, many new Ad sequences have been added to the databases. We suggest that as more sequence information becomes available, additional criteria, based on sequence similarity and genome organization, should be considered when Ads are classified. We have found that with regard to sequence composition and genome organization, HEV is closer to OVa than to other avian Ads like CELO and EDS. This points to the possibility that in the future, the division between mastadenoviruses and aviadenoviruses might be revised.

The determination of the complete HEV genome opens the way for the development of vaccines, using specific viral genes, in expression vectors, to produce subunit vaccines and better understanding of the nature of the virus.

MATERIALS AND METHODS

Virus isolation

Turkeys were exposed to the virulent field strain of HEV and killed after 5 days, and their spleens were removed. Twice, the tissue volume of TCN (10 mM Tris–HCl, pH 7.2, 10 mM $CaCl_2$, 100 mM NaCl) water was added, and the spleens were ground and homogenized. The homogenate was frozen and thawed three times at −70°C and 37°C, respectively, followed by centrifugation at 8000 *g,* at 4°C, for 20 min. The supernatant was mixed with trichlorotrifluoroethane (1:3) and centrifuged at 5000 *g*, at 4°C, for 30 min. The supernatant was collected and added on top of the following gradient prepared in TCN consisting of 12 ml of 46.2% (w/v) CsCl (density, 1.35 g/ml), 12 ml of 35% (w/v) CsCl (density, 1.24 g/ml), and 6 ml of 1 M sucrose. The gradient was centrifuged for 24 h at 85,000 *g*, at 4°C, with an SW28 rotor. The virus was isolated from a white ring formed between the two CsCl layers. This virus band was collected, diluted in TE (10 mM Tris, 1 mM EDTA), and repelleted by centrifugation at 96,000 *g* for 2 h. The pellet was collected, resuspended in distilled water, and dialyzed against TE. The virus was stored at −20°C until use.

Purification of HEV DNA

The virus was incubated for 3 h in a solution containing 0.01 M Tris, 0.01 M NaCl, 0.01 M EDTA, 0.5% SDS, and 50 mg/ml Proteinase K (Sigma). After incubation, the DNA of HEV was extracted by phenol–chloroform, precipitated with ethanol, and electrophoresed in 0.8% agarose gel. The viral DNA (26 kb in size) was visualized with ethidium bromide.

Sequencing procedure

HEV DNA was digested by either *Eco*RI or *Pst*I. The resulting restriction fragments were separated on agarose gel, purified, and cloned into plasmid pBS SK(+) (Stratagene). Cloned DNA fragments that were >3 kb were digested by exonuclease III (Promega) to create a series of nested deletions. The initial sequence information was obtained by sequencing the cloned fragments using two commercially available primers corresponding to the 5′- and 3′-ends of the pBS multiple cloning site (Universal Primers; New England Biolabs). The sequencing was carried out at the Unit for Biological Services of

the Weizmann Institute using the *Taq* Dyedeoxy Terminator system and an ABI 373 automatic sequencing apparatus. From 350 to 500 bp were resolved in a typical sequencing run. Once the initial sequence information had been obtained, primer walking was used to complete the sequencing of each cloned fragment. Gaps between the separate contiges were bridged by polymerase chain reaction (PCR) amplification of the viral DNA and subsequent sequencing of at least three independent PCR products per gap. When necessary, as in the case of the genome ends, the HEV DNA was directly sequenced. Both strands of the entire viral genome were sequenced, and each nucleotide was determined at least three times.

The processing of the raw sequence data, in the form of ABI trace files, and the assembly of the separate readings, first into contiges and then into a complete genome, were conducted with the GAP4 sequence assembly program and other programs that are included in the Staden package (Bonfield *et al.*, 1995).

The complete nucleotide sequence of HEV has been submitted to the GenBank, and its accession number is AF074946.

Sequence analysis

Sequence analysis was performed by means of the University of Wisconsin Genetics Computer Group software package [Wisconsin Package Version 9.1, Genetics Computer Group (GCG), Madison, Wisconsin].

Initial searches for similar sequences were conducted by the BLAST program, and more accurate searches were done by FASTA or TFASTA. Multiple sequence alignments were performed either by Pileup, which is part of the GCG package, or by Clustal X (Higgins and Sharp, 1988). The coding regions were identified by their similarity to sequences in GenBank of human Ad2 (J01917), Ad12 (X73478), Ad40 (L19443), EDS (Y09598), CELO (U46933), ovine Ads (U18755, U40837, U31557, and U40839), and canine Ad (U55001).

ACKNOWLEDGMENTS

We are very grateful to Prof. Shmuel Rozenblatt for his support, and we also thank Ofir Israeli for his technical assistance.

REFERENCES

Beard, C. W., Ball, A. O., Wooley, E. H. and Spindler, K. R. (1990). Transcription mapping of mouse adenovirus type 1 early region 3. *Virology* 175, 81–90.

Bonfield, J. K., Smith, K. F., and Staden, R. (1995). A new DNA sequence assembly program. *Nucleic Acids Res.* 24, 4992–4999.

Boulanger, P., Lemay, P., Blair, G. E., and Russell, W. C. (1979). Characterization of adenovirus protein IX. *J. Gen. Virol.* 44, 783–800.

Boursnell, M. E. G., Green, P. F., Samson, A. C. R., Campbell, J. I. A., Deuter, A., Peters, R. W., Millar, N. S., Emmerson, P. T., and Binns, M. M. (1990). A recombinant fowlpox virus expressing the hemmaglutinin-neuraminidase gene of Newcastle disease virus (NDV) protects chickens against challenge by NDV. *Virology* 178, 297–300.

Chen, M., Mermod, N., and Horwitz M. S. (1990). Protein-protein interactions between adenovirus DNA polymerase and nuclear factor-I mediate formation of the DNA replication preinitiation complex. *J. Biol. Chem.* 265, 18634–18642.

Chiocca, S., Kurzbauer, R., Schaffner, G., Baker, A., Mautner, V., and M. Cotton. (1996). The complete DNA sequence and genomic organization of the avian adenovirus CELO. *J. Virol.* 70, 2939–2949.

Davison, A.J., Telford, E. A. R., Watson, M. S., McBride, K., and Mautner, V. (1993). The DNA sequence of adenovirus type 40. *J. Mol. Biol.* 234, 1308–1316.

Ding, J., McGrath, W. J., Sweet, R. M., and Mangel, W. F. (1996). Crystal structure of the human adenovirus proteinase with its 11 amino acid cofactor. *EMBO J.* 15, 1778–1783.

Domermuth, C.H, and Gross, W. B. (1984). Hemorrhagic enteritis and related infections. *In* "Diseases of Poultry," 8th Ed. (M. S. Hofstad, H. J. Barnes, B. W. Calnck, W. B. Reid, and H. W. Yoder, Jr., Eds.). Iowa State University Press, Ames, Iowa, pp. 511–516.

Dragulev, B. P., Sira, S., Abouhaider, M. G., and Campbell, J. B. (1991). Sequence analysise of putative E3 and fiber genomic regions of two strains of canine adenovirus type 1. *Virology* 183, 298–305.

Freimuth, P., and Anderson, C. W. (1993). Human adenovirus serotype-12 virion precursors pMu and pVI are cleaved at amino-terminal and carboxy-terminal sites that conform to the adenovirus-2 endoproteinase cleavage consensus sequence. *Virology* 193, 348–355.

Green, N. M., Wrigley, N. G., Russel, W. C., Martin, S. R., and Mclachlan, A. D. (1983). Evidence for a repeating cross-β sheet structure in the adenovirus fiber. *EMBO J.* 8, 1357–1365.

Hess, M., Blocker, H., and Brandt, P. (1997). The complete nucleotide sequence of the egg drop syndrome virus: an intermediate between mastadenoviruses and aviadenoviruses. *Virology* 238, 145–156.

Higgins, D. G., and Sharp, P. M. (1988). Clustal: a package for performing multiple sequence alignment on a microcomputer. *Gene* 73, 237–244.

Jones, R. F., Asch, B. B., and Yohn, D. S. (1970). On the oncogenic properties of chicken embryo lethal orphan virus, an avian adenovirus. *Cancer Res.* 30, 1580–1585.

Jucker, M. T., McQuiston, J. R., van den Hurk, J. V., Boyle, S. M., and Pierson, F. W. (1996). Characterization of the haemorrhagic enteritis virus genome and the sequence of the putative penton base and core protein genes. *J. Gen. Virol.* 77, 469–479.

Kleiboeker, S. B. (1995). Sequence analysis of the fiber genomic region of a porcine adenovirus predicts a novel fiber protein. *Virus Res.* 39, 299–309.

Kozarsky, K. F., and Wilson, J. M. (1993). Gene therapy: adenovirus vectors. *Curr. Options Genet. Dev.* 3, 499–503.

Laver, W. G., Younghusband, H. B., and Wrigley, N. G. (1971). Purification and properties of chick embryo lethal orphan virus (an avian adenovirus). *Virology* 45, 598–614.

Lutz, P., and Kedinger, C. (1996). Properties of the adenovirus IVa2 gene product, an effector of late phase dependent activation of the major late promoter. *J. Virol.* 70, 1396–1405.

Mangel, W. F., McGrath, W. J., Toledo, D., and Anderson, C. (1993). Viral DNA and a viral peptide are cofactors of adenovirus virion proteinase activity. *Nature* 361, 274–275.

Mathias, P., Wickham, T., Moore, M., and Nemerow, G. (1994). Multiple adenovirus serotypes use alpha V integrins for infection. *J. Virol.* 68, 6811–6814.

Mei, Y. F., and Wadell, G. (1996). Epitopes and hemagglutination binding domain on subgenus B:2 adenovirus fibers. *J. Virol.* 70, 3688–3697.

Meissner, D. D., Hirsch, G. N., LaRue, E. A., Fulcher, E. A., and Spindler, K. R. (1997). Completion of the DNA sequence of mouse adenovirus type1: sequence of E2, L1 and L2 (18–51 map units). *Virus Res.* 51, 53–64.

Mul, Y. M., and van der Vliet, P. C. (1992). Nuclear factor I enhances

adenovirus DNA replication by increasing the stability of a preinitiation complex. *EMBO J.* 11, 751–760.

Pitcovski, J., Di-Castro, D., Shaaltiel, Y., Azriel, A., Gutter, B., Yarkoni, E., Michael, A., Krispel, S., and Levi, B. Z. (1996). Insect cell derived VP2 of infectious bursal disease virus confer protection against the disease in chickens. *Avian Dis.* 40, 753–761.

Rawlins, D. R., Rosenfeld, P. J., Wides, R. G., Challenberg, M. D., and Kelly., T. J., Jr. (1984). Structure and function of the adenovirus origin of replication. *Cell* 37, 309–319.

Salas, M. (1991). Protein-priming of DNA replication. *Ann. Rev. Biochem.* 60, 39–71.

Sarma, P. S., Huebner, R. J., and Lane, W. T. (1965). Induction of tumors in hamsters with an avian adenovirus. *Science* 149, 1108.

Shenk, T. (1996). Adenoviridae: the viruses and their replication. *In* "Virology" (B. N. Fields, D. M. Knipe, and P. M. Howley, Eds.) Lippincott-Raven, New York, pp. 2111–2148.

Shinagawa, M., Yoichi, I., Matsuda, A., Tsukiama, T., and Sato, G. (1987). Phylogenetic relationship between adenoviruses as inferred from nucleotide sequences of inverted terminal repeats. *Gene* 55, 85–93.

Smart, J. E., and Stillman, B. W. (1982). Adenovirus terminal protein precursor. *J. Biol. Chem.* 25, 13499–13506.

Stillman, B. W., Lewis, J. B., Chow, L. T., Mathews, M.B, and Smart. J. E. (1981). Identification of the gene and mRNA for the adenovirus terminal protein precursor. *Cell* 23, 497–508.

Stillman, B. W., Topp, W. C., and Engler, J. A. (1982). conserved sequences at the origin of adenovirus DNA replication. *J. Virol.* 44, 530–537.

Stouten, P. F. W., Sander, C., Ruigrok, R. W. H., and Cusack, S. (1992). New triple-helical model for the shaft of the adenovirus fiber. *J. Mol. Biol.* 226, 1073–1084

Suresh, M., and J. M. Sharma. (1996). Pathogenesis of type II avian adenovirus infection in turkeys: *in vivo* immune cell tropism and tissue distribution of the virus. *J. Virol.* 70, 30–36.

Temperley, S. M., and Hay, R. T. (1992). Recognition of the adenovirus type 2 origin of DNA replication by the virally encoded DNA polymerase and preterminal proteins. *EMBO J.* 11, 761–768.

van den Hurk, J. V. (1992). Characterization of the structural proteins of hemorrhagic enteritis virus. *Arch. Virol.* 126, 195–213.

van Oostrum, J., and Burnett, R. M. (1985). Molecular composition of the adenovirus type 2 virion. *J. Virol.* 56, 439–448.

Vrati, S., Brookes, D. E., Strike, P., Khatri, A., Boyle, D. B., and Both, G. W. (1996). Unique genome arrangement of an ovine adenovirus: identification of new proteins and proteinase cleavage sites. *Virology* 220, 186–199.

Webster, A., Leith, I. R., Nicholson, J., Hounsell, J., and Hay, R. T. (1997). Role of preterminal protein in adenovirus replication. *J. Virol.* 71, 6381–6389.

Yamanouchi, K., Inui, K., Sugimoto, M., Asano, K., Nishimaki, F., Kiching, R. P., Takamatsu, H., and Barrett, T. (1993). Immunisation of cattle with a recombinant vaccinia vector expressing the haemagglutinin gene of rinderpest virus. *Vet. Record* 13, 152–156.