Avian encephalomyelitis virus is a picornavirus and is most closely related to hepatitis A virus

Philip Marvil, Nick J. Knowles, A. P. Adrian Mockett, Paul Britton, T. David K. Brown and Dave Cavanagh

1 Institute for Animal Health, Compton Laboratory, Compton RG20 7NN, UK
2 Institute for Animal Health, Pirbright Laboratory, Pirbright GU24 0NF, UK
3 Intervet UK Ltd, Houghton, Huntingdon PE17 2BQ, UK
4 Department of Pathology, Division of Virology, University of Cambridge, Cambridge CB2 1QP, UK

The complete RNA genome of avian encephalomyelitis virus (AEV) has been molecularly cloned and sequenced. This revealed AEV to be a member of the Picornaviridae and consequently it is the first avian picornavirus for which the genome has been sequenced. Excluding the poly(A) tail the genome comprises 7032 nucleotides, which is shorter than that of any mammalian picornavirus sequenced to date. An open reading frame commencing at nucleotide 495 and terminating at position 6896 (6402 nucleotides) potentially encodes a polyprotein of 2134 amino acids. The polyprotein sequence has 39% overall amino acid identity with hepatitis A virus (HAV; genus Hepatovirus), compared to 19 to 21% for viruses from the other five picornavirus genera. Eleven cleavage products were predicted. The highest identity (49%) with HAV was in the P1 region, encoding the capsid proteins. The 5' and 3' untranslated regions (UTRs) comprise 494 and 136 nucleotides, respectively. The 5' UTR is the shortest of any picornavirus sequenced to date and, unlike HAV, it does not contain a long polypyrimidine tract.

Introduction

Picornaviruses are small ssRNA viruses with a genome of 7·1 to 8·9 kb whose non-enveloped capsids have icosahedral symmetry and comprise four structural proteins. They are classified into six genera, Enterovirus, Rhinovirus, Cardiovirus, Aphthovirus, Hepatovirus and Parechovirus which, so far, contain only viruses of mammals (Minor et al., 1995; Pringle, 1997). Viruses of a size, morphology and buoyant density indicative of the family Picornaviridae have been shown to cause, or have been observed in association with, a variety of diseases of chickens, turkeys and ducks (McNulty et al., 1990; Imada & Kawamura, 1997; Woolcock & Fabricant, 1997; McNeilly et al., 1994; Calnek et al., 1997). Moreover, their frequent location in the gut and stability at low pH has led to these viruses being referred to as enteroviruses or enterovirus-like viruses. However, there have been few biochemical investigations and no sequence analyses of these viruses, in part because of the difficulty of growing them in cell culture.

Avian encephalomyelitis (AE) is a disease of young chickens, pheasants, quail and turkeys, observed worldwide (Tannock & Shafren, 1994; McNulty & McFerran, 1996; Calnek et al., 1997). Infection of poultry with AE virus (AEV) by the faecal–oral route is common, field strains being enterotropic. The virus can cause slight reduction in egg production in susceptible laying hens and can be transmitted in embryos. This results in reduced hatching and those infected chicks which do hatch show characteristic tremors and/or ataxia between 1 and 7 days of age. These chicks act as a major source of infection for other chicks and a second wave of disease is often seen approximately 11 days later. The clinical signs are very much age-related as older birds show no neurological signs after infection. Virus used for the preparation of vaccine is usually grown in embryonating chicken
Table 1. Comparison of AEV with representative viruses of the six picornavirus genera

<table>
<thead>
<tr>
<th>Region</th>
<th>Hepatovirus HAV</th>
<th>Enterovirus PV-1</th>
<th>Rhinovirus HRV-1B</th>
<th>Cardiovirus EMCV</th>
<th>Aphthovirus FMDV-O</th>
<th>Parechovirus EV-22</th>
<th>Unclassified ERV-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ UTR</td>
<td>36:0</td>
<td>37:6</td>
<td>34:8</td>
<td>38:1</td>
<td>37:1</td>
<td>38:9</td>
<td>39:9</td>
</tr>
<tr>
<td>P1</td>
<td>49:2</td>
<td>65:2</td>
<td>16:9</td>
<td>37:0</td>
<td>18:2</td>
<td>37:5</td>
<td>20:0</td>
</tr>
<tr>
<td>P2</td>
<td>28:0</td>
<td>51:0</td>
<td>19:9</td>
<td>42:9</td>
<td>18:1</td>
<td>41:9</td>
<td>21:3</td>
</tr>
<tr>
<td>VP0</td>
<td>53:5</td>
<td>71:8</td>
<td>21:1</td>
<td>41:8</td>
<td>19:0</td>
<td>40:0</td>
<td>26:7</td>
</tr>
<tr>
<td>VP1</td>
<td>41:9</td>
<td>59:3</td>
<td>13:0</td>
<td>34:1</td>
<td>14:7</td>
<td>31:2</td>
<td>13:4</td>
</tr>
<tr>
<td>3C</td>
<td>39:5</td>
<td>63:3</td>
<td>18:8</td>
<td>39:8</td>
<td>21:6</td>
<td>44:8</td>
<td>16:4</td>
</tr>
<tr>
<td>3D</td>
<td>39:4</td>
<td>60:6</td>
<td>30:2</td>
<td>50:3</td>
<td>31:5</td>
<td>53:7</td>
<td>27:8</td>
</tr>
<tr>
<td>3’ UTR</td>
<td>49:2</td>
<td>42:3</td>
<td>50:0</td>
<td>50:0</td>
<td>41:9</td>
<td>44:6</td>
<td>45:6</td>
</tr>
</tbody>
</table>

All values are for protein comparisons except for the 5’ and 3’ UTRs which are nucleotide identities. HAV, human hepatitis A virus, strain HM-175; PV-1, poliovirus 1, strain Mahoney; HRV-1B, human rhinovirus 1B; strain B32; EMCV, encephalomyocarditis virus, strain 8; FMDV-O, foot-and-mouth disease virus O; strain Kaufbeuren; EV-22, echovirus 22, strain Harris; ERV-2, equine rhinovirus 2, strain P43e/71. %id, percentage identity; %sim, percentage similarity.

Methods

- **Virus.** The Calnek vaccine strain of AEV, grown in embryonated chicken eggs, was used.

- **Isolation of AEV RNA.** AEV was purified from a vaccine batch of AEV-infected embryos by differential centrifugation in caesium chloride and resuspended in water. The virion RNA was extracted by the proteinase K method (Sambrook et al., 1989).

- **Cloning the genome of AEV.** Standard molecular biological procedures, essentially as described by Sambrook et al. (1989), were used to clone AEV cDNAs. Briefly, cDNA synthesis was performed using a cDNA synthesis kit (Amersham), reverse transcription of AEV RNA being primed by oligo(dT15-18) or by random six residue primers. Second strand synthesis produced double-stranded cDNA which was subsequently ligated into the Smal site of pHCI19. This was used to transform Escherichia coli DH5α. Colonies were probed with virion RNA that had been 3’-end-labelled with [32P]dCTP using reverse transcriptase (Grunstein & Hogness, 1975). Plasmid DNA was prepared from 14 colonies. The cDNAs ranged in size from 0.1 to 2.9 kb. Some overlapped each other whilst three had poly(A) tails. The sequences of the 14 cDNAs were established using the Sequenase version 2.0 DNA sequencing kit (USB/Amersham).

- **AEV sequence-specific oligonucleotides** were designed on the basis of sequences near the ends of the cloned cDNAs and used in RT–PCR with Superscript II reverse transcriptase (Gibco) and Psv2 (Pfu) DNA polymerase (Stratagene). Sixteen clones were obtained and sequenced to bridge the gaps between the initial cDNA clones.

The sequence of the 5’ end of the genome was obtained by first priming an RT reaction with an oligonucleotide (AE-32) situated near the end of the 5’-most cDNA, subsequently shown to be located 322 to 330 nucleotides from the 5’ end of the genome. The single-stranded DNA product was C-tailed by terminal transferase (Promega), precipitated with ethanol and then amplified using a 14-mer oligo(dG) and oligonucleotide AE-32 in a 40 cycle Pfu PCR comprising 94 °C, 55 °C and 72 °C, each for 1 min. Some (1 μl) of the product was amplified in a 30 cycle Pfu PCR using oligo(dG) and either oligonucleotide AE-45 or AE-33 corresponding to positions subsequently identified as being 173 to 190 and 313 to 330 nucleotides from the 5’ end of the genome. cDNAs...
of approximately 230 and 360 bp were generated using oligonucleotides AE-45 and AE-33, respectively, and were ligated into the EcoRV site of pBluescript (Stratagene). Four clones (three and one obtained using oligonucleotides AE-33 and AE-45, respectively) were sequenced.

**Sequence analysis.** Sequence comparisons between AEV and other picornaviruses were made using the GCG 7.0 package (Devereux et al., 1984). Protein sequences of the P1 and 2C + 3CD poly peptides of representatives of the six picornavirus genera, plus AEV and ERV-2 were aligned using CLUSTALW (Thompson et al., 1994). Distance matrices were calculated using the program PROTDIST (Dayhoff option) and phylogenetic trees were constructed using the neighbour-joining algorithm of Saitou & Nei (1987) as implemented in the program NEIGHBOR; both programs were from the PHYLIP package (Felsenstein, 1993). The trees were drawn using the program TREEVIEW (Page, 1996).

The potential secondary structures of the untranslated regions of AEV were examined using the program RNAdraw v1.1 (Matzura & Wennborg, 1996).

**Results and Discussion**

**Features of the AEV genome and deduced polyprotein sequence**

The AEV genome sequence comprised 7032 nucleotides. The presence of a poly(A) tail indicated that the genome was a positive-sense RNA. An open reading frame commenced at nucleotide 495 and terminated at position 6896 (6402 nucleotides), potentially encoding a polyprotein of 2134 amino acids (Fig. 1). The polyprotein of mammalian picornaviruses is cleaved to produce 10 (parechoviruses), 11 (enteroviruses, rhinoviruses and hepatoviruses), 12 (cardioviruses and equine rhinoviruses) or 14 (foot-and-mouth disease viruses) end-products (Rueckert, 1996; Stanway et al., 1994). Cardio- and aphthoviruses have an N-terminal leader (L) polypeptide, which in the latter is a proteinase. Comparisons of the predicted polyprotein sequence of AEV with representatives of each of the picornavirus genera showed that it was most closely related to hepatitis A virus (Table 1). Therefore, the AEV polyprotein sequence was analysed for potential cleavage sites, based on an alignment with the HAV sequence. This analysis predicted that the AEV polyprotein would be cleaved into 11 products, assuming a cleavage of VP0 and no N-terminal L polypeptide (Fig. 1). Preceding and following the polyprotein open reading frame were 5’ and 3’ untranslated regions (UTRs) comprising 494 and 136 nucleotides, respectively.

**Analysis of the AEV untranslated regions**

The 5’ 150 nucleotides of the HAV RNA genome fold into a secondary structure unrelated to that of other picornaviruses. Instead of forming a cloverleaf-like structure (Andino et al., 1993; Walker et al., 1995), they apparently fold into three separate stem–loops, I, IIA and IIB, which are followed by the polypyr imidine tract pY1 (Brown et al., 1991; Shaffer & Lemon, 1995). It has also been proposed that the stability of stem–loops IIA and IIB is increased by the interaction of their loop sequences with nucleotides located 5’ of each stem, thus forming pseudo-knots (Shaffer et al., 1994). AEV lacks a long polypyr imidine tract and analysis of the 5’ 150 nucleotides using the program RNAdraw (Matzura & Wennborg, 1996) revealed five potential stem–loop structures (data not shown). Folding of the complete 5’ UTR of AEV still produced the first two stem–loops (bases 2 to 68). However, the remainder of the sequence folded to produce a complex structure with no obvious similarity to the predicted structure of HAV (data not shown).

The 3’ UTR of HAV was far shorter than that of AEV and could be folded into two simple stem–loops (data not shown). In contrast, the 3’ UTR of AEV could be folded into three more complex stem pools which contained one or two unpaired regions (data not shown).

**Polypeptide processing**

In picornaviruses, most polypeptide cleavages are effected by one or more virus-encoded proteases, although in cardio-, aphtho- and equine rhinoviruses the cleavage between P1–2A and 2B is effected by a poorly understood mechanism related to the 2A sequence itself (Donnelly et al., 1997). In all six genera, the 3C proteinase (3Cpro) functions both in cis and in trans and affects all polypeptide cleavages except for those mediated in cis by 2A and L (2A cleaves at its N terminus in entero- and rhinoviruses and L cleaves at its C terminus in aphthoviruses and equine rhinoviruses). The primary cleavage involving 2A differs between the picornavirus genera; in entero- and rhinoviruses the cleavage is between VP1 and 2A while in cardio-, aphtho- and equine rhinoviruses it is between 2A and 2B. In HAV this primary cleavage appears to be between P1–2A and 2B (Anderson & Ross, 1990; Borovec & Anderson, 1993). The predicted polypeptide cleavage sites of AEV and HAV are shown in Fig. 1 and features of the polypeptides are in Table 2.

1. **Primary cleavages (P1–2A/2B and P2/P3).** In HAV both the P1–2A/2B and P2/P3 cleavages are mediated by 3Cpro (Kusov et al., 1992; Schultheiss et al., 1994).

   The cleavage between HAV 2A and 2B was predicted to be between Q*389/G*381 (Cohen et al., 1987; Najarian et al., 1985; Paul et al., 1987). However, more recently it has been shown that the 2A/2B cleavage is between Q*389/A*337 (Martin et al., 1995; Schultheiss et al., 1995). AEV and HAV have very little sequence similarity in this area of the polyprotein and it is difficult to predict where the 2A/2B cleavage might take place. We have suggested Q*276/G*271 (which is a dipeptide probably used at the 3C/3D site; see below). However, an E/G pair 36 residues downstream could be an alternative 2A/2B cleavage site.
Fig. 1. For legend see facing page.
AEV is a picornavirus

Fig. 1. Comparison of the deduced amino acid sequence of the AEV polyprotein with that of HAV (HM-175; Cohen et al., 1987). Each sequence is presented in its entirety with dots inserted to maintain alignment. Identical amino acids are indicated by vertical lines. Predicted cleavage sites between the 11 final virus polypeptides are shown. The secondary structure of the 3C protease (Allaire et al., 1994; Bergmann et al., 1997) is indicated below that sequence: ‘--‘, α-helix; ‘---‘, β-sheet; ‘loop’ or ‘knob’, significant surface protrusions. Other sequence features and motifs discussed in the text are shown in reverse contrast.

AEV GELRAVTHVLNLKRLDKNPAPILVAFETVLYIYNQCSKFSLASMAIAVVLKCELGI 1176
HAV SPLRDCIAVHQQKLNLQSNQMTARCEPVLYCDKNGQSLTALATKICHYGV 1253
AEV SHVEGITYTIPNFMDWGYACLQSVDQHMTGDTDAPSPLRNLKMANLEK 1236
HAV EPEKNIYTVPSADYWGYSCQGKQKTQDAEDMFDCQLYSVPMLRMSLEE 1313
AEV KGTPFNSGPADNLSLHCPKTCYTDIAIRRHLHKVVKSPKEEF...HAMLDVAKAK 1294
HAV KGRHFSPSFIATNSWNSPSPTVYKIEADRLRHLKKEFVPSNPKNPSDMLNVLNLR 1373
AEV KAGYCNLDLDFPQISLDASTPSVQDVLVLMLHTVQDKTLMDIIOYVQA...
HAV TNDAIKDMSCVL...IMDGHNVSNMLLSSLVTMEIRFQNMPFLWSSQGISDDNN 1429
AEV <3A>.SNPREYPTQMTHEGNSGKVLYN...LFEKIKTFVYGAVLVSYSVFPAYHMKHMFP 1405
HAV DSAYEFQGPSGPSNKLSGLFQGSPVNTKHHVVAAGLVILVGSFVGYFVHKFSRE 1489

3A >> 3B (VPg)

HAV RDQQHDSAYMAAIKLPLVRVLEQGS...DAQVDVSDSNVHGNLVQGVGQPNEAIRHWLNYL
HAV EEPAPAMVYTVPTKVQVLADPDESQSTLEIAGLRVLKNQVFQGGQKENCVRWN

3B >> 3C (pro)

HAV ALGKVVDLWVFPSYAKMEMKYFPNFRTYYSIASGNVQLQSDVFGQDVNLKV
HAV B1 —— C1 —— D1 —— E1a —— E1b
HAV PRLPKRDITRRHNEETIRNEAESGPGTCLCTHLQFLTSRNLKEAATYVCESDDG
HAV PTPKFRDIQHPFIFKQGDLPRANLVLATLTGTVNTMILEQPKMEKATYHKHNDG
domain connection

HAV VRTDJSVGRSWSKAACTVAGGLGAVTSNNKMQIVQIVVAGGAHASSRVTKMEIE
HAV TTQDLTVQOQKRSNVEGPGQLSQNLSSQNIQZLAVANMLVAILQVTQEMFO
HAV —— B3 —— B4 —— B5 —— B6
HAV —— C2 —— E5 —— E6 —— E7 —— E8
HAV —— D2 —— F1 —— F2 —— F3 —— F4
HAV —— D3 —— F5 —— F6 —— F7 —— F8
HAV <3B>

3C << 3D (pol)

HAV MLKTRACQRWITKEFVEIKVSQGKTHYKSLPDYPCQEIVKCPKTLF.YQKEDIVM
HAV NIDKKIESQRQIMKVEFTQCSMNVMVSKTILFRK3PIYHIHDKTMNPANPAMPSKAESIDMA
HAV —— X ——

3C >> 3D (pol)

HAV VMLAKYSSPIVESAPLATGTVVEATYNRMRWSFEPFRO.LTYDECNIEGLDAILDKTS
HAV VMLSKSFLPVEPEPDYKEASIFYQNKVIQKVQTLVVDPLFDLDMLMTSAPGIDAINMSSP
HAV GFYPNTLGRKSLLINGK.....MAQLQDVQKEVMDLHNRNSOQVFPTTCAKDELR
HAV GFYPQVQEKETKRDLWLMLNGGJLVPHquQRLPFTVNMECSDLVQVPTTCQPKDLR
HAV PLSKVMLGTIRAIECPVSFSTILFRILGQAALQISHQPGHFTIGAVGDQDWHCMWY
HAV PLEKLVEKTRAIDACPDLYSLRMYQWPASYFHLNPFGHTQVAGIDPRQWMDELFK
HAV SIVTOCLDVLWFGQNYDASLSPFPIMHYAGRVLQYOICQGDLFPRLMPEVPQSNVQHGS
HAV TMIRGDVLDDLDSFADASLSFPIMREARGSMLSEGSTPSHFGTALINTIYSHYHLYN
HAV MRYVHHSGNQSGTAPSLNLSINNYINHVLCAKISFVPSFKILYYCDEVLCLL
HAV CCYICGSMIPNLCBLNINCNWLWYYFSLQPKGSPVCFQALKPQYQDVLIVF
HAV KKEY.LDQKSFPSLSPVQSLELLGLSPTADMEKVVTIVPKHMPSLQFRFTVDWESICHP
HAV SRTQDLNLDLQIGKTVDEFKLLGNTATSDKKNVQPLKPSLTLKFSNMLVEDR1.RP
HAV RISSETVYSMLAWSKDNSSDMKLDILTTSIWPMPHGCPRKYYTFTCLWTLRGVCVBGILYTP
HAV AISEKTIWLISIWWRSNANAEPEQNLNENAAQWFAFMNHGCEYFQKFFYTFQSCLKEEMYEYRLK
HAV TYKELEVRDVLRYFIRDDSF 2134
HAV SYDVWRMRP...YDQCFICDLS. 2227

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 24 Nov 2018 08:30:25
Table 2. Genome features of AEV

<table>
<thead>
<tr>
<th>Gene</th>
<th>Nucleotide sequence</th>
<th>Amino acid sequence</th>
<th>Predicted N-terminal cleavage sites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Start</td>
<td>End</td>
<td>Length</td>
</tr>
<tr>
<td>5' UTR</td>
<td>1</td>
<td>494</td>
<td>494</td>
</tr>
<tr>
<td>Polypeptide</td>
<td>495</td>
<td>6896</td>
<td>6402</td>
</tr>
<tr>
<td>P1</td>
<td>495</td>
<td>2765</td>
<td>2271</td>
</tr>
<tr>
<td>P2</td>
<td>2766</td>
<td>4535</td>
<td>1770</td>
</tr>
<tr>
<td>P3</td>
<td>4536</td>
<td>6896</td>
<td>2361</td>
</tr>
<tr>
<td>VP0</td>
<td>495</td>
<td>1220</td>
<td>726</td>
</tr>
<tr>
<td>VP4</td>
<td>495</td>
<td>551</td>
<td>57</td>
</tr>
<tr>
<td>VP2</td>
<td>552</td>
<td>1220</td>
<td>609</td>
</tr>
<tr>
<td>VP3</td>
<td>1221</td>
<td>1955</td>
<td>735</td>
</tr>
<tr>
<td>VP1</td>
<td>1956</td>
<td>2765</td>
<td>810</td>
</tr>
<tr>
<td>2A</td>
<td>2766</td>
<td>2804</td>
<td>39</td>
</tr>
<tr>
<td>2B</td>
<td>2805</td>
<td>3557</td>
<td>753</td>
</tr>
<tr>
<td>2C</td>
<td>3558</td>
<td>4535</td>
<td>978</td>
</tr>
<tr>
<td>3A</td>
<td>4536</td>
<td>4730</td>
<td>195</td>
</tr>
<tr>
<td>3B</td>
<td>4731</td>
<td>4793</td>
<td>66</td>
</tr>
<tr>
<td>3C</td>
<td>4794</td>
<td>5438</td>
<td>645</td>
</tr>
<tr>
<td>3D</td>
<td>5439</td>
<td>6896</td>
<td>1458</td>
</tr>
<tr>
<td>3' UTR</td>
<td>6897</td>
<td>7032</td>
<td>136</td>
</tr>
</tbody>
</table>

* Confirmed by N-terminal sequencing (Gauss-Müller et al., 1986, 1991; Harmon et al., 1992; Tesar et al., 1994).

The P2/P3 cleavage in HAV has been confirmed, by N-terminal sequencing, as being between Q1412/G1423 (Harmon et al., 1992); in AEV the Q1437/S1438 dipeptide is the most likely equivalent site.

(2) Secondary cleavages. The VP0/VP3 cleavage in HAV is between Q245/M246 and in AEV it is predicted to occur at Q242/M243 in a region of high identity between the two viruses. The VP3/VP1 cleavage has been identified by terminal sequencing of the mature HAV polypeptide as Q1081/V1082 (Gauss-Müller et al., 1986) and in AEV it is predicted to be at Q1467/G1468 although primary sequence identity in this region is low.

In HAV the cleavage between VP1 and 2A has variously been predicted to be between Q745/S746 (Schultheiss et al., 1994), E791/S792 (Cohen et al., 1987) or Q836/A837 (Najarian et al., 1985; Paul et al., 1987). The first of these seems unlikely since it lies upstream of a sequence, PRA751-753, commonly found in picornavirus VP1 sequences and which immediately follows β-sheet I. Jia et al. (1993) reported that the cleavage site lay about 40 amino acids downstream of the previously predicted E791/S792 and suggested that Q836/A837 may be the junction. However, it was later shown that this is the 2A/2B cleavage site (Martin et al., 1995; Schultheiss et al., 1995). It seems probable that, in HAV, E791/S792 is the junction between VP1 and 2A. In AEV an E/S dipeptide is also present at positions 757/758 and we predict that this is the most likely VP1/2A cleavage site.

In HAV the 2B/2C cleavage is mediated by 3CPPRO and occurs at residues Q1087/S1088 (Kusov et al., 1992; Schultheiss et al., 1994). Although there is little similarity between AEV and HAV in this region, alignment of the two sequences suggests that a Q/G at positions 1021/1022 could be used by the former.

The AEV 3B or VPg polypeptide is predicted to lie between two Q/S cleavages at positions 1412/1413 and 1433/1434, resulting in a protein two amino acids shorter than its HAV equivalent. The cleavage between 3C and 3D is again predicted to lie in a region of little similarity between AEV and HAV at Q1461/C1464.

It would thus appear that all the polyprotein cleavages in AEV (except that between VP4 and VP2, if it exists) are probably mediated by 3CPPRO as is the case for HAV.

(3) Cleavage of VP0. Picornaviruses, except for echovirus 22 and 23 (Stanway et al., 1994), undergo an autocatalytic cleavage following encapsidation of the viral RNA; thus VP0 is cleaved to give VP4 and VP2. In HAV inefficient cleavage of VP0 has been reported (Bishop & Anderson, 1993). However, no VP4 polypeptide has ever been identified. Alignment of the AEV VP0 with that of HAV shows very little sequence identity in the first 35 amino acid residues. Consequently,
prediction of a maturation cleavage site, if it exists, is difficult; we suggest that L<sup>19</sup>/N<sup>20</sup> may be used.

(4) Myristoylation of P1. In most picornviruses the P1 precursor polypeptide is covalently bound by its N-terminal glycine residue (when the methionine is removed) to a molecule of myristic acid via an amide linkage (Chow et al., 1987; Paul et al., 1987). Consequently, the cleavage products VP0 and VP4, which contain the P1 N terminus, are also myristoylated. This myristoylation is carried out by myristoyl transferase which recognizes an eight amino acid signal beginning with glycine. In picornviruses a five residue consensus sequence motif, GxxxT/S, has been identified (Palmenberg, 1989). HAV has such a sequence but it commences five or seven residues into the proposed capsid region (Fig. 1). HAV is reported to use two different AUG codons to initiate translation (Tesar et al., 1992). It has been postulated that a short leader peptide could be released by cleavage between Q<sup>6</sup>/G<sup>7</sup>, leaving the sequence GIFQT at the terminus. However, it has been shown that this cleavage does not occur and also that no myristoylation takes place (Tesar et al., 1993). In the parechoviruses (echovirus 22 and 23) myristoylation consensus sequences occur 13 residues into VP0. However, although this polypeptide is apparently not myristoylated, there does appear to be some type of molecule blocking its N terminus (Stanway et al., 1994). In AEV the first sequence which fits the myristoylation motif is GPDRT at residues 29 to 33. However, the last three amino acids of this sequence are completely conserved with HAV and would probably lie within VP2 (Fig. 1). We would therefore predict that the AEV VP0 is not myristoylated.

Comparison of AEV with mammalian picornviruses

The genome of AEV (7032 nucleotides) is smaller than that of any of the mammalian picornviruses (7102 to > 8828 nucleotides; Skern et al., 1985; Wutz et al., 1996). The 5’ UTR of AEV, at 494 nucleotides, is considerably shorter than those of other picornviruses (610 to 1200) while the 3’ region, at 494 nucleotides, is considerably shorter than those of other picornviruses (50). Consequently, the cleavage products VP0 and VP4, which contain the P1 N terminus, are also myristoylated. This myristoylation is carried out by myristoyl transferase which recognizes an eight amino acid signal beginning with glycine. In picornviruses a five residue consensus sequence motif, GxxxT/S, has been identified (Palmenberg, 1989). HAV has such a sequence but it commences five or seven residues into the proposed capsid region (Fig. 1). HAV is reported to use two different AUG codons to initiate translation (Tesar et al., 1992). It has been postulated that a short leader peptide could be released by cleavage between Q<sup>6</sup>/G<sup>7</sup>, leaving the sequence GIFQT at the terminus. However, it has been shown that this cleavage does not occur and also that no myristoylation takes place (Tesar et al., 1993). In the parechoviruses (echovirus 22 and 23) myristoylation consensus sequences occur 13 residues into VP0. However, although this polypeptide is apparently not myristoylated, there does appear to be some type of molecule blocking its N terminus (Stanway et al., 1994). In AEV the first sequence which fits the myristoylation motif is GPDRT at residues 29 to 33. However, the last three amino acids of this sequence are completely conserved with HAV and would probably lie within VP2 (Fig. 1). We would therefore predict that the AEV VP0 is not myristoylated.

The complete polyprotein had 39-3% identity with that of HAV (492 to > 8828 nucleotides; Skern et al., 1985; Wutz et al., 1996). The 5’ UTR of AEV, at 494 nucleotides, is considerably shorter than those of other picornviruses (610 to 1200) while the 3’ UTR (136 nucleotides) is not unusual (42 to 164 nucleotides). In addition, the 5’ UTR of AEV did not possess a long poly pyrimidine tract which in HAV lies between 5’ end replication-associated structures (including two pseudo-knots) and an internal ribosome entry site (IRES) (Shaffer et al., 1994).

The deduced amino acid sequences of the AEV polyprotein and the predicted cleavage products were compared with the corresponding proteins of seven picornviruses, including representatives of the six mammalian genera and the unclassified equine rhinovirus type 2 (ERV-2; Wutz et al., 1996).

The complete polyprotein had 39-3% identity with that of HAV and 19-2 to 21-7% identity with representatives of the five other genera and ERV-2 (Table 1). Most striking was the high identity (49-2%) between the P1 precursor proteins, comprising the capsid proteins, of AEV and HAV, extending to 53-5% for the VP0 component (Fig. 1). The lower identity (41-9%) between the VP1 proteins is in keeping with the observation that this is the most variable capsid protein among picornviruses. The predicted lengths of the AEV structural poly peptides (VP0, 242 aa; VP3, 245 aa; VP1, 270 aa) are very similar to those of HAV (VP0, 245 aa; VP3, 246 aa; VP1, 300 aa) and the three dimensional capsid structures are likely to be similar. Recently, a linear antigenic epitope has been discovered on VP3 of HAV (Bosch et al., 1998). Convalescent human sera recognized a synthetic peptide (FWRGDLVFDFQV; located at residues 355 to 366 in Fig. 1). The replacement of the arginine, glycine, or aspartic acid at positions 357, 358, or 359, respectively by other amino acids induced the loss of synthetic peptide recognition by human convalescent sera as did shorter versions of the peptide. Antibodies generated by this peptide bound to intact HAV and neutralized its infectivity. Antipeptide antibodies inhibited convalescent serum binding to HAV and two HAV monoclonal antibodies completely inhibited the binding of antipeptide antibodies to HAV. The equivalent sequence in AEV is highly conserved with respect to HAV (Fig. 1) having substitutions only at aa positions 360 (L to I) and 363 (D to E). This high degree of conservation suggests that HAV antibodies to this region may cross-react with AEV.

The lowest identity between AEV and HAV (28-0%) was seen in the P2 region. This resulted principally from the very low similarities between the 2A and 2B proteins; the 2C protein, which is involved in RNA synthesis, had 40-7% identity. The 2C poly peptides of both AEV and HAV contain the NTP binding motifs GxGxGKS/T (domain A) and byhyhyxxD (in which hy is any hydrophobic residue; domain B) present in putative helicases and all picornavirus 2Cs (Gorbalenya et al., 1988, 1989; Fig. 1).

Identity between the P3 proteins, comprising 3A, 3B, 3C and 3D, of AEV and HAV was lower than that seen when comparing the P1 region. Little is known about the function of the 3A poly peptide. However, all picornavirus 3A proteins contain a putative transmembrane x-helix. Although primary sequence identity is low in this protein AEV and HAV both have a hydrophobic stretch at positions 1378 to 1398 and 1462 to 1482, respectively (Fig. 1). The genome-linked poly peptide, VPg, which is encoded by the 3B region, shared only six amino acids with that of HAV. However, the third residue was a tyrosine, consistent with its linkage to the 5’ end of the virus genome (Rothberg et al., 1978). The three dimensional structure of the HAV 3C cysteine protease has recently been solved and the active-site residues identified (Allaire et al., 1994; Bergmann et al., 1997). The cysteine at 1691 is the nucleophile, H<sup>1691</sup> is the general base and the specificity for glutamine residues is defined mainly by H<sup>1716</sup>; all three residues are conserved in the AEV sequence (Fig. 1). The 3D poly peptide is the major component of the RNA dependent RNA polymerase and both AEV and HAV contain motifs conserved in picornavirus like virus RNA dependent RNA polymerases; PSG, YGDD and FLKR (Argos et al., 1984; Fig. 1).
Fig. 2. Phylogenetic relationship of AEV with representative species of the six mammalian picornavirus genera. The relationship was examined by the neighbour-joining algorithm using amino acid similarities of (a) the capsid precursor polypeptide P1 and (b) the combined 2C, 3C and 3D non-structural protein region. The bars represent distance as units of the expected fraction of amino acids changed (Felsenstein, 1993). The nucleotide sequence databank accession numbers of the viruses used are: human hepatitis A virus (HAV), M14707, K02990, M20273, X75215 and M34084; simian HAV (SHAV), DO0924 and M59286; echovirus type 22 (EV-22), L02971; echovirus type 23 (EV-23), AJ005695; equine rhinovirus (ERV) type 2, X96871; encephalomyocarditis virus (EMCV), M81861; Theiler’s murine encephalomyelitis virus (TMEV), M20562; ERV type 1 (ERV-1), X96870; foot-and-mouth disease virus (FMDV) type A, M10975; FMDV type O, X00871; FMDV type SAT2 was provided by A. M. Q. King; human rhinovirus (HRV) type 2, X02316; HRV type 14, K02121; coxsackievirus (CV) type A9, D00627; CV type A16, U05876; CV type A21, D00538; CV type B5, X67706; poliovirus (PV) type 1, J02281; human enterovirus (EV) type 70, D00820; and bovine enterovirus (BEV) type 1, D00214.
Phylogenetic analysis

The relationships between the proteins of AEV and those of the mammalian picornaviruses were also examined by the neighbour-joining algorithm using amino acid similarities of P1, the capsid protein precursor (Fig. 2a), and the combined 2C, 3C and 3D proteins (Fig. 2b), which were the most highly conserved sequences in the non-structural protein-coding region. This confirmed that, in respect of both the structural and non-structural proteins, AEV was clearly more closely related to HAV virus than to members of the other picornavirus genera.

Taxonomic status

AEV is currently not assigned to a genus within the Picornaviridae (Minor et al., 1995). As with several other avian viruses with picornaviral morphology AEV has been referred to as an ‘enterovirus’. Our data demonstrate unequivocally that AEV is a member of the Picornaviridae, but it is equally clear that AEV should not be assigned to the genus Enterovirus. Whether AEV should be assigned to the genus Hepatovirus, currently comprising human and simian HAVs, or to a new genus is a matter for debate.

We would like to thank W. Baxendale, Intervet UK Ltd, for financial support for this project. We are also indebted to Drs F. Ghazi and G. Stanway for allowing us to use the echovirus 23 sequence prior to publication and to Dr A. M. Q. King for access to the FMDV-SAT2 sequence.

References


Received 27 July 1998; Accepted 4 November 1998