ORIGINAL ARTICLE

Sequence and phylogenetic analysis of nucleocapsid genes of porcine epidemic diarrhea virus (PEDV) strains in China

Zhili Li · Feng Chen · Yao Yuan · Xiduo Zeng · Zhongyan Wei · Ling Zhu · Baoli Sun · Qingmei Xie · Yongchang Cao · Chunyi Xue · Jingyun Ma · Yingzuo Bee

Received: 24 May 2012/Accepted: 29 November 2012/Published online: 7 February 2013 © The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Porcine epidemic diarrhea virus (PEDV) causes acute diarrhea and dehydration with high mortality rates in swine. It has become increasingly problematic in China. Since the nucleocapsid (N) protein is highly conserved, it is a candidate protein for early diagnosis and vaccine development. In this study, the N genes of 15 PEDV strains were amplified by RT-PCR and cloned into the pMT-19T vector, sequenced, and compared to each other as well as to PEDV reference strains. The nucleotide sequences of the N gene of the Chinese PEDV strains consist of 1326 nucleotides and encode a 441-aa-long peptide. The nucleotide sequences of the fifteen PEDV strains in our study were 96.1-100 % identical to each other, and the deduced amino acid sequences were 94.8-100 % identical. Sequence comparison with other PEDV strains selected from Gen-Bank revealed that their nucleotide sequences were 94.2-99.7 % identical to those of the Chinese PEDV strains, and their deduced amino acid sequences were 94.1-99.5 % identical. In addition, the fifteen strains showed a high degree of nucleotide sequence identity to the early domestic strains (98.4-99.7 %) except the LZC strain, but less sequence identity to the vaccine strain (CV777) used in

B. Sun \cdot Q. Xie \cdot J. Ma (\boxtimes) \cdot Y. Bee (\boxtimes)

College of Animal Science, South China Agricultural University, Tianhe District, Wushan Road, Guangzhou, 510642 Guangdong, People's Republic of China e-mail: majy2400@scau.edu.cn

Y. Bee e-mail: yingzuobi@163.com

Y. Cao · C. Xue

China (94.7-97.7 %). Phylogenetic analysis showed that the Chinese PEDV strains are composed of a separate cluster including three early domestic strains (JS-2004-02, LJB/03 and DX) but differ genetically from the vaccine strain (CV777) and the early Korean strains (Chinju99 and SM98).

Introduction

Porcine epidemic diarrhea (PED), caused by porcine epidemic diarrhea virus (PEDV), is an acute, highly contagious, and devastating enteric disease that is characterized by severe enteritis and diarrhea, with high mortality rates in suckling pigs [1]. PED was first reported in England in 1971 [2]. Since then, PED occurs in most swine-raising countries in Europe, as well as China, Korea, Thailand, and Japan [3–6]. However, this disease is becoming a big concern especially in Asia, where outbreaks are often more acute and severe than those observed in Europe [7, 8]. Since the beginning of October 2010, a porcine epidemic diarrhea epizootic has been occurring in China, affecting pigs of all ages but characterized by high mortality rates among suckling piglets. The outbreak has been prevalent nationwide and has caused huge economic losses [9–11]. Most of the affected farms have lost 100 % of their newborn piglets, usually within 7 days, but sometimes even within only a few hours of birth. Few sows or boars show any clinical signs, which is inconsistent with a previous report of an outbreak in Thailand in 2007 that was characterized by pigs of all ages being infected and showing different degrees of diarrhea and anorexia [4]. In addition, in China, sporadic outbreaks of PED have been seen year round, and not just in the winter months.

Z. Li \cdot F. Chen \cdot Y. Yuan \cdot X. Zeng \cdot Z. Wei \cdot L. Zhu \cdot

State Key Laboratory of Biocontrol, College of Life Sciences, Sun Yat-sen University, Guangzhou 510275, People's Republic of China

Porcine epidemic diarrhea virus (PEDV) is an enveloped, single-stranded RNA virus, belonging to the order Nidovirales, family Coronaviridae, genus Alphacoronavirus [12–14]. The genome is comprised of a 5' untranslated region (UTR), a 3' UTR, and at least seven open reading frames (ORFs) that encode four structural proteins, (spike [S], envelope [E], membrane [M], and nucleocapsid [N]) and three non-structural proteins (replicase 1a and 1b and ORF3) [15–17]. It has been reported that the N protein binds to viral RNA, providing a structural basis for the helical nucleocapsid, which is a basic phosphoprotein associated with the genome [18]. In addition, the N protein is thought to be important in inducing cell-mediated immunity in the host [19]. The N protein participates in transcription of the viral genome, the formation of the viral core, and packaging of viral RNA [20, 21]. In the early stages of PEDV infection, a pig can produce high levels of antibodies against the N protein. Since the N protein is highly conserved, it is the best candidate protein for use as an antigen for early diagnosis reagents and vaccine development [22-24].

The purpose of the present study was to investigate the genetic characteristics of the N gene between 2010 and 2012 during PED outbreaks in different region of China. In this study, RNA was extracted directly from the feces or intestinal contents of piglets infected with PEDV. The N genes were cloned and sequenced, and the sequences were submitted to GenBank and compared with the N genes of other PEDV strains. In addition, the N protein motifs (including phosphorylation sites and hydrophilic regions) were identified. These data provide additional molecular epidemiological information on PEDV circulating in China and provide a basis for further development of diagnostic reagents and methods as well as assisting in vaccine selection.

Materials and methods

Sample collection

Porcine samples (including feces and intestinal contents) from piglets with severe watery diarrhea, dehydration and high mortality were collected from 55 farms in five provinces in China during February 2010 to March 2012. These samples were confirmed to be positive for PEDV by reverse transcription polymerase chain reaction (RT- PCR) [9].

Primer design for RT- PCR

In order to determine the sequences of the N gene, primers were designed based on known published sequences in GenBank (CV777 and Brl/87). The sense primer was 5'- TGCGGTTCTCACAGATAGTG-3', and the antisense primer was 5'-AAGTCGCTAGAAAAACACTCAGTA AT-3'. The size of the amplified product was predicted to be 1380 bp.

RT-PCR

Viral RNA was extracted from samples using TRIzol Reagent (Invitrogen, CA, USA), resuspended in nuclease-free water, and kept at -70 °C until further use. Reverse transcription was performed at 50 °C for 30 min in a reaction mixture consisting of 2.5 µl RNA (0.2 µg), 1 µl primer (10 pmol), 1 µl Prime Script One Step Enzyme mix, 8 µl RNase-free H₂O, and 12.5 µl 2 × One Step Buffer. The cycling conditions for the PCR were 94 °C for 2 min, followed by 32 cycles of denaturation (94 °C for 10 s), annealing (58 °C for 30 s), and extension (72 °C for 1 min), followed by a final extension at 72 °C for 7 min. Both the reverse transcription and the polymerase chain reaction were conducted using a PrimeScript One Step RT-PCR Kit (TaKaRa, Japan).

Cloning and sequencing

The amplified PCR products were subjected to gel electrophoresis, excised from the agarose gel and purified using an Agarose Gel DNA Purification Kit (TaKaRa, Japan). The PCR products were cloned into the pMD19-T vector according to the manufacturer's instructions (TaKaRa, Japan), three clones were sent to Shanghai Sangon Bioengineering Ltd. to be sequenced in both directions for each fragment, and the sequence was analyzed.

Genome and amino acid analysis

Fifteen sequences were selected randomly and aligned using ClustalX software (version 1.83), Bioedit and DNASTAR to examine their genetic diversity. Phylogenetic trees were constructed by the neighbour-joining (NJ) method using Molecular Evolutionary Genetics Analysis (MEGA) software (version 4.0). Bootstrap values were estimated for 1000 replicates. The reference strains used for sequence alignment, sequence analysis, and phylogenetic analysis with the Chinese PEDV strains are shown in Table 1.

Protein sequence analysis

The hydrophilic regions of the deduced amino acid sequences were analysed using DNASTAR software. The phosphorylation sites were predicted using NetPhis and NetPhos K analysis tools available at http://www.cbs.dtu.dk/services/NetPhos/ and http://www.cbs.dtu.dk/services/NetPhosK/.

Table 1 Details of the strains and the reference strains used in this study

Strain	Geographic origin	Accession no.	Strain	Geographic origin	Accession no.
LJB/03	Heilongjiang, China, 2006	DQ072726	CH-JKA1- 2011	Sanshui, Guangdong Province, China, 2011	JN173296
JS-2004-2	Jiangsu, China, 2004	AY653206	CH-QY1- 2011	Qingyuan, Guangdong Province, China, 2011	JN173297
LZC	Gansu, China, 2006	EF185992	CH-QY2- 2011	Qingyuan, Guangdong Province, China, 2011	JN173298
DX	Gansu, China, 2007	EU031893	CH-SWK1- 2011	Xinfeng, Jiangxi Province, China, 2011	JN173300
Chinju99	Korean, 2002	AF237764	CH-SWK6- 2011	Xinfeng, Jiangxi Province, China, 2011	JN173301
SM98	Korean, 2010	GU937797	CH-GD1- 2010	Yunfu, Guangdong Province, China, 2010	JN173277
CV777	England, 2001	AF353511	CH-GD2- 2010	Yunfu, Guangdong Province, China, 2010	JN173278
CH-GDXS4- 2010	Foshan, Guangdong Province, China, 2010	JN173285	CH-SIC1- 2011	Deyang, Sichuan Province, China, 2011	JQ390539
CH-GX1-2011	Guilin, Guangxi Province, China, 2011	JN173302	CH-SIC2- 2011	Deyang, Sichuan Province, China, 2011	JQ390540
CH-GX2-2011	Guilin, Guangxi Province, China, 2011	JN173303	CH-JS1-2011	Huaian, Jiangsu Province, China, 2011	JQ390543
CH-GDHSY- 2011	Yangjiang, Guangdong Province, China, 2011	JN173275	CH- JS2- 2011	Huaian, Jiangsu Province, China, 2011	JQ390544

Results

Homology analysis of the N gene

Sequence homology results were based on the fifteen Chinese PEDV strains and seven commonly used strains published in GenBank (Table 1). The fifteen Chinese field PEDV strains had 96.1-100 % nucleotide and 94.8-100 % deduced amino acid sequence identity to each other. Sequence comparison with the other seven selected strains of PEDV revealed that the Chinese field PEDV strains had nucleotide sequence identities of 94.2-99.7 % and deduced amino acid sequence identities of 94.1-99.5 %. In addition, the fifteen strains showed a high degree of nucleotide sequence identity to the early domestic strains (98.4-99.7 %) except the LZC strain, but less identity to the vaccine strain (CV777) used in China (94.7-97.7 %).

Sequence analysis of the N gene

Sequence analysis revealed that none of the fifteen strains had insertions or deletions in the N gene. In each case, the N gene had an ORF of 1326 nucleotides coding for a 441-amino-acid protein. The alignment with the Chinju99 strain indicates that the N gene is highly conserved except for some point mutations located at amino acid positions 84, 123, 142, 144, 206, 242 to 255 292, 317, 350, 369 and 394 to 413 (shown in Table 2). The Chinese strains had 15 to 18 amino acids mismatched when compared to

Chinju99. Interestingly, the sequences of the fifteen strains were highly conserved in the 5' region (bases 1 to 252), while some variation was observed in the 3' region (bases 1050 to 1233). This finding indicates that the N-terminal part of the protein is more conserved than the C-terminal part. In addition, there were four highly conserved regions at bp 1-83, 517-616, 950-1020 and 1106-1178.

Prediction of phosphorylation sites and analysis of hydrophilic regions of the N protein

The deduced amino sequence of the N gene of strain CH-GX1-2011 was randomly selected to be analyzed for specific motifs. The analysis indicated that the protein had seven potential asparagine (N)-linked glycosylation sites, consistent with the number seen in the Chinju99, LJB/03 and DX strains. Moreover, the CH-GX1-2011 strain had seven potential protein kinase C phosphorylation sites, nine casein kinase II phosphorylation sites, one tyrosine kinase phosphorylation site, and two cAMP- and cGMP-dependent protein kinase phosphorylation sites. In addition, a large hydrophilic region was identified in the central region of the protein (Fig. 1).

Phylogenetic analysis of the N gene

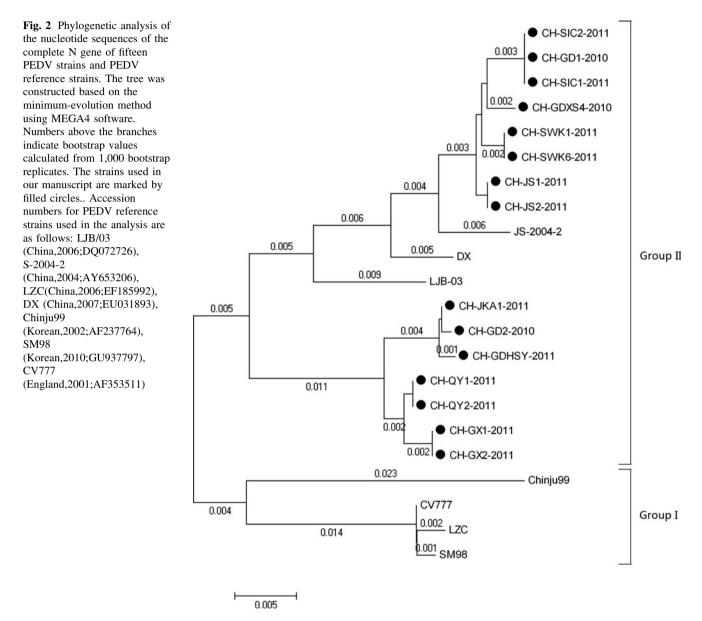
Phylogenetic analysis based on nucleotide and encoded amino acid sequences of the N gene confirmed that all field strains fell into two groups (Fig. 2). Group I consisted of Chinju99 (Korean), CV777 (Europe), LZC (China), and

ŝ
E
Ē
0 0
d to
ğ
E
å
E
ŏ
e strains c
ā
Str
e strains c
es
. <u>E</u>
5
ē
f the Chinese
of the (
S 0
5
Ξŧ.
nt
Ē
It
·5
Â
<u>p</u>
a
2
÷Ħ
e amino acid point mutation
ø
2 The amino acid point mutations of
able 2
Ĕ

Table 2 The amino acid point mutations of the Chinese strains compared to Chinju99) acid poir	it mutatior	s of the (Chinese st	rains com	pared to (Chinju99											
Strains	Position	Positions of amino acid point mutations	o acid poi	int mutatic	suc													
	84	123	142	144	206	242	246	249	252	255	292	317	350	369	394	397	400	413
Chinju99	G	К	A	٧	S	Н	Н	0	K	N	R	Ρ	I	А	I	Q	Е	Λ
CH-SWK1-2011	$\mathbf{G}{\rightarrow}\mathbf{A}$	$\mathrm{K}{\rightarrow}\mathrm{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S}\! o\! \mathbf{N}$	H→L	Ч→Q	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	$\mathrm{K}{\rightarrow}\mathrm{R}$	$\mathbf{N} \! \rightarrow \! \mathbf{S}$	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{A}$	Q→L		$V\!\rightarrow\!A$
CH-SWK6-2011	G→A	$\mathbf{K}{\rightarrow}\mathbf{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S} \! \rightarrow \! \mathbf{N}$	H→L	Ч→Q	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	$\mathrm{K}{\rightarrow}\mathrm{R}$	$\mathbf{N} \! \! \rightarrow \! \mathbf{S}$	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{A}$	Q→L		$V\!\rightarrow\!A$
СН-QY1-2011	$\mathbf{G}{\rightarrow}\mathbf{A}$	$\mathrm{K}{\rightarrow}\mathrm{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S}\! ightarrow \mathbf{N}$	H→L	Ч→Q	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	ı	ı	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{T}$	Q→L	E→D	
СН-QY2-2011	$\mathbf{G}{\rightarrow}\mathbf{A}$	$\mathrm{K}{\rightarrow}\mathrm{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S}\! o\! \mathbf{N}$	H→L	Ч→Q	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	ı	ı	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{T}$	Q→L	E→D	
CH-JKA1-2011	G→A	ı	$A\!\rightarrow\! T$	$\mathbf{V} \! \rightarrow \! \mathbf{S}$	$\mathbf{S} \! \rightarrow \! \mathbf{N}$	H→L	Ч→Q	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	ı	ı	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{T}$	Q→L	E→D	
CH-GDXS4-2010	G→A	$\mathbf{K}{\rightarrow}\mathbf{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S} \! \rightarrow \! \mathbf{N}$	H→L	H→L	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	$\mathrm{K}{\rightarrow}\mathrm{R}$	$\mathbf{N}\! \! \rightarrow \! \mathbf{S}$	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{A}$	Q→L		$V\!\rightarrow\!A$
CH-GDHSY-2011	G→A	ı	$A\!\rightarrow\! T$	$\mathbf{V} \! \rightarrow \! \mathbf{S}$	$\mathbf{S} \! \rightarrow \! \mathbf{N}$	H→L	Ч→Q	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	ı	ı	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{T}$	Q→L	E→D	
CH-SIC1-2011	$\mathbf{G}{\rightarrow}\mathbf{A}$	$\mathbf{K}{\rightarrow}\mathbf{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S} \! \rightarrow \! \mathbf{N}$	H→L	Ч→Q	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	$\mathrm{K}\!\rightarrow\!\mathrm{R}$	N→S	$R\!\rightarrow\!A$	$\mathrm{P}\!\rightarrow\!\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathbf{I}\!\rightarrow\!\mathbf{A}$	Q→L		$V\!\rightarrow\!A$
CH-SIC2-2011	$\mathbf{G}{\rightarrow}\mathbf{A}$	$\mathbf{K}{\rightarrow}\mathbf{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S} \! \rightarrow \! \mathbf{N}$	H→L	Ч→Q	K	$\mathrm{K}\!\rightarrow\!\mathrm{R}$	N→S	$R\!\rightarrow\!A$	$\mathrm{P}\!\rightarrow\!\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathbf{I}\!\rightarrow\!\mathbf{A}$	Q→L		$V\!\rightarrow\!A$
CH-GD1-2010	$\mathbf{G}{\rightarrow}\mathbf{A}$	$\mathbf{K} {\rightarrow} \mathbf{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S}\! \rightarrow\! \mathbf{N}$	H→L	Ч→О	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	$\mathrm{K}\!\rightarrow\!\mathrm{R}$	$\mathbf{N}\! \! \rightarrow \! \mathbf{S}$	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathbf{I}\!\rightarrow\!\mathbf{A}$	Q→L		$V\!\rightarrow\!A$
CH-GD2-2010	$\mathbf{G}{\rightarrow}\mathbf{A}$	ı	$A\!\rightarrow\! T$	$\mathbf{V} \! \rightarrow \! \mathbf{S}$	$\mathbf{S}\! \rightarrow\! \mathbf{N}$	H→L	Ч→О	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	ı	ı	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{T}$	Q→L	E→D	ı
CH-GX1-2011	$\mathbf{G}{\rightarrow}\mathbf{A}$	ı	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S}\! \rightarrow\! \mathbf{N}$	H→L	Ч→О	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	ı	ı	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{T}$	Q→L	E→D	ı
CH-GX2-2011	$\mathbf{G}{\rightarrow}\mathbf{A}$	ı	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S}\! \rightarrow\! \mathbf{N}$	H→L	Ч→О	$\mathbf{Q}\!\rightarrow\!\mathbf{K}$	ı	ı	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	$\mathrm{I}{\rightarrow}\mathrm{T}$	$\mathbf{A}\!\rightarrow\!\mathbf{S}$	$\mathrm{I}\!\rightarrow\!\mathrm{T}$	Q→L	E→D	ı
CH-JS1-2011	$\mathbf{G}{\rightarrow}\mathbf{A}$	$\mathbf{K} {\rightarrow} \mathbf{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$\mathbf{S}\! \rightarrow\! \mathbf{N}$	H→L	Ч→О	$Q\!\!\rightarrow\!K$	$\mathrm{K} {\rightarrow} \mathrm{R}$	$\mathbf{N}\! \! \rightarrow \! \mathbf{S}$	$R\!\rightarrow\!A$	$\mathbf{P}{\rightarrow}\mathbf{S}$	I→T	$\mathbf{A} {\rightarrow} \mathbf{S}$	I→T	Q→L	ı	ı
CH- JS2-2011	$G \! \rightarrow \! A$	$\mathrm{K}{\rightarrow}\mathrm{N}$	$A\!\rightarrow\! T$	$V\!\rightarrow\!A$	$S\! ightarrow N$	H→L	Ч→Q	Q→K	$K \!\rightarrow\! R$	N→S	$R\!\rightarrow\!A$	$\mathrm{P}{\rightarrow}\mathrm{S}$	I→T	$A {\rightarrow} S$	I→T	Q→L		



Fig. 1 Hydrophilic regions in the central region of the N protein of the CH-GX1-2011 strain



SM98 (Korean), while group II contained 18 Chinese strains. The Chinese PEDV strains had a close relationship to the three domestic strains (JS-2004-02, LJB/03 and DX) and clustered together.

Discussion

In the present study, the N genes of Chinese PEDV field strains isolated between 2010 and 2012 were amplified by

RT-PCR, cloned and sequenced to determine the genetic characteristics of viruses causing PED outbreaks in China. The results confirmed that the N gene had an ORF of 1326 nucleotides, coding for a protein of 441 amino acids. None of the Chinese strains were found to have sequence insertions or deletions in their N genes. Sequence comparison with other PEDV strains selected from GenBank indicated that the N genes of the Chinese strains were highly conserved, even though these strains originated from different geographic regions. The alignment also showed that the N

gene sequences have a high degree of nucleotide sequence identity. This could be useful information for the development of genetically engineered N proteins for vaccine development and prevention of PEDV infections.

The N protein is a phosphorylated structural protein that is associated with the viral genome and is abundant in virus-infected cells [25]. Therefore, the appearance of the N protein indicates replication of PEDV, and this can be used for early and accurate detection of virus replication in infected cells [23, 26]. Previous studies have shown that the N protein of the Chinju99 isolate has seven potential T- or S-linked phosphorylation sites [27]. In this study it was revealed that the CH-GX1-2011 strain (representing all of the 15 Chinese strains) had the same number of T- or S-linked phosphorylation sites as Chinju99 despite that fact that there were 15-18 mismatched amino acids when compared to Chinju99, LJB/03 and DX [28, 29]. Moreover, the CH-GX1-2011 strain had seven potential protein kinase C phosphorylation sites, nine casein kinase II phosphorylation sites, one tyrosine kinase phosphorylation site, and two cAMP- and cGMP-dependent protein kinase phosphorylation sites. There are larger hydrophilic regions in the center of the N protein that might play a role in transcription and replication of the viral genome.

Phylogenetic trees were constructed and analyzed using nucleotide and deduced amino acid sequences of the N gene. Similarities and differences among the PEDV strains were observed, which helped in elucidating the phylogenetic relationship between the Chinese PEDV strains and the reference PEDV strains. The results showed that the Chinese PEDV strains characterized in this study make up a separate cluster that includes three other Chinese strains (JS-2004-02, LJB/03 and DX), while they differed genetically from the vaccine strain (CV777) and the early Korean strains (Chinju99 and SM98).

In conclusion, the N genes of Chinese PEDV strains isolated in 2010-2012 during PED outbreaks were sequenced and compared to other reference strains. The N genes were found to be closely related to the N gene of earlier domestic PEDV isolates (JS-2004-02, LJB/03 and DX). The N gene was highly conserved but still had some unique point mutations as well as conserved regions. It is hoped that these data may add to other molecular epidemiological studies of PEDV in China and neighboring countries. Furthermore, it may also lay the foundation for further development and selection of PEDV vaccines.

Acknowledgments This work was supported by two grants from the Important Science and Technology Specific Projects of Guangdong Province (No. 2011B090400414 and No. 2011B020306005). We would like to thank Dr. Lianxiang Wang, Dr. Jianqiang Niu and Dr. Deming Wu for their support and comments during the preparation of this manuscript. We thank Guangdong Wen's Foodstuffs Group Co., Ltd. (China) for providing animal samples and viruses. **Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Pijpers A, van Nieuwstadt AP, Terpstra C, Verheijden JH (1993) Porcine epidemic diarrhoea virus as a cause of persistent diarrhoea in a herd of breeding and finishing pigs. Vet Rec 132: 129–131
- Pensaert MB, de Bouck P (1978) A new coronavirus-like particle associated with diarrhea in swine. Arch Virol 58:243–247
- Chen J, Wang C, Shi H, Qiu H, Liu S, Chen X, Zhang Z, Feng L (2010) Molecular epidemiology of porcine epidemic diarrhea virus in China. Arch Virol 155:1471–1476
- Puranaveja S, Poolperm P, Lertwatcharasarakul P, Kesdaengsakonwut S, Boonsoongnern A, Urairong K, Kitikoon P, Choojai P, Kedkovid R, Teankum K, Thanawongnuwech R (2009) Chinese-like strain of porcine epidemic diarrhea virus, Thailand. Emerg Infect Dis 15:1112–1115
- Sueyoshi M, Tsuda T, Yamazaki K, Yoshida K, Nakazawa M, Sato K, Minami T, Iwashita K, Watanabe M, Suzuki Y (1995) An immunohistochemical investigation ofporcine epidemic diarrhea. J Comp Pathol 113(1):59–67
- Song DS, Kang BK, Oh JS, Ha GW, Yang JS, Moon HJ, Jang YS, Park BK (2006) Multiplex reverse transcription-PCR for rapid differential detection of porcine epidemic diarrhea virus, transmissible gastroenteritis virus, and porcine group A rotavirus. J Vet Diagn Invest 18:278–281
- Martelli P, Lavazza A, Nigrelli AD, Merialdi G, Alborali LG, Pensaert MB (2008) Epidemic of diarrhoea caused by porcine epidemic diarrhoea virus in Italy. Vet Rec 162:307–310
- Song Daesub, Park Bongkyun (2012) Porcine epidemic diarrhea virus: a comprehensive review of molecular epidemiology, diagnosis, and vaccines. Virus genes 44:167–175
- Li ZL, Zhu L, Ma JY, Zhou QF, Song YH, Sun BL, Chen RA, Xie QM, Bee YZ (2012) Molecular characterization and phylogenetic analysis of porcine epidemic diarrhea virus (PEDV) field strains in south China. Virus Genes 45(1):181–185
- Fan JH, Zuo YZu, Li JH, Pei LH (2012) Heterogeneity in membrane protein genes of porcine epidemic diarrhea viruses isolated in China. Virus Genes. doi:10.1007/s11262-012-0755-4
- Sun RQ, Cai RJ, Chen YQ, Liang PS, Chen DK, Song CX (2012) Outbreak of porcine epidemic diarrhea in sucking piglets, China. Emerg Infect Dis 18(1):161–163
- Sergeev OV (2009) Porcine epidemic diarrhea. Voprosy virusologii 54:4–8
- Brian DA, Baric RS (2005) Coronavirus genome structure and replication. Curr Topic Microbiol Immunol 287:1–30
- Bridgen A, Kocherhans R, Tobler K, Carvajal A, Ackermann M (1998) Further analysis of the genome of porcine epidemic diarrhoea virus. Adv Experiment Med Biol 440:781–786
- Duarte M, Laude H (1994) Sequence of the spike protein of the porcine epidemic diarrhoea virus. J Gen Virol 75(Pt 5):1195– 1200
- Chen JF, Sun DB, Wang CB, Shi HY, Cui XC, Liu SW, Qiu HJ, Feng L (2008) Molecular characterization and phylogenetic analysis of membrane protein genes of porcine epidemic diarrhea virus isolates in China. Virus Genes 36:355–364
- Utiger A, Tobler K, Bridgen A, Ackermann M (1995) Identification of the membrane protein of porcine epidemic diarrhea virus. Virus Genes 10:137–148

- 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
- Saif LJ (1993) Coronavirus immunogens. Vet Microbiol 37(3-4):285-297
- 20. Lee C, Park CK, Lyoo YS, du Lee S (2008) Genetic differentiation of the nucleocapsid protein of Korean isolates of porcine epidemic diarrhoea virus by RT-PCR based restriction fragment length polymorphism analysis. Vet J 178:138–140
- Li BX, Ge JW, Li YJ (2007) Porcine aminopeptidase N is a functional receptor for the PEDV coronavirus. Virology 365: 166–172
- Duarte M, Gelfi J, Lambert P, Rasschaert D, Laude H (1993) Genome organization of porcine epidemic diarrhoea virus. Adv Expt Med Biol 342:55–60
- Li BX, Ma GP, Ge JW, Li YJ (2009) Porcine aminopeptidase N is a functional receptor for the PEDV coronavirus. Chin J Virol 25:220–225
- Tobler K, Bridgen A, Ackermann M (1993) Sequence analysis of the nucleocapsid protein gene of porcine epidemic diarrhoea virus. Adv Expt Med Biol 342:49–54

- 25. 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/Japan Soc Vet Sci 61:827–830
- 26. Liu DQ, Ge JW, Qiao XY, Jiang YP, Liu SM, Li YJ (2012) Highlevel mucosal and systemic immune responses induced by oral administration with Lactobacillus-expressed porcine epidemic diarrhea virus (PEDV) S1 region combined with Lactobacillusexpressed N protein. Appl Microbiol Biotechol 93:2437–2446
- Lee HK, Yeo SG (2003) Cloning and sequence analysis of the nucleocapsid gene of porcine epidemic diarrhea virus Chinju99. Virus Genes 26:207–212
- Junwei G, Baoxian L, Lijie T, Yijing L (2006) Cloning and sequence analysis of the N gene of porcine epidemic diarrhea virus LJB/03. Virus Genes 33:215–219
- 29. Li JQ, Liu JX, Lan X, Cheng J, Wu R, Lou ZZ, Yin XP, Li XR, Li BY, Yang B, Li ZY (2009) Cloning the structure genes and expression the N gene of porcine epidemic diarrhea virus DX. Virologica Sinica 24(3):179–186