Fusion protein gene nucleotide sequence similarities, shared antigenic sites and phylogenetic analysis suggest that phocid distemper virus type 2 and canine distemper virus belong to the same virus entity

Ilona K. G. Visser,1,2 Roger W. J. van der Heijden,3 Marco W. G. van de Bildt,2 Marcel J. H. Kenter,2 Claes Örvell4 and Albert D. M. E. Osterhaus2,3,*

1 Seal Rehabilitation and Research Centre, Hoofdstraat 94a, 9968 AG Pieterburen, 2 Laboratory of Immunobiology, National Institute of Public Health and Environmental Protection, P.O. Box 1, 3720 BA Bilthoven and 3 Division of Virology, Department of Infectious Diseases and Immunology, Veterinary Faculty, State University of Utrecht, Yalelaan 1, 3508 TD Utrecht, The Netherlands and 4 Central Microbiological Laboratory of Stockholm City Council, Department of Virology, S-107-26 Stockholm, Sweden

Nucleotide sequencing of the fusion protein (F) gene of phocid distemper virus-2 (PDV-2), recently isolated from Baikal seals (Phoca sibirica), revealed an open reading frame (nucleotides 84 to 2075) with two potential in-frame ATG translation initiation codons. We suggest that the second in-frame ATG triplet at positions 264 to 266 initiates the translation, resulting in a protein of 537 amino acid residues with a calculated $M_r$ of 63035. The putative F1/F2 cleavage site, located approximately 100 amino acid residues from the N terminus, is identical to those of the F proteins of phocid distemper virus-1 (PDV-1) isolated from European harbour seals (Phoca vitulina) and of canine distemper virus (CDV). A full scale comparison of morbillivirus F genes reveals that the conserved F0 extracellular protein-encoding region contains a large number of non-expressed mutations, suggesting that this part of the protein is under strong functional constraints. Phylogenetic analysis of morbillivirus F gene nucleotide sequences revealed a closer evolutionary relationship between PDV-2 and CDV than between PDV-1 and PDV-2. These data were supported by cross-reactivity patterns of PDV-2 and CDV obtained with monoclonal antibodies to structural proteins of PDV-1 and CDV, and suggest that PDV-2 is a strain of CDV, resulting from a trans-species infection.

Before the recognition of morbillivirus-related epizootics among aquatic mammals, four members of the genus Morbillivirus of the Paramyxoviridae family had been identified which induce serious disease in their natural host species: measles virus (MV) of humans, rinderpest virus (RPV) of cattle and other artiodactyls, pesto-despetits-ruminants virus of small ruminants and canine distemper virus (CDV) of dogs (Kingsbury et al., 1978).

Since 1988 we and others have isolated morbilliviruses from several aquatic mammal species. Phocid distemper virus type 2 (PDV-2) caused an epizootic among Baikal seals (Phoca sibirica) in Siberia that started in 1987 (Grachev et al., 1989; Osterhaus et al., 1989), and phocid distemper virus type 1 (PDV-1) caused an epizootic among harbour seals (Phoca vitulina) in north-west Europe from 1988 (Mahy et al., 1988; Osterhaus & Vedder, 1988). In addition, dolphin morbillivirus (DMV) was the basis of an epizootic among Mediterranean striped dolphins (Stenella coeruleoalba) that started in 1990 (Domingo et al., 1990; Visser et al., 1990), and from 1988 onward we and others have isolated a porpoise morbillivirus (PMV) from harbour porpoises (Phocoena phocoena) in north-west Europe (McCullough et al., 1991; Visser et al., 1993). Of these newly recognized morbilliviruses, PDV-1 has been characterized most extensively (Kövamees et al., 1991; Barrett et al., 1992; Blixenkrone-Möller et al., 1992a; Sharma et al., 1992) and was recently accepted as a new member of the genus Morbillivirus (Pringle, 1991). However, the phylogenetic relationships between PDV-2, DMV and PMV on the one hand, and the morbilliviruses of terrestrial mammals on the other, have not been studied.

Morbilliviruses consist of a helical nucleocapsid structure of negative single-stranded RNA of about 16000 nucleotides wrapped in a nucleoprotein (NP).
surrounded by the phosphoprotein (P), the large (L) protein and the matrix (M) protein. The M protein is associated with the viral envelope that contains the two glycosylated transmembrane proteins, the haemagglutinin (H) and fusion (F) proteins. All six structural proteins are encoded by separate transcription units, or genes, organized within the viral genome as 3'NP-P(C/V)-M-F-H-L 5' (Barrett et al., 1991). Although parts of the P gene are conserved among the morbilliviruses, the F protein encoding region has been shown to be most conserved, followed by the M, H, NP and P genes (Barrett et al., 1991). The conservation of these proteins has also been shown by cross-reactivities of polyclonal and monoclonal antibodies (MACs) raised against the different morbillivirus proteins (Örvell & Norrby, 1974; Örvell et al., 1985, 1990; Sheshberadaran et al., 1986).

A panel of MACs raised against the F, H and NP proteins of PDV-1 (Örvell et al., 1990) was tested for cross-reactivities with PDV-2 and four different CDV strains in an indirect ELISA (Visser et al., 1990; Blixenkrone-Möller et al., 1992b). All five PDV-1 F MACs tested cross-reacted with all viruses used. Only one of the 11 PDV-1 H MACs (1.069D9 H1) tested cross-reacted with all viruses. Another H-specific MAB (1.122D11 H6) showed cross-reactivity with PDV-2 and two of the three CDV strains (Convac and Onderstepoort). None of the three PDV-1 NP MACs tested showed any cross-reactivity with PDV-2 or CDV. A similar result has been described with a panel of F, H, NP and P MACs raised against the Convac strain of CDV (Örvell et al., 1985; Visser et al., 1990; Blixenkrone-Möller et al., 1992b). Collectively these data suggest that the F protein is well conserved between these viruses and that the PDV-2 exhibits a pattern that would fit within the pattern found for CDV strains.

To reveal the evolutionary relationship between PDV-2 and the other morbilliviruses, we determined the nucleotide sequence of the conserved PDV-2 F gene. Total RNA was isolated by the RNAzol method (Linna/Biotex Laboratories) from PDV-2-infected Vero cell cultures that showed extensive cytopathic changes (Visser et al., 1990). Single-stranded cDNA was synthesized using Moloney murine leukaemia virus H reverse transcriptase superscript (Gibco BRL) and an oligo(dT) primer. The PDV-2 F gene was subsequently amplified by PCR in a DNA thermal cycler (Perkin-Elmer Cetus) using CDV F gene-specific oligonucleotide primers: 5' primer (5' CCGGACTCCAGGGTCTCAGGAGATCGTGGTAAC 3') corresponding to the first 17 nucleotides of the CDV F sequence as published by Barrett et al. (1987) (SstI restriction site underlined) and a 3' primer (5' CGCACTAGTATGATGATGCTTTTCCGGTCCG 3') corresponding to the inverted complementary sequence of the CDV F nucleotide positions 2169 to 2185 (Barrett et al., 1987) (SpeI restriction site underlined). The amplified PCR product was digested with SstI and SpeI, and size-selected on a 1% ethidium bromide-agarose gel. The purified product (2.16 kb) was ligated into an SstI/SpeI-digested Bluescript phagemid vector and transformed into CaCl2-treated competent XL-1 Blue bacteria. Several recombinant clones were isolated and sequenced in both directions with non-radioactive T3 and T7 oligonucleotide primers (Applied Biosystems) using the 370A Automated Sequencer (Applied Biosystems). The nucleotide sequence of the PDV-2 F gene was aligned with the F gene sequences of other morbilliviruses (Richardson et al., 1986; Barrett et al., 1987; Tsuikiyama et al., 1988; Kövamees et al., 1991) using the program PILEUP of the GCG software package (Devereux et al., 1984).

Analysis of the PDV-2 F nucleotide sequence revealed that the largest open reading frame (ORF) starts at the ATG at positions 84 to 86 and ends at position 2075 (Fig. 1). Three additional ATG codons were found upstream of the F0-encoding region, one in-frame ATG at
Table 1. Nucleotide and amino acid sequence similarity between the F proteins of different morbilliviruses

<table>
<thead>
<tr>
<th>Origin of sequence</th>
<th>Nucleotide and amino acid sequence similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PDV-2 F</td>
</tr>
<tr>
<td>PDV-2 F</td>
<td>70*</td>
</tr>
<tr>
<td>PDV-1 F</td>
<td>79 (86)*</td>
</tr>
<tr>
<td>CDV F</td>
<td>91 (97)</td>
</tr>
<tr>
<td>MV F</td>
<td>69 (69)</td>
</tr>
<tr>
<td>RPV F</td>
<td>67 (67)</td>
</tr>
</tbody>
</table>

* Numbers above the diagonal indicate similarities (%) between the F gene nucleotide sequences. Numbers below the diagonal indicate similarities (%) between the amino acid sequences. Numbers in parentheses indicate the similarity excluding areas preceding the putative F0 proteins (amino acids 136 to 662). Morbillivirus F sequences used: PDV-2 (this work); PDV-1 (Kövamees et al., 1991); CDV (Barrett et al., 1987); MV (Richardson et al., 1986); RPV (Tsukiyama et al., 1988).

positions 264 to 266 and two out-of-frame ATGs at positions 226 to 228 and 406 to 408. Neither of the flanking regions of the two in-frame ATG codons were in good context for ribosomal binding (Kozak, 1986). The first out-of-frame ATG codon at 226 to 228 would be the most favourable for translation initiation. However, this initiation codon is followed by a stop codon at nucleotides 265 to 267, positioned downstream in the same reading frame, and is not likely to initiate translation of the F0 coding region. In order to define the position and length of the putative signal sequence, we performed a computer analysis based on the method of von Heijne (1986). This analysis revealed that the most likely signal sequence encoded the residues at amino acids 126 to 135. This is in agreement with a putative signal sequence for CDV F previously described by Barrett et al. (1987). Furthermore, Evans et al. (1990) showed that the ATG codon at positions 464 to 466 is required for in vivo translation of the CDV F protein. However, this translation initiation codon as described for CDV (Evans et al., 1990) is not conserved in the PDV-2 sequence, which contains an ATA codon at this position (464 to 466). To exclude the possibility that this was due to a reverse transcriptase or Taq polymerase error, we isolated a new batch of total RNA from PDV-2-infected Vero cells and cloned the region of nucleotides 336 to 551 using PCR with a 5' primer (ATAGAGCTCGGTTGGTCATCACACACCCAGAG) located 128 nucleotides upstream of the above predicted signal sequence and a 3' primer (ATAGTTATGTTTTGACATGTCATGATCCGATATAATCC) located 60 nucleotides downstream of the putative signal sequence. Several recombinant clones were sequenced in both directions and the sequences obtained were completely identical to those analysed before. This confirms that the PDV-2 F gene lacks the ATG triplet at positions 464 to 466, and this represents a significant difference between PDV-2 and CDV. In addition, the possibility cannot be excluded that one of the two out-of-frame ATG codons can be utilized in vivo as an in-frame translation initiation codon as a result of RNA ‘recoding’ (Gesteland et al., 1992; Jacks et al., 1988).

Comparison of the PDV-2 F nucleotide sequence with the sequences of F genes of CDV, PDV-1, MV and RPV revealed the highest similarity with the CDV F gene (Table 1). The nucleotide sequence alignment (Fig. 1) reveals that as well as the transmembrane coding region the F0 protein sequence contains a large number of silent mutations, suggesting that the protein is under strong functional constraints.

Phylogenetic and bootstrap analyses were performed using positions 494 to 2072 of our PDV-2 and previously published MV, RPV, CDV and PDV-1 F0-encoding sequences (Richardson et al., 1986; Barrett et al., 1987; Tsukiyama et al., 1988; Kövamees et al., 1991). Genetic distances between the nucleotide sequences were calculated according to the two-parameter method (Kimura, 1980). These analyses showed that all morbillivirus F nucleotide sequences of carnivores, including seals, constitute a separate group, as proposed earlier in part by Norrbjörk et al. (1992). Within this group the PDV-2 F gene sequence is more closely related to the CDV F gene sequence than to the PDV-1 F gene sequence. This suggests an interspecies exchange of morbilliviruses between carnivores. These analyses also confirmed the close evolutionary relationship between the MV and RPV F genes as was suggested before on the basis of antigenic cross-reactivity and sequence analysis studies previously performed on these viruses (Richardson et al., 1986; Hsu et al., 1988). This indicates that MV and RPV evolved from a common ancestral virus. It is interesting to note that we have recently shown that the cetacean morbilliviruses DMV and PMV are antigenically most related to the ruminant morbilliviruses, suggesting an evolutionary relationship between the viruses infecting these animal species (van Bressem et al., 1993; Visser et al., 1993).

The F protein of morbilliviruses is synthesized as an inactive precursor protein (F0) attached to a signal sequence. After removal of the signal sequence, the F0 protein is cleaved post-translationally by cellular proteases into two disulphide-linked F1 and F2 protein subunits (Scheid & Choppin, 1977). The M, of the PDV-2 F0 protein can be determined by Western blot analysis, and was shown to be approximately 63K as was found for CDV F0 (Rima, 1983). The putative PDV-2 F1/F2 cleavage site is situated at about 100 amino acids from the F0 N terminus. The amino acid sequences of the putative F1/F2 cleavage sites of CDV, PDV-1 and PDV-2 are identical (RRQRR) as shown in Fig. 2 (Kövamees...
The basis of similarity to other morbilliviruses for which this position was determined (Barrett et al., 1987; Buckland et al., 1987). Within the proposed F0 protein, 13 cysteine residues were detected, of which 12 are completely conserved among all morbillivirus F proteins (Fig. 2). The potential glycosylation sites of the PDV-2 F0 protein correspond to those earlier described for the CDV and PDV-1 F0 proteins (Fig. 2) (Barrett et al., 1987; Kováimees et al., 1991). Deducing amino acid sequence similarities of the PDV-2 F protein with other morbillivirus F proteins are given in Table 1.

In the present study we provide evidence for a close evolutionary relationship between PDV-2 and CDV by showing a largely similar pattern of reactivities with mAbs raised against the F protein of PDV-1, and a high degree of F gene nucleotide sequence similarity between CDV-F and PDV-2 F. Phylogenetic and bootstrap analyses of the presently known morbillivirus F gene sequences strongly support the close relationship between PDV-2 and CDV, suggesting a recent species exchange of morbillivirus in carnivores. The phylogenetic analysis also suggests that either MV or RPV results from a trans-species infection of an ancestral morbillivirus between humans and ruminants. Whether trans-species exchanges between different carnivores commonly occur cannot be addressed with these data but requires sequence analysis of F genes of other morbilliviruses. Such additional F gene sequence information for different CDV and PDV-2 isolates may reveal whether PDV-2 should be considered a separate member of the genus *Morbillivirus*, distinct from CDV, or is a strain of CDV as the phylogenetic and sequence analyses suggest.

The authors kindly acknowledge Ms C. Kruysen and Ms M. Eskens for help in preparing the manuscript and Professor M. Nei for the NJTREE program.

References


isolates of morbilliviruses from different carnivores. Archives of Virolology 123, 279–294.


(Received 10 March 1993; Accepted 13 May 1993)