第19卷5期 中国病毒学 19(5):481-486
2004年10月 VIROLOGICA SINICA October 2004

犬冠状病毒大熊猫株纤突蛋白全基因的克隆与序列分析*

乔 军[1,3]，夏咸柱[1**]，胡桂学[2]，谢之景[1]，闫 芳[1]，杨松涛[1]，黄 耕[1]

（1.中国人民解放军军事医学科学院军事兽医研究所，吉林长春 130062；2.吉林农业大学动物科技学院，吉林长春 130118；3.塔里木农垦大学动物科技学院，新疆阿拉尔 843300）

摘要：在国内首次对犬冠状病毒大熊猫野毒株（CCV DXMV）纤突蛋白基因进行了克隆和序列测定。该基因全长4362 bp，编码1453个氨基酸，N端前18个氨基酸为推测的信号肽序列，后1435个氨基酸构成成熟蛋白。与GenBank中已发表的11个CCV毒株S基因相比，S基因核苷酸序列同源性在40.2%～99.5%之间；推导的氨基酸序列同源性在15.9%～99.0%之间。DXMV株S基因变异区主要集中在该基因前1/2处，其中350-370、439-478、1718-1818三个区域碱基变异较大，而1060-1700区却十分保守。基于S全基因及其蛋白的聚类分析表明，DXMV株与K378、NVSL和US patent株亲源关系最近。推导的DXMV株S蛋白氨基酸序列潜在的N-联糖基化位点与CCV强毒V54相同，为34个，比Insavc-1弱毒多一个；其中第566-568位糖基化位点为多数强毒拥有而弱毒没有的。另外，DXMV株S蛋白疏水性及抗原表位与其它毒株有一定的差异，这些差异对DXMV株致病性和免疫原性等影响尚待进一步的研究。

关键词：犬冠状病毒大熊猫株；纤突蛋白基因；克隆；序列分析

中图分类号：S852.65 **文献标识码**：A **文章编号**：1003-5125(2004)05-0481-06

Cloning and Sequence Analysis of Spike Protein Gene of Canine Coronavirus Giant Panda's Isolate*

QIAO Jun[1,3], XIA Xian-zhu[1**], HU Gui-xue[2], XIE Zhi-jing[1], YAN Fan[1], YANG Song-tao[1], HUANG Gen[1]

(1. Institute of Military Veterinary, Academy of Military Medicine of PLA, Changchun 130062, China; 2. Department of Animal Science and Technology, Jinlin Agricultural University, Changchun 130118, China; 3 Department of Animal Science and Technology, Tarim Agricultural University, Alar Xinjiang 843300, China.)

Abstract: Spike (S) gene of *Canine coronavirus* giant panda's isolate (CCV strain DXMV) was amplified and cloned into pGEM-T for sequencing. The complete length of spike gene of strain DXMV was 4362 bp, which encoded 1453 amino acids. The initiative 18 amino acids were signal peptide. The homologies of nucleic acid and amino acid of spike among strains of CCV were 40.2-99.5% and 15.9-99.0%, respectively. The variation mainly exist in front half second region of S gene, especially in the regions of 350-370, 439-478 and 1718-1818 nucleotide acids, while the region 1060-1700 nucleotide acids shows high conservative. Phylogenetic tree based on S gene and deduced amino acid sequence of 12 different strains of CCV showed that strain DXMV had closer relationships with wild strain K378, NVSL and US patent. There are 34 N-glycosylation sites in deduced S protein of strain DXMV and V54, but 33 for strain

收稿日期：2004-03-22， 修回日期：2004-04-26
* 基金项目：国家自然科学基金（30000123）
作者简介：乔军（1971-），男，新疆籍，博士研究生，主要从事分子病毒学研究。
** 通讯作者：夏咸柱（1939-），男，江苏省籍，中国工程院院士，主要从事动物病毒学研究。
Received date: 2004-03-22, Accepted date: 2004-04-26
* Foundation item: National Science Foundation of China (No. 30000123)
** Corresponding author. Phone: 0431－6986748, E-mail: xia_xianzhu@yahoo.com.cn

Insavc-l. Noticeably, the potential glycosylation site 566-568 was owned by DXMV and most virulent strains but not for weak virulent strains. Moreover, there were some differences in the hydrophobicity and antigenic index between DXMV and other strains of CCV. The effects of the differences in S protein on pathogenicity and immunogenicity should be further studied.

Key words: *Canine coronavirus* giant panda's isolate; Spike protein gene; Cloning; Sequence analysis

CLC number: S852.65 **Document code:** A **Article ID:** 1003-5125(2004)05-0481-06

According to serum proofs of neutralization test and analysis of phylogenetic tree, coronaviruses can be divided into four serogroups. *Canine coronavirus* (CCV), an important causative agent of canine enteritis, is subordinate to serogroup I. It can infect all kinds of dogs, resulting in diarrhea, loss of appetite and weight[1,2]. What's more, it even infect giant panda (*Ailuropoda melanoleuca*) and others wildlife animals [3-6]. In 1997, our laboratory isolated a coronavirus from the liver of acutely died giant panda of Wolong Reserve with CRFK cell line, and verified that it was a strain of CCV by RT-PCR and sequencing. However, so far, no one had sequenced the complete spike (S) gene of giant panda's isolate of CCV. In order to study the novel molecular biological characteristics of giant panda's isolate and relationships with other strains of CCV, the following experiments have been carried out.

1 Materials and Methods

1.1 CCV giant panda's isolate (strain DXMV)

The strain DXMV was isolated from the liver of an acutely died giant panda with feline kidney CRFK cell line by the culture of monolayer cell and adaptation of continuous passages, which had been identified by conventional methods, RT-PCR and sequencing.

1.2 Main reagents

M-MLV, RNasin, T_4 DNA ligase and dNTPs were purchased from MBI company. RNA Extraction kit was from QIAGEN company. Taqplus DNA polymerase, DNA gel extraction kit were gained from Shanghai biotechnological Co.Ltd. pGEM-T vector from Promega company. Restriction endonucleases and other regents were obtained from Takara Co.Ltd .

1.3 Specific primers

According to spike gene sequence of CCV strain V54 reported by GenBank, a pair of specific primers was designed and used to amplify S gene of strain DXMV. Upstream primer: 5'-CAGGTACCACCATG ATTGTGCTTACATTGTGCC-3';Downstream primer: 5'-GAGCGGCCGCACATTTTGAATTCAGTGAAC ATGA-3'. In order to further clone, *Kpn* I and *Not*I sites were added in the 5'end of upstream and downstream primers, respectively.

1.4 Isolation of total RNA

Strain DXMV was cultivated in monolayer cell of CRFK. When cells appeared cytopathic effect, the culture was harvested and used to isolate the total RNA. RNA Extraction kit was used as described as the kit instruction of menu to isolate total RNA. The isolated RNA was finally dissolved in 20μL RNase-free water.

1.5 Reverse transcription and PCR amplification

Reverse transcript was carried out according to the reference[7]. PCR volume was 50μL with RT product 2μL, 10×PCR buffer 5μL, 25 mmol/L $MgCl_2$ 4μL, 25 pmol/μL upstream primer 1μL, downstream primer 1μL, 2.5mmol/L dNTPs 4μL, H_2O 32.7μL and 5U/μL Taqplus polymerase 0.3 μL. The condition of PCR reaction was as follows: 96℃, 200s; 94℃, 40s; 54℃, 50s; 72℃, 280s; 35 cycles; then 600s extension. The PCR products were detected in 0.7% gel electrophoresis and photographed in UV.

1.6 Cloning, sequencing and sequence analysis

The PCR product was recovered by DNA gel extraction kit and ligatured with pGEM-T overnight. Then the ligatured products were transformed into competence *E.coli* JM109. The recombinants were screened by blue/white plaque, PCR and restriction enzyme digestion reaction. The correct recombinant was named pTS and delivered to Takara Co.Ltd for sequencing. The sequencing result was analyzed by molecular software of DNAstar and DNAMAN.

2 Results

2.1 Amplification of S gene by PCR

The electrophoresis results showed that the length

of PCR product was 4362 bp, which coincided with the expectant length (Fig. 1)

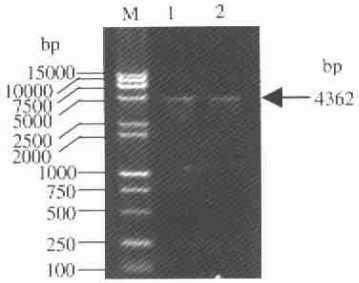

Fig. 1 Amplification of S gene

M, D, L-15000+2000; 1, CCV strain V54; 2, CCV strain DXMV.

2.2 Identification of recombinant pTS

The fragment amplified from the template pTS were tallied with expectant results. The interesting gene fragments were also gained with restriction enzyme digestion reaction. Two recombinants were screened from 12 white plaques (Fig.2).

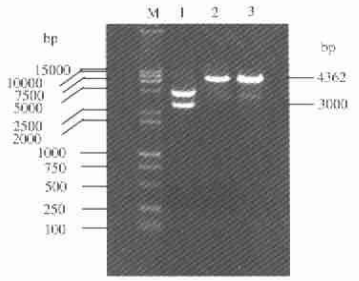

Fig. 2　Identification of recombinant pTS by enzyme digestion

M, D, L-15000+2000; 1, pTS/*Kpn*I+*Not*I; 2, pTS/*Hin*dIII; 3, pTS/ *Not* I.

2.3 cDNA sequence of S protein gene and deduced amino acid sequence

The sequencing results showed that the complete length of spike protein gene of strain DXMV was 4362bp and encoded 1453 amino acids (GenBank accession number, AY436637). The initiative 18 amino acids were signal peptide (Fig.3). Compared with S gene of other strains of CCV, the variation parts mainly existed in front half region of S gene, especially in the regions of 350-370, 439-478 and 1718-1818 nucleotide acids. However, the region of 1060-1700 nucleotide acids shows high conservativeness. There are 34 potential N-glycosylation sites in deduced amino acid sequence of S protein for strain DXMV and V54, but 33 for strain Insavc-1. Noticeably, the potential glycosylation site in 566-568 amino acids was owned by giant Panda's isolate and most CCV virulent strains but not for weak virulent strains. The potential glycosylation sites on 537-539, 557-559, 566- 568, 14-16 and 94-96 amino acids, respectively, played important roles in the formation of the subsites Aa, Ab, Ac, B1 and B2 according to the interrelated analysis based on S protein of transmissible gastroenteritis virus (TGEV). Additionally, the sequence MKRSGYGQPIA in 537-547 amino acids of N-terminal of deduced S protein was found, which was entirely same as the sequence forming the subsite Ac of TGEV.

2.4 Analysis of phylogenetic tree

Phylogenetic tree based on S genes and deduced amino acid sequences of 12 different strains showed that there were two different serotypes for CCV, respiratory-type (CrCV) and enteron-type. Two serotypes showed low identity. Enteron-type CCV can also be divided into genotype I (DXMV, US patent, k378, CCV6, Insavc-1, V54, 5821, NVSL and BGF10) and genotype II (Elmo02, 2303). The strain DXMV was subordinate to genotype I and had closer relationships with high virulent strain K378, NVSL and US patent, but far distance with virulent strain Elmo02, 2303. Compared S genes with other strains, the homologies of nucleotide and amino acid of spike among CCV strains were 40.2%～99.5% and 15.9%～99.0%, respectively. (All strains were gathered from GenBank. All sequences' accession numbers were as follows: CCV DXMV, AY436637; US patent, AR206117; k378, A22732; CCV6, A22882; CCV Insavc-1, D 13096; CCV V54, A22886; CCV 5821, AB017789; CrCV, CQ772298; NVSL, AF116244; BGF10, AY342160; Elmo02, AY307020; 2303, AY307021).

2.5 Analyses of predicted hydrophobicity and antigenic index

Hydrophobicity and antigenic index of amino acid sequence of spike protein of DXMV strain were analyzed by DNAMAN and DNAstar software. The results showed it had two apparent hydrophobic regions, 1-18 and 1405-1425 amino acids, which possible consisted of signal peptide and transmembrane region, respectively (arrow note). In addition, there are some differences in predicted antigenic index of spike protein between strain DXMV and Insavc-1, which possibly results in differences in their pathogenicity and immunogenicity (Fig.6).

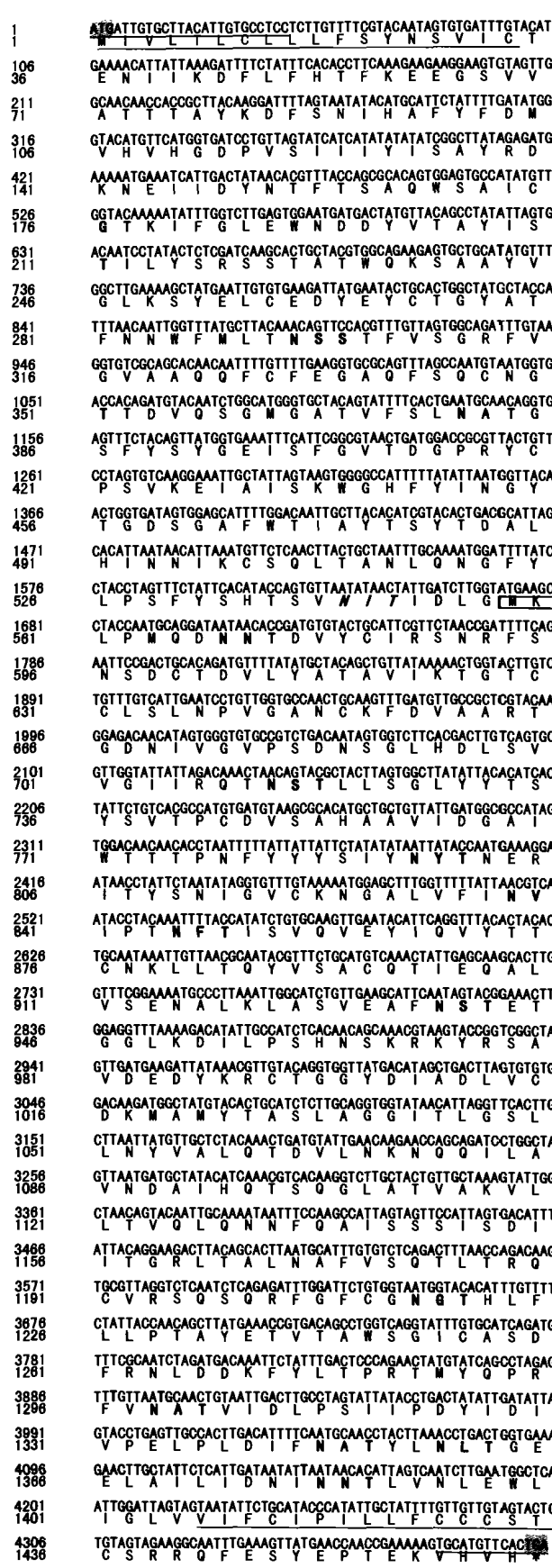

Fig. 3 Nucleotide sequence of S gene of CCV strain DXMV and deduced amino acid sequence

The initiation and stop codon were shadowed. The upstream and downstream primers were underlined in nucleotide sequence. Signal peptide and transmembrane region were underlined in deduced amino acid sequence. Potential N-glycosylation sites were noted with boldface type. The necessary glycosylation which consisting of antigen site A and B were marked by italics. The deduced subsite Ac was covered with box.

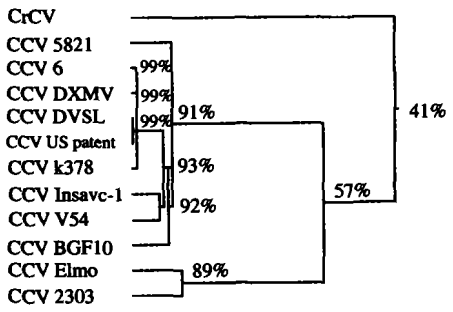

Fig. 4 Phylogenetic tree of CCV based on homology of S gene

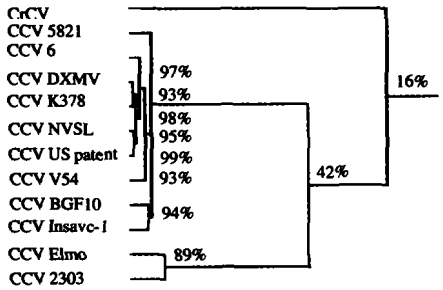

Fig. 5 Phylogenetic tree of CCV based on homology of deduced S protein

Numbers in figure 4 and 5 denote the homology of S gene and deduced amino acid sequence of 12 different strains.

3 Discussion

With the rapid development of canine and wildlife animal industry, CCV is seriously imperiling the security of canines and some precious wildlife animals such as giant pandas. Mainka *et al*[2] (1994), who carried out serological survey of giant panda in the Wolong Reserve in China, found that three in eight giant panda's sera are CCV antibody positive, which indicated that CCV may infect panda and induce specific neutralization antibody. In our study, we found that the variation region chiefly existed in the front half of S gene, especially in the regions of 350-370, 439-478 and 1718-1818 nucleotide acids which encoded the globe of S protein, while the sequence of back half of S gene exhibited high conservativeness.

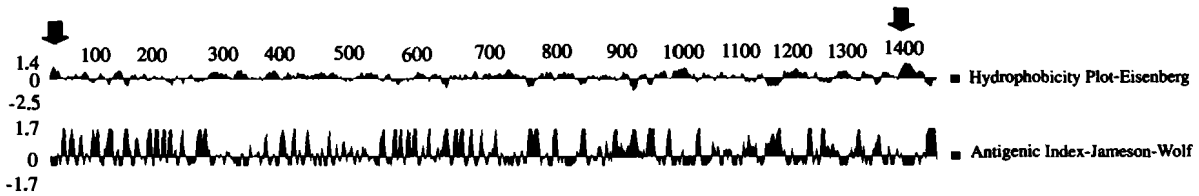

Fig. 6 Analysis of hydrophobicity and antigenic index plots of S protein of CCV strain DXMV by DNAstar software

Spike, a main structural protein of CCV which is on the surface of virion, can induce body's immune system to produce neutralization antibody against CCV. Swine TGEV, another detailedly studied member of serogroup I coronavirus, had close relationships with CCV in antigenicity[8, 9,15]. Correa *et al* (1988) verified that there were four antigenic sites A, B, C and D on the spike protein by identification of monoclonal antibody against TGEV[10]. Falima *et al* (1991) further found that antigenic site A was composed of three subsites Aa, Ab and Ac and the amino acid residues 534, 591 and 543 possibly played important roles in the formation of these subsites, respectively[11]. Moreover, the amino acid residue 586 can also influence the forma- tion of subsite Aa and Ab, which denoted that antigenic site A had complex threedimensional structure. Delmas *et al* (1990) found that the subsite Ac of TGEV can be signified by 537-MKRSGYG QPIA-547[12]. Our studies on spike protein of DXMV strain coincided with Delmas's studies. Additionally, Simpkins and Christine *et al* found that glycosylation was necessary to form the correct structure for antigenic site A and B[13,14]. According to present studies on TGEV, spike protein of CCV was deduced to have similar antigenic structure and functions. It not only can specifically recognize host-cell receptor (Aminopeptidase-N, APN), resulting in fusion of viral envelope with cell membrane, but also can induce specific neutralization antibody for carrying antigenic determinant of B cells. Meanwhile, it possibly plays an important role in pathogenesis because it has close relationships with viral virulence and cell tropism.

Although there was only one serotype for enteron-type CCV, genetic differences had been found among different strains. Pratelli et al found that there existed two distinct genotypes of CCV in pups and two genotypes can be simultaneously detected in the fecal

samples of dogs with diarrhea [16,17]. Two prokaryotic expressed proteins of S gene of genotype I can only strongly react with the sera-positive for genotype I, whereas the sera-positive for genotype II showed low there exist in reactivity in the ELISA assay[18], which denotes that antigenic differences of S protein between genotype I and genotype II. In our studies, phylogenetic tree also indicated that there was distinct genetic diversity among different strains. Whether genetic variability of S gene may result in the changes of host range needs to be studied further. Giant pandas are precious wildlife animals in China. Although government has tooks effective measures to protect giant panda in breeding and reproduction, prevention and control of infectious disease, especially virus disease, is still a arduous and pressing task.

Reference

[1] Ying Z, Liu J H. Animal virology [M] (second edition). Beijing: Science press. 1997, 671-703.

[2] Mochizuki M, Hashimoto M, Ishida T, *et al.* Recent epidemiological status of canine viral enteric infections and Giardia infection in Japan [J]. J Vet Med Sci. 2001, 63:573-575.

[3] Mainka S A, Qiu X, He T, *et al.* Serologic survey of giant pandas (*Ailuropoda melanoleuca*), and domestic dogs and cats in the Wolong Reserve, China [J]. J Wildl Dis. 1994, 30:86-89.

[4] Marsilio F, Tiscar P G, Gentile L, *et al.* Serologic survey for selected viral pathogens in brown bears from Italy [J]. J Wildl Dis. 1997, 33: 304-307.

[5] Kimber K R, Kollias G V, Dubovi E J. Serologic survey of selected viral agents in recently captured wild North American river otters (Lontra canadensis)[J]. J Zoo Wildl Med. 2000, 31: 168-175.

[6] Zamke R L, Evermann J, Ver Hoef J M, *et al.* Serologic survey for canine coronavirus in wolves from Alaska [J]. J Wildl Dis. 2001, 37: 740-745.

[7] Pratelli A, Tinelli A, Decaro N, *et al.* PCR assay for the detection and the identification of atypical canine coronavirus in dogs[J]. J Virol Methods. 2002, 106:209-213.

[8] Laude H, Rasschaert D, Delmas B, *et al.* Molecular biology of transmissible gastroenteritis virus [J]. Vet Virol. 1990, 23: 147-154.

[9] Sanchez C, Jimenez G, Laviada M D, *et al.* Antigenic homology among coronavirus related to transmissible gastroenteritis virus[J]. Virology. 1990, 174: 410-413.

[10] Correa I, Jimenez G, Sune C, *et al.* Antigenic structure of the E2 glycoprotein from transmissible gastroenteritis coronavirus[J]. Virus Res. 1988, 10: 77-94.

[11] Falima G, Willewp A, Correa C, *et al.* Residues involved in the antigenic sites of transmissible gastroenteritis virus S glycoprotein [J]. J Virol. 1991, 183: 225-238.

[12] Delmas B, Rasschaert D, Gtodet M, *et al.* Four major antigenic site of the coronavirus transmissible gastroenteritis virus are located on the amino-terminal half of spike glycoprotein S [J]. J Gen Virol. 1990, 71: 1313-1323.

[13] Simpkins R A, Weilnau P A, Bias J, *et al.* Antigenic variation among transmissible gastroenteritis virus (TGEV) and porcine respiratory coronavirus strains detected with monoclonal antibodies to the S protein of TGEV[J]. Am. J. Vet. Res. 1992, 53: 1253-1258.

[14] Christine K, Graham D, Yolken R, *et al.* Point mutations in the S protein connect the sialic acid binding activity with the enteropathogenicity of TGEV[J]. J Virol. 1997, 71 (4): 3285-3287.

[15] Wesley R D. The S gene of canine coronavirus, strain UCD-1, is more closely related to the S gene of transmissible gastroenteritis virus than to that of feline infectious peritonitis virus[J]. Virus Res. 1999, 61(2):145-152.

[16] Pratelli A, Martella V, Decaro N, *et al.* Genetic diversity of a canine coronavirus detected in pups with diarrhea in Italy[J]. J Virol Methods. 2003, 110(1): 9-17.

[17] Pratelli A, Decaro N, Tinelli A, *et al.* Two genotypes of canine coronavirus simultaneously detected in the fecal samples of dogs with diarrhea[J]. J Clin Microbiol. 2004, 42:1797-1799.

[18] Pratelli A, Elia G, Decaro N, *et al.* Cloning and expression of two fragments of the S gene of canine coronavirus type I[J]. J Virol Methods. 2004, 117:61-65.