Complete nucleotide sequence of avian paramyxovirus type 6 isolated from ducks

P.-C. Chang,1 M.-L. Hsieh,1 J.-H. Shien,2 D. A. Graham,3 M.-S. Lee2 and H. K. Shieh1,2

1,2 Graduate Institute of Veterinary Microbiology1 and Department of Veterinary Medicine2, National Chung Hsing University, Taichung, Taiwan, Republic of China
3 Veterinary Science Division, Department of Agriculture and Rural Development for Northern Ireland, Stoney Road, Stormont, Belfast BT4 3SD, UK

There are nine serotypes of avian paramyxovirus (APMV). Only the genome of APMV type 1 (APMV-1), also called Newcastle disease virus (NDV), has been completely sequenced. In this study, the complete nucleotide sequence of an APMV-6 serotype isolated from ducks is reported. The 16236 nt genome encodes eight proteins, nucleocapsid protein (NP), phosphoprotein (P), V protein, matrix protein (M), fusion protein (F), small hydrophobic (SH) protein, haemagglutinin–neuraminidase (HN) protein and large (L) protein, which are flanked by a 55 nt leader sequence and a 54 nt trailer sequence. Sequence comparison reveals that the protein sequences of APMV-6 are most closely related to those of APMV-1 (NDV) and -2, with sequence identities ranging from 22 to 44%. However, APMV-6 contains a gene that might encode the SH protein, which is absent in APMV-1, but present in the rubulaviruses simian virus type 5 and mumps virus. The presence of an SH gene in APMV-6 might provide a link between the evolution of APMV and rubulaviruses. Phylogenetic analysis demonstrates that APMV-6, -1, -2 (only the F and HN sequences were available for analysis) and -4 (only the HN sequences were available for analysis) all cluster into a single lineage that is distinct from other paramyxoviruses. This result suggests that APMV should constitute a new genus within the subfamily Paramyxovirinae.

Introduction

Members of the family Paramyxoviridae are enveloped, negative-stranded RNA viruses that infect a great variety of mammalian and avian species (Lamb & Kolakofsky, 1996). According to current taxonomy and nomenclature, the family Paramyxoviridae consists of two subfamilies. The subfamily Paramyxovirinae contains three genera, Rubulavirus, Respirovirus (formerly known as the genus Paramyxovirus) and Morbillivirus. The subfamily Pneumovirus contains two genera, Pneumovirus and Metapneumovirus (Rima et al., 1995; Mayo & Pringle, 1998; Pringle, 1998). There are nine serotypes of avian paramyxovirus (APMV). APMV type 1 (APMV-1), also called Newcastle disease virus (NDV), is assigned into the genus Rubulavirus.

NDV remains one of the most important pathogens of poultry (Alexander, 1997). Outbreaks of NDV have been causing severe economic losses in many countries (Alexander, 1995; Lomniczi et al., 1998; Yang et al., 1999) and at least three panzootics of NDV have been recognized (Alexander, 1988). In addition to NDV, APMV-2 and -3 might also lead to severe disease. For example, APMV-2 infections have been associated with severe respiratory disease and a drop in egg production in turkeys (Bankowski et al., 1995; Lipkind et al., 1995) and APMV-3 infections might also be responsible for problems in egg production of turkeys (Alexander et al., 1983; Alexander, 2000). In comparison, infections of APMV-4 to -9 appear to be apathogenic, except that APMV-6 might cause a mild respiratory disease and problems in egg production in turkeys (Alexander, 1997). Cross-reaction tests of haemagglutination inhibition (HI) and neuraminidase inhibition (NI) show that APMV could be divided into two subgroups: the first subgroup contains APMV-2 and -6 and the second subgroup contains APMV-1, -3, -4, -7, -8 and -9 (Lipkind & Shihmanter, 1986).

The nucleotide sequence of the complete genome of NDV has been determined (Krishnamurthy & Samal, 1998; de Leeuw & Peeters, 1999). The genome of NDV is 15186 nt in length and encodes at least seven proteins, nucleocapsid protein (NP),...
phosphoprotein (P), V protein, matrix protein (M), haemagglutinin–neuraminidase (HN) protein, fusion protein (F) and large protein (L). NDV is currently classified into the genus *Rubulavirus*, but unlike the rubulaviruses simian virus type 5 (SV-5) and mumps virus (MuV), NDV does not contain the small hydrophobic (SH) protein. Phylogenetic analyses based on nucleotide or protein sequences, together with other biological properties of NDV, suggest that NDV (and probably other APMV) should be assigned to a new genus or subfamily (de Leeuw & Peeters, 1999; Seal et al., 2000). Although this assignment is evident for NDV, it remains questionable for other APMV serotypes (APMV-2 to -9), as little information is available concerning the sequences of APMV-2 to -9. To date, only the sequences of the F and HN genes of APMV-2 and the HN gene of APMV-4 are deposited in GenBank (accession numbers D13977, D14030 and D14031, respectively; all unpublished). Here, we report the complete nucleotide sequence of an APMV-6 serotype and compare its sequence with those of other paramyxoviruses in order to understand the genomic structure and taxonomic position of APMV.

**Methods**

- **Viruses and serological studies.** Virus was isolated from domestic ducks by inoculating tissue homogenates into 9- to 11-day-old embryonated Muscovy duck eggs (Alexander, 1989). Pathogenicity tests, including the mean time of death, intracerebral pathogenicity index (ICPI) and intravenous pathogenicity index (IVPI), were conducted in embryonated specific-pathogen-free (SPF) hens’ eggs, as described previously (Alexander, 1989). HI tests were carried out on this virus isolate, together with reference antigens to APMV-1 (chicken/Ulster2C), APMV-2 (chicken/California/Yucapa/56), APMV-3 (turkey/England/1087/82), APMV-4 (duck/Hong Kong/D3/75), APMV-6 (duck/Hong Kong/199/77), APMV-7 (dove/Tennessee/4/75), APMV-8 (goose/Delaware/1053/76), APMV-9 (duck/New York/22/78) and mono-specific polyclonal chicken antiserum, as described previously (Graham et al., 1999). Antiserum were homologous in all cases, except for APMV-1 and -3, when antiseras were raised against APMV-1/chicken/Herst/33/56 and APMV-3/parakeet/Netherlands/449/75 were used.

- **Virus purification and RNA isolation.** Virus was purified by discontinuous sucrose gradient centrifugation. In brief, 400 ml allantoic fluid was clarified by centrifugation at 5000 rpm for 30 min at 4 °C in a JA-10 rotor (Beckman). Supernatant was collected and overlaid onto a 20% sucrose solution and centrifuged at 25000 rpm for 4 h at 4 °C in an SW-28 rotor (Beckman). The pellet was collected and resuspended in a total of 10 ml TNE buffer (100 mM Tris, pH 7-2, 100 mM NaCl, 1 mM EDTA). The resuspended virus was overlaid onto a 20–50% discontinuous sucrose gradient and centrifuged at 35000 rpm for 2 h at 4 °C in an SW-41 rotor (Beckman). Virus bands were collected and diluted by adding 4 vols of TNE buffer. The virus was then pelleted by centrifugation at 35000 rpm for 2 h at 4 °C in an SW-41 rotor. The virus pellet was resuspended in 1 ml of TNE buffer and stored at −20 °C. Viral RNA was extracted using Trizol reagent (Life Technologies), according to the manufacturer’s instructions.

- **Genome ‘walking’ by RT–PCR.** The NP gene was used as the first start-point for walking and sequencing of the APMV-6 genome. Based on the consensus sequences of the NP genes of rubulaviruses, two primers, NP (+), 5’ AAYRCSGKMTKRCWSSCWTTCTT, and NP (−), 5’ GWSWYYCWCCATKGAW (Y = C/T, R = A/G, S = G/C, K = G/T, M = A/C and W = A/T), were designed. These primers amplify a 236 bp fragment from APMV-6. The sequence of this 236 bp fragment was then used as the first start-point for walking the APMV-6 genome. The second start-point for genome walking was a primer design based on the consensus sequences of the HN genes of different strains of NDV. This primer, 5’ CTGCCCTCGCGCCCCCATGAG, anneals specifically to the corresponding region of the HN gene of APMV-6 and serves as the second start-point for genome walking. Two strategies were used for genome walking. The first strategy, ‘targeted gene walking PCR’ (Parker et al., 1991), is based on the use of two primers. The first primer is specific and anneals to a known sequence, whereas the second is a nonspecific walking primer, which anneals to an unknown site. We found the following nonspecific primers useful for walking the genome of APMV-6: NSP1 (5’ AKCRAAGACCC), NSP2 (5’ AGTAGAAAC-AAGG), NSP3 (5’ ACAAGGTTGAGG), NSP4 (5’ GTTTTTCTTAA) and NSP5 (5’ ATACGGGTAGAA). The second strategy used for genome walking is ‘SMART PCR cDNA synthesis’ (Clontech). In this strategy, first-strand cDNA is synthesized by SuperScript II reverse transcriptase (Life Technologies) using a primer specific to APMV-6. SuperScript II adds a poly(dC) tail to the 3′ end of the newly synthesized cDNA. RT–PCR is then carried out using the SMART II primer (5’ AACGAGTGTAAACACCGGAGTCCCGGG) coupled with the primer used for first-strand cDNA synthesis.

RT–PCR was carried out in a 25 μl reaction mixture containing 2.5 μl of 10 X reaction buffer, 2.5 μl dNTPs (2 mM each of four dNTPs), 0.2 μl AMV reverse transcriptase (50 U/μl), 0.3 μl RNase inhibitor (40 U/μl), 0.5 μl Taq DNA polymerase (9 U/μl), 1 μl of each primer (10 pmol each of four dNTPs), 1 μl of RNA template and 17 μl of water. All reagents were purchased from Promega. The general conditions for RT–PCR were 42 °C for 50 min (reverse transcription), 95 °C for 3 min, 35 cycles of 95 °C for 40 s (denaturation), 45 °C for 40 s (annealing) and 72 °C for 1 min (extension), followed by 72 °C for 7 min (final extension).

- **3′- and 5′-RACE (rapid amplification of cDNA ends).** The sequences of the 3′- and 5′-termini of the viral genome were amplified by 3′- and 5′-RACE. 3′-RACE was carried out as described previously (Schutze et al., 1995; de Leeuw & Peeters, 1999). In brief, 100 pmol of the primer ALG3 (5′ CACCAATCTATGACGCTGTTCTT), which is specific for the HN gene of APMV-6 and serves as the second start-point for genome walking, was ligated to the 3′ end of the viral RNA (5 μg) by T4 RNA ligase (50 U) (New England Biolabs) in a 25 μl reaction mixture at 25 °C for 4 h. The ligated product was purified using the High Pure PCR Product Purification kit (Roche) and then used as a template for RT–PCR with the primers ALG4 (5′ GAAGGATCCAGAATCGATAG), which is complementary to ALG3, and SP3 (5′ AGTGGAAACCGATGTCTG), which is specific for the NP gene of APMV-6. 5′-RACE was carried out using the 5′-RACE kit (Roche). In brief, first-strand cDNA was synthesized by AMV reverse transcriptase using the primer SP1 (5′ CCGAGCCTG-ACGATTGATG), which anneals to sequences located 61 nt downstream of the termination codon of the L gene of APMV-6. First-strand cDNA was then purified using the High Pure PCR Product Purification kit and a poly(A) tail was added to the 3′ end of the cDNA using terminal transferase and dATP. The poly(A)-tailed cDNA was then used as the template for PCR with the primers oligo(dT)-anchor (5′ ACCACCGG-TATCGATCTGACGCTGTTTTTTTTTTTTTTTTTTTTT) (V = A/C/G) and SP2 (5′ GGTATCGGTATCGGAGATTA), which anneals to sequences located 144 nt downstream of the termination codon of the L gene of APMV-6.

- **Cloning and sequencing of RT–PCR products.** PCR products were purified using the Geneclean III kit (BIO101) and cloned into the...
Viruses used as antigen were APMV-1 HI tests of APMV-6/duck/Taiwan/Y1/98

Table 1. HI tests of APMV-6/duck/Taiwan/Y1/98

<table>
<thead>
<tr>
<th>Antiserum</th>
<th>APMV-1</th>
<th>APMV-2</th>
<th>APMV-3</th>
<th>APMV-4</th>
<th>APMV-6</th>
<th>APMV-7</th>
<th>APMV-8</th>
<th>APMV-9</th>
<th>Y1/98</th>
</tr>
</thead>
<tbody>
<tr>
<td>APMV-1</td>
<td>512</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td>APMV-2</td>
<td>–</td>
<td>512</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>32</td>
</tr>
<tr>
<td>APMV-3</td>
<td>–</td>
<td>–</td>
<td>256</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td>APMV-4</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>256</td>
<td>–</td>
<td>1024</td>
<td>–</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td>APMV-6</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1024</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1024</td>
</tr>
<tr>
<td>APMV-7</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4096</td>
<td>–</td>
<td>–</td>
<td>128</td>
</tr>
<tr>
<td>APMV-8</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>2048</td>
<td>–</td>
<td>32</td>
</tr>
<tr>
<td>APMV-9</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>512</td>
<td>8</td>
</tr>
<tr>
<td>Y1/98</td>
<td>32</td>
<td>8</td>
<td>8</td>
<td>32</td>
<td>512</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>512</td>
</tr>
</tbody>
</table>

Results

Virus characterization

As a result of an influenza virus surveillance program in Taiwan in 1998, a virus causing haemagglutination was isolated from domestic ducks. This virus was shown by electron microscopy to be a paramyxovirus (data not shown). HI tests using reference antigens and antisera to APMV-1 to -9 demonstrate that the virus belongs to the APMV-6 serotype (Table 1), designated APMV-6/duck/Taiwan/Y1/98. For simplicity, the virus is abbreviated in this work to APMV-6. The duck from which APMV-6 was isolated appeared to be healthy. Moreover, both the ICPI and the IVPI of APMV-6 in SPF chickens were found to be zero. These results suggest that APMV-6 is apathogenic. Similar observations have been reported for other APMV-6 serotypes isolated from ducks (Shortridge et al., 1980; Marius-Jestin et al., 1987).

Cloning and sequencing of the APMV-6 genome

A total of 22 overlapping cDNA clones, covering the entire genome of APMV-6, was obtained by either target gene walking RT–PCR or 5’- and 3’-RACE (see Methods). Sequences compiled from these clones show that APMV-6 is 16236 nt long (GenBank accession number AY029299). This genome is larger than those of most other members of the subfamily Paramyxovirinae, for which the genome size is approximately 15 500 nt. The number 16 236 is a multiple of 6; therefore, APMV-6 conforms to the ‘rule of six’, which plays an important role in the replication of paramyxoviruses (reviewed by Kolakofsky et al., 1998).

Leader and trailer sequences of APMV-6

The 3’ and 5’ ends of the APMV-6 genome comprise the leader and trailer regions; the leader sequence is 55 nt long and the trailer sequence is 54 nt long (Fig. 1A, B). The length of the leader sequence, 55 nt (Fig. 1A), is highly conserved among members of the subfamily Paramyxovirinae, confirming that APMV-6 belongs to this subfamily. In comparison, the length of the trailer is variable (Fig. 1B). APMV-6 contains a 54 nt trailer, which is within the typical range (40–60 nt) of most members of the subfamily Paramyxovirinae (Fig. 1B). The 3’ terminus of the leader (UGGU) and the 5’ terminus of the trailer (ACCA) sequences are identical in all Paramyxovirinae (Fig. 1A, B; shaded sequences). For APMV-6, the termini of
Fig. 1. Alignment of the (A) leader and (B) trailer sequences of APMV-6 with those of other members of the subfamily Paramyxovirinae. Sequences are presented in the 3′→5′ direction. Sequences conserved in the 3′ and 5′ ends of all members of the subfamily Paramyxovirinae are shaded. Sequences conserved in the 3′ and 5′ ends of APMV-6, NDV and members of the Morbillivirus and Respirovirus genera are underlined. The mRNA start position (gene start) is boxed. (C) Complementary nucleotides between the 3′ and 5′ ends of APMV-6 are indicated by asterisks.

the leader (UGGUUUGU) and trailer (ACCAAACAA) sequences are similar to those of NDV and members of the genera Morbillivirus [canine distemper virus (CDV), measles virus (MeV) and rinderpest virus (RPV)] and Respirovirus [bovine and human parainfluenza virus (PIV) type 3 and Sendai virus (SeV)] (Fig. 1A, B; underlined sequences), but distinct from members of the genus Rubulavirus (hPIV-2, MuV and SV-5, SV-41). Complementarity is found between the 3′ and 5′ ends of APMV-6; 22 of the first 24 terminal nucleotides are exactly complementary to each other (Fig. 1C), indicating that the promoters for genomic and anti-genomic replication are located in the first 24 terminal nucleotides.
Table 2. Gene start, gene end and intergenic regions of APMV-6

Sequences are presented in the 3′ → 5′ direction.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Gene start (phase*)</th>
<th>Gene end</th>
<th>Intercistronic sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>CUC₃UUC (2)</td>
<td>AAUUCU₆</td>
<td>GUUCCCA</td>
</tr>
<tr>
<td>P</td>
<td>CUC₃UUC (2)</td>
<td>AAUUA₄</td>
<td>CA</td>
</tr>
<tr>
<td>M</td>
<td>CUC₃UUG (2)</td>
<td>AAUUGU₄</td>
<td>AGUCAAAGCACAUUCAGCAUAUGACGUCCCCAGGUUAUGGAGGGUGGUUAUCUGCGA</td>
</tr>
<tr>
<td>F</td>
<td>CUC₃UUC (2)</td>
<td>AAUAGU₄</td>
<td>AAGGCUCUUAUUAGCGUCUCGGUUAGAUUGAUUGGUCCUUUUUUGCGGA</td>
</tr>
<tr>
<td>SH</td>
<td>CUC₃UUC (2)</td>
<td>AAUAUA₄</td>
<td>GUAUUACCGA</td>
</tr>
<tr>
<td>HN</td>
<td>CUC₃UUC (4)</td>
<td>AAUAU₆</td>
<td>GCCCUCGCACAAUUUGUAUCGCUCUAAAGCUGCCGCUGAGCGGUACGCUACGUCA</td>
</tr>
<tr>
<td>L</td>
<td>CUC₃UUC (4)</td>
<td>AAUAU₆</td>
<td>2–63 nt</td>
</tr>
<tr>
<td>Consensus</td>
<td>CUC₃₄UUC</td>
<td>AAAUUN₁₋₄U₅₋₇</td>
<td></td>
</tr>
</tbody>
</table>

* Subunit hexamer phasing position (6n+x) at the gene start site.

Table 3. Characteristics of proteins encoded by APMV-6

<table>
<thead>
<tr>
<th>Gene</th>
<th>Start codon*</th>
<th>Stop codon†</th>
<th>No. of residues</th>
<th>Calculated molecular mass (kDa)</th>
<th>Best match‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>128</td>
<td>1525</td>
<td>465</td>
<td>51.2</td>
<td>NP of NDV (44)</td>
</tr>
<tr>
<td>P</td>
<td>1687</td>
<td>2979</td>
<td>430</td>
<td>46.3</td>
<td>P of NDV (22)</td>
</tr>
<tr>
<td>M</td>
<td>3435</td>
<td>4735</td>
<td>366</td>
<td>40.4</td>
<td>M of NDV (27)</td>
</tr>
<tr>
<td>F</td>
<td>4598</td>
<td>6265</td>
<td>555</td>
<td>59.8</td>
<td>F of APMV-2 (44)</td>
</tr>
<tr>
<td>SH</td>
<td>6542</td>
<td>8232</td>
<td>142</td>
<td>15.6</td>
<td></td>
</tr>
<tr>
<td>HN</td>
<td>7122</td>
<td>8570</td>
<td>613</td>
<td>67.9</td>
<td>HN of APMV-2 (41)</td>
</tr>
<tr>
<td>L</td>
<td>9278</td>
<td>16003</td>
<td>2241</td>
<td>251.9</td>
<td>L of NDV (38)</td>
</tr>
</tbody>
</table>

* Location of the first base in the first ATG.
† Location of the third base in the stop codon.
‡ Protein sequence identities obtained by BLAST homology searching (%).

Gene start, gene end and intergenic sequences of APMV-6

The sequences of the gene start, gene end and intergenic regions of APMV-6 are summarized in Table 2. The conserved sequence for the gene start of APMV-6 is CUC₃₄UU, which shows little homology to the gene starts of other paramyxoviruses (Kolakofsky et al., 1998). The subunit hexamer phasing positions of the gene starts of APMV-6 are 2, 2, 2, 2, 2, 4 and 4 (Table 2), which are also unique among members of the family Paramyxoviridae. In comparison, the gene end of APMV-6, AAUUN₁₋₄U₅₋₇, exhibits some degree of homology to that of NDV (AAUCU₆₋₇) (Krishnamurthy & Samal, 1998). The length of the intergenic sequences of APMV-6 ranges from 2 to 63 nt (Table 2), which is comparable to those of NDV, which range from 1 to 47 nt. All intergenic sequences of APMV-6 end with an A (Table 2), which is similar to those observed in NDV. No other conserved sequence is found in the intergenic sequences of APMV-6. The lack of conservation in both the length and the sequence of the intergenic regions is common in APMV-6, NDV, MuV and other rubulaviruses, but differs entirely from members of the Respirovirus and Morbillivirus genera, which have conserved trinucleotide intergenic sequences (Kolakofsky et al., 1998).

Protein-coding regions

Seven ORFs are found within the genome of APMV-6. Six of the seven ORFs find their best matches in the NP, P, M, F, HN and L proteins of NDV or APMV-2 by BLAST searching (Table 3), suggesting that APMV-6 is most closely related to NDV and APMV-2. The sequence identities between APMV-6 and NDV range from 44 (NP) to 22% (P), whereas the identities between APMV-6 and -2 are 44 (F) and 41% (HN).
(Table 3). Although the sequences of only the F and HN genes for APMV-2 are available for comparison, APMV-6 still find their best homologue in APMV-2, but not in NDV (Table 3). This result suggests that APMV-6 is most closely related to APMV-2, followed by NDV. The fifth ORF, located between the HN and F genes of APMV-6, encodes a protein of 142 residues. BLAST searching revealed no significant sequence homology between this protein and any protein in GenBank. However, the location of this protein (between the HN and F genes), is reminiscent of the SH protein found in the rubulaviruses SV-5 and MuV.

F protein cleavage site

Alignment of the F protein cleavage sites of APMV-6, -2 and NDV is shown in Fig. 2(A). For NDV, virulent strains have two pairs of basic residues (RKRQKR) at the cleavage site, whereas avirulent strains have two single basic residues (GEGRL) at the cleavage site (Collins et al., 1993). In addition, virulent strains of NDV have an F residue at the N terminus of the F1 protein, whereas avirulent strains have an L residue at the corresponding position (Collins et al., 1993). As shown in Fig. 2(A), the F protein cleavage site of
APMV-6 contains only a single basic residue (R) at the cleavage site; moreover, the N terminus of the F1 protein of APMV-6 is an L residue. This finding indicates that APMV-6 is an avirulent strain, which is consistent with the result of pathogenicity studies. Fig. 2(A) also shows the F protein cleavage site of APMV-2, for which two single basic residues are found, at the cleavage site; the N terminus of the F1 protein is an F residue. This result implies that APMV-2 might bear some characteristics of virulent viruses.

**P/V protein**

The P gene of APMV-6 encodes a protein of 430 amino acids (46–3 kDa). A putative RNA editing (insertion) site is found 470 nt downstream of the initiation codon (ATG) of the P protein. Sequences of the editing site of APMV-6, UUUUCCC, conform well to the conserved sequence, U(U(C)-UCCC, found in all members of the subfamily Paramyxovirinae (Fig. 2B; shaded sequences) (Steward et al., 1993). For APMV-6, sequences of the editing site (UUUUUCCC) are similar to those of NDV and members of the genera Morbillivirus and Respirovirus (Fig. 2B; underlined sequences), but distinct from members of the genus Rubulavirus. The addition of one residue (G) into the editing site of APMV-6 leads to a frame-shift translation from the P protein to the V protein (P to V editing). The sequence of the putative V protein of APMV-6 exhibits a high degree of sequence homology at the C-terminal region to those of other Paramyxovirinae (Fig. 2C). A total of 18 residues, including seven C residues, is conserved among all members of the subfamily Paramyxovirinae. The C residues are suggested to constitute the zinc-binding domain of the V protein.

**SH protein**

Hydrophilicity analysis shows that the putative SH protein of APMV-6 contains a hydrophobic region near the terminus of the protein (Fig. 3). This hydrophobic region is about 30 residues in size, which is similar to the general size of the hydrophobic regions of other paramyxoviruses (Fig. 3). However, the SH protein of APMV-6 exhibits two unique features. First, it contains a second large hydrophobic region of about 20 residues downstream of the first hydrophobic region (Fig. 3). The presence of two hydrophobic regions, both of 20–30 residues, is unique among paramyxoviruses. Second, the full size of the SH protein of APMV-6 is 142 residues, which is larger than the general size (44–81 residues) of the paramyxovirus SH protein, but is smaller than the unusual large size (174 residues) of the SH protein of avian pneumo-
virus (APV), also known as turkey rhinotracheitis virus (TRTV) (Ling et al., 1992). Two potential N-linked glycosylation sites, located at residues 90 and 105, are found in the SH protein of APMV-6. Both sites are on the C-terminal side of the hydrophobic region, which is consistent with the notion that the C-terminal side of the SH protein is located on the outside.
Fig. 4. Phylogenetic analysis of the NP, P, M, F, HN and L proteins of members of the family Paramyxoviridae. Trees were constructed by the neighbour-joining method using the CLUSTAL program in LASERGENE (DNASTAR). Bootstrap values calculated using the PHYLIP software with 500 bootstrap replicates are presented for selected nodes. The absence of bootstrap values at some nodes of the P and M trees is due to the inconsistency of the topology of trees generated by the CLUSTAL and PHYLIP methods. The absence of bootstrap values at branches containing APV is because APV was used as the outgroup for the generation of the PHYLIP trees; thus, the bootstrap value cannot be determined. Sequence distances, which were generated using the PAM250 matrix in LASERGENE, are indicated by the scale. Viruses analysed include hPIV-1, hPIV-4a, hPIV-4b, NDV strains Beaudette C and Lasota, phocine distemper virus (PDV), Hendra virus (Hendra) and Nipah virus (Nipah).
of the cell membrane (Hiebert et al., 1988; Collins & Mottet, 1993).

Phylogenetic analysis

Phylogenetic trees of the NP, P, M, F, HN and L proteins of members of the family Paramyxoviridae are shown in Fig. 4. In all trees, the genera Rubulavirus, Morbillivirus, Respirovirus and Pneumovirus separate into distinct clusters. Moreover, APMV serotypes are more closely related to the genus Rubulavirus than to viruses of any other genera. Furthermore, for all genes analysed, APMV-6 forms a single cluster with NDV, APMV-2 and -4 (Fig. 4). In the trees of the NP, P, M, F and L proteins, the bootstrap values (74–100) for the cluster of APMV serotypes are statistically significant (> 70). In the phylogenetic tree of the HN protein, although the bootstrap value of 59 is relatively low, NDV, APMV-2, -4 and -6 still form a single cluster. Therefore, the overall result of phylogenetic analysis indicates that APMV forms a unique cluster within the subfamily Paramyxovirinae. This finding is consistent with the very recent taxonomy approved by the ICTV executive committee, which assigned APMV into the new genus Avulavirus – avian rubulavirus. Note also that, in the trees of the HN protein where the sequences of APMV-2 and -4 are available for comparison, APMV-6 first forms a cluster with APMV-2, which then joins with NDV and APMV-4 (Fig. 4). This result indicates that APMV-6 is more closely related to APMV-2 than to NDV and APMV-4.

Discussion

The genome of APMV-6 exhibits some unique features among members of the family Paramyxoviridae. First, the size of APMV-6 is 16 236 nt, which is larger than the general genome size (15 500 nt) of most other members of the family Paramyxoviridae, but is still smaller than Hendra virus, a newly emerged virus that has a genome size of 18 234 nt (Wang et al., 2000). Second, the gene starts of APMV-6 begin uniformly with a C residue, whereas those of most other members of the family Paramyxoviridae begin with a U residue (Kolakofsky et al., 1998). Moreover, the gene start of APMV-6 contains a stretch of 5–6 C residues, which is also unique among members of the family Paramyxoviridae. Third, the subunit hexamer phasing positions of the gene starts of APMV-6 are 2, 2, 2, 2, 4 and 4, which are different from both NDV (2, 4, 4, 4, 3 and 6) and other members of the family Paramyxoviridae (de Leeuw & Peeters, 1999). Finally, the putative SH protein of APMV-6 contains two hydrophobic regions, each between 20 and 30 residues in size, which contrasts to the single hydrophobic region found in the SH proteins of other paramyxoviruses (Steward et al., 1993). The SH protein is a type II integral membrane protein, of which the hydrophobic region constitutes the transmembrane domain (Hiebert et al., 1988; Collins & Mottet, 1993); theoretically, a single hydrophobic region is enough for integration of the protein into the plasma membrane. Why APMV-6 contains two hydrophobic regions and whether the second hydrophobic region plays a role in the integration of the SH protein remains unknown.

NDV is classified into the genus Rubulavirus. However, phylogenetic analysis and other biological properties all suggest that NDV should be assigned to a new genus (de Leeuw & Peeters, 1999; Seal et al., 2000). This assignment is evident for NDV, but remains undecided for other APMV, primarily due to the lack of APMV sequence data. In this report, we show the complete nucleotide sequence of APMV-6 and demonstrate that NDV, APMV-2, -4 and -6 form a unique phylogenetic cluster within the subfamily Paramyxovirinae. This result might provide an important basis for the reclassification of all APMV serotypes into a new genus, separate from the genus Rubulavirus. As well as the result of phylogenetic analysis, there is additional evidence supporting this separation. First, the sequences of the 5′- and 3′-termini and the editing site of the APMV-6 P and V proteins are all similar to those of NDV and other members of the genera Morbillivirus and Respirovirus, but distinct from those of the genus Rubulavirus. Second, NDV and APMV-6 edit their P gene mRNA from P to V; this strategy is similar to that employed by respiroviruses and morbilliviruses, but differs from that of rubulaviruses, which edit their mRNA from V to P (Steward et al., 1993). Third, the hosts of NDV and other APMV are birds, whereas those of rubulaviruses and all other paramyxoviruses are mammals.

Although APMV should form a new genus, the new genus is still closer to the genus Rubulavirus than to other genera within the subfamily Paramyxovirinae. This conclusion is based on the following observations. First, in all of the phylogenetic trees examined in this work, APMV form a cluster that is closer to the genus Rubulavirus than to the Morbillivirus and Respirovirus genera. Second, the lack of conservation in length and sequence of intergenic regions of the viral genome is common in NDV, APMV-6 and members of the genus Rubulavirus, but entirely different from members of the Respirovirus and Morbillivirus genera, which have conserved trinucleotide intergenic sequences (Kolakofsky et al., 1998). Finally, APMV-6 contains the SH gene, which is present in SV-5 and MuV (genus Rubulavirus), but absent in all members of the Respirovirus and Morbillivirus genera. Therefore, although APMV should constitute a new genus, the genus is still closely related to the genus Rubulavirus. This notion justifies, to some extent, the previous taxonomic classification of NDV into the genus Rubulavirus, and also justifies the current name of the new genus, Avulavirus. The presence of the SH gene in APMV-6 suggests that APMV-6 might play an intermediate role between the evolution of APMV and members of the genus Rubulavirus.

We show by phylogenetic analysis that APMV-6 is more similar to APMV-2 than to NDV and APMV-4. A previous study using HI and NI tests has shown that APMV serotypes
are divided into two subgroups: the first group contains APMV-0 and -2 and the second group contains the remaining APMV serotypes (Lipkind et al., 1995). Therefore, the result of grouping by phylogenetic analysis is consistent with that determined by traditional serological study.

Target gene walking PCR has been used successfully for the amplification of unknown DNA sequences adjacent to known DNA sequences (Parker et al., 1991). Two primers are used for this type of PCR: one is a sequence-specific primer, which hybridizes to known sequences, and the other is a nonspecific walking primer, which hybridizes to unknown sequences. PCR using the two primers allows the amplification of an unknown sequence adjacent to the site of a known sequence. It was found that Taq polymerase can initiate PCR from the nonspecific walking primer, as long as there is a partial homology at the 3' end of the primer (Parker et al., 1991). In this report, we show that the same strategy can be used for walking on RNA as well as DNA. Moreover, walking can be directed either upstream or downstream of the known RNA sequence. It was found that Taq polymerase, can initiate DNA synthesis from a primer that bears a sequence. It was found that

References


P.-C. Chang and others


Received 18 April 2001; Accepted 29 May 2001