

Genetic variation of nucleocapsid genes of porcine epidemic diarrhea virus field strains in China

Jianfei Chen · Xiaozhen Liu · Hongwu Lang ·
Zhongtian Wang · Da Shi · Hongyan Shi ·
Xin Zhang · Li Feng

Received: 14 November 2012 / Accepted: 4 December 2012 / Published online: 5 February 2013
© Springer-Verlag Wien 2013

Abstract Porcine epidemic diarrhea virus (PEDV) poses a great threat to the Chinese swine industry. During 2006–2011, 74.0 % (94/127) of samples from 32 swine-raising farms in 15 provinces were positive for PEDV by reverse transcription nested polymerase chain reaction (RT-nested PCR). The sequences of nucleocapsid (N) genes of 32 representative field strains were determined, and the open reading frames (ORFs) of these N genes were 1326 nucleotides long. The N gene sequences were found to be more than 95 % identical to each other. The number of potential phosphorylation sites in the N protein varied from 5 to 12. A phylogenetic tree based on the N genes showed that the 32 Chinese field strains formed three groups.

Porcine epidemic diarrhea virus (PEDV), a member of genus *Alphacoronavirus*, family *Coronaviridae*, which together with the families *Arteriviridae* and *Roniviridae*, constitute the order *Nidovirales*, has a single-stranded positive-sense RNA genome of approximately 28 kb that is infectious. PEDV was first reported in Belgium and the United Kingdom in 1978 [1]. In China, PEDV was first

confirmed by fluorescent antibody test and serum neutralization in 1984 [2]. Although a bivalent attenuated vaccine against TGEV and PEDV is being used in China [3], PEDV occurs frequently on many swine-raising farms in China.

Coronavirus nucleocapsid (N) proteins vary from 377 to 455 amino acids in length, are highly basic, and have a high (7 to 11 %) serine content. These serines are potential targets for phosphorylation. Antigenic studies have shown that the N protein is one of the immunodominant antigens in members of the family *Coronaviridae* [4]. The N protein of infectious bronchitis virus (IBV) is a relevant target for immune recognition in both mice and chickens [5].

The sequences of N genes of PEDV strains CV777 and Br1/87 were first determined in 1993, and the deduced amino acid sequences of these strains are 441 amino acids in length [6]. A nested polymerase chain reaction (PCR) based on a partial sequence of the N gene was designed to detect PEDV in Japan in 1999 [7]. The purpose of this study was to investigate the sequence diversity of N genes of PEDV field strains during 2006–2011 in China. Because the detection of PEDV in feces or contents of the small intestine might be affected by the reliability and sensitivity of the technique [8], a reverse transcription nested PCR (RT-nested PCR) was established to amplify the full-length N genes of field strains in this study.

One hundred twenty-seven porcine fecal samples or contents of the small intestine were collected from piglets showing watery diarrhea and dehydration on 32 swine-raising farms in 15 provinces in China from January 2006 to August 2011. All of the samples were diluted with phosphate-buffered saline to make 10 % (v/v) suspensions. The suspensions were vortexed for 1 min and clarified by centrifugation for 10 min at 5,000 rpm. The supernatants were collected for RT-nested PCR.

J. Chen and X. Liu contributed equally to this work.

J. Chen · X. Liu · D. Shi · H. Shi · X. Zhang · L. Feng (✉)
Division of Swine Infectious Diseases, State Key Laboratory
of Veterinary Biotechnology, Harbin Veterinary Research
Institute, Chinese Academy of Agricultural Sciences,
Harbin 150001, People's Republic of China
e-mail: fl@hvri.ac.cn

H. Lang · Z. Wang
China Institute of Veterinary Drug Control,
Beijing 100081, People's Republic of China

Viral RNA was extracted from the supernatants using TRIzol Reagent (Invitrogen Corp., Carlsbad, USA), and the first-strand complementary DNA (cDNA) was synthesized with M-MLV reverse transcriptase (Promega, USA) using a specific primer (N1L, 5'-TCAAATACCTGGCACGCT CT-3') according to the manufacturer's instructions.

Two pairs of primers (N1U, 5'-TATAAGGTTGCTAC TGGCGT-3' and N1L, 5'-TCA AATACCTGGCACGCT CT-3'; N2U, 5'-GTCAAACACGGCGACTATT-3' and N2L, 5'-TGGCACTACCCTGGAACATA-3') for RT-nested PCR were designed and synthesized according to the corresponding sequence of CV777 (AF353511). The outer span, including primers N1U and N1L, was 1843 bp, and the inner span, including primers N2U and N2L, was 1465 bp. The fragments containing the full-length N gene were amplified from 94 samples using the inner primers (N2U/N2L). The overall detection rate of PEDV in the samples was 74.0 % (94/127). TGEV was also detected by RT-PCR in our lab, and the overall detection rate was 36.5 % in pigs with diarrhea in China.

PCR products were excised from 1.0 % agarose gels, purified using an AxyPrep™ DNA Gel Extraction Kit (Axygen Scientific, inc., USA), and cloned into the T-tailed

vector pMD18-T, and these clones were introduced into JM109 competent cells (TaKaRa, China) by transformation. Three recombinant DNA clones were sequenced by the dideoxy nucleotide chain terminator method. All of the sequences in this study have been deposited in the GenBank database. The field strains and their accession numbers are as follows: CH/IMB/06 (FJ473387), CH/HNCH/06 (FJ473388), CH/JSX/06 (FJ473389), CH/HLJH/06 (FJ473390), CH/IMT/06 (FJ473391), CH/SHH/06 (FJ473392), CH/HLJM/07 (FJ473393), CH/HNH/07 (FJ473394), CH/GSJ/07 (HM210880), CH/JL/09 (HM210881), CH/GDS/09 (HM210882), CH/HLJQ/2010 (HQ455345), CH/HLJHG/2010 (HQ455346), CH/HNZZ/2011 (JN601052), CH/BJYQ-1/2011 (JN601053), CH/BJYQ-2/2011 (JN601054), CH/FJND/2011 (JN601055), CH/GDQY-1/2011 (JN601056), CH/GDQY-2/2011 (JN601057), CH/GDQY-3/2011 (JN601058), CH/HLJHG/2011 (JN601059), CH/SDRZ-1/2011 (JN601060), CH/SDRZ-2/2011 (JN601061), CH/GXNN/2011 (JN601062), CH/BJSY/2011 (JQ735953), CH/HLJHRB/2011 (JQ743650), CH/HLJHH/2011 (JQ743651), CH/GXWP/2011 (JQ743652), CH/XJUrumsqi/2011 (JQ743653), CH/GXQZ/2011 (JQ743654), CH/ZJHZ/2011 (JQ743655), CH/GXWM/2011 (JQ743656).

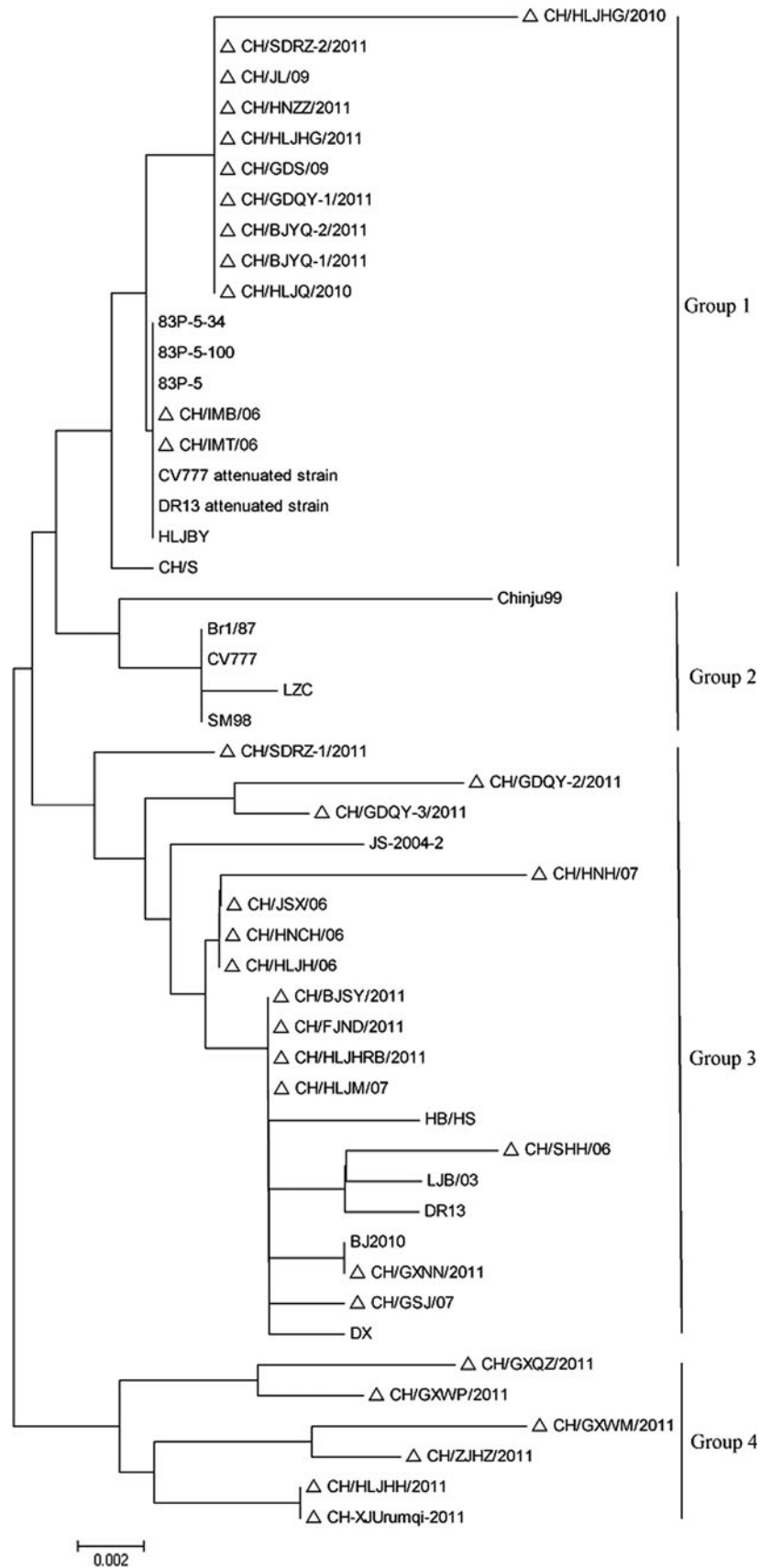
Table 1 The predicted phosphorylation sites of N proteins of Chinese field strains and CV777

Strain	Predicted phosphorylation site																Total
	¹⁴³ S	¹⁴⁵ S	¹⁴⁷ S	¹⁴⁹ S	¹⁵¹ S	¹⁵⁷ S	¹⁵⁹ S	¹⁶¹ S	¹⁶³ S	¹⁷⁵ S	¹⁸⁷ S	²⁵³ S	²⁵⁵ S	²⁶⁰ T	²⁶⁶ S	²⁷⁰ S	
CH/GDQY-2/2011	-	-	+	-	+	-	-	+	+	+	-	-	-	-	-	-	5
CH/HLJHG/2010	-	-	+	+	+	-	+	+	+	-	-	-	-	-	-	+	7
CH/GDQY-3/2011	-	-	+	-	+	-	-	+	+	+	+	-	-	-	-	+	7
CV777, CH/IMB/06, CH/IMT/06, CH/GSJ/07, CH/JL/09, CH/GDS/09, CH/HLJQ/2010, CH/BJYQ-1/2011, CH/BJYQ-2/2011, CH/GDQY-1/2011, CH/HLJHG/2011, CH/HLJHH/2011, CH/HNZZ/2011, CH/SDRZ-1/2011, CH/SDRZ-2/2011, CH/XJUrumsqi/2011, CH/ZJHZ/2011	-	-	+	+	+	-	+	+	+	+	-	-	-	-	-	+	8
CH/HNH/07	-	-	+	+	+	-	+	+	+	+	-	+	+	-	-	-	9
CH/HNCH/06, CH/JSX/06	-	-	+	+	+	-	+	+	+	+	-	+	+	-	-	+	10
CH/HLJH/06	-	-	+	+	+	-	+	+	+	+	-	+	+	-	-	+	10
CH/SHH/06	-	-	+	+	+	-	+	+	+	-	-	+	+	+	+	+	11
CH/GXWM/2011	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	+	11
CH/HLJM/07, CH/BJSY/2011, CH/FJND/2011, CH/GXNN/2011, CH/GXQZ/2011, CH/GXWP/2011, CH/HLJHRB/2011	-	-	+	+	+	-	+	+	+	+	-	+	+	+	+	+	12

+ Phosphorylation site present

- No phosphorylation site

Fig. 1 Phylogenetic analysis of the nucleotide sequences of N genes of PEDV field strains. The tree was constructed based on the neighbor-joining method using the MEGA 5.05 software. The scale bar indicates the branch lengths for 0.2 % nucleotide differences. The Chinese PEDV strains are marked with triangles



The N genes of the 32 field strains were found to contain a single open reading frame (ORF) consisting of 1326 nucleotides. There were no nucleotide deletions or insertions in the ORFs of the N genes. All of the N genes had a hexamer motif (CTAAAC), which is the transcription-regulating sequence (TRS), located in the nine nucleotides upstream of the initiator ATG, as recognized in a previously study [9]. There was a nine-nucleotide conserved sequence (AGAAACTTT) between the TRS and the start codon of the N gene. The N genes of the 32 field strains showed 95.3–100 % sequence identity to each other. They showed lower sequence identity to the field strain LZC (95.0–97.4 %) than to other Chinese reference field strains (95.6–99.7 %). They showed 95.9–100 % sequence identity to three attenuated strains (CV777, DR13 and 83P-5), whose N genes were 100 % identical to each other.

The N proteins of the 32 field strains were predicted to be 441 amino acids in length, with 7.3–8.4 % (32–37 serines) serine content. Phosphorylated N protein binds to viral RNA with a higher binding affinity than non-viral RNA, suggesting that phosphorylation of the N protein determines the recognition of viral RNA [10]. The phosphorylation sites of the N proteins of the field strains were predicted by the Web tool DISPHOS (<http://www.ist.temple.edu/DISPPOS>), which uses disorder information to detect phosphorylation sites. Only residues with a prediction value >0.5 are considered to be phosphorylated. The number of predicted phosphorylation sites of N proteins varied from 5 to 12. Of the 32 field strains, one strain had 5 predicted phosphorylation sites, two strains had 7, sixteen had 8, one had 9, three had 10, two had have 11 and seven strains had 12 (Table 1). The deduced amino acid sequences of 32 field strains showed 95.0–100 % sequence identity to each other. They showed 95.0–98.0 and 96.2–100.0 % sequence identity to CV777 and attenuated strains, respectively.

Like the spike glycoprotein gene [11, 12], the N gene is an important component in the phylogenetic analysis of the epidemiological situation of coronaviruses in the field [13, 14]. A phylogenetic tree based on the N gene was constructed using MEGA 5.05 software [15], and the tree showed that the PEDV strains were divided into four groups (Fig. 1). The 32 field strains were divided into three groups. Twelve field strains in group 1 have a close relationship to CH/S, 83P-5. Fourteen field strains in group 3 have a close relationship to five Chinese reference strains (LJB/03, JS-2004-2, DX, BJ2010, HB/HS). Six field strains in group 4 are genetically different from other Chinese field strains and may represent a novel genotype of PEDV. It is notable that these field strains have 4–6 nucleotides that are different from the other strains at nt 1290–1298.

In summary, PEDV has a high prevalence rate in swine herds. The field strains are genetically diverse in their N

genes, both among themselves and as compared with reference strains. Phylogenetic analysis indicated that there are three genotypes of PEDV prevailing in China. Moreover, the sequences of N genes will enrich the information of the sequence database and form the basis for further functional exploration of PEDV.

Acknowledgments This work was supported by grants from National Natural Science Foundation of China (30901081, 31172350), State Key Laboratory of Veterinary Biotechnology (SKLVB201218), Natural Science Foundation for Distinguished Young Scholars of Heilongjiang Province (JC201118), and Higher School Science and Technology Innovation Team Project of Heilongjiang Province (2011TD001).

References

1. Pensaert MB, de Bouck P (1978) A new coronavirus-like particles associated with diarrhea in swine. *Arch Virol* 58:243–247
2. Xuan H, Xing D, Wang D, Zhu W, Zhao F, Gong H (1984) Study on the culture of porcine epidemic diarrhea virus adapted to fetal porcine intestine primary cell monolayer. *Chin J Vet Sci* 4(3): 202–208
3. Tong Y, Feng L, Li W, Zhu Y, Wang M, Ma S (1999) Development of Bi-combined attenuated vaccine against transmissible gastroenteritis virus and porcine epidemic diarrhea virus. *Chin J Pre Vet Med* 21(6):406–410
4. Lecomte J, Cainelli-Gebara V, Mercier V, Mansour S, Talbot PJ, Lussier G, Oth D (1987) Protection from mouse hepatitis virus type 3-induced acute disease by an anti-nucleoprotein monoclonal antibody. *Arch Virol* 97:123–130
5. Boots HAM, Benaissa-Trouw JB, Hesselink W, Rijke E, Schrier C, Hensen JE (1992) Induction of anti-viral immune responses by immunization with recombinant-DNA encoded avian coronavirus nucleocapsid protein. *Vaccine* 10:119–124
6. Bridgen A, Duarte M, Tobler K, Laude H, Ackermann M (1993) Sequence determination of the nucleocapsid protein gene of the porcine epidemic diarrhoea virus confirms that this virus is a coronavirus related to human coronavirus 229E and porcine transmissible gastroenteritis virus. *J Gen Virol* 74(9):1795–1804
7. Kubota S, Sasaki O, Amimoto K, Okada N, Kitazima T, Yasuhara H (1999) Detection of porcine epidemic diarrhea virus using polymerase chain reaction and comparison of the nucleocapsid protein genes among strains of the virus. *J Vet Med Sci* 61: 827–830
8. Guscetti F, Bernasconi C, Tobler K, Van Reeth K, Pospischil A, Ackerman M (1998) Immunohistochemical detection of porcine epidemic diarrhea virus compared to other methods. *Clin Diag Lab Immuno* 5:412–414
9. Duarte M, Tobler K, Bridgen A, Rasschaert D, Ackermann M, Laude H (1994) Sequence analysis of the porcine epidemic diarrhea virus genome between the nucleocapsid and spike protein genes reveals a polymorphic ORF. *Virology* 198:466–476
10. Chen H, Gill A, Dove BK, Emmett SR, Kemp CF, Ritchie MA, Dee M, Hiscox JA (2005) Mass spectroscopic characterization of the coronavirus infectious bronchitis virus nucleoprotein and elucidation of the role of phosphorylation in RNA binding by using surface plasmon resonance. *J Virol* 79:1164–1179
11. Park SJ, Moon HJ, Yang JS, Lee CS, Song DS, Kang BK, Park BK (2007) Sequence analysis of the partial spike glycoprotein gene of porcine epidemic diarrhea viruses isolated in Korea. *Virus Genes* 35:321–332

12. Lee D, Park CK, Kim SH, Lee C (2010) Heterogeneity in spike protein genes of porcine epidemic diarrhea viruses isolated in Korea. *Virus Res* 149:175–182
13. Park JY, Pak SI, Sung HW, Kim JH, Song CS, Lee CW, Kwon HM (2005) Variations in the nucleocapsid protein gene of infectious bronchitis viruses isolated in Korea. *Virus Genes* 31:153–162
14. Kuo SM, Wang CH, Hou MH, Huang YP, Kao HW, Su HL (2010) Evolution of infectious bronchitis virus in Taiwan: characterisation of RNA recombination in the nucleocapsid gene. *Vet Microbiol* 144(3–4):293–302
15. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *MBE* 28:2731–2739