Comparative analysis of the gene encoding the nucleocapsid protein of dolphin morbillivirus reveals its distant evolutionary relationship to measles virus and ruminant morbilliviruses

Merete Blixenkrone-Møller, Gert Bolt, Elisabeth Gottschalk and Marcel Kenter

1 Laboratory of Virology and Immunology, The Royal Veterinary and Agricultural University, Bülowsvej 13, Frederiksberg C, DK-1870, Denmark and 2 Department of Virology, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

A morbillivirus of uncertain origin recently killed hundreds of Mediterranean dolphins. This is the first report of the nucleotide and deduced amino acid sequence of a dolphin morbillivirus (DMV) gene. The sequence of the nucleocapsid (N) gene including boundaries was determined. When the DMV N gene coding region was compared with the corresponding sequences of other morbilliviruses a distant evolutionary relationship between these viruses and DMV was apparent. Phylogenetic analysis of the sequence data provided further evidence that DMV is not closely related to any known morbillivirus, whereas phocine distemper virus exhibits a relatively close relationship to canine distemper virus.

Within the last few years previously unknown morbilliviruses have caused epidemics among aquatic wildlife in European waters. Thus, phocine distemper virus (PDV) killed more than 18000 seals in northern Europe in 1988 to 1989 (Kennedy et al., 1988a; Osterhaus et al., 1988) and also posed a threat to land-living carnivores (Blixenkrone-Møller et al., 1992b). More recently another highly pathogenic morbillivirus of uncertain origin caused epidemics among striped dolphins (Stenella coeruleoalba) in the Mediterranean (Domingo et al., 1990; Van Bressem et al., 1993). Morbillivirus infections have also been reported in harbour porpoises (Phocoena phocoena) in the North Sea (Kennedy et al., 1988b).

The major aims of the present study were to characterize the dolphin morbillivirus (DMV) at the genomic level and to search for clues to its origin.

The Morbillivirus genus comprises five previously established members: canine distemper virus (CDV) of canids and their relatives, PDV recognized in seals and mink, measles virus (MV) of man, rinderpest virus (RPV) and peste-des-petits-ruminants virus (PPRV) of ruminants. All of these may cause severe disease in their respective hosts. In studies using monoclonal antibodies morbillivirus isolates from dolphin and porpoise proved closely related but distinct from previously established members of the genus (Welsh et al., 1992; Visser et al., 1993). Genome studies also suggested that the cetacean morbilliviruses should be considered a new distinct entity of the genus (Barrett et al., 1993).

The nucleocapsid (N) protein is the principal component of the helical capsids wrapping the negative sense genomic RNA of morbilliviruses. We report here the nucleotide sequence and deduced amino acid sequence of the N gene of DMV. The N gene was selected for use in phylogenetic analysis because it is the only gene for which the entire coding sequence has been published for all known morbilliviruses (Diallo et al., 1994). By aligning the sequence data obtained for DMV with those of the other morbilliviruses new clues as to their evolutionary relationships were obtained.

A morbillivirus was isolated from a striped dolphin stranded on the Mediterranean coast of Spain in 1990 (Domingo et al., 1990). (Tissue samples were kindly donated by M. Domingo, University of Barcelona, Spain.) A lung tissue homogenate was inoculated into primary canine kidney epithelial cell cultures and then co-cultivated with Vero cells essentially as described previously (Blixenkrone-Møller et al., 1992b). Upon the fifth passage in Vero cell cultures c.p.e. appeared with swollen and rounded cells and syncytiat cells. The unique antigenic make-up of this DMV isolate as compared to other morbilliviruses has been established using panels of PDV and CDV monoclonal antibodies (Blixenkrone-Møller, 1993).

Poly(A)+ RNA was purified from Vero cells infected with the fifth viral passage of the DMV isolate. A cDNA
library was constructed from this RNA in the pSport-1 vector and propagated in *Escherichia coli* DH5α (Superscript plasmid system, BRL). The DMV library was blotted onto nylon membranes and screened with a nearly full-length clone of the PDV N gene (N10) (Blixenkrone-Møller et al., 1992a). The probe was labelled with [32P]dCTP according to the technique of Feinberg & Vogelstein (1983) and hybridization was performed under low stringency conditions as previously described (Blixenkrone-Møller et al., 1992a). Two positive clones, N1 and N4, later determined to contain nearly full-length clone of the PDV N gene (N10) (Blixenkrone-Møller et al., 1992a) were sequenced completely on both strands using the dyeoxide-chain termination method (Sanger et al., 1977) and primers synthesized according to the sequences obtained.

The sequences of the N gene boundaries were obtained from polymerase chain reaction (PCR) products using genome-sense RNA as a template for initial reverse transcription. For amplification of the DMV genome adjoining the 3’ end of the N gene an antigenome-sense primer, A3 (5’ ACCAACAAAGTTTG 3’), representing the very 3’-end of the MV genome (Blumberg et al., 1988) and a genome-sense primer, A4, from nucleotide (nt) 171 to 187 (Fig. 1) were selected. Similarly, the location of the A1/A2 primer set synthesized to generate and amplify DNA products of the genome junction between the N gene and the phosphoprotein (P) gene is indicated in Fig. 1 (A1, nt 1423 to 1439; A2, nt 1803 to 1819). The A2 primer was based on sequences of the 5’ end of the P messenger of DMV obtained from the present library by probing with a full-length clone of the MV P gene (Schmid et al., 1987) as described above. (The cDNA clone, peV1, was kindly donated by R. Cattaneo, University of Zürich, Switzerland.)

Synthesis of cDNA and subsequent PCR was done in a single tube reaction essentially as described by Doherty et al. (1989). For initial reactions the A1 and A3 primers were used for priming Moloney murine leukemia virus reverse transcriptase under the conditions specified by the manufacturer (First-Strand Synthesis kit, Pharmacia). Subsequent amplifications were performed with the two selected primer pairs and AmpliTaq DNA

Fig. 1. Nucleotide sequence of the DMV N gene and flanking sequences including the start of the P gene. The positive antigenome sequence is displayed as determined from cDNA clones and PCR products. The locations of the primers used to obtain sequences at the gene boundaries are overlined. The sequence in lower case letters represents the first 15 nucleotides of the MV leader sequence (Blumberg et al., 1988).
Table 1. Gene-end, intergenic and gene-start sequences of N and P genes of DMV, MV, RPV, CDV and PDV*

<table>
<thead>
<tr>
<th>Leader</th>
<th>N gene</th>
<th>Intergenic sequence</th>
<th>Gene-end</th>
<th>Gene-start</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMV</td>
<td>ATTACA</td>
<td>CTT</td>
<td>AGGACAAAGT</td>
<td>P gene</td>
</tr>
<tr>
<td>MV</td>
<td>GTTATA</td>
<td>CTT</td>
<td>AGGACCCAGGT</td>
<td></td>
</tr>
<tr>
<td>RPV</td>
<td>ATTATA</td>
<td>CTT</td>
<td>AGGACCCAGGT</td>
<td></td>
</tr>
<tr>
<td>CDV</td>
<td>ATTATA</td>
<td>CTT</td>
<td>AGGACCCAGGT</td>
<td></td>
</tr>
<tr>
<td>PDV</td>
<td>ATTATA</td>
<td>CTT</td>
<td>AGGACCCAGGT</td>
<td></td>
</tr>
</tbody>
</table>

* Data are compiled from this study, Baron et al. (1993), Curran et al. (1992) and Sidhu et al. (1993).

Fig. 2. Alignment of N protein sequences of all known morbilliviruses. DMV, dolphin morbillivirus; PPRV, pestes-des-petits-ruminants virus; RPV, rinderpest virus; MV, measles virus; CDV, canine distemper virus; PDV, phocine distemper virus. Asterisks represent amino acid identity. The nucleotide and amino acid sequences were obtained from the EMBL database. DMV, this paper; PPRV, accession no. X74443; PDV, accession no. X75717. The MV consensus sequence was taken from Cattaneo et al. (1989); RPV was taken from Kamata et al. (1991); CDV was taken from Rozenblatt et al. (1985) and Sidhu et al. (1993). Multiple alignments were carried out using the CLUSTAL program of the PC GENE package (Bairoch, 1992).

The DMV N gene sequence is shown in Fig. 1 in antigenome (message) sense, together with the derived amino acid sequence of the coding region. The single long DMV gene open reading frame encodes an amino acid sequence of 523 residues; this is the same length as determined for the corresponding proteins of CDV and PDV, whereas two amino acid residues are missing when polymerase according to the supplier's instructions (GeneAmp, Perkin Elmer Cetus). The DNA was amplified during 35 cycles of 1 min at 94 °C, 1 min at 50 °C and 2 min at 72 °C. The PCR products were cloned into the pGEM-T vector and propagated in JM109 (Promega) prior to nucleotide sequencing on both DNA strands using T7 and SP6 primers (Promega).
Fig. 3. Phylogenetic analysis of the nucleotide sequences of the coding regions of morbillivirus N proteins. Alignment of the nucleotide sequences was done by use of the PILEUP program of the GCG package for the VAX computer (Devereux et al., 1984) and by hand (data not shown). The building of the dendrogram was carried out using the MEGA program version 1.01 (Kumar et al., 1993). The branching order represents possible evolutionary relationships and the branch lengths are proportional to the genetic distances between the sequences and the hypothetical common ancestors that existed at the nodes in the tree. The figure above the PPRV branch refers to the branch length. The percentages indicate the bootstrap P values after 2000 replications with a random seed number of 2397 (Blair Hedges, 1992).

compared to the N proteins of RPV, PPRV and MV. The calculated Mr value of 57 520 for the N protein of DMV is in close agreemen-t with those calculated for other morbilliviruses. The nucleotide sequence of the coding region of the N gene of DMV displays between 67 and 69% pairwise identity when aligned with the corresponding sequences of other morbilliviruses (data not shown).

The intergenic triplet, CUU (in the positive antigenome sense), known to be highly conserved between morbilliviruses is present at the terminus of the N gene (Fig. 1, Table 1). The base triplet at the N–P junction is flanked by sequences easily recognizable due to their similarity to the gene boundary sequences described for other morbilliviruses (Table 1).

The alignments of the N protein sequence determined for DMV with those of other morbilliviruses are given in Fig. 2. The N-terminal domain of the protein (residues 1 to 398), representing around 80% of the molecule, is relatively well conserved between DMV and other morbilliviruses, whereas no or very little similarity is found in the C-terminal part of the proteins. Despite this variable region, hydropathy profiles of the N protein of DMV showed strong similarities to those of the other morbilliviruses indicating conserved overall structural features (data not shown). Furthermore, morbillivirus N proteins have three clusters (residues 173 to 182, 272 to 279, and 332 to 341) of sequence identity between members of the Paramyxovirus and Morbillivirus genera, indicating their phylogenetic relationship (Morgan, 1991). From the alignments given in Fig. 2 DMV displays a nearly equidistant relationship to all the members of the morbillivirus group. Thus DMV exhibits an amino acid sequence identity of 71% with MV and the ruminant morbilliviruses compared to 68 to 69% when aligned to PDV and CDV, respectively.

Phylogenetic analyses were performed on the nucleotide sequence of the coding region of the DMV N gene and the corresponding sequences of all presently recognized members of the Morbillivirus genus. The sequences were aligned using the PILEUP program of the GCG package (Devereux et al., 1984) and by hand. The phylogenetic tree shown in Fig. 3 was constructed with the MEGA program (Kumar et al., 1993) using a distance matrix calculated according to the formula of Kimura (1980) and the neighbour-joining method (Saitou & Nei, 1987). The robustness of the predicted phylogenetic tree was statistically tested using the bootstrap method (Felsenstein, 1985; Blair Hedges, 1992). From the P values given in Fig. 3 it is apparent that CDV and PDV form a phylogenetic entity within the Morbillivirus genus, whereas the grouping of MV together with RPV and any other groupings within the genus have a markedly lower confidence level. Overall, the deduced evolutionary relationships support and extend recent findings that DMV is phylogenetically distinct from previously established members of the Morbillivirus genus (Barrett et al., 1993). The phylogenetic analysis further shows that DMV is not closely related to any known morbillivirus, whereas PDV exhibits a relatively close relationship to CDV. In the dendrogram (Fig. 3) the DMV lineage seems to have diverged early in morbillivirus evolution and a direct epidemiological link between the recent European morbillivirus epidemics among seals and those of dolphins appears most unlikely. Interestingly, this viral phylogeny
tallies with that of the host species involved. Current evidence suggests that modern-day cetaceans have developed from land-based ruminants (Martin, 1990), whereas seals seem to have arisen from a terrestrial carnivorous, otter-like ancestor (King, 1983). Serological investigations dating back to 1989 provide evidence for DMV-like morbillivirus infections in various cetacean species of European waters (Ross et al., 1992); however, the exact origin of virulent cetacean morbillviruses remains unknown.

The extensive migration pattern of aquatic wildlife (Kennedy et al., 1989; Markussen & Have, 1992), which may be influenced by environmental disturbances, could be the common factor leading to the introduction of previously unknown morbilliviruses into susceptible host populations among European aquatic wildlife. Moreover, it is reasonable to assume that spontaneous mutations play a crucial role in expanding the host range of morbilliviruses and fostering the survival of these viruses in nature.

We thank Ms T. Dannemann for excellent technical assistance, and Dr J. Christensen for scientific discussion. We are grateful to Dr M. Domingo for dolphin tissue samples, Dr R. Cattaneo for cDNA of measles virus P gene transcript, and Drs T. Barrett and M. Baron for scientific discussions and part of the phylogenetic analyses. This work was supported by the Danish Agricultural and Veterinary Research Council and Knud Højgaard Foundation.

References


(Received 28 March 1994; Accepted 16 June 1994)