Genetic diversity of the attachment (H) protein gene of current field isolates of canine distemper virus

Gert Bolt,1 Tove Dannemann Jensen,1 Elisabeth Gottschalck,1 Peter Arctander,2 Max J. G. Appel,3 Robin Buckland4 and Merete Blixenkrone-Møller1

1 Laboratory of Virology and Immunology, Department of Veterinary Microbiology, The Royal Veterinary and Agricultural University, Bülowsvej 13, 1870 Frederiksberg C, Denmark
2 Department of Population Biology, Zoological Institute, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen, Denmark
3 The James Baker Institute for Animal Health, New York State College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
4 Institut Pasteur de Lyon, Immunologie et Stratégie Vaccinale, Avenue Tony-Garnier, 69365 Lyon Cedex 07, France

To characterize the variability of recent field isolates of canine distemper virus (CDV) from different hosts and geographical areas, we conducted nucleotide sequence analysis of the gene encoding the haemagglutinin (H), the attachment protein of this virus. Pronounced differences between field isolates were revealed in comparison to the Convac and Onderstepoort vaccine strains. The diversity of CDV appeared to exceed that determined for measles virus. Phylogenetic analysis also separated the field isolates of CDV from the vaccine strains and provided evidence for the existence of different contemporary genotypes of CDV. Isolates from a Greenlandic sledge dog and a Siberian seal formed a distinct lineage. The remaining isolates formed a group. This group contained two European isolates from mink and ferret, a single lineage comprising three European dog isolates, and another separate lineage of North American isolates from dog, javelina, raccoon and captive leopards.

Canine distemper virus (CDV) is a highly contagious viral pathogen which can cause lethal systemic disease in dogs and other carnivores. CDV belongs to the genus *Morbillivirus* within the family *Paramyxoviridae*, which includes other contagious mammalian pathogens, namely phocine (seal) distemper virus (PDV), measles virus (MeV), rinderpest virus (RPV), peste-des-petits-ruminants virus (PPRV) and cetacean morbillivirus. Furthermore, a newly isolated equine virus has been suggested to be a new member of the genus *Morbillivirus* (Murray et al., 1995).

CDV has a broad host-range among canids and related species. In addition, fatal CDV infection has been reported in large felids (Appel et al., 1994), javelinas (collared peccaries; Appel et al., 1991) and Lake Baikal seals (Visser et al., 1990).

Like MeV, CDV is a monotypic virus as defined by polyclonal antisera and a single exposure to these viruses normally confers long-lasting immunity. In general, the introduction of live attenuated CDV vaccines in the 1950s and their extensive use have drastically reduced the incidence of canine distemper in dogs. However, canine distemper outbreaks, in which previously immunized dogs become infected, have recently been observed (Blixenkrone-Møller et al., 1993; C. E. Kommonen, National Veterinary and Food Research Institute, Helsinki, Finland, personal communication). Similarly, in the USA, measles epidemics among vaccinated young human adults have been documented in recent years (Rota et al., 1992). This raises the question of whether the vaccines currently used efficiently protect against present-day circulating wild-types.

We have previously examined the antigenic basis of vaccine strains and wild-type isolates of CDV by use of a large panel of monoclonal antibodies against the N, P, F and H viral components (Blixenkrone-Møller et al., 1992, 1993). The highest antigenic variation was found in the H protein, whereas the F and P proteins were affected to a much lower extent.

To elucidate the genetic basis of the H protein diversity of CDV, we conducted nucleotide sequence analysis of recent field isolates. Our results provide evidence for pronounced genetic diversity between field isolates of CDV, all of which

Author for correspondence: Merete Blixenkrone-Møller.
Fax + 45 35 28 27 42. e-mail mbm@kvl.dk

The H protein gene sequences reported in this paper have been submitted to the EMBL/GenBank databases and assigned the following accession numbers: Z47759 (Mink/DK86), Z47760 (Dog/GR88), Z47761 (Dog/DK91, B+C; Z47762 (Dog/US89), Z47763 (Leopard/US91), Z47764 (Javelina/US89) and Z47765 (Racoon/US89).
The Onderstepoort strain is used in the present study. 

Field isolates of CDV compared in the present study

<table>
<thead>
<tr>
<th>Isolate name*</th>
<th>Location</th>
<th>Isolate description</th>
<th>H gene sequence†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dog/GR88</td>
<td>Northern Greenland</td>
<td>Blixenkrone-Møller et al. (1992)</td>
<td>Present study</td>
</tr>
<tr>
<td>Dog/US89</td>
<td>Colorado, USA</td>
<td>Blixenkrone-Møller et al. (1992)</td>
<td>Present study</td>
</tr>
<tr>
<td>Dog/DK91B</td>
<td>Copenhagen, Denmark</td>
<td>Blixenkrone-Møller et al. (1993)</td>
<td>Present study</td>
</tr>
<tr>
<td>Dog/DK91C</td>
<td>Copenhagen, Denmark</td>
<td>Blixenkrone-Møller et al. (1993)</td>
<td>Present study</td>
</tr>
<tr>
<td>Dog/Ger90</td>
<td>Bielefeld, Germany</td>
<td>Harder et al. (1993)</td>
<td>Present study</td>
</tr>
<tr>
<td>Mink/DK86</td>
<td>Jutland, Denmark</td>
<td>Blixenkrone-Møller et al. (1992)</td>
<td>Present study</td>
</tr>
<tr>
<td>Ferret/Ger89</td>
<td>Northern Germany</td>
<td>Harder et al. (1993)</td>
<td>Present study</td>
</tr>
<tr>
<td>Raccoon/US89</td>
<td>Michigan, USA</td>
<td>Blixenkrone-Møller et al. (1992)</td>
<td>Present study</td>
</tr>
<tr>
<td>Javelina/US89b</td>
<td>Arizona, USA</td>
<td>Appel et al. (1991)</td>
<td>Present study</td>
</tr>
<tr>
<td>Leopard/US91a</td>
<td>Illinois, USA</td>
<td>Appel et al. (1994)</td>
<td>Present study</td>
</tr>
<tr>
<td>Leopard/US92c</td>
<td>California, USA</td>
<td>Appel et al. (1994)</td>
<td>Present study</td>
</tr>
<tr>
<td>Seal/Si88</td>
<td>Lake Baikal, Siberia</td>
<td>Visser et al. (1990)</td>
<td>Present study</td>
</tr>
</tbody>
</table>

* a, isolated from a subacute fatal case in January 1991; b, isolated from an acute fatal case in April 1991; c, isolated from a wildlife collared peccary; d, isolated from a captive black leopard; e, isolated from a captive Chinese leopard, accession no. Z54156 (A92-27/4); f, isolated from a wildlife freshwater seal, accession no. X84998 (PDV-2).  
† a, accession no. X85000 (5804/Han90); b accession no. X84999 (1493/Han89).

exhibited a substantial genetic distance from the vaccine strains.

The H gene nucleotide sequences of eight field isolates of CDV were determined. The origins of the isolates collected between 1986–1991 are given in Table 1. After one to six passages in Vero cells all isolates were plaque purified three times to provide stocks for RNA extraction. For each isolate, an ORF of 1821 nucleotides was found, predicting a H protein of 607 amino acids. The Convac strain also has 607 residues, but the deduced H protein of the Onderstepoort strain appears to be three amino acids shorter (Fig. 1).

Sequence data were analysed using the PC/GENE software package (IntelliGenetics) for MS-DOS. The positions of the start codons of the H and L gene open reading frames (ORFs) and of the stop codon of the F gene ORF were perfectly conserved as were the F–H and H–L intergenic triplets. For all field isolates, an ORF of 1821 nucleotides was found, predicting a H protein of 607 amino acids. The Convac strain also has 607 residues, but the deduced H protein of the Onderstepoort strain appears to be three amino acids shorter (Fig. 1).

The hydrophathy profiles of the predicted H proteins from the field isolates are very similar to those of the vaccine strains (data not shown). 11 of 12 cysteine residues, 28 of 48 glycine residues and 27 of 41 proline residues are completely conserved among the 14 H proteins (Fig. 1). Four potential N-linked glycosylation sites (N–X–S/T) at amino acids 19, 149, 422 and 587 are shared among all 14 H proteins. The Convac vaccine strain and the field isolates share three potential N-linked glycosylation sites, of which two (at amino acids 391 and 456) are present in all field isolates, while the third (at amino acid 603) is present in only seven of the field isolates. Remarkably, one potential N-linked glycosylation site at amino acid 309 is unique to all field isolates (Fig. 1). This finding confirms and extends previous studies by Harder et al. (1996). Interestingly, a potential H protein glycosylation site not present in attenuated MeV strains has also been detected in recent MeV field isolates (Rota et al., 1992). Whether extra glycosylation
CDV H genes of field isolates

Fig. 1. Amino acid sequences of the deduced CDV H proteins from the field isolates analysed in the present study, from Leopard/US92 (Harder et al., 1996), Dog/Ger90, Ferret/Ger89 and Seal/Si88 (Mamanev et al., 1995) and from the Onderstepoort (Barrett et al., 1987; Curran et al., 1991; accession no. D00758) and Convac (Kovalev et al., 1991a; accession no. Z35493) vaccine strains. Only differences from the Onderstepoort strain are indicated for the field isolates. The potential membrane-spanning hydrophobic region is marked by overlining. Underlining marks potential N-linked glycosylation sites. For details of the isolates, refer to Table 1.

Recent investigations have shown that amino acids 451 and 481 of the MeV H protein play key roles in the interaction with the CD46 receptor (Lecouturier et al., 1996). Further studies may determine whether the same region of the CDV H protein is involved in interaction with the as yet unknown CDV receptor(s).

In the coding region, 36.3 ± 1.8% (mean ± SD) of the nucleotide differences between the field isolates and the vaccine
strains involve amino acid differences. Overall, the most
distinct sequence divergence (7–10% at the nucleotide level
and 8–11% at the amino acid level) is revealed in the
comparison made between the group of field isolates on the
one hand and the vaccine strains on the other. Corresponding
divergences determined within the group of field isolates is
within a range of 0–7%. For comparison, the H gene
divergence between CDV and its closest relative, PDV, is
much greater (33–35% at the nucleotide level and 25–27% at
the amino acid level; Curran et al., 1992; Kovamees et al.,
1991 b). It should be noted that we did not find indications of
a skewed relationship of PDV to either the vaccine strains or
the field isolates of CDV, and in the following phylogenetic
analysis, we used the PDV H gene sequence as outgroup
(Scotland, 1992). The CDV isolates analyzed appeared to be
more heterogeneous and to differ more from currently used
vaccine strains than comparable MeV isolates and vaccine
strains, since the H protein divergence determined for MeV
field isolates and vaccine strains only reached 3% (Rota et al.,
1992). An appealing explanation for the pronounced genetic
variation of CDV as compared to MeV may be the multiple
interspecies transmissions and possible adaptations involved in
the epizootiology of CDV.

Phylogenetic analysis of the 14 CDV sequences was carried
out with the maximum parsimony method using the PAUP
version 3.1.1 software for Macintosh (Swofford, 1993). Two
different minimum length trees, representing the minimum
number of genetical events matching the observed nucleotide
differences were found by branch and bound search. The
robustness of the data was estimated by the bootstrap method
(Felsenstein, 1985), and a 75% majority rule consensus
bootstrap tree was constructed (Fig. 2). There is no conflict
between the two most parsimonious trees and the consensus
bootstrap tree.

The phylogenetic analysis, based on nucleotide sequences
of the entire long ORF of the CDV H gene, provides the
necessary informative sites for sound identification of distinct
contemporary CDV genotypes. The two vaccine strains,
possibly representing CDV genotypes circulating at the time
of their isolation, are separated from the group of recent field
isolates and form a single lineage. Among the contemporary
field isolates, the isolate from a virgin soil outbreak in a sled
dog population in remote Inuit settlements of arctic Northern
Greenland, and an isolate from an outbreak among freshwater
seals of Lake Baikal in Siberia form a distinct lineage. The
remaining 10 field isolates group together. Within this group,
the isolates from European dogs form a single lineage and the
North American isolates from dog, javelina, raccoon and
captive leopards form another separate lineage. The branches
of the isolates from European mustelid species have low
bootstrap values (58%), and their further relationship to the
two above lineages of European and American isolates could
not be determined.

In a previous study of CDV based on a 429 bp fragment of
the phosphoprotein gene of eight field isolates but also
including a more limited number of H gene sequences, the
 genetic differences of the isolates were found to correlate with
their geographical origin (Harder et al., 1996). Similar geo-
 graphically associated genotypes have been described for
MeV and RPV (Chamberlain et al., 1993; Rima et al., 1995).

In the present phylogeny of the H protein gene, all North
American isolates from domestic dog, wildlife species and
captive large felids comprise a single genotype. However, the
present study further reveals the existence of distinct geno-
types of CDV which do not seem to correlate exclusively to
the geographical origin of the isolates. Thus, a distinct lineage
comprising two isolates from Asia and Greenland with no
evident epidemiological link is identified.

The separation of the European mustelid isolates from the
established European dog isolate lineage could not be resolved,
and it is possible that more than one lineage of CDV has
circulated recently in Europe.

The H gene sequences of the field isolates compared in this
study were all amplified from CDV isolates after a limited
number of passages in cell culture. However, since all isolates
were passaged in Vero cells, we cannot rule out the possibility
that certain H protein amino acids have been changed during
in vitro cultivation of the isolates. In a comparable study by
Rota et al. (1994), however, only a few nucleotide differences
of the H gene sequences were identified in vaccine strains,
which had undergone extensive passages in vitro, as compared

![Fig. 2. Phylogenetic analysis of the CDV H genes and flanking regions.
The corresponding region of the closely related morbillivirus phocine
distemper virus (PDV) (Kovamees et al., 1991 b; accession no. Z369779)
was used as outgroup (Scotland, 1992). The tree is a 75% majority rule
consensus bootstrap tree found by the maximum parsimony method using
the PAUP version 3.1.1 software (Swofford, 1993). Bootstrap values
(Felsenstein, 1985) of each grouping are shown as percentages. The
length and consistency index (Siebert, 1992 and references therein) of
the tree are 606 and 0.693, respectively. Branch lengths are not related
to phylogenetic distances.](image-url)
to the progenitor Edmonton virus isolate. Unfortunately, sequence data for the progenitor virus of the currently used vaccine strains of CDV are not available. Therefore, the present study does not allow us to judge whether the H gene sequence of the Onderstepoort vaccine strain is a valid representative of the progenitor field virus which dates back to North American distemper outbreaks in the 1930s.

Previous analyses of the H genes from MeV field isolates indicated that during the last 20 years, the evolution away from the MeVs isolated in the 1950s had accelerated. This was suggested to be connected with increased immunological pressure due to vaccination campaigns (Rota et al., 1992). The present data on the CDV H gene do not allow for a similar evaluation. However, we have found no evidence suggesting that antigenic drift has played a decisive role in the apparent vaccination failures observed among Danish dogs during a recent distemper epidemic (Blixenkrone-Møller et al., 1992, 1993). Because of the broad host-range of CDV including a vast wildlife reservoir, a strong vaccine pressure appears less likely than for MeV.

This study was supported by the Danish Agricultural and Veterinary Research Council (Grant no. 13-4450-1). We thank T. Barrett and I. R. Pedersen for valuable comments on the manuscript.

References


Received 5 July 1996; Accepted 23 October 1996