Virus Genes (2014) 49:45–57
DOI 10.1007/s11262-014-1062-z

Genetic characterization of an isolate of canine distemper virus from a Tibetan Mastiff in China

Weike Li · Tiansong Li · Yuxiu Liu · Yuwei Gao · Songtao Yang · Na Feng · Heting Sun · Shengle Wang · Lei Wang · Zhigao Bu · Xianzhu Xia

Received: 4 February 2014 / Accepted: 12 March 2014 / Published online: 2 April 2014

Abstract Canine distemper (CD) is a highly contagious, often fatal, multisystemic, and incurable disease in dogs and other carnivores, which is caused by canine distemper virus (CDV). Although vaccines have been used as the principal means of controlling the disease, CD has been reported in vaccinated animals. The hemoagglutinin (H) protein is one of the most important antigens for inducing protective immunity against CD, and antigenic variation of recent CDV strains may explain vaccination failure. In this study, a new CDV isolate (TM-CC) was obtained from a Tibetan Mastiff that died of distemper, and its genome was characterized. Phylogenetic analysis of the H gene revealed that the CDV-TM-CC strain is unique among 20 other CDV strains and can be classified into the Asia-1 group with the Chinese strains, Hebei and HLJ1-06, and the Japanese strain, CYN07-hV. The H gene of CDV-TM-CC shows low identity (90.4 % nt and 88.9 % aa) with the H gene of the classical Onderstepoort vaccine strain, which may explain the inability of the Tibetan Mastiff to mount a protective immune response. We also performed a comprehensive phylogenetic analysis of the N, P, and F protein sequences, as well as potential N-glycosylation sites and cysteine residues. This analysis shows that an N-glycosylation site at aa 108-110 within the F protein of CDV-TM-CC is specific for the wild-type strains (5804P, A75/17, and 164071) and the Asia-1 group strains, and may be another important factor for the poor immune response. These results provide important information for the design of CD vaccines in the China region and elsewhere.

Keywords Canine distemper virus (CDV) · VerodogSLAM cells · Phylogenetic analysis · Virulence · Genotype

Weike Li and Tiansong Li contributed equally to this paper.

W. Li · H. Sun
Wildlife Resources College, Northeast Forestry University, Harbin 150040, China
e-mail: weikeduo0535@163.com

T. Li
College of Chemistry and Biology, Beidhua University, Jinlin 132013, China

Y. Liu · Y. Gao · S. Yang · N. Feng · S. Wang · L. Wang · X. Xia (✉)
Institute of Military Veterinary, Academy of Military Medical Sciences, No. 666, Liuying Xilu, Jingyue Economic Development Zone, Changchun 130122, China
e-mail: xia_xzh@yahoo.com.cn

Z. Bu (✉)
Key Laboratory of Veterinary Public Health of Ministry of Agriculture, State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, 427 Maduan Street, Harbin 150001, China
e-mail: zgbu@yahoo.com

Introduction

Canine distemper (CD), one of the most well-known infectious diseases of wild and domestic Canidae, is a highly contagious, often fatal, multisystemic, and incurable viral disease that mainly affects the respiratory, gastrointestinal, and central nervous systems and is caused by canine distemper virus (CDV) [1]. Vaccination remains the principal means of controlling the disease. Alteration of the antigenicity of new live modified vaccines has greatly helped to control CD [2, 3]. Although the use of vaccines has reduced the incidence of CD over the past decades, it has been reported that CD (as a sporadic, endemic, or outbreak disease) can also affect vaccinated animals [4–8].

Analysis of CDV strains from various animal samples has demonstrated an important relationship with the H gene/glycoprotein, which has changed by genetic/antigenic drift. As the key protein for CDV, H is used for attachment to cell receptors as the first step of infection and mediates adequate host immune response [9]. The H protein is considered to have the highest antigenic variation and can reflect genetic changes in comparative studies of CDV strains [10–13]. This variation may affect neutralization-related sites with disruption of important epitopes. Analysis of CDV strains from different animal species and geographical settings has revealed that the geographic pattern is an important factor in the genetic/antigenic drift affecting the H gene/glycoprotein of CDV [14–20]. Therefore, the H gene may be used for identification and phylogenetic classification of CDV strains, which have been identified into seven major genetic lineages, namely America-1 and -2, Asia-1 and -2, Arctic-like, Europe, and wild-life [21, 22], as well as an indication of the antigenic response of the virus.

Three other proteins, the nucleocapsid (N) protein, the phosphoprotein (P) protein, and the fusion (F) protein, also have important roles for CDV and could provide additional sources of antigenic variability among strains. The N protein has immunosuppressive properties and is the major component of the CDV virion. The N-terminal domain of the N protein is generally well conserved, while the C-terminal end is poorly conserved and is considered hypervariable. The C-terminal tail of the N protein also contains the majority of its phosphorylation sites and antigenic sites [23, 24]. During active infection, antibodies made against the N protein in the host are predominant and account for most of the complement-fixing antibody [25, 26]. The P protein is relatively well conserved and plays a vital role in transcription and replication [27]. This protein is an essential component of the viral RNA phosphoprotein complex (vRNAP) [28] and also function as a chaperone for the N protein. The F protein is a type I integral membrane protein that mediates viral penetration by fusion between the virion envelope and the host cell plasma membrane at neutral pH. It is synthesized as an inactive precursor, F0, and must be proteolytically cleaved to produce the functionally active fusion protein, which consists of disulfide-linked F1 and F2 polypeptides [29]. Like the H protein, the F protein has high antigenic variation.

In this study, the wild-type CDV-TM-CC strain was isolated from the spleen of a 1-year-old Tibetan Mastiff that developed clinical signs of CD after having received all standard vaccines. To determine whether this occurrence may be explained by variations in specific nucleotide or amino acid residues of the CDV circulating in China, we sought to genetically characterize the CDV-TM-CC strain.

Table 1 Primers used for amplification of the full-length genome of the CDV-TM-CC strain

Primer	Sequence (5′ → 3′)	Position
A1F	ACCAGAAAAAGTTGGCTATG	1
AIR	GCFGTTFCACCCATCFGTTG	1036
A2F	GAACAAGCCTAGAAFTGCTG	842
A2R	TGAGGGCTTTGAGGCATTCC	1858
A3F	FTCAAGACCAGFGFTACAFC	1683
A3R	GCAGGTFGGCGGACATFCTC	2627
A4F	TCTGCAGTTCCCACGCAATC	2427
A4R	CCTGCAGTTCGCCTTACCAC	3748
A5F	GFGAFAGAFCCAGGACTCGG	3540
A5R	CTCACCTAATTCFGCFFTGG	5495
A6F	FAFGGTGCATFGGAATAGCC	5296
A6R	CCTACATCTAACCTCTCAAG	6649
AT	GTGTAAGTGTTATAGCACAG	6452
A7R	AAGGGTTCCCATGACGTTFG	8271
ASF	TGGATCAAGTTGAAGAGGTG	8061
ASR	CGTGCATACTCTAAAATAGC	9127
A9F	GGACTTAGGTATGATGACTG	8928
A9R	CCFGAGAGAGAAGGFGATGT	9932
A10F	FTAAFGTGTTAGTGTCFCGG	9631
A10R	CAFFFFAGCGAATAGCCTCC	10676
A11F	ATTCCTCAGGTACAACCCAC	10478
A11R	GGTCAAACCCTTTCTCAATG	11682
A12F	CCACAGAATACTTTGTAGCC	11428
A12R	AFCGAFGAGCGGAFCAFGCC	12413
A13F	CTTTAGCCGCTTTCTTGATG	12160
A13R	ACAGTCTAAATAAGTGCTCG	13080
A14F	CCGCTCATCGATCAAGACTC	12402
A14R	CTATTGATATCATCATCGCC	13438
A15F	GACTTTATAACCAGAGTCAC	13261
A15R	ACATGCAGFAGAGFTGATCC	14321
A16F	AFCACTGAFGCFGFTGGATG	14100
A16R	ACCAGACAAAGCTGGGTATG	15690
B1F	ACCAGAAAAAGTTGGCT	1
B1R	TCCGTTGTCTTGGATGCT	1315
B2F	GCTCTGGAGTTATGCTAT	1100
B2R	TTAACCTCTTCACCGCTG	2162
B3F	CCTACCATGTCAGCAAAG	1817
B3R	AAGCGATAGTGAAAGCAG	3378
B4F	TACCGCACCTTCCAAAGC	3168
B4R	GCACTCTAATCTCCATAG	4331
B5F	CGATGTACTGGTAAGATG	4212
B5R	ACCTAATTCTGCTFTGGT	5492
B6F	AATCCAATGCAACCAACT	5251
B6R	AGCAGGGCCTAAGGCAAC	6623
B7F	AGTCCTGATAAATTGCTG	6489
B7R	GACTCGAATCTTFTGCG	7828
BSF	TGGAGCTACTACTFCA	7645
B8R	ATACGGACTAAATTCTCAAC	8874

Table 1 continued

Primer	Sequence (5′ → 3′)	Position
B9F	AGATGTCCTTACTGAGTC	8584
B9R	GAGAGAGAAGGTGATGTC	9929
B10F	ACTTTGCACGTCAGAATC	9786
B10R	CCTAGCTATACTTFTCAG	11585
B11F	TTGGCAGCACATGAAAGC	11298
BUR	GACCCTTCGCCAACTTAG	12483
B12F	GAGGGAGGCAATTGCTGG	12242
B12R	TGGGAGTCAAGAGAGGGT	13705
B13F	GGTCAATGTGCTGCAATC	13491
B13R	TCTCTCGTTGTCCAGTTC	14499
B14F	TACGTCTGAAGAACATGG	14363
B14R	ACCAGACAAAGCTGGGTA	15690

Two sets of forward (F) and reverse (R) primers were used (sets A and B) for verification of the sequence in redundant amplicons

Materials and methods

Cells and viruses

VerodogSLAM cells constitutively expressing the CDV receptor dog signaling lymphocyte activation molecule (SLAM) were cultured in Dulbecco's modified Eagle medium (DMEM; Gibco) supplemented with 10 % heat-inactivated fetal bovine serum (FBS) with an additional 8 μg of G418 per ml.

The wild-type CDV-TM-CC strain was originally isolated from spleen homogenate (10 % w/v suspension) from a Tibetan Mastiff that succumbed to naturally infection. Virus was propagated in VerodogSLAM cells and stored at −80 °C.

RNA preparation and RT-PCR

Total RNA was prepared from VerodogSLAM cells infected with CDV-TM-CC according to the manufacturer's instructions (Total RNA Kit I, OMEGA). The reverse transcription reactions were performed using M-MLV Reverse Transcriptase (Invitrogen) with oligo d(T) and random primers. According to the complete consensus genomic sequence of CDV (GenBank), two sets of primers were designed to amplify the entire genome (Oligo6.0 design software), as shown in Table 1. Sequences were assembled and compared using DNA sequence analysis software (DNAStar), and the complete consensus genomic sequence was determined. PCR amplification was carried out using Phusion High-Fidelity DNA Polymerase (New England BioLabs). Clones (amplicons emcompassing the full-length CDV-TM-CC genome) were obtained

Table 2 Nucleotide sequence accession numbers of CDV strains used in this study

Strain	Country	Organ	GenBank	Reference
00-2601	USA	Raccoon brain	AY443350.1	[40]
01-2689	USA	Raccoon brain	AY649446.1	Unpublished
007Lm	Japan	Canine pulmonary	AB474397.1	Unpublished
011C	Japan	Canine cerebellum	AB476401.1	[41]
50Con	Japan	Canine conjunctiva	AB476402.1	[41]
55L	Japan	Canine lung	AB475099.1	[41]
98-2646	USA	Raccoon brain	AY542312.2	Unpublished
98-2654	USA	Raccoon brain	AY466011.2	[40]
5804P	USA	–	AY386316.1	[42]
164071	USA	Blood lymphocytes	EU716337.1	Unpublished
A75/17	Switzerland	–	AF164967.1	Unpublished
CDV3	China	Mink	EU726268.1	Unpublished
CYN07-hv	Japan	Macaca fascicularis	AB687721.2	[7, 43]
Hebei	China	Mink lung	KC427278.1	Unpublished
HLJ1-06	China	Fox	JX681125.1	[44]
M25CR	Japan	Canine cerebrum	AB475097.1	[41]
Onderstepoort	USA	–	AF014953.1	[34, 45]
Phoca/Caspoian/2007	Kazakhstan	Vero cells	HM046486.1	Unpublished
Shuskiy	Kazakhstan	–	HM063009.1	Unpublished
Snyder Hill	Canada	Cerebrospinal fluid	AF259552.1	[16]

A GenBank number is provided for each of the strains of CDV that were compared with CDV-TM-CC in this study. The geographical location of strain isolation and the species/organ of isolation are also indicated, as well as the clade into which the strains are categorized

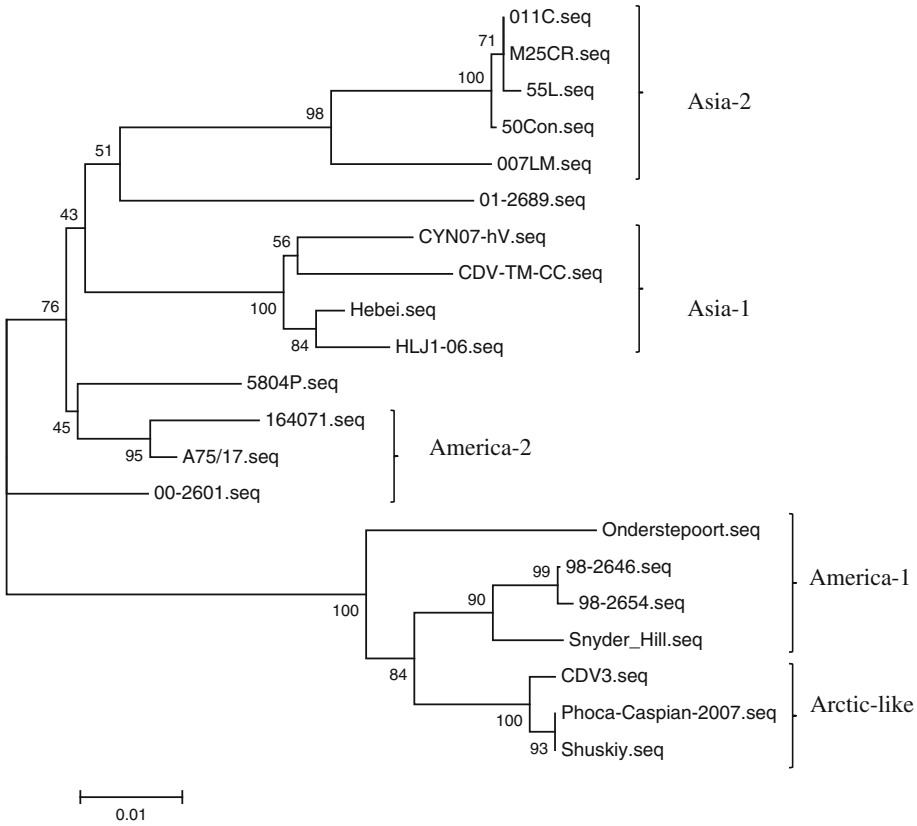


Fig. 1 Phylogenetic relationship between different CDV strains on the basis of the amino acid alignment of the H protein of the CDV-TM-CC strain in comparison with other strains from GenBank. Results show that CDV-TM-CC is an Asia-1 strain

from thirty RT-PCR reactions using CDV-specific oligonucleotides.

Sequence alignment and phylogenetic analysis

To genetically characterize the CDV-TM-CC strain, the deduced amino acid sequence was compared to F and H gene fragments of the variant field isolates shown in Table 2. A phylogenetic tree was constructed based on the deduced amino acid sequences in supplementary Table 1 using MEGA 5.0, and multiple sequence alignment was carried out using ClustalW. Statistical significance of the phylogeny was estimated by bootstrap analysis over a 1,000 pseudoreplicate data set.

Results

Phylogenetic analysis of the nucleotide and deduced amino acid sequences of the H gene of CDV-TM-CC

The wild-type CDV-TM-CC strain was isolated from the spleen of a 1-year-old Tibetan Mastiff in Jilin province that had succumbed to CD after having received all standard vaccines (6 weeks first immunization, 8 weeks second immunization, 10 weeks third immunization with Distemper, adenovirus type 2, parvovirus, parainfluenza quadruple vaccine; Canine coronavirus disease killed virus vaccine portion, USA). The virus was propagated in VerodogSLAM cells and the virulence of the strain was confirmed (data not shown). To identify sequence features that may explain the failure of the vaccine strain to protect the dog against CD, we sequenced the entire genome, using two sets of overlapping primers (Table 1).

Within the CDV genome, the H gene is a major causative disease determinant and also has one of the highest rates of mutation. Consequently, the phylogenetic relationship of CDV strains is often based on the deduced amino acid sequence of the H protein. The H gene of the CVT-TM-CC strain has 1,824 nucleotides and the inferred protein sequence has 607 amino acids, similar to the other CDV strains. Amino acid analysis of the H protein from CDV-TM-CC and 20 other CDV strains in GenBank (Table 2) identified seven clades of CDV strains (America-1, America-2, Asia-1, Asia-2, Europe, Arctic-like, and Europe wildlife). CDV-TM-CC was classified into the Asia-1 group with the strains CYN07-hV (Japan), Hebei (China), and HLJ-06 (China) (Fig. 1). CDV-TM-CC has high conservation of both the H gene nucleotide sequence (97.4 % nt identity with CYN07-hV, 98.1 % with HLJ1-06, and 98.2 % with Hebei) and the amino acid sequence (97.4 % aa identity with CYN07-hV, 97.4 % with HLJ1-06, and

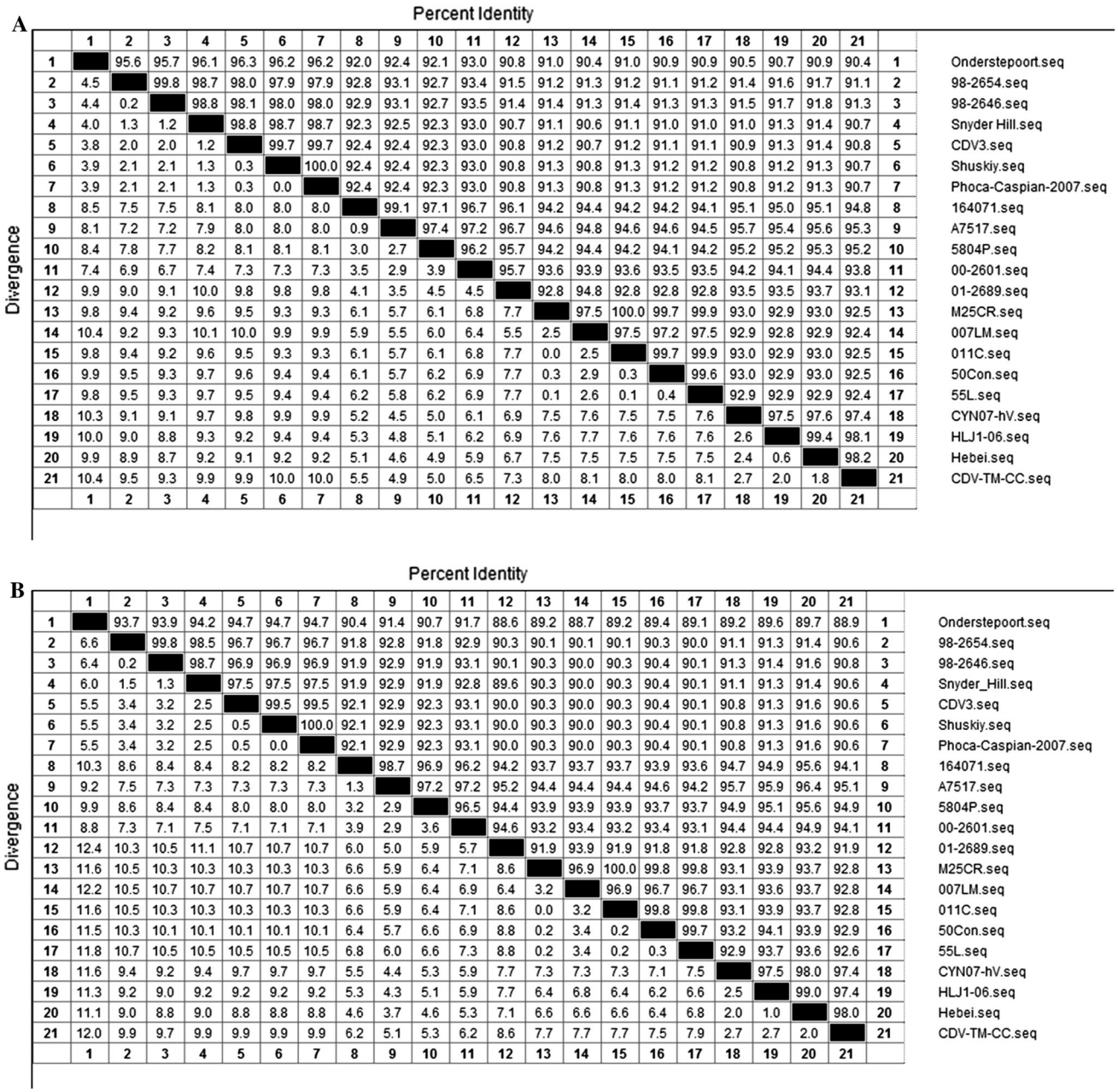


Fig. 2 Alignment of the nucleotide and amino acid sequences of H protein of different CDV strains. The percent identity to and divergence from the nucleotide sequence (**a**) and protein sequence (**b**) of CDV-TM-CC is shown

98.0 % with Hebei). On the other hand, the America-1 group (strains Onderstepoort, 98-2654, 98-2646, and Snyder Hill) was less similar to CDV-TM-CC in both the nucleotide sequence (90.4 % nt identity with Onderstepoort, 91.1 % nt with 98-2654, 91.3 % with 98-2646, and 90.7 % with Snyder Hill) and the amino acid sequence (88.9 % identity with Onderstepoort, 90.6 % with 98-2654, 90.8 % with 98-2646, and 90.6 % with Snyder Hill) (Fig. 2a, b).

N-linked glycosylation sites of the H protein of the CDV-TM-CC strain

Glycosylation is an important factor in determining the antigenicity of many proteins [30]. Prediction of the glycosylation sites of the H gene (http://www.cbs.dtu.dk/services/NetNGlyc/, NetNGlyc 1.0 Servera) identified a total of eight potential glycosylation sites at positions ^{19}NSS, ^{149}NFT, ^{309}NGS, ^{391}NQT, ^{422}NIS, ^{456}NGT, ^{584}NIT, and

^{587}NST (Fig. 3). Among them, six glycosylation sites (positions 19–21, 149–151, 391–393, 422–424, 456–458, and 587–590) were conserved for all of the CDV strains. Notably, the 309–311 N-glycosylation site is specific for virulent strains [14, 18] with the exception of A75/17. The 584 N-glycosylation site is specific for the Asian-1 strains, suggesting that it was acquired later [18, 20]. CDV-TM-CC has both of these predicted glycosylation sites, which could explain its virulence properties.

Phylogenetic analyses of the amino acid sequence of the N and P proteins

To determine whether the conservation of CDV-TM-CC also extends to other proteins within the virus, we assessed the similarity of the N and P proteins. Consistent with the results for the H protein, the homology of the deduced CDV-TM-CC amino acid sequence of the N protein to the Asia-1 strains (CYN07-hV, HLJ1-06, and Hebei) was high with 98.7-98.9 % identity, as shown in Fig. 4. The N protein sequence of CDV-TM-CC also showed 98.1 % identity with the Asia-2 group (strains M25CR, 007Lm, 011C, 50Con, and 55L), and 97.5 % identify with the Onderstepoort strain. Moreover, CDV-TM-CC had high similarity (98.5, 98.7, and 97.9 % identity) with wild-type strains 164071, A75/17, and 5804P. The lowest homology of the CDV-TM-CC N protein sequence (96.6–96.8 % aa identity) was found with Arctic-like strains CDV3, Shuskiy, and Phoca-Caspian-2007. This relatively high similarity between the N protein of CDV-TM-CC and other CDV strains is consistent with the generally high conservation among N proteins.

The phylogenetic relationship of CDV-TM-CC based on the deduced amino acid sequence of the P protein was also analyzed (Fig. 5). Similar to the results for the H protein, CDV-TM-CC classified into the Asia-1 group, but was in a separate branch from the classical Onderstepoort vaccine strains. These results verify the classification of CDV-TM-CC as an Asia-1 group virus.

Phylogenetic analysis of the signal peptide region of the F protein

The signal peptide is a short amino acid sequence at the N-terminus of the majority of newly synthesized proteins that are destined towards the secretory pathway and is a highly divergent region [31]. Analysis of the 1–135 aa signal peptide region of the F protein of CDV-TM-CC demonstrated the same set of amino acid variations in comparison with the Onderstepoort strain as for the other Asia-1 strains (CYN07-hV, HLJ1-06, and Hebei): ^{8}S/^{8}K, ^{11}T/^{11}P, ^{19}R/^{19}P, ^{81}G/^{81}R, ^{101}Q/^{101}W, ^{102}I/^{102}F, and ^{112}S/^{112}A (Fig. 6). Among the Asia-2 strains (M25CR, 007Lm, 011C, 50Con, and 55L), variations in comparison with the Onderstepoort strain were found in ^{30}T/^{30}S, ^{53}S/^{53}A, ^{55}R/^{55}W, ^{59}S/^{59}Y, ^{62}N/^{62}K, ^{99}R/^{99}K, ^{110}I/^{110}V, and ^{111}N/^{111}K. Additionally, both the Asia-1 and Asia-2 strains had clade-specific amino acid variation in ^{21}P/^{21}Q. Moreover, the CDV-TM-CC strain had characteristic additional variations in ^{107}P/^{107}Y and ^{116}C/^{116}Y. Therefore, the signal peptide region of CDV-TM-CC has both Asia group-specific and individual variations.

Phylogenetic analysis of the F2 and F1 regions

Among the CDV strains, amino acid variation was also found in ^{208}K/^{208}N in the F2 region (aa 136–224) for the Asia-1 group. Generally, there was high conservation within the hydrophobic fusion peptide (FP) domain at the N-terminus of the membrane anchored F1 subunit, with the exception of ^{233}A/^{233}V in the 98-2654 and 98-2646 strains (Fig. 6). Amino acid variations between the Asia-1 and Asia-2 groups were also found in a region between the helical bundles (HB) and heptad repeats B (HRB) at ^{394}V/^{394}S, ^{429}R/^{429}K, and ^{466}L/^{466}I; within the trans-membrane (TM) domain at ^{627}C/^{627}Y, ^{634}Q/^{634}R, and ^{637}H/^{637}F; and within the cytoplasmic tail (CT) domain, at ^{656}R/^{656}K. Among the Asia group strains, the HRA (aa 250-307) and HB (aa 328-374) domains were highly conserved, with the exception of a ^{280}Q/^{280}A variation in the HRA domain. Likewise, the amino acids were highly conserved in the HRB (aa 557-601) domain in all CDV strains except for Hebei (^{583}D/^{583}N) and 5804P (^{587}V/^{587}I). Common amino acid changes in other regions of CDV strains in comparison to the Onderstepoort strain were found at ^{317}K/^{317}R and ^{556}S/^{556}G.

N-linked glycosylation sites and cysteine residues of the F protein

The potential N-glycosylation sites (N-X-S/T) of the F protein were highly conserved at ^{141}NLS, ^{173}NVS, ^{179}NCT, and ^{517}NQS in the F1 region among all CDV strains as reported previously [32–34] (Fig. 6). Moreover, the Asia-1 group (strains CYN07-hV, HLJ1-06 Hebei, and CDV-TM-CC) had specific potential N-glycosylation sites at ^{62}NRT and ^{108}NAT in the signal peptide region, with the exception of the CDV-TM-CC strain, which had the sequence ^{62}NKT. Five of these six potential glycosylation sites of the CDV-TM-CC strain were at the same positions within the known virulent CDV strains (A75/17, 5804P and 164071) at aa 108–110, 141–143, 173–175, 179–181, and 517–519, whereas ^{62}NKT was unique for CDV-TM-CC, and ^{62}NRT and ^{38}NIT were unique for 5804P.

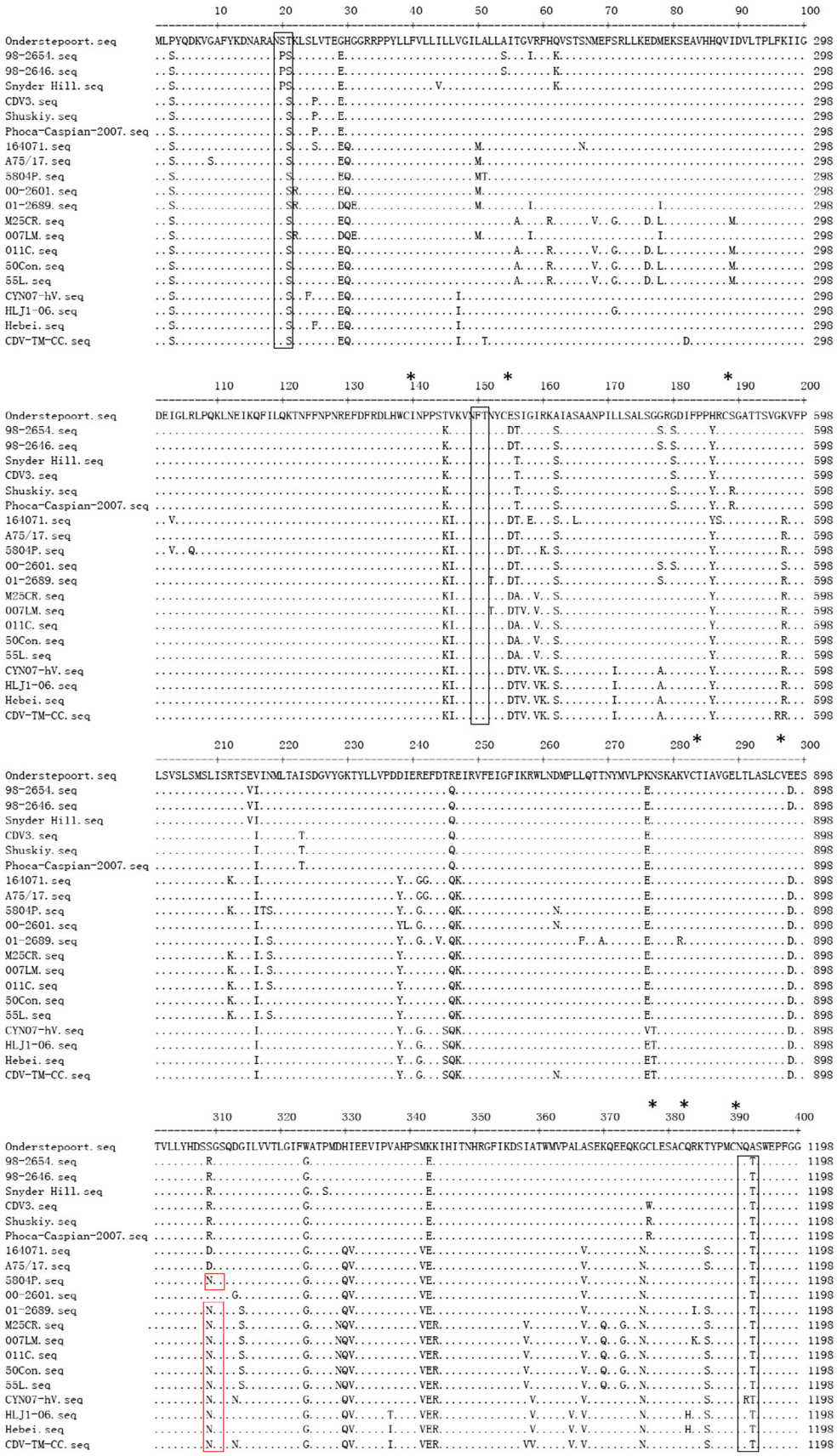
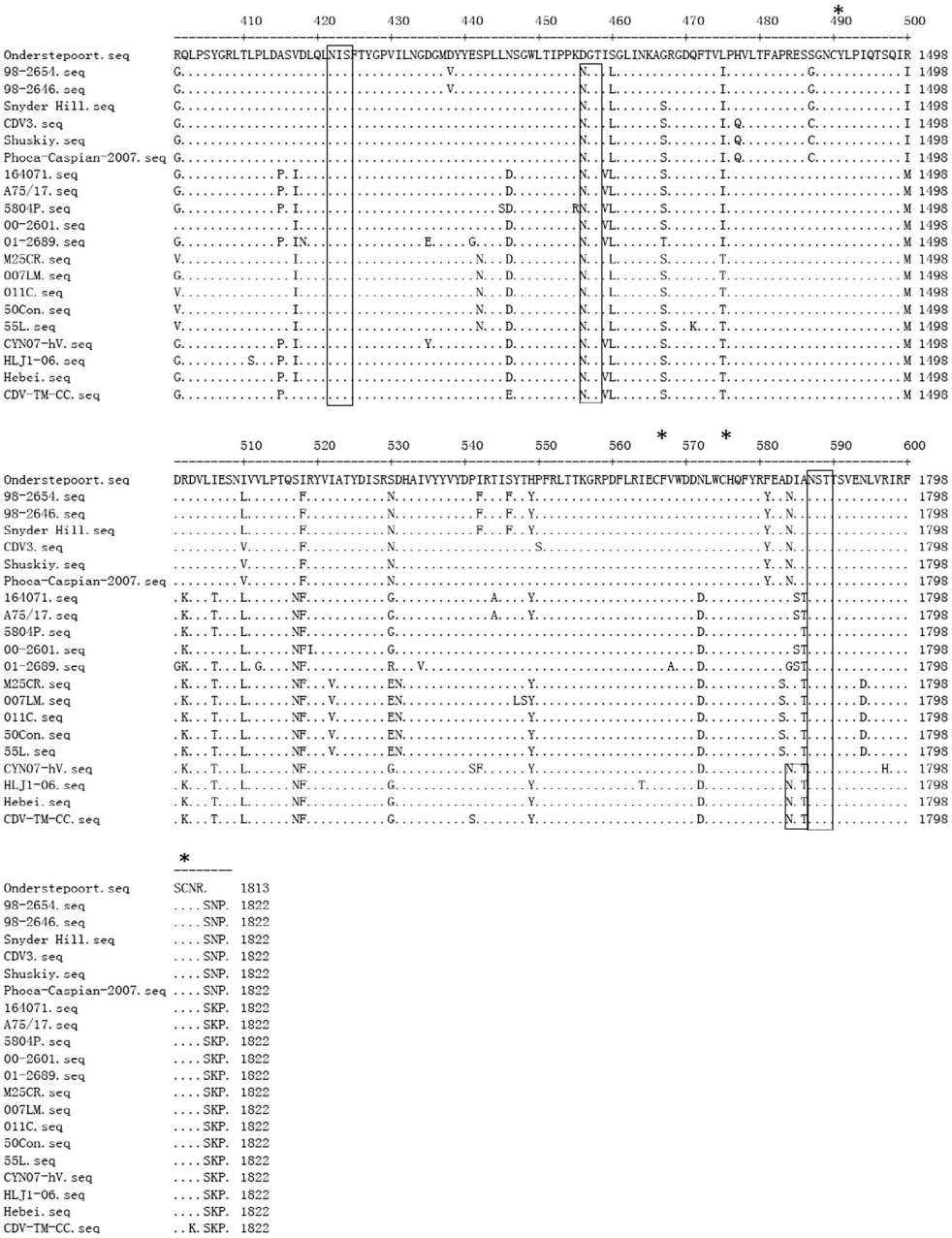


Fig. 3 Conservation of the amino acid sequence of the H protein. The protein sequences of each of the strains in this study are compared to the sequence of the vaccine strain, Onderstepoort. Potential N-link glycosylation sites (N-X-S/T) are *boxed*. Among them, the 309–311 N-glycosylation site, which is specific for the wild-type strain, is *boxed* in *red*, and the 584–586 N-glycosylation site, acquired in the Asian-1 strains, is *boxed* in *blue*. The position of cysteine residues is indicated by *asterisks*

Fig. 3 continued



Cysteine is an α-amino acid that plays an important role in intramolecular disulfide bond formation and the steric structure of proteins. In the F protein of CDV-TM-CC, a total of 16 cysteine residues were detected. Among them, 14 residues (aa 123, 132, 180, 307, 446, 455, 470, 478, 502, 507, 509, 531, 628, and 629) were located at identical positions in all CDV strains; however, several amino acid(s) were characteristic to individual strain(s), such as ^{67}R/^{67}C in the America group (strains Onderstepoort, 98-2654, 98-2646, Snyder Hill, CDV3, Shuskiy and Phoca-Caspian-2007) and ^{116}Y/^{116}C in CDV-TM-CC. The presence of amino acid variations, as well as specific N-glycosylation sites and cysteine residues within the F protein, could affect the immune response to CDV-TM-CC.

Discussion

Improved vaccination has reduced the frequency and magnitude of CD [35]. Distemper vaccination failures are uncommon, but outbreaks of CD continue to occur among vaccinated individuals and populations [4, 5, 36, 37]. The most common factor in CD occurrence is a lack of

Percent Identity

Divergence

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
1		97.9	97.9	99.0	97.9	98.1	97.9	98.1	98.3	97.7	97.9	98.5	97.9	97.9	97.9	97.9	97.9	97.9	97.7	97.7	97.5	1	Onderstepoort.seq
2	1.9		99.8	97.9	96.9	97.1	96.9	98.7	98.7	98.9	99.0	98.9	97.7	97.7	97.7	97.7	97.7	97.9	98.1	98.1	97.9	2	98-2654.seq
3	1.9	0.0		97.9	96.9	97.1	96.9	98.7	98.7	98.9	99.0	98.9	97.7	97.7	97.7	97.7	97.7	97.9	98.1	98.1	97.9	3	98-2646.seq
4	0.8	1.9	1.9		97.9	98.1	97.9	98.1	98.3	97.7	97.9	98.5	97.9	97.9	97.9	97.9	97.9	97.5	97.7	97.7	97.5	4	Snyder Hill.seq
5	1.9	2.9	2.9	1.9		99.6	99.4	97.1	97.3	96.8	96.9	97.5	96.9	96.9	96.9	96.9	96.9	96.6	96.8	96.8	96.6	5	CDV3.seq
6	1.7	2.7	2.7	1.7	0.2		99.6	97.3	97.5	96.9	97.1	97.7	97.1	97.1	97.1	97.1	97.1	96.8	96.9	96.9	96.8	6	Shuskiy.seq
7	1.9	2.9	2.9	1.9	0.4	0.2		97.1	97.3	96.8	96.9	97.5	96.9	96.9	96.9	96.9	96.9	96.6	96.8	96.8	96.6	7	Phoca-Caspian-2007.seq
8	1.7	1.2	1.2	1.7	2.7	2.5	2.7		99.2	98.5	98.9	99.4	98.3	98.3	98.3	98.3	98.3	98.5	98.7	98.7	98.5	8	164071.seq
9	1.5	1.2	1.2	1.5	2.5	2.3	2.5	0.6		98.7	99.0	99.6	98.5	98.5	98.5	98.5	98.5	98.7	98.9	98.9	98.7	9	A7517.seq
10	2.1	1.0	1.0	2.1	3.1	2.9	3.1	1.4	1.2		99.4	98.9	97.7	97.7	97.7	97.7	97.7	97.9	98.3	98.1	97.9	10	5804P.seq
11	1.9	0.8	0.8	1.9	2.9	2.7	2.9	1.0	0.8	0.4		99.2	98.1	98.1	98.1	98.1	98.1	98.3	98.7	98.5	98.3	11	00-2601.seq
12	1.4	1.0	1.0	1.4	2.3	2.1	2.3	0.4	0.2	1.0	0.6		98.7	98.7	98.7	98.7	98.7	98.9	99.0	99.0	98.9	12	01-2689.seq
13	1.9	2.1	2.1	1.9	2.9	2.7	2.9	1.5	1.4	2.1	1.7	1.2		99.8	99.8	99.8	99.8	97.7	97.9	97.9	98.1	13	M25CR.seq
14	1.9	2.1	2.1	1.9	2.9	2.7	2.9	1.5	1.4	2.1	1.7	1.2	0.0		99.8	99.8	99.8	97.7	97.9	97.9	98.1	14	007LM.seq
15	1.9	2.1	2.1	1.9	2.9	2.7	2.9	1.5	1.4	2.1	1.7	1.2	0.0	0.0		99.8	99.8	97.7	97.9	97.9	98.1	15	011C.seq
16	1.9	2.1	2.1	1.9	2.9	2.7	2.9	1.5	1.4	2.1	1.7	1.2	0.0	0.0	0.0		99.8	97.7	97.9	97.9	98.1	16	50Con.seq
17	1.9	2.1	2.1	1.9	2.9	2.7	2.9	1.5	1.4	2.1	1.7	1.2	0.0	0.0	0.0	0.0		97.7	97.9	97.9	98.1	17	55L.seq
18	1.9	1.9	1.9	2.3	3.3	3.1	3.3	1.4	1.2	1.9	1.5	1.0	2.1	2.1	2.1	2.1	2.1		98.9	98.9	98.7	18	CYN07-hV.seq
19	2.1	1.7	1.7	2.1	3.1	2.9	3.1	1.2	1.0	1.5	1.2	0.8	1.9	1.9	1.9	1.9	1.9	1.0		99.0	98.9	19	HLJ1-06.seq
20	2.1	1.7	1.7	2.1	3.1	2.9	3.1	1.2	1.0	1.7	1.4	0.8	1.9	1.9	1.9	1.9	1.9	1.0	0.8		98.9	20	Hebei.seq
21	2.3	1.9	1.9	2.3	3.3	3.1	3.3	1.4	1.2	1.9	1.5	1.0	1.7	1.7	1.7	1.7	1.7	1.2	1.0	1.0		21	CDV-TM-CC.seq
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		

Fig. 4 Alignment of the amino acid sequences of the N protein of different CDV strains. The percent identity to and divergence from the protein sequence of CDV-TM-CC is shown

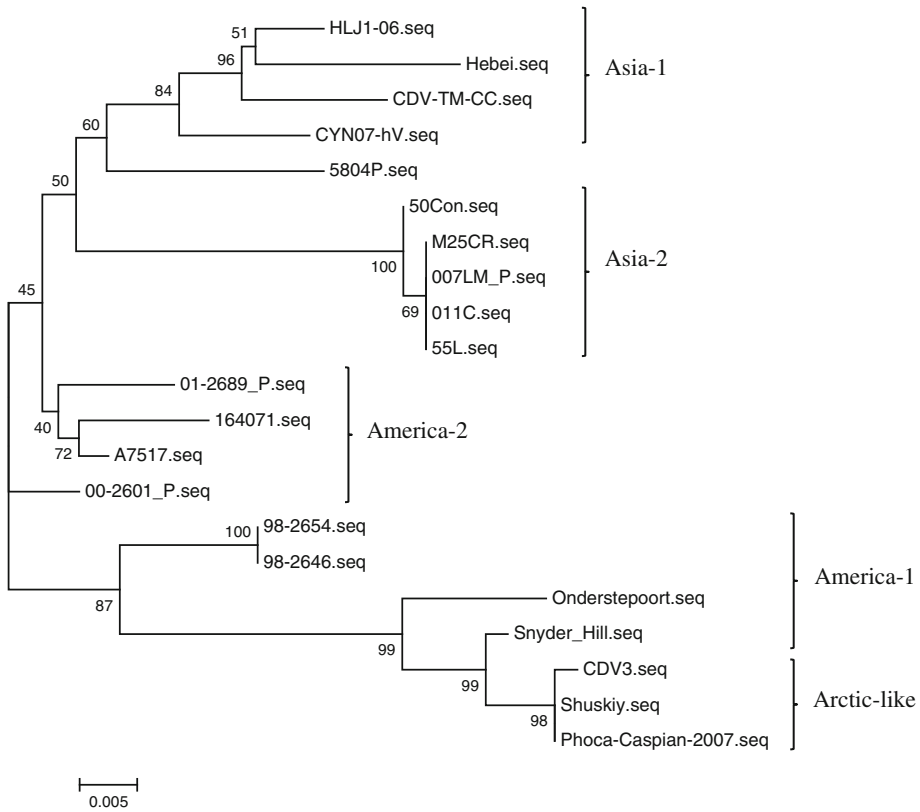


Fig. 5 Phylogenetic analysis based on the amino acid sequences of the P protein. Results verify the categorization of CDV-TM-CC as an Asia-1 strain

appropriate vaccination, while the residual virulence of the vaccine strains is another possibility. Furthermore, genetic/antigenic drift of wild-type CDV strains driven mainly by geographic variation is an important factor in vaccine failure. In this study, we isolated a new virulent CDV strain, CDV-TM-CC, from a Tibetan Mastiff that died due to natural infection in Jilin province. According to the complete consensus genomic sequence of CDV, at least

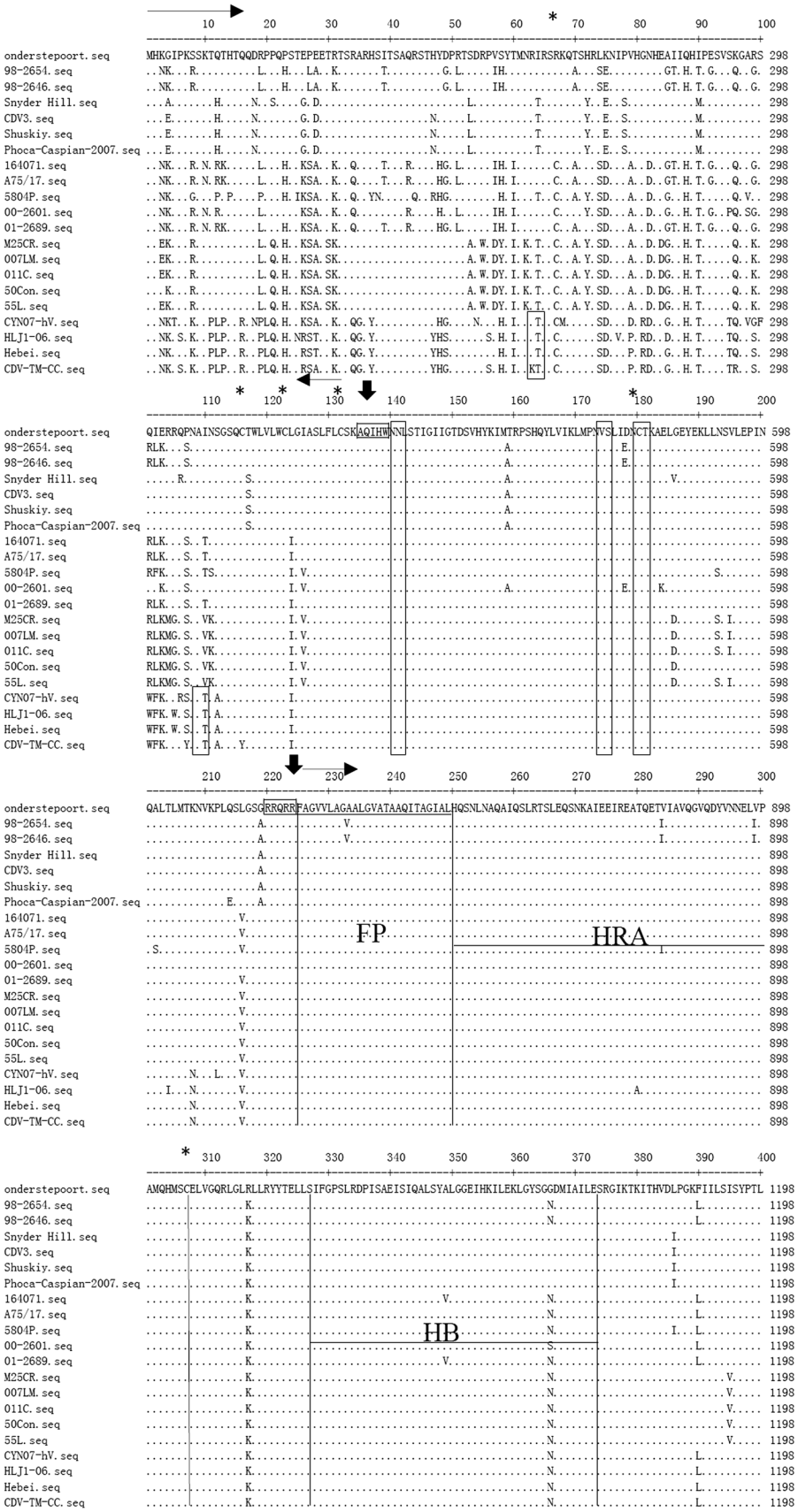
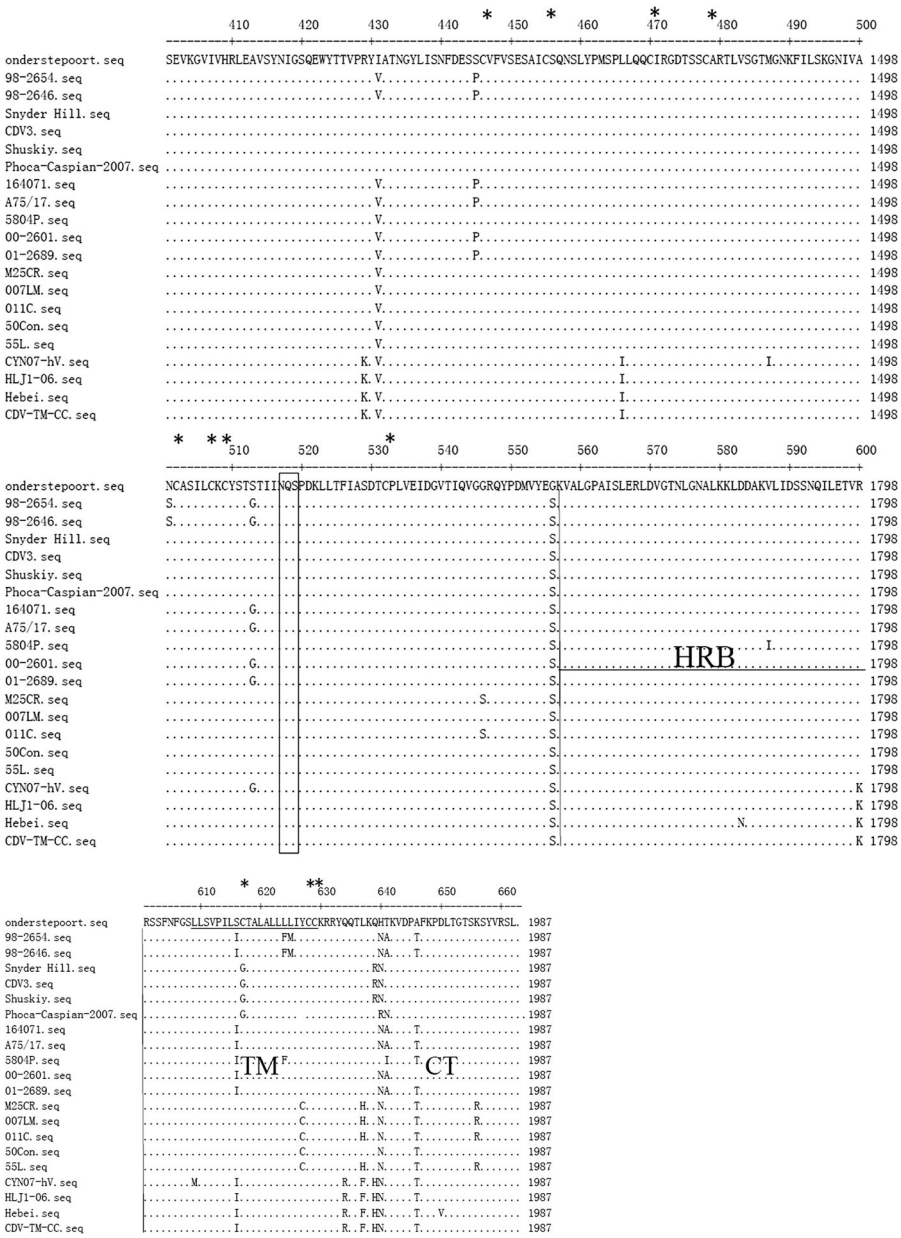


Fig. 6 Alignment of the amino acid sequence of the F protein of different CDV strains. The sequences of the F proteins of each of the strains in this study are compared to the sequence of the vaccine strain, Onderstepoort. The potential N-linked glycosylation sites (N-X-S/T) are *boxed*. The positions of the fusion peptide (FP), heptad repeats (HRA and HRB), helical bundles (HB), transmembrane (TM) and cytoplasmic tail (CT) regions are indicated. The positions of cysteine residues are marked by *asterisks*, and cleavage sites are marked by *arrows*

Fig. 6 continued



two sets of primers were designed within different CDV genes, leading to the identification of the complete CDV-TM-CC sequence.

The H protein, a major structural protein of CDV, mediates host selection and pathogenicity, and the rate of genetic variation for its gene is greater than for other genes. With geographically distinct lineages, many studies have demonstrated that phylogenetic analysis can be carried out in accordance with the deduced amino acid sequences of the H protein [14, 18, 21, 38]. In this study, phylogenetic analysis based on the H protein identified seven clades of CDV strains (America-1, America-2, Asia-1, Asia-2, Europe, Arctic-like, and Europe wild-life), and CDV-TM-CC was classified into the Asia-1 group, with the highest identity to the Chinese strains, HLJ1-6 and Hebei, and the Japanese strain, CYN07-hV. Potential N-glycosylation sites may differ for the H protein of the wild-type and vaccine strains of CDV. Usually, only 4–7 potential sites are found within vaccine strains (such as Onderstepoort), in comparison with 8–9 sites in wild-type CDV strains (for example, 5804P). In particular, the 309–311 N-glycosylation site, which is specific for the wild-type strain [14, 18], is suggestive of the pathogenicity of the CDV-TM-CC strain. Furthermore, the 584–586 N-glycosylation site has been acquired in the Asian-1 strains [18, 20]. Further study may determine whether these differences in glycosylation

may contribute to the failure of the vaccine to protect against virulent strains of CDV.

The N protein is a highly conserved immunogenic protein that can elicit cellular and humoral immunity [39]. Based on sequence differences between the gene of the wild strains and vaccine strain, the N protein may affect the seroprotection rate of the host and lead to immune failure. Like the H protein, the N protein of CDV-TM-CC showed the highest homology with the Asia-1 group. High homology was also observed with the Asia-2 group (strains M25CR, 007Lm, 011C, 50Con, and 55L) and wild-type strains (164071, A75/17, and 5804P). Moreover, the lowest homology was found between CDV-TM-CC and the Onderstepoort strain. Variation in the immunodominant epitope of the virus may change the structure, and therefore, we can speculate that the T cell-mediated immune response may be altered by variations in this protein. The P gene is extremely well conserved and, therefore, is particularly important in the phylogenetic classification. Based on the phylogenetic relationship of the deduced amino acid sequence of the P protein, CDV-TM-CC was also classified into the Asia-1 group. These results highlight the importance of considering the geographical setting to control the occurrence of the disease in a more efficient manner.

The F protein is a surface glycoprotein that mediates viral entry into the host cell by fusion of the virion envelope and the host cell plasma membrane at a neutral pH. Within the F protein, the signal peptide region (aa 1–135) has the lowest amino acid homology, especially at positions 13–37 and 72–112. However, our analysis shows that the signal peptide region is relatively well conserved among the Asia-1 group, except for specific individual amino acids, indicating that the signal peptide of the F protein is geographically distinct. In addition, three amino acids specific to the CDV-TM-CC strain (^{62}K, ^{107}Y, and ^{116}Y) are located in the signal peptide region. The previous study reported that the amino acids ^{208}K and ^{216}L are specific for the CDV vaccine strains; however, we also found ^{208}K in the wild-type strains in the America group (A75/17, 164071, and 5804P) and Asia-2 group (011C, M25CR, 55L,50Con, and 007Lm). The F protein of the CDV-TM-CC strain has six potential glycosylation sites. Among them, differences were found to reside mainly in the signal peptide region, but no clear rule could obviously explain the differences in the wide-type and vaccine strains or the geographical variation, including the occurrence of a strain-specific site (62–64 NKT) for CDV-TM-CC. Four additional potential glycosylation sites were recognized at positions 141–143, 173–175, 179–181 in the F2 region and 517–519 in the F1 region, as reported previously [32–34]. The 108–110 N-glycosylation site is specific for the wild-type strains (5804P, A75/17, and 164071) and the Asia-1 group (Hebei, HLJ1-06, and CYN07-hV), and may be another important factor in vaccination failure. The fusion peptide (FP) domain also was found to be highly conserved among all CDV strains, except for ^{233}A/^{233}V in 98-2654 and 98-2646. In short, the genetic/antigenic drift observed in the currently circulating CDV strains should be considered as a possible factor leading to the resurgence of CD cases. Analysis of CDV strains detected globally and from a variety of host species will provide a more in-depth understanding of the global ecology of CDV and will provide the basis for the improvement of current CDV vaccines.

Conclusion

The wild-type CDV-TM-CC strain, originally isolated from spleen homogenate from a fully vaccinated Tibetan Mastiff in China, was classified into the Asia-1 group cluster of CDV strains based on the sequence of its H protein and verified by the sequence of its P protein. Variations in specific amino acid residues, N-glycosylation sites, and cysteine residues throughout the CDV-TM-CC genome may explain the failure of the dog to mount vaccine-mediated protection against CD. These results provide the foundations for the global improvement in current CDV vaccines.

Acknowledgments This work was supported by Ecology of Zoonoses and Research of Infection and Immunity mechanisms (2012CB722501).

References

1. M. Vandevelde, A. Zurbriggen, Vet. Microbiol. **44**, 271–280 (1995)
2. M.J. Appel, B.A. Summers, Vet. Microbiol. **44**, 187–191 (1995)
3. M.J. Appel, *Virus Infections of Carnivores* (Elsevier Science Publishers B V, Amsterdam, 1987), pp. 133–159
4. O. Glardon, R. Stockli, Schweiz. Arch. Tierheilkd. **127**, 707–716 (1985)
5. M. Blixenkrone-Moller, V. Svansson, P. Have, C. Orvell, M. Appel, I.R. Pedersen, H.H. Dietz, P. Henriksen, Vet. Microbiol. **37**, 163–173 (1993)
6. R.K. Maes, A.G. Wise, S.D. Fitzgerald, A. Ramudo, J. Kline, A. Vilnis, C. Benson, J. Vet. Diagn. Investig. **15**, 213–220 (2003)
7. K. Sakai, T. Yoshikawa, F. Seki, S. Fukushi, M. Tahara, N. Nagata, Y. Ami, T. Mizutani, I. Kurane, R. Yamaguchi, H. Hasegawa, M. Saijo, K. Komase, S. Morikawa, M. Takeda, J. Virol. **87**, 7170–7175 (2013)
8. E. Garde, G. Pérez, G. Acosta-Jamett, B. Bronsvoort, Animals **3**, 843–854 (2013)
9. P. Yuan, T.B. Thompson, B.A. Wurzburg, R.G. Paterson, R.A. Lamb, T.S. Jardetzky, Structure **13**, 803–815 (2005)
10. M. Blixenkrone-Moller, V. Svansson, M. Appel, J. Krogsrud, P. Have, C. Orvell, Arch. Virol. **123**, 279–294 (1992)
11. T.C. Harder, K. Klusmeyer, H.R. Frey, C. Orvell, B. Liess, J. Virol. Methods **41**, 77–92 (1993)

12. K. Iwatsuki, S. Tokiyoshi, N. Hirayama, K. Nakamura, K. Ohashi, C. Wakasa, T. Mikami, C. Kai, Vet. Microbiol. **71**, 281–286 (2000)
13. C. Orvell, M. Blixenkrone-Moller, V. Svansson, P. Have, J. Gen. Virol. **71**(Pt 9), 2085–2092 (1990)
14. G. Bolt, T.D. Jensen, E. Gottschalck, P. Arctander, M.J. Appel, R. Buckland, M. Blixenkrone-Moller, J. Gen. Virol. **78**(Pt 2), 367–372 (1997)
15. M.A. Carpenter, M.J. Appel, M.E. Roelke-Parker, L. Munson, H. Hofer, M. East, S.J. O'Brien, Vet. Immunol. Immunopathol. **65**, 259–266 (1998)
16. L. Haas, W. Martens, I. Greiser-Wilke, L. Mamaev, T. Butina, D. Maack, T. Barrett, Virus Res. **48**, 165–171 (1997)
17. T.C. Harder, M. Kenter, H. Vos, K. Siebelink, W. Huisman, G. van Amerongen, C. Orvell, T. Barrett, M.J. Appel, A.D. Osterhaus, J. Gen. Virol. **77**(Pt 3), 397–405 (1996)
18. K. Iwatsuki, N. Miyashita, E. Yoshida, T. Gemma, Y.S. Shin, T. Mori, N. Hirayama, C. Kai, T. Mikami, J. Gen. Virol. **78**(Pt 2), 373–380 (1997)
19. V. Martella, A. Pratelli, F. Cirone, N. Zizzo, N. Decaro, A. Tinelli, M. Foti, C. Buonavoglia, Mol. Cell. Probes **16**, 77–83 (2002)
20. M. Mochizuki, M. Hashimoto, S. Hagiwara, Y. Yoshida, S. Ishiguro, J. Clin. Microbiol. **37**, 2936–2942 (1999)
21. V. Martella, F. Cirone, G. Elia, E. Lorusso, N. Decaro, M. Campolo, C. Desario, M.S. Lucente, A.L. Bellacicco, M. Blixenkrone-Moller, L.E. Carmichael, C. Buonavoglia, Vet. Microbiol. **116**, 301–309 (2006)
22. A.J. McCarthy, M.A. Shaw, S.J. Goodman, Proc. Biol. Sci. **274**, 3165–3174 (2007)
23. C.H. Hsu, D.W. Kingsbury, Virology **120**, 225–234 (1982)
24. K.W. Ryan, A. Portner, K.G. Murti, Virology **193**, 376–384 (1993)
25. M. Graves, D.E. Griffin, R.T. Johnson, R.L. Hirsch, I.L. de Soriano, S. Roedenbeck, A. Vaisberg, J. Virol. **49**, 409–412 (1984)
26. E. Norrby, Y. Gollmar, Infect. Immun. **6**, 240–247 (1972)
27. J. Curran, J.B. Marq, D. Kolakofsky, Virology **189**, 647–656 (1992)
28. S.M. Horikami, J. Curran, D. Kolakofsky, S.A. Moyer, J. Virol. **66**, 4901–4908 (1992)
29. A. Scheid, P.W. Choppin, Virology **57**, 475–490 (1974)
30. Takeuchi H., and Haltiwanger R.S. Role of Glycosylation of Notch in Development. Elsevier, New York, pp. 638–645, %@ 1084–9521 (2010)
31. R.K. Plemper, A.L. Hammond, R. Cattaneo, J. Biol. Chem. **276**, 44239–44246 (2001)
32. K. Iwatsuki, N. Miyashita, E. Yoshida, Y.S. Shin, K. Ohashi, C. Kai, T. Mikami, J. Vet. Med. Sci. **60**, 381–385 (1998)
33. I.K. Visser, R.W. van der Heijden, M.W. van de Bildt, M.J. Kenter, C. Orvell, A.D. Osterhaus, J. Gen. Virol. **74**(Pt 9), 1989–1994 (1993)
34. T. Barrett, D.K. Clarke, S.A. Evans, B.K. Rima, Virus Res. **8**, 373–386 (1987)
35. G. Chappuis, Vet. Microbiol. **44**, 351–358 (1995)
36. T. Mori, Y.S. Shin, M. Okita, N. Hirayama, N. Miyashita, T. Gemma, C. Kai, T. Mikami, J. Gen. Virol. **75**(Pt 9), 2403–2408 (1994)
37. N.T. Lan, R. Yamaguchi, A. Inomata, Y. Furuya, K. Uchida, S. Sugano, S. Tateyama, Vet. Microbiol. **115**, 32–42 (2006)
38. J.J. Zhao, X.J. Yan, X.L. Chai, V. Martella, G.L. Luo, H.L. Zhang, H. Gao, Y.X. Liu, X. Bai, L. Zhang, T. Chen, L. Xu, C.F. Zhao, F.X. Wang, X.Q. Shao, W. Wu, S.P. Cheng, Vet. Microbiol. **140**, 34–42 (2010)
39. Cardoso A.I., Blixenkrone-Moller M., Fayolle J., Liu M., Buckland R., and Wild T.F., Virology 225, 293–299, %@ 0042–6822 (1996)
40. J.A. Lednicky, T.P. Meehan, M.J. Kinsel, J. Dubach, L.L. Hungerford, N.A. Sarich, K.E. Witecki, M.D. Braid, C. Pedrak, C.M. Houde, J. Virol. Methods **118**, 147–157 (2004)
41. S. Sultan, N. Charoenvisal, N.T. Lan, R. Yamaguchi, K. Maeda, K. Kai, Virol. J. **6**, 157 (2009)
42. V. von Messling, C. Springfeld, P. Devaux, R. Cattaneo, J. Virol. **77**, 12579–12591 (2003)
43. K. Sakai, N. Nagata, Y. Ami, F. Seki, Y. Suzaki, N. Iwata-Yoshikawa, T. Suzuki, S. Fukushi, T. Mizutani, T. Yoshikawa, N. Otsuki, I. Kurane, K. Komase, R. Yamaguchi, H. Hasegawa, M. Saijo, M. Takeda, S. Morikawa, J. Virol. **87**, 1105–1114 (2013)
44. Jiang Q., Hu X., Ge Y., Lin H., Jiang Y., Liu J., Guo D., Si C., and Qu L., Genome announcements 1 (2013)
45. M.S. Sidhu, W. Husar, S.D. Cook, P.C. Dowling, S.A. Udem, Virology **193**, 66–72 (1993)