

## Original Article

# Phylogenetic features of hemagglutinin gene in canine distemper virus strains from different genetic lineages

Peng Liao<sup>1,2\*</sup>, Li Guo<sup>1\*</sup>, Yongjun Wen<sup>1\*</sup>, Yangling Yang<sup>1</sup>, Shipeng Cheng<sup>1</sup>

<sup>1</sup>Institute of Special Economic Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, Changchun, China; <sup>2</sup>Institute of Subtropical Agriculture, Chinese Academy of Sciences, Changsha, China. \*Equal contributors.

Received February 3, 2015; Accepted April 3, 2015; Epub April 15, 2015; Published April 30, 2015

**Abstract:** In the present study, the genotype of two Canine distemper virus (CDV) strains, namely, ZJJ-SD and ZJJ-LN, were investigated, based on the whole hemagglutinin (HA) gene. The CDV strains were obtained from two foxes in Shandong Province and Liaoning Province in 2011. Phylogenetic analyses were carried out for 260 CDV strains worldwide, and a statistical analysis was performed in the amino acid substitutions at positions 530 and 549 of the HA protein. Phylogenetic analyses revealed that the two strains, ZJJ-SD and ZJJ-LN, belonged to the CDV Asia I lineage. Site 530 of HA protein was found to be relatively conserved within CDV lineages in different host species by combining the genetic sequence data with the published data from 260 CDV strains worldwide. The data analysis showed a bias toward the predicted substitution Y549H for the non-dog strains in Asia I and Europe lineages. The ratio of site 549 genetic drift in the HA gene were significantly different between dogs and non-dogs in the two lineages. The strain ZJJ-SD, from wild canid, has an Y549H substitution. It is one of three Y549H substitution for wild canids in Asia I lineages. Site 530 of HA protein was not immediately relative to CDV genetic drift from dogs to non-dogs. Statistical analysis indicated that non-dog strains have a high probability to contain Y549H than dog strains in Asia I and Europe lineages. Thus, site 549 is considered important in genetic drift from dogs to non-dogs, at least in Asia I and Europe lineages.

**Keywords:** Canine distemper virus (CDV), hemagglutinin (HA), phylogenetic analysis, genetic drift

## Introduction

Canine distemper virus (CDV) is an enveloped, single-stranded, negative RNA virus from the *Morbillivirus* genus and family Paramyxoviridae. CDV is prevalent among the dog population and causes disease in several types of carnivores worldwide [1-3]. The CDV hemagglutinin (HA) gene encodes the HA protein, which has a significant role in virus invasion. This protein is an important factor that binds the cell receptor-signaling lymphocytic activation molecule SLAM (CD150) in all morbilliviruses [3-5].

Experiments have shown that the CDV HA protein is the most variable protein in any member of the genus *Morbillivirus*. This finding could explain why CDV has a broader host range compared with other morbilliviruses [2]. The SLAM binding region for CDV is important when CDV enters into the host cells [5]. Phylogenetic analyses of the sequence data of strains from

domestic dogs and non-dog species indicate that positive selection drives amino acid substitutions at positions 530 and 549 within the SLAM binding region of the HA gene [6]. Although six different amino acids [aspartic acid (D), glutamic acid (E), glycine (G), asparagine (N), arginine (R), and serine (S)] were observed at site 530, majority of CDV strains retrieved from domestic dogs had either 530G or 530E. However, strains from terrestrial non-dog carnivores and one strain from an aquatic carnivore (Baikal seal, *Pusa sibirica*) had R, D, and N residues [3, 6]. In different hosts, CDV strains retrieved from domestic dogs typically showed tyrosine (Y) at site 549. However, 7 out of 12 strains from non-dog hosts encoded histidine (H) at site 549 [3, 6]. The wild-type domestic dog CDV strain in ferrets (*Mustela putorius*) showed a genetic drift in Y549H, which is also 106 expected in non-dog hosts [7]. Analysis of amino acid substitutions at the known functional positions 530 and 549 within the SLAM bind-

ing region of the HA gene indicated that the significant differences between the prevalence of 549Y and 549H in wild canid and non-canid strains suggested a degree of virus adaptation to the categories of the host [3].

Most phylogenetic analyses that focused on the HA gene from CDV strains from different parts of the world have shown that CDV strains usually clustered within different geographical areas (America I, America II, Asia I, Asia II, Europe, European wildlife, Arctic, South America, and Southern Africa) [8-10]. These results indicate that the HA gene has undergone genetic drift in different geographical regions [8].

In the current study, the genotype of two CDV strains, ZJJ-SD and ZJJ-LN, were analyzed. Phylogenetic analyses revealed that the strains examined belonged to the CDV Asia I lineage. The combination of the genetic sequence data with the published data from 262 CDV strains worldwide confirmed the importance of amino acid substitutions in the HA protein at position 549 (Y to H/L) to the spread of domestic-dog-adapted CDV strains to other animals and vaccine strains, especially in Asia I and Europe lineages. Strains from non-dog hosts were biased toward the predicted substitution Y549H in Asia I lineages Europe and vaccine strain lineages. Site 530 of the HA protein was found to be conserved within CDV lineages from different host species, but has undergone genetic drift from G/E to A in the Asia I lineage (GenBank No. HM623893.1 and HM623891.1), V in the European wildlife lineage (GenBank No. JN153021.1 and JN153022.1), and K in the Asia II lineage (GenBank No. AY297453.1).

### Materials and methods

#### *Clinical specimens*

The viscera of the foxes, which were suspected to have CDV infection, were collected from the provinces of China. CDV-positive samples were identified by RT-PCR, which amplifies a 335 bp long fragment of the nucleoprotein gene [11].

#### *RT-PCR*

The total RNA was extracted from 100 mg of homogenized tissue samples (lung, spleen, and liver) by using the TRIzol reagent (Invitrogen).

The extracted RNA was eluted using 30  $\mu$ L of diethyl pyrocarbonate-treated water. Reverse transcription was performed in a final volume of 20  $\mu$ L containing 12  $\mu$ L of RNA solutions, 4  $\mu$ L of 5 $\times$ RT buffer, 1  $\mu$ L of dNTP (10 mM each), 1  $\mu$ L (50 pmol) of random hexamers, 1  $\mu$ L of reverse transcriptase (AMV) (TaKaRa), and 1  $\mu$ L of RNase inhibitor, and was incubated at 42°C for 1 h and at 70°C for 15 min. PCR amplification was performed with primers targeting the HA gene, H-F (5'-TTAGGGCTCAGGTAGTCCA-3'), and H-R (5'-CTAAGKCCAATTGARA TGTGT-3'; K = G/T, R = A/G). PCR amplicons of the expected size (1879 bp) were obtained. After pre-denaturation at 95°C for 5 min, the PCR amplification cycle was optimized at 94°C for 45 s; 52°C for 45 s; and 72°C for 2 min, for 35 cycles with a final extension step at 72°C for 10 min. Approximately 10  $\mu$ L of the amplified product was visualized using gel electrophoresis in 1.0% agarose gel in the presence of ethidium bromide. Bands were visualized by ultraviolet light transillumination.

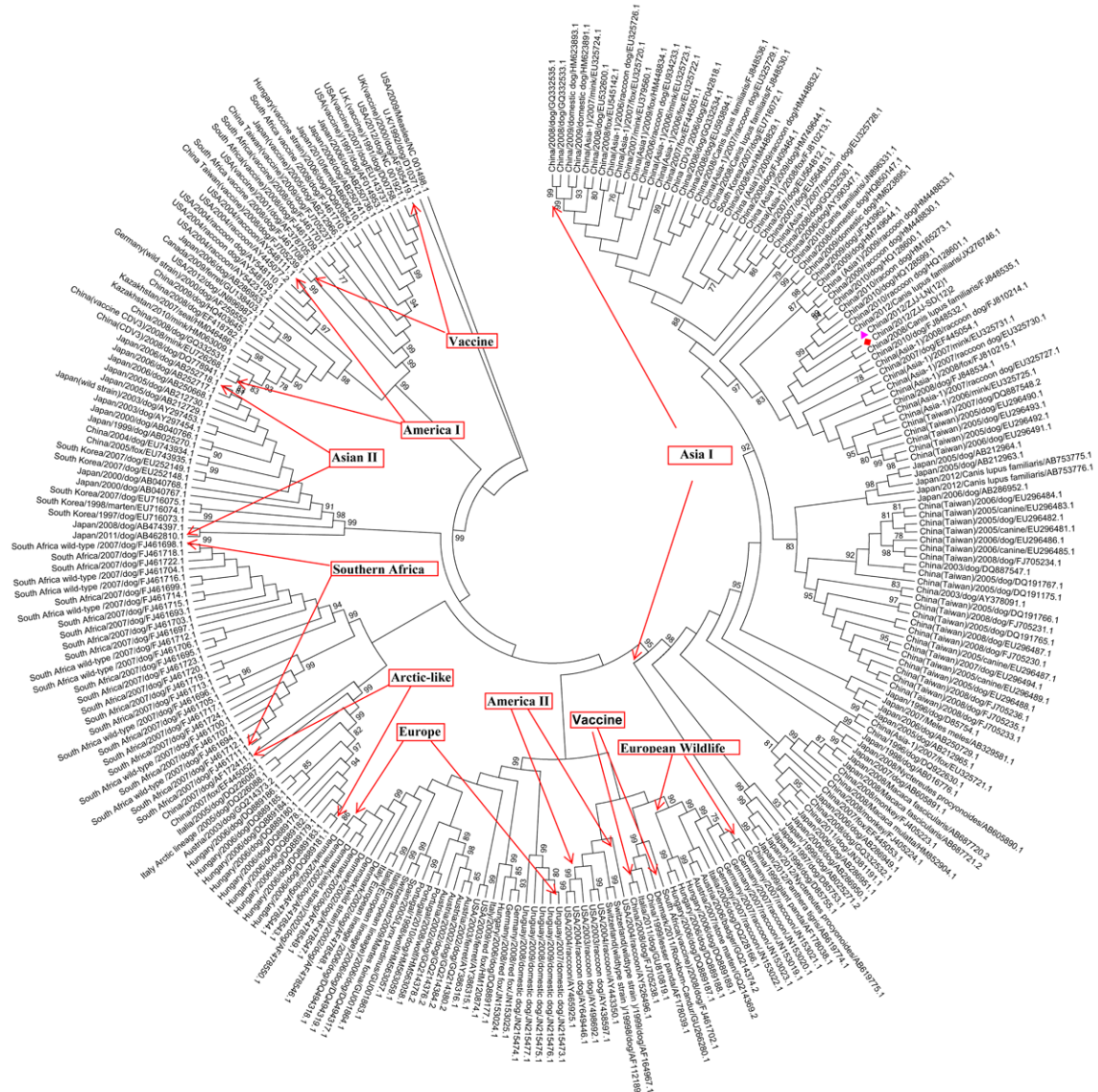
#### *Data set collection and phylogenetic analysis*

Sequences were extracted from the GenBank database. Datasets for the analysis included CDV HA sequences acquired from dog and non-dog hosts. Datasets were constructed for 262 HA genes with accession numbers shown in [Table S1](#) and [Figure 1](#). Sequences were aligned using Clustal X [12]. A neighbor-joining tree was generated with MEGA 5 [13]: (i) using 2 complete HA gene sequences obtained in the current study; and (ii) 260 published HA sequences from GenBank. Bootstrap consensus trees were inferred from 1000 replicates using the close-neighbor-interchange algorithm. The two complete CDV HA gene sequences used in the current study were provided by Jianjun Zhao. These sequences were recovered from two domesticated foxes caught in Shandong Province and Liaoning Province, China, in 2011.

#### *Analysis of amino acid sites*

The aa present at sites 530 and 549 on the HA protein was identified from a total of 262 CDV HA genes. The sequence data from numerous CDV outbreaks in a wide variety of carnivore species and locations would have been ideal. However, available sequence data were strongly biased to CDV strains from domestic dogs,

# HA gene analysis in CDV



**Figure 1.** Phylogenetic relationships of canine distemper virus strains obtained from two foxes (presented in bold) caught in China with the published strains based on the sequences of the hemagglutinin gene using the ClustalW program and the MEGA 5.0 software package. Neighbor-joining algorithm was used to generate the tree. Statistical support for nodes was provided by bootstrapping 1000 replicates. Only bootstrap values > 75% were shown. Triangles (▲) (China/2012/ZJJ-LN (12) 1, collected in Liaoning Province) and rhombuses (◆) (China/2012/ZJJ-SD (12) 2, collected in Shandong Province) indicate the two Chinese wild-type CDV strains analyzed in this study. For each strain, the country of origin, year, host species, and GenBank accession number are given. Marked areas indicate strains within lineages. The 262 strains included within each lineage are shown in [Table S1](#).

wild canids from Asian fur farms, and multiple sequence data from genetically identical strains from a relatively limited number of hosts. The ratio of amino acid drift in the HA protein at positions 530 and 549 was analyzed. Meanwhile, the genetic drifts of the HA gene in CDV virus strains from different genetic lineages were evaluated.

## Results

Complete CDV HA gene sequences were obtained from two domesticated foxes caught in China. Phylogenetic analyses of the HA sequences produced almost identical phylogenetic trees when the neighbor-joining method was used (**Figure 1**). Phylogenetic analyses

## HA gene analysis in CDV

revealed that the two strains belonged to the CDV Asia I lineage. The combination of the genetic sequence data with the published data from 260 CDV strains worldwide confirmed the bias of amino acid substitutions in the HA protein at position 549 (Y to H/L/S) in CDV strains from dogs and non-dogs and vaccines ([Tables S1](#) and [S2](#)). 3 out of 34 wild canids and 2 out of 13 non-canids have Y549H substitution but at the same time 3 out of 67 domestic dogs have the same substitution for Asia I lineage. The first two genetic variation rates are much higher than the latter. The same is in Europe lineage. Although, strains from dog and non-dog hosts were all exist substitution Y549H or Y549L or Y549S except no data available at all for wild canids and non-canids in Asia II and Southern Africa lineages, wild canids in America I, II and European wild life lineages or non-canids in Arctic lineage. In this paper, site 549 of the HA protein in one of two strains, ZJJ-SD, belonged to the wild canids with drift to H. That means a higher genetic drift ratio Y549H from dog to non-dog strains at least in Asia I and Europe lineages. Site 530 of the HA protein was conserved within CDV lineages from different host species. On the basis of the statistical analysis result ([Table S2](#)), site 530 of the HA protein was not directly related to the genetic drift of hemagglutinin gene in CDV strains from dog to non-dog species. However, this site has undergone a genetic drift toward A in the Asia I lineage (GenBank No. HM623893.1 and HM623891.1), V in the European wildlife lineage (GenBank No. JN153021.1 and JN153022.1), and K in the Asia II lineage (GenBank No. AY297453.1).

The evaluation results showed that the comparison of 262 CDV strains recovered worldwide illustrated that the residue at site 530 was generally conserved within the different CDV lineages regardless of host species (e.g., G in Asia I and in Europe; D in European wildlife; N in Arctic, South Africa, and America I; and E in Asia II, S in Vaccine) ([Table S2](#)). Across the globe, strains from non-canid wildlife species showed five different residues at site 530 (G, D, N, V, and R), whereas the domestic dog strains showed eight residues (G and E to R/D/N/V/A/K in [Table S2](#)) at this site. All the CDV strains were included in the Asia I lineage, regardless of host species code for G. In the European wildlife lineage, domestic dogs and non-dog species predominantly code D. The statistical

analysis result also indicated that site 530 was not directly related to the genetic drift of hemagglutinin gene in CDV strains from different genetic lineages. Site 530 of the genetic variation showed geographical differences.

Global comparison of the aa present at HA gene site 549 in 262 CDV strains revealed a highly significant probability of Y (rather than H) at this site in all the strains, except vaccine strains. Data restriction was observed in areas in which both 549Y and 549H CDV strains have been shown to circulate (Asia I, Asia II, Europe, European wildlife, America I, and America II). A significant bias to 549Y strains recovered from domestic dogs was also observed. Likewise, geographic clade is a variable element. Geographic clade has more significant effect on the existence of H or Y at site 549 compared with host species (e.g., both "Europe" and "European wildlife" have a greater mix of H and Y strains in canids than other clades). These results indicated that dog strains have a high probability to contain 549Y, whereas the non-dog strains showed a bias toward H at site 549. Nevertheless, a trend toward 549Y exists in Asia I and Asia II lineages. The ratio of genetic drift in dogs from Asia I was 4.5% (3/67), whereas that in canids and non canids was 8.8% (3/34) and 15.4% (2/13). The ratio of genetic drift in dogs from Europe was 9.5% (2/21), whereas that in canids and non canids was 8.8% (3/34) and 15.4% (2/13). The difference was highly significant ( $P < 0.01$ ). Thus, site 549 of HA protein is important to the genetic drift of hemagglutinin gene in CDV strains from dog to non-dog species, at least in Asia I and Europe lineages.

### Discussion

The strains recovered from the two foxes were identified under the Asia I wildlife lineage through phylogenetic analysis. Most of the HA proteins in the Asia I lineage were found to have residue 549Y. The HA gene of numerous strains in the Asia I and II lineages have been described, and the comparable sequence data from many of the strains in the Asia I and II lineages are available [11, 14]. Genotypes of sufficient CDV strains recovered from the foxes in China were also available to provide some general conclusions. The results suggest that a previously unrecognized CDV lineage is circulating in China and that the CDV strains are evolving independently within the area.

## HA gene analysis in CDV

Domestic-dog-adapted CDV strains were also found to spread readily to other carnivores. Currently, no evidence presented that carnivores in the Arctic or Southern Africa encountered CDV strains with 549H. These problems, including the small sample sizes and the likely dependence between identical strains, make statistical analyses of residues at these sites challenging. Given the strong effect of lineage on site 530, the diversity of residues at this site was reported only in strains from three categories of host: domestic dogs, wild canids, and non-canids.

One of the key factors influencing the ability of CDV strains to spread from domestic dog hosts to non-dog carnivore hosts is the amino acid substitutions at sites 530 and 549 in the SLAM binding region of the CDV HA protein [6]. The present results covered 260 CDV strains worldwide and confirmed the importance of amino acid substitutions in the HA protein at position 549 (Y to H/L) to the spread of domestic-dog-adapted CDV strains to other animals or vaccine strains. Site 530 of the HA protein in 262 CDV strains was found to be conserved within CDV lineages in different host species. However, this site has undergone genetic drift toward either A in Asia I lineages (HM623893.1 and HM623891.1), V in European wildlife lineages (JN153021.1 and JN153022.1), and K in Asia II lineages (AY297453.1). The HA protein in the two fox strains described had residues 549Y and 549H (Tables S1 and S2). The patterns of 549Y and 549H in wild canid and non-canid species were significantly different. On the basis of the data analysis, the ratio of site 549 genetic drift in the HA gene from dogs and non-dogs have a significant difference. This finding indicates that the virus-host adaptations were dependent on the residue at this site or at least for hosts within the broad categories. The present null hypothesis assumes that domestic dogs, wild-canids, and non-canid hosts will likely encounter strains with Y or H at site 549 and will be equally susceptible to both types of strains. However, most domestic dogs have been infected with CDV strains by other domestic dogs. Therefore, their chance of infection with a 549H strain was low. CDV transmission in wild carnivore species could also occur between individuals within a species and could be influenced by factors such as population size, degree of sociality, and ranging patterns [15]. Strains from non-dog hosts were biased

toward the predicted substitution Y549H in Asia I and Europe lineages. An unreported genetic drift was also observed in the HA gene of CDV at position 549, showing Y549H and Y549S in Asia I lineages, Y549L in vaccine strain lineages. The currently available data suggest that both domestic dogs and wild canid species will more likely be infected by CDV strains with 549Y than 549H in the SLAM binding region of the HA protein. However, determining whether the probability of infection of non-canid species will significantly be altered by the presence of either H or Y at site 549 is difficult because of insufficient data. Likewise, CDV strains recovered from wild canid species and non-canid species differed in the prevalence of H or Y at site 549. This finding raises questions on the effect of the residue at this site based on the evaluation of CDV strains to canid versus non-canid host species.

In conclusion, the study revealed that site 530 of the HA protein was not immediately relative to CDV genetic drift from dogs to non-dogs. The analysis indicated that dog strains will more likely have 549Y than 549H, whereas the non-dog strains showed different biases toward H or Y at site 549. Thus, site 549 is considered important in genetic drift from dogs to non-dogs, at least in Asia I and Europe lineages.

### Acknowledgements

This research was supported by the National Natural Science Foundation of China (grant: 31402088).

### Disclosure of conflict of interest

None.

**Address correspondence to:** Dr. Peng Liao, Institute of Subtropical Agriculture, Chinese Academy of Sciences, 644 Yuandaer Road, Changsha 410125, China. Tel: +86-731-84619703; Fax: +86-731-84612685; E-mail: liaopeng@isa.ac.cn; Shipeng Cheng, Institute of Special Economic Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, 4899 Juye Avenue, Changchun 130112, China. Tel: +86+431-81919802; Fax: +86+431-81919802; E-mail: tcscsp@126.com

### References

- [1] Appel MJ and Summers BA. Pathogenicity of morbilliviruses for terrestrial carnivores. *Vet Microbiol* 1995; 44: 187-191.

## HA gene analysis in CDV

- [2] Pomeroy LW, Bjornstad ON and Holmes EC. The evolutionary and epidemiological dynamics of the paramyxoviridae. *J Mol Evol* 2008; 66: 98-106.
- [3] Nikolin VM, Wibbelt G, Michler FU, Wolf P and East ML. Susceptibility of carnivore hosts to strains of canine distemper virus from distinct genetic lineages. *Vet Microbiol* 2012; 156: 45-53.
- [4] Tatsuo H, Ono N and Yanagi Y. Morbilliviruses use signaling lymphocyte activation molecules (CD150) as cellular receptors. *J Virol* 2001; 75: 5842-5850.
- [5] von Messling V, Oezguen N, Zheng Q, Vongpunsawad S, Braun W and Cattaneo R. Nearby clusters of hemagglutinin residues sustain SLAM-dependent canine distemper virus entry in peripheral blood mononuclear cells. *J Virol* 2005; 79: 5857-5862.
- [6] McCarthy AJ, Shaw MA and Goodman SJ. Pathogen evolution and disease emergence in carnivores. *Proc Biol Sci* 2007; 274: 3165-3174.
- [7] von Messling V, Springfield C, Devaux P and Cattaneo R. A ferret model of canine distemper virus virulence and immunosuppression. *J Virol* 2003; 77: 12579-12591.
- [8] Martella V, Cirone F, Elia G, Lorusso E, Decaro N, Campolo M, Desario C, Lucente MS, Bellicchio AL, Blixenkron-Moller M, Carmichael LE and Buonavoglia C. Heterogeneity within the hemagglutinin genes of canine distemper virus (CDV) strains detected in Italy. *Vet Microbiol* 2006; 116: 301-309.
- [9] Panzera Y, Calderon MG, Sarute N, Guasco S, Cardeillac A, Bonilla B, Hernandez M, Francia L, Bedo G, La Torre J and Perez R. Evidence of two co-circulating genetic lineages of canine distemper virus in South America. *Virus Res* 2012; 163: 401-404.
- [10] Woma TY, van Vuuren M, Bosman AM, Quan M and Oosthuizen M. Phylogenetic analysis of the haemagglutinin gene of current wild-type canine distemper viruses from South Africa: lineage Africa. *Vet Microbiol* 2010; 143: 126-132.
- [11] Tan B, Wen YJ, Wang FX, Zhang SQ, Wang XD, Hu JX, Shi XC, Yang BC, Chen LZ, Cheng SP and Wu H. Pathogenesis and phylogenetic analyses of canine distemper virus strain ZJ7 isolate from domestic dogs in China. *Virol J* 2011; 8: 520.
- [12] Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997; 25: 4876-4882.
- [13] Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011; 28: 2731-2739.
- [14] Zhao JJ, Yan XJ, Chai XL, Martella V, Luo GL, Zhang HL, Gao H, Liu YX, Bai X, Zhang L, Chen T, Xu L, Zhao CF, Wang FX, Shao XQ, Wu W and Cheng SP. Phylogenetic analysis of the haemagglutinin gene of canine distemper virus strains detected from breeding foxes, raccoon dogs and minks in China. *Vet Microbiol* 2010; 140: 34-42.
- [15] Guiserix M, Bahi-Jaber N, Fouchet D, Sauvage F and Pontier D. The canine distemper epidemic in Serengeti: are lions victims of a new highly virulent canine distemper virus strain, or is pathogen circulation stochasticity to blame? *J R Soc Interface* 2007; 4: 1127-1134.

## HA gene analysis in CDV

**Table S1.** Amino acids present at position 530 and 549 within the SLAM binding region of the hemagglutinin protein of 262 canine distemper virus (CDV) strains

	530	549	Origin/year/species/Accession number
			Asia I
			Domestic dogs
1	G	Y	China/2008/dog/GQ332535.1
2	G	Y	China/2008/dog/GQ332533.1
3	A	Y	China/2009/domestic dog/HM623893.1
4	A	Y	China/2009/domestic dog/HM623891.1
5	G	Y	China/2008/dog/EU532600.1
6	G	Y	China/2008/dog/GQ332534.1
7	G	Y	China/2008/dog/EU593894.1
8	G	Y	China/2008/Canis lupus familiaris/FJ848536.1
9	G	Y	China/2008/Canis lupus familiaris/FJ848530.1
10	G	Y	South Korea/2007/dog/EU716072.1
11	G	Y	China/2008/dog/FJ409464.1
12	G	Y	China/2009/dog/HM749644.1
13	G	Y	China/2007/dog/EU564812.1
14	G	Y	China/2007/dog/EU564813.1
15	G	Y	China/2008/dog/GQ332530.1
16	G	Y	China/2006/dog/AY390347.1
17	G	Y	China/2010/Canis familiaris/JN896331.1
18	G	Y	China/2008/domestic dog/HQ850147.1
19	G	Y	China/2009/domestic dog/HM623895.1
20	G	Y	China/2009/dog/JF343962.1
21	G	Y	China/2009/dog/HM749644.1
22	G	Y	China/2010/dog/HQ128600.1
23	G	Y	China/2010/dog/HQ128599.1
24	G	Y	China/2012/Canis lupus familiaris/JX276746.1
25	G	Y	China/2008/Canis lupus familiaris/FJ848535.1
26	G	Y	China/2010/dog/FJ848532.1
27	G	Y	China/2007/dog/EF445054.1
28	G	Y	China/2008/dog/FJ848534.1
29	G	Y	China (Taiwan)/2007/dog/DQ887548.2
30	G	Y	China (Taiwan)/2005/dog/EU296490.1
31	G	Y	China (Taiwan)/2005/dog/EU296493.1
32	G	Y	China (Taiwan)/2005/dog/EU296492.1
33	G	Y	China (Taiwan)/2006/dog/EU296491.1
34	G	Y	Japan/2005/dog/AB212964.1
35	G	Y	Japan/2005/dog/AB212963.1
36	G	Y	Japan/2012/Canis lupus familiaris/AB753775.1
37	G	Y	Japan/2012/Canis lupus familiaris/AB753776.1
38	G	S	Japan/2006/dog/AB286952.1
39	G	Y	China (Taiwan)/2006/dog/EU296484.1
40	G	Y	China (Taiwan)/2005/canine/EU296483.1
41	G	H	China (Taiwan)/2005/dog/EU296482.1
42	G	H	China (Taiwan)/2005/canine/EU296481.1
43	G	Y	China (Taiwan)/2006/dog/EU296486.1
44	G	Y	China (Taiwan)/2006/canine/EU296485.1

## HA gene analysis in CDV

45	G	Y	China (Taiwan)/2008/dog/FJ705234.1
46	G	Y	China/2003/dog/DQ887547.1
47	G	Y	China (Taiwan)/2005/dog/DQ191767.1
48	G	Y	China (Taiwan)/2005/dog/DQ191175.1
49	G	Y	China/2003/dog/AY378091.1
50	G	Y	China (Taiwan)/2005/dog/DQ191766.1
51	G	Y	China (Taiwan)/2008/dog/FJ705231.1
52	G	Y	China (Taiwan)/2005/dog/DQ191765.1
53	G	Y	China (Taiwan)/2008/dog/EU296487.1
54	G	Y	China (Taiwan)/2008/dog/FJ705230.1
55	G	Y	China (Taiwan)/2005/canine/EU296487.1
56	G	Y	China (Taiwan)/2007/dog/EU296494.1
57	G	Y	China (Taiwan)/2005/canine/EU296489.1
58	G	Y	China (Taiwan)/2005/dog/EU296488.1
59	G	Y	China (Taiwan)/2008/dog/FJ705236.1
60	G	Y	China (Taiwan)/2008/dog/FJ705235.1
61	G	Y	China (Taiwan)/2008/dog/FJ705233.1
62	G	Y	Japan/1996/dog/D85754.1
63	G	Y	Japan/2006/dog/AB250729.1
64	G	Y	Japan/2005/dog/AB212965.1
65	G	Y	China/1996/dog/DQ922630.1
66	G	Y	Japan/1998/dog/AB016776.1
67	G	H	Japan/2007/dog/AB605891.1
			Asia I
			Wild canids
68	G	Y	China/2006/racc. dog/EU325726.1
69	G	Y	China/2007/fox/EU325720.1
70	G	Y	China/2009/fox/HM448834.1
71	G	Y	China/2006/racc. dog/EU934233.1
72	G	Y	China/2006/fox/EU325722.1
73	G	Y	China/2007/fox/EF445051.1
74	G	Y	China /2006/dog/EF042818.1
75	G	Y	China/2007/racc. dog/EU325729.1
76	G	H	China/2008/fox/HM448829.1
77	G	Y	China/2009/racc. dog/HM448832.1
78	G	Y	China/2008/fox/FJ810213.1
79	G	Y	China/2007/racc. dog/EU325728.1
80	G	Y	China/2009/racc. dog/HM448833.1
81	G	Y	China/2009/racc. dog/HM448830.1
82	G	Y	China/2010/raco. dog/HM165273.1
83	G	Y	China/2010/racc. dog/HQ128601.1
84	G	Y	<b>China/2012/ ox/ZJJ-LN (12) 1</b>
85	G	H	<b>China/2012/fox /ZJJ-SD (12) 2</b>
86	G	Y	China/2008/racc. dog/FJ810214.1
87	G	Y	China/2007/racc. dog/EU325730.1
88	G	Y	China/2008/fox/FJ810215.1
89	G	Y	China/2007/racc. dog/EU325727.1
90	G	Y	China/2007/fox/EU325721.1
91	G	Y	China/2007/fox/EF445053.1



## HA gene analysis in CDV

92	G	Y	Japan/2006/dog/AB286949.1
93	G	Y	China/2008/dog/GQ332532.1
94	G	Y	China/2011/dog/JN381191.1
95	G	Y	Japan/2006/dog/AB286951.1
96	G	Y	Japan/2006/dog/AB286950.1
97	G	Y	Japan/1999/dog/AB025271.2
98	G	Y	Japan/1997/dog/D85753.1
99	G	Y	Japan/1996/dog/D85755.1
			Asia I
			Non-canids
100	G	Y	China /2007/mink/EU325724.1
101	G	H	China/2007/mink/EU379560.1
102	G	Y	China/2006/mink/EU325723.1
103	G	Y	China/2007/mink/EU325731.1
104	G	Y	China/2006/mink/EU325725.1
105	G	H	Japan/2007/Meles meles/AB329581.1
106	G	H	Japan/2008/Nyctereutes procyonoides/AB605890.1
107	G	Y	Japan/2008/Macaca fascicularis/AB687720.2
108	G	Y	Japan/2008/Macaca fascicularis/AB687721.2
109	G	Y	China/2008/monkey/FJ405223.1
110	G	Y	China/2008/Macaca mulatta/HM852904.1
111	G	Y	China/2008/monkey/FJ405224.1
112	G	Y	Japan/2012/Nyctereutes procyonoides/AB619775.1
113	G	Y	Japan/2012/Panthera tigris/AB619774.1
114	G	Y	China/1999/giant panda/AF178038.1
			Europe wildlife
			Domestic dogs
115	N	H	Italia/2005/dog/DQ228166.1
116	D	Y	Hungary/2006/dog/DQ889189.1
117	D	Y	Hungary/2006/dog/DQ889188.1
118	D	Y	Hungary/2006/dog/DQ889187.1
			Europe wildlife
			Non-canids
119	V	H	Germany/2007/raccoon/JN153021.1
120	D	H	Germany/2007/raccoon/JN153020.1
121	D	H	Germany/2007/raccoon/JN153019.1
122	D	H	Germany/2007/raccoon/JN153023.1
123	V	H	Germany/2007/raccoon/JN153022.1
124	D	H	Austria/2006/badger/GQ214374.2
125	D	H	Austria/2007/stone marten/GQ214369.2
			Vaccine
126	D	Y	South Africa/2008/ vaccine/FJ461702.1
127	D	Y	Denmark/vaccine/GU266280.1
128	D	Y	China/lesser panda/AF178039.1
129	D	Y	Sweden/dog/GU810819.1
130	D	Y	China/2008/ FJ705238.1
131	S	H	China Taiwan/2008/dog/FJ705239.1
132	S	H	South Africa/2008/dog/FJ461708.1
133	S	H	USA/2001/dog/AF378705.1

## HA gene analysis in CDV

134	S	H	South Africa/2008/dog/FJ461709.1
135	S	H	South Africa/2008/dog/FJ461701.1
136	S	H	China Taiwan/2009/dog/FJ705237.1
137	S	H	Japan/2005/dog/AB212966.1
138	S	H	South Africa vaccine/2008/dog/FJ461710.1
139	S	H	Hungary/2006/dog/DQ903854.1
140	S	H	Japan/2010/ferret/AB606410.1
141	S	L	Japan/2006/dog/AB250741.1
142	S	L	Japan/2006/dog/AB250736.1
143	S	H	USA/1997/dog/AF014953.1
144	S	H	USA/2007/dog/EU143737.1
145	S	H	U.K./1991/dog/D00758.1
146	S	H	USA/2012/dog/NC 001921.1
147	S	H	UK/2000/dog/AF305419.1 America II Domestic dogs
148	G	Y	Switzerland/1999/dog/AF164967.1
149	G	Y	Switzerland/19998/dog/AF112189.1 America II Non-canids
150	G	H	USA/2004/raccoon/AY526496.1
151	G	H	USA/2004/raccoon/AY443350.1
152	G	H	USA/2003/racc. dog/AY438597.1
153	R	H	USA/2003/racc. dog/AY498692.1
154	R	H	USA/2004/racc. dog/AY649446.1
155	R	H	USA/2004/raccoon/AY465925.1 Europe Domestic dogs
156	G	Y	Uruguay/2007/domestic dog/JN215473.1
157	G	Y	Uruguay/2009/domestic dog/JN215476.1
158	G	Y	Uruguay/2008/domestic dog/JN215475.1
159	G	Y	Uruguay/2009/domestic dog/JN215477.1
160	G	Y	Uruguay/2008/domestic dog/JN215474.1
161	G	Y	Hungary/2006/dog/DQ889177.1
162	G	Y	Austria/2002/dog/GQ214380.2
163	G	Y	Austria/2002/dog/GQ214384.2
164	G	Y	Austria/2002/dog/GQ214376.2
165	G	Y	Austria/2002/dog/GQ214378.2
166	G	Y	Italia/2006/dog/DQ494317.1
167	G	Y	Italia/2006/dog/DQ494319.1
168	G	Y	Italy/2006/dog/DQ494318.1
169	G	Y	Denmark/2002/dog/AF478548.1
170	G	Y	Denmark/2002/dog/AF478546.1
171	G	Y	Denmark/2002/dog/AF478547.1
172	G	Y	Denmark/2002/dog/AF478549.1
173	G	Y	Denmark/2002/dog/AF478550.1
174	G	H	Denmark/2002/dog/AF478543.1
175	G	H	Denmark/2002/dog/AF478544.1
176	G	Y	Portugal/2010/dog/HM563059.1

## HA gene analysis in CDV

			Europe
			Wild canids
177	G	Y	Germany/2008/red fox/JN153025.1
178	G	Y	Germany/2008/red fox/JN153024.1
179	G	H	Italy/2009/red fox/HM120874.1
180	G	Y	Portugal/2008/wolf/HM563058.1
181	G	Y	Portugal/1998/wolf/HM563057.1
			Europe
			Non-canids
182			
183	G	Y	USA/2003/ferret/ay386315.1
184	G	H	USA/2003/ferret/ay386316.1
185	G	Y	Spain/2005/Lynx pardinus/GU001863.1
186	G	Y	Switzerland/2009/Martes foina/GU001864.1
			Arcitic-Like
			Domestic dogs
187	N	Y	Hungary/2006/dog/DQ889181.1
188	N	Y	Hungary/2006/dog/DQ889183.1
189	N	Y	Hungary/2006/dog/DQ889179.1
190	N	Y	Hungary/2006/dog/DQ889182.1
191	N	Y	Hungary/2006/dog/DQ889178.1
192	N	Y	Hungary/2006/dog/DQ889180.1
193	N	Y	Hungary/2006/dog/DQ889184.1
194	S	Y	Hungary/2006/dog/DQ889185.1
195	S	Y	Hungary/2006/dog/DQ889186.1
196	N	Y	Austria/2003/dog/GQ214373.2
197	N	Y	Italy /2005/dog/DQ226088.1
198	N	Y	Italia/2005/dog/DQ226087.1
199	G	Y	China/2007/dog/AF172411.1
			Arcitic-Like
			Wild canids
200	N	Y	China/2007/fox/EF445052.1
			Southern Africa
			Domestic dogs
201	N	Y	South Africa/2007/dog/FJ461711.1
202	N	Y	South Africa/2007/dog/FJ461712.1
203	N	Y	South Africa/2007/dog/FJ461694.1
204	N	Y	South Africa/2007/dog/FJ461707.1
205	N	Y	South Africa/2007/dog/FJ461700.1
206	N	Y	South Africa/2007/dog/FJ461724.1
207	N	Y	South Africa/2007/dog/FJ461717.1
208	N	Y	South Africa/2007/dog/FJ461705.1
209	N	Y	South Africa/2007/dog/FJ461696.1
210	N	Y	South Africa/2007/dog/FJ461713.1
211	N	Y	South Africa/2007/dog/FJ461719.1
212	N	Y	South Africa/2007/dog/FJ461720.1
213	N	Y	South Africa/2007/dog/FJ461723.1
214	N	Y	South Africa/2007/dog/FJ461695.1
215	N	Y	South Africa/2007/dog/FJ461706.1

## HA gene analysis in CDV

216	N	Y	South Africa /2007/dog/FJ461712.1
217	N	Y	South Africa/2007/dog/FJ461697.1
218	N	Y	South Africa/2007/dog/FJ461703.1
219	N	Y	South Africa/2007/dog/FJ461693.1
220	N	Y	South Africa/2007/dog/FJ461715.1
221	N	Y	South Africa /2007/dog/FJ461714.1
222	N	Y	South Africa/2007/dog/FJ461699.1
223	N	Y	South Africa /2007/dog/FJ461716.1
224	N	Y	South Africa /2007/dog/FJ461704.1
225	N	Y	South Africa/2007/dog/FJ461722.1
226	N	Y	South Africa/2007/dog/FJ461718.1
227	N	Y	South Africa /2007/dog/FJ461698.1
			Asia II
			Domestic dogs
228	E	Y	Japan/2011/dog/AB462810.1
229	E	Y	Japan/2008/dog/AB474397.1
230	E	Y	South Korea/1997/dog/EU716073.1
231	G	Y	South Korea/1998/marten/EU716074.1
232	G	Y	South Korea/2007/dog/EU716075.1
233	E	Y	Japan/2000/dog/AB040767.1
234	E	Y	Japan/2000/dog/AB040768.1
235	E	Y	South Korea/2007/dog/EU252148.1
236	E	Y	South Korea/2007/dog/EU252149.1
237	R	Y	China/2005/fox/EU743935.1
238	R	Y	China/2004/dog/EU743934.1
239	E	Y	Japan/1999/dog/AB025270.1
240	E	Y	Japan/2000/dog/AB040766.1
241	E	Y	Japan/2003/dog/AY297454.1
242	K	Y	Japan/2003/dog/AY297453.1
243	E	Y	Japan/2005/dog/AB212729.1
244	E	Y	Japan/2005/dog/AB212730.1
245	E	Y	Japan/2006/dog/AB250668.1
246	E	Y	Japan/2006/dog/AB252717.1
247	E	Y	Japan/2006/dog/AB252718.1
			America I
			Domestic dogs
248	N	H	China/2008/dog/DQ778941.1
249	N	H	China/2008/dog/GQ332531.1
250	N	H	China/2008/dog/EF418782.1
251	N	H	China/2009/dog/HQ403645.1
252	N	Y	Germany/2000/dog/AF259552.1
253	N	Y	USA/2012/dog/JN896987.1
254	N	Y	Japan/2006/dog/AB286953.1
			America I
			Non-canids
255	N	H	China/2008/mink/EU726268.1
256	N	H	Kazakhstan/2010/mink/HM063009.1
257	N	H	Kazakhstan/2007/seal/HM046486.1
258	N	Y	Canada/2009/ferret/GU138403.1

## HA gene analysis in CDV

259	N	Y	USA/2004/raccoon/AY542312.2
260	N	Y	USA/2004/racc. dog/AY548109.1
261	N	Y	USA/2004/racc. dog/AY548110.1
262	N	Y	USA/2004/raccoon/AY445077.2
263	N	Y	USA/2004/raccoon/AY548111.1

The strains are organized based on CDV lineages (Europe, European wildlife, Arctic, Asia I and II, America I and II) or geographical area (southern Africa) and within the three categories of terrestrial species. Strains are named by country, year of collection (if known), and GenBank accession number. The strains retrieved in this study are in bold.

**Table S2.** Ratio of genetic drift in hemagglutinin gene

Genotype	Species	530	549
Asia I	Domestic dogs	G 65/67, A 2/67	Y 63/67, H 3/67, S 1/67
	Wild canids	G 34/34	Y 31/34, H 3/34
	Non-canids	G 13/13	Y 11/13, H 2/13
Asia II	Domestic dogs	G 2/20, R 2/20, K 1/20, E 15/20	Y 20/20
	Wild canids	NR	NR
	Non-canids	NR	NR
America I	Domestic dogs	N 7/7	Y 3/7, H 4/7
	Wild canids	NR	NR
	Non-canids	N 9/9	Y 6/9, H 3/9
America II	Domestic dogs	G 2/2	NR
	Wild canids	NR	H 6/6
	Non-canids	G 3/6, R 3/6	Y 19/21, H 2/21
Europe	Domestic dogs	G 21/21	Y 19/21, H 2/21
	Wild canids	G 5/5	Y 4/5, H 1/5
	Non-canids	G 4/4	Y 3/4, H 1/4
European wildlife	Domestic dogs	D 3/4, N 1/4	Y 3/4, H 1/4
	Wild canids	NR	NR
	Non-canids	D 5/7, V 2/7	H 7/7
Southern Africa	Domestic dogs	N 27/27	Y 27/27
	Wild canids	NR	NR
	Non-canids	NR	NR
Arctic	Domestic dogs	N 10/13, S 2/13, G 1/13	Y 13/13
	Wild canids	N 1/1	Y 1/1
	Non-canids	NR	NR
Vaccines		D 5/22, S 17/22	Y 5/22, H 15/22, L 2/22