Molecular analysis of duck hepatitis virus type 1 reveals a novel lineage close to the genus *Parechovirus* in the family *Picornaviridae*

Min-Chul Kim, Yong-Kuk Kwon, Seong-Joon Joh, A. Michael Lindberg, Jun-Hun Kwon, Jae-Hong Kim and Sun-Joong Kim

1National Veterinary Research and Quarantine Service, 480 Anyang 6 dong, Manan-gu, Anyang, Gyeonggi-do 420-824, Republic of Korea
2Department of Chemistry and Biomedical Sciences, University of Kalmar, SE-391 82 Kalmar, Sweden
3College of Veterinary Medicine, Seoul National University, Seoul 151-742, Republic of Korea

Duck hepatitis virus type 1 (DHV-1) was previously classified as an enterovirus, based primarily on observed morphology and physicochemical properties of the virion. The complete nucleotide sequences of two strains (DRL-62 and R85952) of DHV-1 have been determined. Excluding the poly(A) tail, the genomes are 7691 and 7690 nt, respectively, and contain a single, large open reading frame encoding a polyprotein of 2249 aa. The genome of DHV-1 is organized as are those of members of the family *Picornaviridae*: 5′ untranslated region (UTR)–VP0–VP3–VP1–2A1–2A2–2B–2C–3A–3B–3C–3D–3′ UTR. Analysis of the genomic and predicted polyprotein sequences revealed several unusual features, including the absence of a predicted maturation cleavage of VP0, the presence of two unrelated 2A protein motifs and a 3′ UTR extended markedly compared with that of any other picornavirus. The 2A1 protein motif is related to the 2A protein type of the genus *Aphthovirus* and the adjacent 2A2 protein is related to the 2A protein type present in the genus *Parechovirus*. Phylogenetic analysis using the 3D protein sequence shows that the two DHV-1 strains are related more closely to members of the genus *Parechovirus* than to other picornaviruses. However, the two DHV-1 strains form a monophyletic group, clearly distinct from members of the genus *Parechovirus*.

**INTRODUCTION**

The family *Picornaviridae* is divided into nine genera (*Aphthovirus*, *Cardiovirus*, *Enterovirus*, *Hepatovirus*, *Parechovirus*, *Rhinovirus*, *Erbovirus*, *Kobuvirus* and *Teschovirus*) (Stanway *et al.*, 2005). Sequence comparisons of viral genomes are now recognized as the method of choice when classifying viruses into families, genera and species.

The genomes of picornaviruses are positive-stranded RNA molecules that are also used as mRNAs during infection. A picornaviral genome comprises approximately 7200–8500 nt and contains a single, long open reading frame (ORF) encoding a polyprotein of more than 2000 aa. The long ORF is preceded by a 5′ untranslated region (UTR), followed by a short 3′ UTR and a poly(A) tail (Racaniello, 2001). The polyprotein contains the following organization: NH$_2$–(L)–VP0(VP4–VP2)–VP3–VP1–2A–2B–2C–2A–3A–3B–3C–3D–COOH, where all picornavirus genes are orthologues within the family, except L and 2A. The VP1–4 proteins constitute the virion, whilst the other proteins assist with intracellular activities during virus replication.

The picornavirus polyprotein is processed autocatalytically at the conserved interdomain junctions by a proteolytic activity associated with the 3C protease, which, depending on the individual picornavirus, may also be assisted by the proteolytic activities of the L and/or 2A proteases with different specificities (Dougherty & Semler, 1993; Kräusslich & Wimmer, 1988; Palmenberg, 1990; Ryan & Flint, 1997). Additional cleavages of the polyprotein at a few alternative sites may take place, resulting in new products and some intermediate precursors, some of which are stable and/or functionally active (Racaniello, 2001). In most picornaviruses, VP0 has a myristylation motif (GxxxA/T) that is cleaved autocatalytically into VP4 and VP2 proteins during the final stage of virion maturation. The conserved proteins in the family *Picornaviridae* include the multifunctional 2C ATPase (2C$\text{ATP}$$_\text{ase}$), 3C protease (3C$\text{pro}$), 3D RNA-dependent RNA polymerase (3D$\text{pol}$), membrane-associates...
2B and 3A proteins and a small 3B protein (3BVPg) (Racaniello, 2001).

The family Picornaviridae comprises a wide range of important human and animal pathogens. Picornaviruses infect mammals, including humans, and birds (King et al., 2000). Depending on the nature of the individual picornavirus, the infection may cause severe ailments of the gastrointestinal tract and the respiratory, neural, hepatocellular and cardiovascular systems (Hollinger & Emerson, 2001; Pallansch & Roos, 2001).

Duck hepatitis is a fatal, rapidly spreading viral infection of young ducklings characterized primarily by hepatitis. So far, three different viruses, duck hepatitis virus (DHV) types 1, 2 and 3, have been associated with these symptoms. DHV-1 (Wildy, 1971) was classified as an enterovirus and DHV-3 (Haider & Calnek, 1979) was classified as a picornavirus. However, DHV-2 was classified as an astrovirus (Gough et al., 1984). Among DHV types 1, 2 and 3, DHV-1 is now worldwide in distribution and of most economic importance to all duck-growing farms because of the high potential mortality when an infection is not controlled. The virion of DHV-1, 20–40 nm in diameter (Richter et al., 1964), is a non-enveloped, icosahedral particle resistant to ether and chloroform and capable of survival for long periods under usual environmental conditions (Woolcock, 2003). DHV-1 is comparatively stable at temperatures below 40–45°C, but is inactivated rapidly at higher temperatures. The virion has been shown to be stable at pH 3 for 9 h, but not for 48 h (Davis, 1987), and DHV-1 was therefore first classified as an enterovirus-like virus (Wildy, 1971). The growth of DHV-1 is not inhibited in the presence of the nucleoside analogue 5-bromo-2-deoxyuridine, indicating that it has an RNA genome (Tauraso et al., 1969).

The family Picornaviridae was originally classified into four genera based on the antigenic and biophysical properties of the virions (Melnick et al., 1974). Subsequent molecular analysis of the viral genomes supported this classification for the majority of picornaviruses. Likewise, molecular methods also led to the creation of two additional genera, Hepatovirus and Parechovirus, comprising two previously classified enteroviruses that have diverged substantially from other picornaviruses (King et al., 2000). In order to accommodate the molecular diversity of newly recognized picornaviruses, three other genera have recently been introduced (Pringle, 1999; Stanway et al., 2005). The constant increase in the number of picornavirus genera, the current bias towards discovered picornaviruses that infect humans and the huge disparity in the numbers of viruses assigned to different genera (King et al., 2000) indicate that the diversity of the family Picornaviridae is far from being fully described.

In this study, we report and analyse the complete nucleotide sequence of two DHV-1 strains. Genetic and phylogenetic analyses demonstrate clearly that DHV-1 is a member of the family Picornaviridae. The organization of the DHV-1 genome, a map of the polyprotein with suggested processing sites and the unexpected observation of two different 2A protein motifs are described. Comparative and phylogenetic analyses using 3D protein sequences show that DHV-1 constitutes a distinct monophyletic group related to the genus Parechovirus, but with distinct features.

**METHODS**

**Viruses used in the analyses.** Strains DRL-62 and R85952 of DHV-1 were obtained from the ATCC and propagated in 9-day-old embryonated chicken eggs. Previously reported picornavirus sequences used in the comparisons (Table 1) were retrieved from GenBank.

**Isolation of DHV RNA.** DHV-1 viruses (strains DRL-62 and R85952) were purified from infected allantoic fluids by sucrose-gradient ultracentrifugation. Briefly, the allantoic fluids of DHV-1-infected embryos were freeze-thawed three times and were centrifuged at 3000 g for 30 min. The supernatants were concentrated by ultracentrifugation for 8 h at 35 000 r.p.m. by using an SW-41Ti rotor (Beckman Coulter) at 4°C. The pellets were suspended in PBS and loaded onto a 10–40% continuous linear sucrose gradient prepared by using an SW-41Ti rotor. The samples were ultracentrifuged on the gradient for 18 h at 35 000 r.p.m. at 4°C. The pellets were resuspended in PBS and the viral RNA was extracted by using TRizol reagent (Invitrogen) according to the manufacturer’s recommendations.

**Cloning the genome of DHV-1.** Standard molecular biological procedures as described by Sambrook et al. (1989) were used to clone DHV-1 cDNAs. Briefly, reverse transcription was performed by using a cdNA synthesis kit (TaKaRa) and oligo(dT)12–18 or random priming. Double-stranded cdNA was ligated into the Smal site of plasmid pUC18 and the recombinant plasmids were transformed into competent Escherichia coli DH5α. Plasmid DNA was prepared from 30 colonies. Analyses of the insert size of 30 colonies revealed insert sizes between 0.4 and 1.5 kb. Sequencing of the inserted plasmid DNA was performed with the M13 forward and reverse primer set by using an ABI PRISM BigDye Termination cycle sequencing kit (Applied Biosystems) and an ABI 377 automated sequence analyser (Applied Biosystems). The nucleotide sequences of the inserts showed, by using a protein–protein BLAST search at the National Center for Biotechnology Information (NCBI) (Altschul et al., 1997), that 20 of the colonies contained deduced amino acid sequences similar to those of members of the genus Parechovirus. Of 20 colonies with overlapping sequences, two cDNA clones contained a poly(A) sequence and an identical putative 3′ UTR sequence. The remaining gaps were covered by successive rounds of PCR using primers derived from known terminal sequences. The 5′ end of the genome was obtained by using a 5′ rapid amplification of cdNA ends (5′-RACE) kit (TaKaRa) as described by the manufacturer. PCR products were sequenced at least twice for three clones for each product.

**Comparative sequence analysis.** The Vector NTI suite 8 program (InforMax) was used for retrieving the nucleotide sequences of the genes of picornaviruses via the NCBI Entrez server (Table 1). The retrieved sequences were trimmed with the BioEdit program (Hall, 1999) and sequence comparisons between strains DRL-62 and R85952 and other picornaviruses were made by using a sequence-identity matrix in the BioEdit program. Phylogenetic trees were generated by using the neighbour-joining method (Saitou & Nei, 1987) with default settings as implemented in the CLUSTAL_X package (Higgins & Sharp, 1988). Bootstrap probabilities were calculated with 1000 replicates (Felsenstein, 1985). The generated phylogenetic trees were visualized by using the program TreeView (Page, 1996).
RESULTS AND DISCUSSION

Genome features of DHV-1

The genomes of the DRL-62 and R85952 strains of DHV-1 comprised 7691 and 7690 nt, respectively [excluding the poly(A) tail]. The presence of the poly(A) tail in the 3' end of the genome indicated that the DHV-1 genome is of positive-sense polarity and belongs to the family Picornaviridae. The G+C contents of the DHV-1 genomes were 43.3–43.6 mol%, which is similar to the G+C frequency of Ljungan virus (41.6–42.5 mol%), hepatoviruses [44.8 mol% for avian encephalomyelitis virus (AEV)], human parechoviruses (39.0–39.5 mol%) and Human rhinovirus A (39.0 mol%). The 5' and 3' UTRs are separated by an ORF that starts at nt 627 and ends at nt 7373 (6747 nt), encoding a polyprotein of 2249 aa. The 5' UTRs of the DRL-62 and R85952 genomes consist of 626 nt, whilst the 3' UTRs of DRL-62 and R85952 include 318 and 317 nt, respectively, including the amino acid termination codon [the poly(A) tail is not included]. The genomes and polyproteins of the two DHV-1 strains have 97.0% identical nucleotide and 98.7% identical amino acid sequences.

The polyprotein of mammalian picornaviruses is cleaved to produce 10 (human 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). Pairwise comparisons of the polyprotein sequence of DHV-1 with those of representatives of each of the picornavirus genera showed that it is related most closely to the genus Parechovirus, which contains the species Ljungan virus and Human parechovirus. Therefore, the polyprotein sequences of DHV-1 were analysed for potential cleavage sites by alignment with the Ljungan virus and human parechovirus polyprotein sequences. Based on this analysis, the DHV-1 polyprotein is predicted to be processed into 11 final products and the genome organization of the DHV-1 would then be: 5' UTR–VP0–VP3–VP1–2A1–2A2–2B–2C–3A–3B–3C–3D–3' UTR. The major features of the polyproteins encoded by the genomes of the two DHV-1 strains are summarized in Table 2. The genome organization and presence of two different 2A amino acid sequence motifs have, for the Picornaviridae, previously only been described for Ljungan viruses (Johansson et al., 2002). The sequence identity of the DHV-1 genome compared with those of other members of the family Picornaviridae is summarized in Table 3. Although it is clear that DHV-1 has a picornavirus-like organization of its genome, the sequence identity towards other picornaviruses is low. The 3D sequence is frequently used to infer phylogenetic relationships among members of the family Picornaviridae and the closest relatives to DHV-1 are the Swedish strains of Ljungan virus, with only 40% amino acid sequence identity (Table 3).
Nucleotide sequences of the 5’ UTR of DHV-1 strains were obtained by the 5’-RACE method (TaKaRa). The sequences of the two strains, DRL-62 and R85952, were both 626 nt long and the sequences of the 5’ UTR were shorter than corresponding regions of the genomes of members of the genus Parechovirus, Ljungan viruses and human parechoviruses. Sequence identity between the two DHV-1 strains was 96–97% and sequence identities to other picornaviruses in this region were <41% (Table 3). The 5’ UTR of DHV-1 precedes the putative codon for initiation of translation. This codon is located in an optimal Kozak context (GxxAUGG) (Kozak, 1986). The initiation codon of DHV-1 is thus located at nt 627 in the genome. A pyrimidine-rich tract is present between nt 549 and 555 and this is followed by the initiator methionine (AUG) codon, 71 nt after this region.

The 5’ UTR of picornaviruses contains the internal ribosome entry site (IRES), an element that directs translation of the mRNA by internal ribosome binding (Racaniello, 2001). Based on the RNA structure, three major classes of untranslated IRES elements, IRES types I, II and III, have been defined (Beales et al., 2003). The 5’ 626 nt of the DHV-1 RNA genome folds into a secondary structure (Zuker, 2003). Computer-assisted prediction of the secondary structure of the 5’ UTR RNA of DHV-1 produces eight separate stem–loop structures with similarity to the stem–loop structure of the type II IRES (data not shown).

Nucleotide sequences of the 3’ UTR of DHV-1 were generated from two clones containing the 3’ UTR and a poly(A) sequence. The 3’ UTRs of strains DRL-62 and R85952 are 318 and 317 nt, respectively, and the sequence identity of the two strains in this region was 99–100%. 3’ UTRs of picornaviruses are short, ranging in length from 47 nt for human rhinovirus 14 (GenBank accession no. K02121) to 240 nt for Aichi virus (GenBank accession no. NC_001918) and are organized in a higher-order RNA structure thought to be involved in RNA replication (Gromeier et al., 1999; Yamashita et al., 1998). The 3’ UTR sequence of DHV-1 is strikingly longer than the corresponding region of any other picornavirus described previously. The predicted secondary structure, obtained by using the program Mfold (Zuker, 2003), of this region indicated that it is composed of hairpin loop structures containing more than four hairpin loops (data not shown), whereas human parechovirus type 1 (Auvinen & Hyypia, 1990) and HRV-2 (Pöyry et al., 1996), Ljungan virus (Johansson et al., 2003), enteroviruses (Pöyry et al., 1996) and Encephalomyocarditis virus (Cui & Porter, 1995) have one to three hairpin loops in this region of the genome.

Processing of the DHV-1 polyprotein

The picornavirus polyprotein contains interdomain junctions that are, at most sites, processed by 3Cpro protease and the cleavage sites are confined to small regions that contain primary- and tertiary-structure elements recognized by

---

**Table 2. Features of proteins of DHV-1 strain DRL-62**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Nucleotide sequence (nt)</th>
<th>Amino acid sequence (aa)</th>
<th>Predicted N-terminal cleavage sites</th>
<th>Predicted activity or function of viral protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ UTR</td>
<td>1 626 626</td>
<td>– – –</td>
<td>– – –</td>
<td>– – –</td>
</tr>
<tr>
<td>Polyprotein</td>
<td>627 7373 6747</td>
<td>1 2249 2249</td>
<td>– – –</td>
<td>– – –</td>
</tr>
<tr>
<td>P1</td>
<td>627 2819 2193</td>
<td>1 731 731</td>
<td>E/S</td>
<td>–</td>
</tr>
<tr>
<td>P2</td>
<td>2820 5090 2271</td>
<td>732 1488 757</td>
<td>Q/S</td>
<td>–</td>
</tr>
<tr>
<td>P3</td>
<td>5091 7373 2283</td>
<td>1489 2249 761</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>VP0</td>
<td>627 1394 768</td>
<td>1 256 256</td>
<td>–</td>
<td>Capsid protein</td>
</tr>
<tr>
<td>VP3</td>
<td>1395 2105 711</td>
<td>257 493 237</td>
<td>Q/G</td>
<td>Capsid protein</td>
</tr>
<tr>
<td>VP1</td>
<td>2106 2819 714</td>
<td>494 731 238</td>
<td>Q/G</td>
<td>Capsid protein</td>
</tr>
<tr>
<td>2A1</td>
<td>2820 2879 60</td>
<td>732 751 20</td>
<td>E/S</td>
<td>Primary polyprotein processing</td>
</tr>
<tr>
<td>2A2</td>
<td>2880 3734 855</td>
<td>752 1036 285</td>
<td>NPG/P</td>
<td>Inhibition of cell growth</td>
</tr>
<tr>
<td>2B</td>
<td>3735 4091 357</td>
<td>1037 1155 119</td>
<td>Q/S</td>
<td>Membrane-associated</td>
</tr>
<tr>
<td>2C</td>
<td>4092 5090 999</td>
<td>1156 1488 333</td>
<td>Q/S</td>
<td>ATPase</td>
</tr>
<tr>
<td>3A</td>
<td>5091 5369 279</td>
<td>1489 1581 93</td>
<td>Q/S</td>
<td>Membrane-associated</td>
</tr>
<tr>
<td>3B</td>
<td>5370 5471 102</td>
<td>1582 1615 34</td>
<td>Q/S</td>
<td>Initiation of RNA synthesis*</td>
</tr>
<tr>
<td>3C</td>
<td>5472 6014 543</td>
<td>1616 1796 181</td>
<td>Q/S</td>
<td>Protease</td>
</tr>
<tr>
<td>3D</td>
<td>6015 7373 1359</td>
<td>1797 2249 453</td>
<td>Q/G</td>
<td>RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td>3’ UTR</td>
<td>7374 7691 318</td>
<td>(7690)† (317)†</td>
<td>– – –</td>
<td>– – –</td>
</tr>
</tbody>
</table>

*Protein 3BVPg.
†The genomic sequence of strain R85952 contains 7690 nt and the 3’ UTR contains 317 nt.
### Table 3: Comparison of DHV-1 strain DRL-62 with DHV strain R85952 and selected picornaviruses

<table>
<thead>
<tr>
<th>Region</th>
<th>5’ UTR</th>
<th>VP0</th>
<th>VP1</th>
<th>VP2</th>
<th>VP3</th>
<th>VP4</th>
<th>C’</th>
<th>O’</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHV</td>
<td>96-9</td>
<td>23-7</td>
<td>247</td>
<td>62</td>
<td>26</td>
<td>13</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>FMDV</td>
<td>99-6</td>
<td>14-3</td>
<td>14-6</td>
<td>28</td>
<td>18</td>
<td>13</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>EV</td>
<td>98-7</td>
<td>15-5</td>
<td>12-8</td>
<td>16-2</td>
<td>14-7</td>
<td>17</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>HEV-A</td>
<td>97-8</td>
<td>15-0</td>
<td>15-0</td>
<td>17-3</td>
<td>17-3</td>
<td>15-5</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>HRV-A</td>
<td>95-7</td>
<td>15-0</td>
<td>15-0</td>
<td>13-8</td>
<td>13-8</td>
<td>15-1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>R85952</td>
<td>100</td>
<td>38-0</td>
<td>38-0</td>
<td>38-0</td>
<td>38-0</td>
<td>38-0</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>TTV</td>
<td>98-2</td>
<td>14-7</td>
<td>14-7</td>
<td>16-2</td>
<td>16-2</td>
<td>14-7</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>2A</td>
<td>97-8</td>
<td>12-9</td>
<td>12-9</td>
<td>14-2</td>
<td>14-2</td>
<td>16-2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2B</td>
<td>98-5</td>
<td>13-4</td>
<td>13-4</td>
<td>13-7</td>
<td>13-7</td>
<td>16-2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2C</td>
<td>99-0</td>
<td>15-0</td>
<td>15-0</td>
<td>17-3</td>
<td>17-3</td>
<td>15-0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>3A</td>
<td>99-5</td>
<td>14-7</td>
<td>14-7</td>
<td>17-3</td>
<td>17-3</td>
<td>14-7</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>3B</td>
<td>99-2</td>
<td>15-0</td>
<td>15-0</td>
<td>17-3</td>
<td>17-3</td>
<td>15-0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>3C</td>
<td>99-5</td>
<td>15-0</td>
<td>15-0</td>
<td>17-3</td>
<td>17-3</td>
<td>15-0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>3D</td>
<td>99-5</td>
<td>15-0</td>
<td>15-0</td>
<td>17-3</td>
<td>17-3</td>
<td>15-0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>3UTR</td>
<td>99-3</td>
<td>13-4</td>
<td>13-4</td>
<td>13-7</td>
<td>13-7</td>
<td>16-2</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Values are percentage amino acid identities between DRL-62 and R85952 or the selected picornaviruses. Values were calculated with a sequence-identity matrix using CLUSTAL_W in BioEdit. ND, Not determined.

3Cpro (Dougherty & Semler, 1993; Kräusslich & Wimmer, 1988; Palmenberg, 1990; Ryan & Flint, 1997). An alignment of the polyproteins of the DHV-1 strains and their closest recognized relatives, the Ljungan and human parechoviruses, was generated and analysed for putative 3Cpro processing sites of the DHV-1 polyprotein. Predicted cleavage sites of the DHV-1 polyprotein based on processing sites of the Ljungan and human parechoviruses are shown in Table 2. The junction sites appear to be mainly at Q/G and Q/S residues, indicating that these sites may be processed by the virus protease 3C (3Cpro). An exception was observed at the predicted VP1/2A cleavage site (Table 2).

An analysis of the 2A region of the DHV-1 polyprotein revealed that it, surprisingly, contains two different 2A motifs. The complete 2A region of DHV-1 is organized similarly to that reported previously for Ljungan virus. The first motif is a homologue of a 17 aa peptide conserved at the 2A/2B junction in cardio-, erbo-, tescho- and aphthoviruses and corresponds to the 2A1/2A2 junction in the Ljungan virus polyprotein (Johansson et al., 2002). The C terminus of this peptide (2A1) homologue conforms to the amino acid sequence signature G-vExNPG/P, with upper-case letters demonstrating absolutely conserved residues (Fig. 1a) (Beard & Mason, 2000; Cohen et al., 1988; Doherty et al., 1999; Johansson et al., 2002; Li et al., 1996; Pevear et al., 1988; Robertson et al., 1985; Wutz et al., 1996; Zell et al., 2001). This peptide is sufficient to promote separation of the 2A and 2B moieties at the NPG/P junction in mengovirus and Foot-and-mouth disease virus (FMDV) (Donnelly et al., 1997, 2001; Hahn & Palmenberg, 1996; Palmenberg et al., 1992) by a mechanism that is currently the subject of intense debate. In aphthoviruses and probably in erbo- and teschoviruses, this conserved peptide, excluding the most C-terminal Pro residue, is released as a mature 18–21 aa long product through cleavage by 3C at the N terminus and autoprocessing at the C terminus (Donnelly et al., 2001). A similar, 20 aa long peptide could be produced in DHV-1 and Ljungan viruses (Fig. 1a). This model implies that, unlike with all other picornaviruses except Ljungan viruses, the DHV-1 polyprotein is processed at an extra site in a region between VP1 and 2B, which generates two different proteins of the 2A type (2A1 and 2A2), rather than one.

### Genes encoding non-structural proteins

DHV-1 polyproteins contain amino acid sequence motifs present in picornaviral non-structural proteins. Conserved amino acid sequence motifs characteristic of picornavirus 2C^ATPase, 3B^VPg, 3C^pro and 3D^pol (Gorbalenya & Koonin, 1993) are present in the ORF.

The DHV-1 2C protein includes the three conserved ATPase motifs (A, B and C) identified by Koonin & Dolja (1993). The sequence of the highly conserved consensus motif A, GxxGxGKS, which is proposed to be responsible for nucleotide binding (Gorbalenya et al., 1991), was located at position 145–152 within the amino acid sequence. The highly distinct and unique peptide found at this site may reflect a mechanism of cleavage different from that observed in other picornaviruses. This peptide is conserved in all other picornaviruses and is absent from all other picornaviruses except Ljungan virus, which suggests a role for this site in the assembly of the viral core.
The 3B (VPg) peptide includes the conserved tyrosine residue located at position 3; however, only one copy of the gene (the VPg-encoding region) is present. In comparison, three copies of the VPg gene are reported in FMDV (Forss et al., 1984) and in equine rhinovirus 2, there is an additional VPg pseudo-coding region in addition to that encoding the normal product (Wutz et al., 1996).

Examination of the sequence of the virus protease 3C (3C(pr)) of DHV-1 revealed that it contains the catalytic triad, histidine, aspartic acid and cysteine, that has been observed in other viruses of this family. These occur at amino acid positions 38, 69 and 144, respectively. The prediction that aspartic acid forms part of the catalytic triad would fit DHV-1 into the lineage ‘B’ group of viruses, all of which contain this residue at the equivalent position (Ryan & Flint, 1997). This lineage, established by sequence alignments, contains the cardioviruses, aphthovirus, hepatoviruses, parechoviruses and DHV-1. In contrast, lineage ‘A’ contains the enteroviruses and human rhinoviruses, all of which have a glutamate in place of the catalytic aspartic residue. The cysteine residue was found to be contained in the motif GxCG, which is considered to be part of the active site of this enzyme in cardioviruses and Aichi virus (Yamashita et al., 1998).

The 3D protein is an RNA-dependent RNA polymerase and picornavirus 3D sequences contain evolutionarily conserved motifs. Of the eight described conserved motifs (I–VIII) in the 3D region of positive-strand RNA viruses (Koonin, 1991; Koonin & Dolja, 1993), seven (I–VII) of the eight motifs were well conserved in the DHV-1 3D sequence. Among the other members of the family Picornaviridae included in the analysis, these seven motifs of DHV-1 were also well conserved. Acid motif V (SGx3Tx3N), core motif VI (YGDD) for nucleotide binding and motif VII (FLKR) for catalytic function were located at aa 281–290, 317–320 and 365–368 in the 3D sequence, respectively.

Likewise, the sequence conservation that suggests a particular function is also evident for the two different 2A moieties. The N-terminal moiety, 2A1, of DHV-1 is related to the NPGP family of 2A proteins (Fig. 1a) and the C-terminal moiety, 2A2, belongs to the H-NC family of 2A proteins encoded by human parecho- and Ljungan viruses (Fig. 1b), kobuviruses and AEV (Hollinger & Emerson, 2001; Marvil et al., 1999; Yamashita et al., 1998). In this region, a putative transmembrane domain in the C-terminal region, made up of an approximately 18 aa long hydrophobic region, is present in DHV-1 strains DRL-62 and R85952. *Each non-event count of 10.

**Fig. 1.** Multiple sequence alignments of the 2A1 and 2A2 proteins of DHV-1 strains. Black and grey backgrounds highlight alignment columns with 100% and >60% conserved residues, respectively, as defined in the GeneDoc default similarity groups (Nicholas et al., 1997). (a) Alignment of the DHV-1 2A1 and the C-terminal regions of the picornavirus NPGP 2A protein family (grey box). (b) Alignment of the DHV-1 2A2 and the picornavirus H-NC 2A protein family. The conserved H-box, NC-box and a putative transmembrane domain (grey box) made up of an approximately 18 aa long hydrophobic region are observed in the DHV-1 strains DRL-62 and R85952. *Each non-event count of 10.
involved in aspects of virus–host interaction related to cell growth control (Hughes & Stanway, 2000).

Interestingly, the 2A2 region of DHV-1 contains a 161 aa extension compared with the closest known relatives, Ljungan viruses, with the same gene organization. The function of the 161 aa extension of DHV-1 is not known, but this sequence is similar to a plant prototype gene that is found in Arabidopsis thaliana and designated AIG1, a protein that functions in plant resistance to bacteria as an avrRpt2-induced gene (Reuber & Ausubel, 1996). The protein AIG1 is the prototype member of a recently described GTPase family. The AIG1 domain in the Pfam database (Bateman & Haft, 2002) (accession no. PF04548) exhibits the five motifs G1–G5 (Sprang, 1997) characteristic of GTP/GDP-binding proteins. Motifs G1–G3 in the 161 aa extension of DHV-1 were well conserved compared with other AIG1 domains and the motif G4–G5 of DHV-1 was less conserved (data not shown).

**Capsid proteins**

In most picornaviruses, the P1 precursor polypeptide is bound covalently by its N-terminal glycine residue (when the methionine or L protein is removed) to a molecule of myristic acid via an amide linkage (Chow et al., 1987; Paul et al., 1987). The myristylation reaction is carried out by myristoyl transferase, which recognizes an 8 aa signal motif beginning with glycine. In the majority of picornaviruses, a 5 aa consensus sequence motif, GxxxT/S, has been identified in VP0 (Palmenberg, 1989). Although the putative myristylation sequence occurs at aa 13–17, GxxxxT/S, in human parechoviruses and at aa 29–33, GPDRT, in AEV, the motif is probably not used for myristylation of VP0. The DHV-1 VP0 contains only one putative myristylation signal, GAVES, at aa 31–35. The lack of a myristylation signal motif at the N-terminal end of VP0 indicates that, as proposed for human parechoviruses, Ljungan viruses (Johansson et al., 2002) and AEV, the virion morphogenesis of DHV-1 probably does not rely upon myristylation of VP0.

Unlike other picornaviruses, DHV-1, Ljungan viruses and human parechoviruses contain 20, 18 and 21–25 aa long extensions, respectively, that are enriched with basic residues in the N terminus of VP3 (Johansson et al., 2002). This region has been shown to be highly immunogenic in human parechoviruses (Joki-Korpela et al., 2000), but no data are, so far, available for DHV-1 or Ljungan viruses. As has been demonstrated previously for other picornaviruses, the VP1 gene showed the greatest sequence variation of orthologous genes and proteins between the two DHV-1 strains and towards other picornaviruses. In the C terminus of the VP1 region, human parechoviruses 1 and 2 contain an Arg–Gly–Asp (RGD) motif important for cell-surface interactions (Pulli et al., 1997; Stanway et al., 1994; Triantafilou et al., 2000), but the RGD motif was not present in human parechovirus 3 (Ito et al., 2004). The DHV-1 VP1 does not contain an RGD motif, but has an insertion of 43–48 aa compared with human parechoviruses, unique in sequence composition and similar in size to an insert present in the Ljungan virus VP1.

![Fig. 2. Phylogenetic tree showing the relationship between DHV-1 strains (DRL-62 and R85952) and other picornaviruses.](http://vir.sgmjournals.org)
Phylogenetic analysis

The polyprotein of a picornavirus is processed to produce 10–14 end products through a cascade of proteolytic events brought about by virus-encoded proteases (Palmenberg, 1990; Ryan & Flint, 1997). Among the end products, the 3D and 2C proteins are two of the most conserved proteins among the family Picornaviridae, probably due to functional constraints (Gromeier et al., 1999). Therefore, the amino acid sequences of the 3D and 2C proteins of members of the family Picornaviridae are frequently used for inferring phylogenetic relationships between viruses belonging to different genera within the family (Domingo et al., 1995; Gromeier et al., 1999; Pöry et al., 1996; Rodrigo & Dopazo, 1995). It has also been shown that the VP1 sequence, which shows considerable genetic variation, is suitable for phylogenetic analysis within a picornavirus genus and its species (Mulders et al., 2000; Oberste et al., 1999a, b; Santti et al., 2000).

To study how DHV-1 strains DRL-62 and R85952 are related to other members of the family Picornaviridae, the 3D amino acid sequences from the two strains were included in phylogenetic analyses with type species of the different picornavirus genera (Table 1). The phylogenetic relationship inferred by using the 3D protein is shown in Fig. 2. The two DHV-1 strains constitute a monophyletic group related to, but separated from, members of the genus Parechovirus (Ljungan virus and Human parechovirus). However, the low sequence identity to the genus Parechovirus (34–49–40 % in 3D and 29–4–32·8 % in 2C; Table 3) marks a clear distance between DVH-1 and the genus Parechovirus.

Taxonomic status

DHV-1 has a worldwide distribution and was first isolated from the liver of infected ducklings on Long Island, NY, USA, in 1949 (Levine & Fabricant, 1950). Based on physicochemical properties of the virions, the virus was preliminarily classified as an enterovirus (Wildy, 1971).

Another virus (AEV) was also previously referred to as a putative ‘enterovirus’ like DHV-1; however, analysis of the AEV genome demonstrated that AEV is a member of the family Picornaviridae, but related more closely to Hepatitis A virus, and it was therefore classified into the genus Hepatovirus (Marvil et al., 1999). The previous classification of picornaviruses by morphology and physicochemical properties is limited and current taxonomic classification is based primarily on sequence identity, organization and type of genome (Stanway et al., 2005).

Our data demonstrate clearly that DHV-1 is a member of the family Picornaviridae and it is of no doubt that this virus is related more closely to members of the genus Parechovirus than to the genus Enterovirus. The low sequence identity between the DHV-1 strains and the closest relatives of the genome Parechovirus, Ljungan viruses and human parechoviruses, suggests, however, that DHV-1 could belong to a novel genus within the family Picornaviridae.

ACKNOWLEDGEMENTS

The authors would like to thank Hyuk-Man Kwon for propagation of DHV samples and Dr Bang-Hun Hyun for support of the genetic analysis. This work was supported financially by the National Veterinary Research and Quarantine Service (NVRQS), Ministry of Agricultural and Forestry (MAF), Republic of Korea. A. M. L. was supported by the Knowledge Foundation, Sweden, and the Research Programme in Medical Bioinformatics, Karolinska Institute, Stockholm, Sweden.

REFERENCES


Gorbalenya, A. E. & Koonin, E. V. (1993). Comparative analysis of the amino acid sequences of the key enzymes of the replication and


