

# Molecular characterization and phylogenetic analysis of membrane protein genes of porcine epidemic diarrhea virus isolates in China

Jian-Fei Chen · Dong-Bo Sun · Cheng-Bao Wang · Hong-Yan Shi ·  
Xiao-Chen Cui · Sheng-Wang Liu · Hua-Ji Qiu · Li Feng

Received: 3 November 2007 / Accepted: 27 December 2007 / Published online: 24 January 2008  
© Springer Science+Business Media, LLC 2008

**Abstract** Six porcine epidemic diarrhea viruses (PEDVs) were isolated from the fecal samples of piglets infected with PEDV in 2006 in China. The membrane (M) protein genes of six PEDV isolates were amplified by reverse transcriptase-polymerase chain reaction (RT-PCR), then cloned, sequenced, and compared with each other as well as those ten PEDV reference strains. The M protein genes of six Chinese PEDV isolates consisted of 692 nucleotides containing a single open reading frame (ORF) of 681 nucleotides, which encoded a 226aa-long peptide. The conserved intergenic motif (ATAAAC), as previously recognized in Br1/87, was found in the 5 nucleotides upstream of the initiator ATG of M protein genes of six Chinese PEDV isolates. The hexamer motif was also found in CV777, JMe2, LZC, and QH. The M protein of six isolates had three main transmembrane domains (aa20–38, aa43–65, aa75–97). The M protein of one isolate, CH/IMT/06, had one potential glycosylation site, but those of the other five isolates had two. The glycosylation sequence Asn-Phe-Thr was highly conserved in the M proteins of six PEDV isolates. The six PEDV isolates showed nucleotide sequence homology between 98.8 and 100% and deduced amino acid sequence homology between 98.2 and 100%

with each other. The nucleotide and amino acid identity of M protein genes between the six PEDV isolates and ten reference PEDV strains varied from 97.2 to 99.4% and 96.9 to 100%, respectively. On the basis of the phylogenetic relationship of M protein genes, six Chinese PEDV isolates composed of a separate cluster including one Chinese strain JS-2004-02, however, not including the Chinese strain LJB/03. These results demonstrated that there was a new genotype of PEDV prevailing in China.

**Keywords** Porcine epidemic diarrhea virus · M protein genes · Transmembrane domains · Glycosylation site · Phylogenetic relationship · Genotype

## Introduction

Porcine epidemic diarrhea (PED), caused by porcine epidemic diarrhea virus (PEDV), is an acute and highly contagious enteric disease in swine. The disease is characterized by severe enteritis and watery diarrhea, leading to death with a mortality rate of up to 80%. PED was first reported in England in 1971 [1], since then it has been reported in many other swine-producing countries [2]. In China, PED was first reported in 1976 [3]. So far, it has been prevailing in many provinces and become one of the most important viral diarrhea diseases in China, resulting in heavy economic losses in swine industry.

PEDV was first reported in Belgium and the United Kingdom in 1978 [4, 5]. It is a member of Group 1a, genus *Coronavirus*, family *Coronaviridae* within the Order *Nidovirales*. PEDV is an enveloped, single-stranded, positive-sense RNA virus. The genome of it is approximately 28 kb in length and encodes four mainly structural proteins, which are the spike (S) protein, the membrane (M) protein,

J.-F. Chen · D.-B. Sun · C.-B. Wang · H.-Y. Shi · X.-C. Cui ·  
H.-J. Qiu · L. Feng (✉)

Division of Swine Infectious Diseases, National Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, 427 Maduan Street, Harbin 150001, People's Republic of China  
e-mail: fl@hvri.ac.cn

S.-W. Liu  
Division of Avian Infectious Diseases, National Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, 427 Maduan Street, Harbin 150001, People's Republic of China

the envelope (E) protein, and the nucleocapsid (N) protein [6–8].

PEDV occurs frequently in swine-raising farms in spite of using the vaccine strategy at present, damage caused by PEDV infection is continuous and serious in many provinces in China. According to an accurate analysis on genome of prevalent PEDVs, we can prevent outbreaks of the PEDV-induced diarrhea more effectively and choose the correct PEDV vaccine strain. By the phylogenetic analysis of the partial spike glycoprotein genes, the Korean PEDV isolates were divided into three groups (G1, G2, G3), which had three subgroups (G1-1, G1-2, G1-3) [9]. Like S glycoprotein gene, M protein gene is also an important component in the phylogenetic analysis of the epidemiological situation of PEDV in the field. On the basis of sequence analysis of M protein gene, Korean PEDV isolates were shown to be different from foreign PEDV strains [10]. Therefore, the purpose of the present study was to investigate the diversity among the Chinese PEDV isolates and was to find out more prevalent PEDVs in China by the sequence analysis and the phylogenetic relationship of the M protein genes with Chinese PEDV isolates and reference PEDV strains.

## Materials and methods

### Isolates

Porcine fecal samples had been taken from piglets showing watery diarrhea and dehydration on six different swine-raising farms of five provinces in China. These fecal samples had been confirmed positive for PEDV by Anigen Rapid PED Ag Test Kit (Animal Genetics Inc., Korea). PEDV positive fecal samples were diluted with phosphate-buffered saline (PBS; 0.1 M, pH 7.2) to be 10% (V/V) suspensions. The suspensions were vortexed and clarified by centrifugation for 10 min at 5,000 rpm. The supernatants were collected for the amplification of M protein gene by RT-PCR. The isolates, used for nucleotide sequence analysis, amino acid sequence analysis and phylogenetic analysis in this study, were described in Table 1.

### Extraction of viral RNA

The supernatant (200  $\mu$ l) containing PEDV was mixed with 400  $\mu$ l TRIzol Reagent (Invitrogen Corp., Carlsbad, CA) and RNA was isolated according to the manufacturer's instructions. The RNA was air-dried for 10 min and redissolved in 15  $\mu$ l diethyl-pyrocabonate (DEPC)-treated deionized water.

### Primers

Published primers [11] designed based on the published sequence of M gene were used for generating the full M genes of PEDV isolates. Briefly, P1 (5'-CCCCAGTACTG TTATTGACGTATAAAC-3') and P2 (5'-GTTTAGACT AAATGAAGCACTTTC-3') were used for PCR to amplify the M gene. The size of amplified products was 715 bp. P3 (5'-GCCATAAAGTTTCTGTTTAGACTAA-3') was used as the primer for the synthesis of complementary DNA.

### Synthesis of cDNA

For reverse transcription, 15  $\mu$ l of extracted viral RNA was mixed with 1  $\mu$ l P3 and incubated at 70°C for 10 min followed by 2 min on ice. After adding a reaction mixture consisting of 5  $\mu$ l 5 $\times$  First Strand Buffer (Fermentas International Inc., Ltd., USA), 2  $\mu$ l 2.5 mM dNTPs (TaKaRa Biotechnology Co., Ltd., Dalian, China), 100 units of M-MLV reverse transcriptase (Fermentas International Inc., Ltd., USA), 20 units RNAsin (TaKaRa Biotechnology Co., Ltd., Dalian, China), the mixture was incubated at 37°C for 2 h. The reaction was terminated by heating at 98°C for 6 min and chilling on ice. The cDNA was either stored at -70°C or amplified immediately.

### PCR

In PCR, P1 and P2 were used to amplify the M genes of six isolates. Exactly, 1  $\mu$ l cDNA was mixed with a reaction

**Table 1** Six Chinese PEDV isolates used for sequence alignment, sequence analysis, and phylogenetic analysis

Isolates	Samples	Date of isolation	Geographic origin (province)	Accession numbers
CH/HLJH/06	Feces	January, 2006	Harbin (Heilongjiang)	EU033964
CH/HNCH/06	Feces	February, 2006	Changsha (Hunan)	EU033963
CH/IMB/06	Feces	February, 2006	Baotou (Inner Mongolia)	EU033962
CH/IMT/06	Feces	March, 2006	Tongliao (Inner Mongolia)	EU033965
CH/JSX/06	Feces	February, 2006	Xuzhou (Jiangsu)	EU033967
CH/SHH/06	Feces	January, 2006	Shanghai (Shanghai)	EU033966

mixture containing 2.5  $\mu$ l 10 $\times$  Ex Taq buffer (TaKaRa Biotechnology Co., Ltd., Dalian, China), 2  $\mu$ l 2.5 mM dNTPs (TaKaRa Biotechnology Co., Ltd., Dalian, China), 0.5  $\mu$ l each specific primer (10 pmol), 0.25  $\mu$ l Ex Taq DNA polymerase (TaKaRa Biotechnology Co., Ltd., Dalian, China), and 18.25  $\mu$ l sterile deionized water. The amplification was performed with 30 cycles of denaturation at 95°C for 1 min, annealing at 50°C for 1 min, and extension at 72°C for 1 min, followed by a final extension step of 72°C for 10 min. PCR products were analyzed on 1% agarose gels and were purified using AxyPrep™ Gel Extraction Kit (Axygen Scientific, Inc., USA) according to the manufacturer's instructions.

### Cloning and sequencing

Purified PCR products were cloned into a T-tailed vector, pMD18-T (TaKaRa Biotechnology Co., Ltd., Dalian, China) and transformed using JM109 competent cells (TaKaRa Biotechnology Co., Ltd., Dalian, China) according to the manufacturer's instructions. Exactly, for cloning of cDNA, 9  $\mu$ l of purified PCR products, 1  $\mu$ l of pMD18-T (50 ng/ $\mu$ l), and 10  $\mu$ l of 2 $\times$  Solution I were mixed gently and incubated for 6 h at 16°C. The ligation-reaction mixture was then subjected to the transformation protocol, which renders cells competent through heat-shock. For transformation, a tube of 100  $\mu$ l JM109 competent cells was thawed on ice following which 20  $\mu$ l of ligation-reaction mixture was added to the tube, mixed gently and incubated on ice for 30 min. The tube was heated in a 42°C water bath for 60 s and incubated on ice for 2 min immediately, then 800  $\mu$ l room temperature SOC medium was added to the tube and mixed gently, incubated at 37°C for 1 h with vigorous shaking. 200  $\mu$ l of each transformation mixture was immediately plated onto LB agar plate containing ampicillin. The plate was incubated at room temperature until the transformation mixture had been absorbed into the agar, following which it was inverted and then incubated at 37°C overnight. Colonies grown in LB agar plate were cultured in LB medium with vigorous

shaking at 37°C overnight, and DNA was extracted using the AxyPrep™ Plasmid Miniprep Kit (Axygen Scientific, Inc., USA). PCR and restriction enzyme digestion, with enzymes such as EcoRI, followed by electrophoresis through a 1.0% agarose gels was employed for identification of recombinant DNA clones.

Three independent DNA clones of each isolate were sequenced by using M13 sequencing primers by Shanghai Sangon Biological Engineering Technology & Services CO., Ltd. (China). All sequencing reactions were performed in duplicate and all sequences were confirmed by sequencing both strands.

### Sequence analysis

Nucleotide and deduced amino acid sequences of six PEDV isolates were aligned, edited and analyzed with DNAMAN software (Version 5.2.2, Lynnon BioSoft, Inc., USA), and MegAlign software of Lasergene (version 7.1.0, DNASTAR Inc., USA). A phylogenetic tree was then generated using an alignment of M glycoprotein gene nucleotide sequences with reference PEDV strains by Jotun Hein Method in MegAlign software. The reference strains used for sequence alignment, sequence analysis, and phylogenetic analysis with Chinese PEDV isolates were described in Table 2.

## Results

### Sequence analysis

The M protein genes of six Chinese PEDV isolates consisted of 692 nucleotides containing a single open reading frame (ORF) of 681 nucleotides, which encoded a 226aa-long peptide. There were no nucleotides deletion or insertion in the ORFs of six isolates. Six isolates had a conserved intergenic motif (ATAAAC) of 5 nucleotides upstream of the initiator ATG, as previously recognized in Br1/87 [11]. The hexamer motif was also found in CV777,

**Table 2** Ten PEDV reference strains used for sequence alignment, sequence analysis, and phylogenetic analysis

Strains	Countries	Accession numbers	Strains	Countries	Accession numbers
Br1/87	British	Z24733	KPED-9 <sup>a</sup>	South Korea	AF015888
CV777	Belgium	AF353511	KPED-9F <sup>b</sup>	South Korea	AF019893
Chinju99	South Korea	DQ845249	LJB/03	China	AY608890
JMe2	Japan	D89752	LZC	China	EF185992
JS-2004-2	China	AY653205	QH	China	AY974335

<sup>a</sup> The live virus vaccine

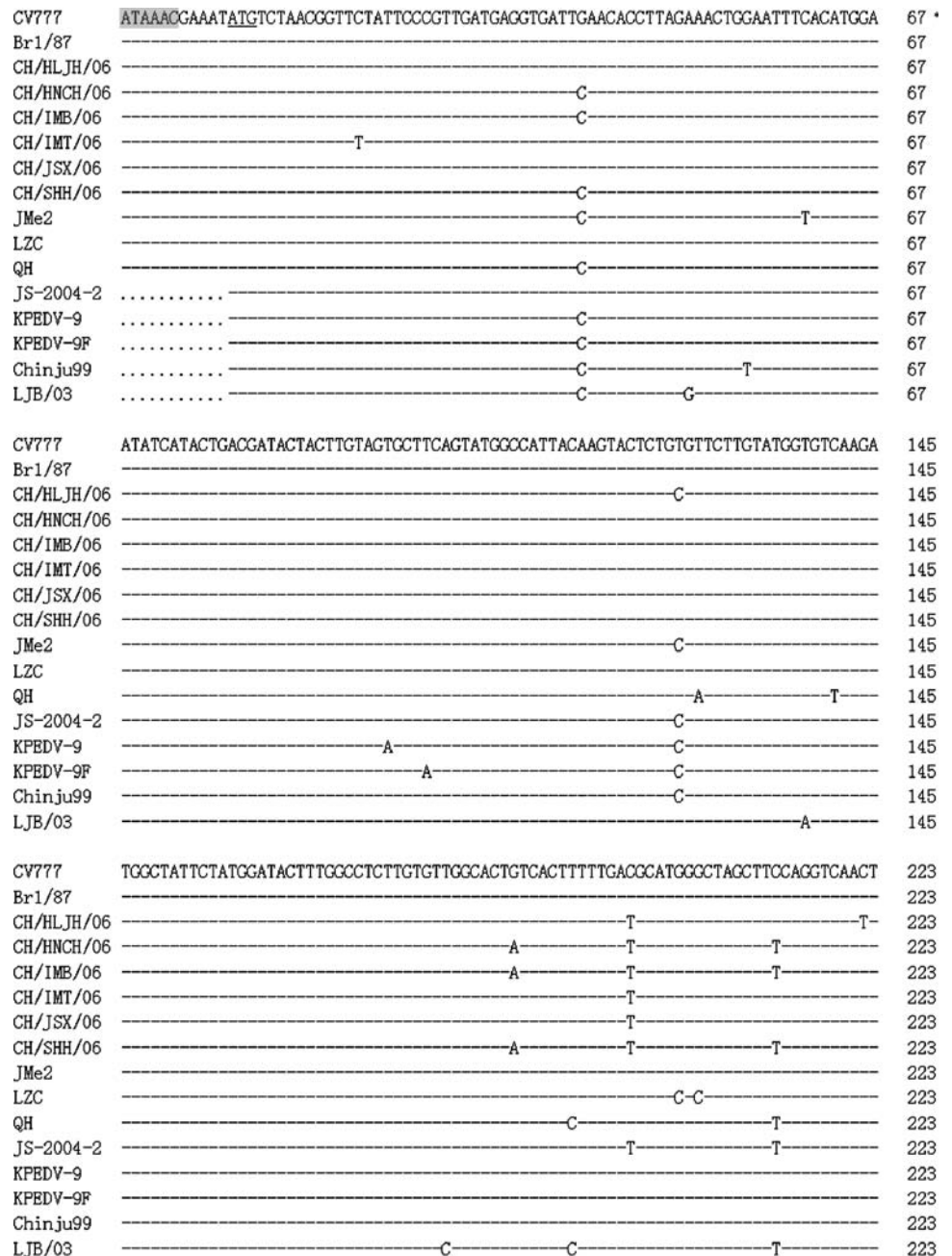
<sup>b</sup> The field virus

JMe2, LZC, and QH. There was a five-nucleotide conserved sequence (GAAAT) between the conserved intergenic motif and the initiator ATG of M protein genes of six isolates. The five-nucleotide conserved sequence was found in the corresponding regions in CV777, JMe2, LZC, and QH, too. Compared to PEDV CV777, CH/JSX/06 had 4 nucleotide mutations, CH/IMT/06 had 6 nucleotide mutations, CH/HLJH/06 had 7 nucleotide mutations, CH/IMB/06, CH/HNCH/06, and CH/SHH/06 had 8 nucleotide mutations (Fig. 1).

The nucleotide changes would lead to changes in the predict amino acid sequences of PEDV isolates. Compared

to PEDV CV777, CH/IMB/06, CH/HNCH/06, and CH/SHH/06 had 1 amino acid change (from E to Q at 13), CH/IMT/06 had 2 amino acid changes (from S to F at 5, from N to D at 167), CH/HLJH/06 had 2 amino acid changes (from V to A at 42, from A to S at 214). Although CH/JSX/06 had nucleotide changes, it had no amino acid change (Fig. 2). The M proteins of six Chinese PEDV isolates were 226 amino acid residues in length. The M proteins of five isolates had two potential N-linked glycosylation sites in their exposed amino-terminal domain, excepting for that of CH/IMT/06, which had only one N-linked glycosylation site because 1 nucleotide change (from C to T at 14) (Fig. 1)

**Fig. 1** Comparison of the nucleotide sequences of the membrane glycoprotein genes of Chinese PEDV isolates and PEDV reference strains by DNAMAN software. The dashes (–) indicate regions where the sequences are identical to those of CV777. The dots (·) indicate unavailable sequence. The asterisk (\*) indicates the numbers are calculated from the start codon. The conserved intergenic motif is in shade, the start codon and the stop codon is underlined, respectively



**Fig. 1** continued

CV777	GGGTCTTTTTCGCTTTTCAGCATCCTTATGGCTTGCATCACTCTTATGCTGGGATAATGTATTTTGTCAATAGCATT	301
Br1/87	-----	301
CH/HLJH/06	-----C-----	301
CH/HNCH/06	-----C-----	301
CH/IMB/06	-----C-----	301
CH/IMT/06	-----	301
CH/JSX/06	-----C-----	301
CH/SHH/06	-----C-----	301
JMe2	-----	301
LZC	-----	301
QH	-----C-----	301
JS-2004-2	-----T-----C-----	301
KPEDV-9	-----C-----	301
KPEDV-9F	-----	301
Chinju99	-----	301
LJB/03	-----C-----C-----C-----	301
CV777	GGTTGTGGCGCAGGACACATTCTTGGTGGTCTTTCAATCCTGAAACTGAOCCGCTTCTCACTACTTCTGTGATGGGCC	379
Br1/87	-----	379
CH/HLJH/06	-----	379
CH/HNCH/06	-----C-----	379
CH/IMB/06	-----C-----	379
CH/IMT/06	-----	379
CH/JSX/06	-----	379
CH/SHH/06	-----C-----	379
JMe2	-----	379
LZC	-----	379
QH	-----C-----A-----	379
JS-2004-2	-----A-----	379
KPEDV-9	-----	379
KPEDV-9F	-----	379
Chinju99	-----	379
LJB/03	-----A-----	379
CV777	GACAGGCTGCATTCCAGTGCCTTGGAGCACCAACTGGTGTAAACGCTAACACTCCTTAGTGGTACATTGCTTGTAGAGG	457
Br1/87	-----T-----	457
CH/HLJH/06	-----	457
CH/HNCH/06	-----	457
CH/IMB/06	-----	457
CH/IMT/06	-----	457
CH/JSX/06	-----	457
CH/SHH/06	-----	457
JMe2	-----	457
LZC	-----	457
QH	-----	457
JS-2004-2	-----	457
KPEDV-9	-----AC-----	457
KPEDV-9F	-----AC-----	457
Chinju99	-----T-----	457
LJB/03	-----	457

destroying *N*-linked glycosylation site through amino acid change (from S to F at 5) (Fig. 2) (predicted by <http://www.cbs.tdu.dk/CBS> Prediction Servers/NetNGlyc 1.0 Server). The glycosylation sequence Asn-Phe-Thr (NFT) was highly conserved in the M proteins of six isolates. The M proteins of six isolates were highly hydrophobic, and had four hydrophobic domains alternating with short hydrophilic regions in the amino-terminal half of the protein. The locations of three main transmembrane regions were aa20–38, aa43–65 and aa75–97, respectively, but the location of the fourth hydrophobic domain was not shown (predicted by <http://www.cbs.tdu.dk/CBS> Prediction Servers/TMHMM

Server v. 2.0). In the center of the M protein, directly adjacent to the third hydrophobic domain, was a stretch of eight amino acids (SWWSFNPE) (Fig. 2), which was well conserved in the coronavirus M proteins [12].

#### Sequence homology analysis

Nucleotide and deduced amino acid sequence homology results are described in Table 3. The six Chinese PEDV isolates have 98.8–100% DNA sequence identities with each other. Among the isolates, CH/IMB/06, CH/HNCH/

Fig. 1 continued

CV777	GCTATAAGGTTGCTACTGGCGTACAGGTAAGTCAATTACCTAATTTTCGTCACAGTCGCCAAGGCCACTACAACAATTG	535
Br1/87	-----	535
CH/HLJH/06	-----	535
CH/HNCH/06	-----	535
CH/IMB/06	-----	535
CH/IMT/06	-----G G-----	535
CH/JSX/06	-----	535
CH/SHH/06	-----	535
JMe2	-----	535
LZC	-----	535
QH	-----	535
JS-2004-2	-----	535
KPEDV-9	-----	535
KPEDV-9F	-----G CG-----	535
Chinju99	-----	535
LJB/03	-----	535
CV777	TCTACGGACGTGTTGGTCGTTCAAGTCAATGCTTCATCTGGCACTGGTGGGTTTCTATGTCCGGTCAAACACGGCGG	613
Br1/87	-----	613
CH/HLJH/06	---T-----	613
CH/HNCH/06	---T-----	613
CH/IMB/06	---T-----	613
CH/IMT/06	---T-----	613
CH/JSX/06	---T-----	613
CH/SHH/06	---T-----	613
JMe2	-----A-----	613
LZC	-----CCG-----	613
QH	-----A-----	613
JS-2004-2	---T-----T-----	613
KPEDV-9	-----A-----	613
KPEDV-9F	-----A-----	613
Chinju99	---T-----A-----	613
LJB/03	-----G-----A-----	613
CV777	ACTATTCAGCTGTGAGTAATCCGAGTCCGGTTCTCACAGATAGTGAGAAAGTGCTTCATTACTCTAA	681
Br1/87	-----	681
CH/HLJH/06	---C-----T-----	681
CH/HNCH/06	---C-----	681
CH/IMB/06	---C-----	681
CH/IMT/06	---C-----	681
CH/JSX/06	---C-----	681
CH/SHH/06	---C-----	681
JMe2	---C-----	681
LZC	-----TT-----	681
QH	---C-----	681
JS-2004-2	---C-----T-----	681
KPEDV-9	---C-----	681
KPEDV-9F	---C-----G-----	681
Chinju99	---C-----C-----	681
LJB/03	---C-----C-----	681

06, and CH/SHH/06 have 100% sequence homologies with each other, while CH/IMT/06 has 98.8% homology with them. Six isolates have 97.2–99.4% DNA sequence identities with ten PEDV reference strains. CH/HLJH/06 has the lowest sequence identities (97.2%) with strain LJB/03, and CH/JSX/06 has the highest sequence identities (99.4%) with strain CV777. The nucleotide identities of six isolates and four Chinese PEDV strains, JS-2004-2, LJB/03, LZC, and QH, are less than 90.0%. The isolates have 98.2–100% homologies with deduced amino acid sequences of each other and have 96.9–100% homologies with deduced amino acid sequences of ten reference strains.

#### Phylogenetic relationship

The Chinese PEDV isolates and ten PEDV reference strains were divided into three clusters based on the phylogenetic relationship of the nucleotide sequences of M protein genes (Fig. 3). One cluster (I) comprised three Korean strains (KPEDV-9, KPEDV-9F, Chinju99), two Chinese strains (LJB/03, QH), and one Japanese strain JMe2. The second cluster (II) consisted of CV777, Br1/87 and one Chinese strain LZC. The third cluster (III) was composed of six Chinese isolates and one Chinese strain, JS-2004-02.

**Fig. 2** Comparison of the deduced amino acid sequences of the membrane glycoprotein genes of Chinese PEDV isolates and reference PEDV strains by DNAMAN software. The dashes (–) indicate regions where the sequences are identical to those of CV777. The potential glycosylation sites predicted by NetNGlyc 1.0 Server are shaded. Three transmembrane domains (aa20–38, aa43–65, and aa75–97) predicted by TMHMM Server v. 2.0 are indicated by single underline, respectively. The conserved stretch of eight amino acids (SWWSFNPE) is indicated by double underline

CV777	<u>MSNGSIPVDEVIEHLRNWNFTWNIILTLVLLVQLQYGHYKYSVFLYGVKMAILWILWPLVLALSPLDAWASFPQVNVV</u>	76
QH	-----Q-----I-----	76
CH/IMT/06	-----F-----	76
CH/SHH/06	-----Q-----	76
CH/JSX/06	-----	76
CH/IMB/06	-----Q-----	76
CH/HLJH/06	-----A-----	76
CH/HNCH/06	-----Q-----	76
Br1/87	-----	76
Chinju99	-----Q-----C-----A-----	76
JMe2	-----Q-----A-----	76
JS-2004-2	-----A-----	76
KPEDV-9	-----Q-----E-----A-----	76
KPEDV-9F	-----Q-----A-----	76
LJB/03	-----Q-----S-----	76
LZC	-----SP-----	76
CV777	<u>FFAFSILMACITLMLWIMYFVNSIRLWRRTHS<u>SWWSFNPE</u>DALLTTSVMGRQVCIPVLGAPTGVTLTLLSGTLLVE</u>	152
QH	-----	152
CH/IMT/06	-----	152
CH/SHH/06	-----	152
CH/JSX/06	-----	152
CH/IMB/06	-----	152
CH/HLJH/06	-----	152
CH/HNCH/06	-----	152
Br1/87	-----F-----	152
Chinju99	-----L-----	152
JMe2	-----	152
JS-2004-2	-----	152
KPEDV-9	-----I-----D-----	152
KPEDV-9F	-----D-----	152
LJB/03	S-----T-----	152
LZC	-----	152
CV777	<u>GYKVATGVQVSQLPNPFVTVAKATTTIVYGRVGRSVNASSGTGWAFYVRSKHGDYSAVSNPSAVLTDSEKVLHLV</u>	226
QH	-----	226
CH/IMT/06	-----D-----	226
CH/SHH/06	-----	226
CH/JSX/06	-----	226
CH/IMB/06	-----	226
CH/HLJH/06	-----S-----	226
CH/HNCH/06	-----	226
Br1/87	-----	226
Chinju99	-----V-----	226
JMe2	-----	226
JS-2004-2	-----S-----	226
KPEDV-9	-----	226
KPEDV-9F	-----R-----	226
LJB/03	-----T-----	226
LZC	-----P-----I-----	226

**Discussion**

The M protein is a multispinning membrane protein with a small, amino-terminal domain located on the exterior of the virion, or, intracellularly, in the lumen of the endoplasmic reticulum (ER). The ectodomain is followed by three transmembrane segments and then a large carboxy terminus comprising the major part of the molecule. This latter domain is situated in the interior of the virion or on the cytoplasmic face of intracellular membranes. The M protein is the most abundant viral constituent and is almost

invariably glycosylated. Intriguingly, whereas the group 1 and group 3 coronaviruses—with transmissible gastroenteritis virus (TGEV) and infectious bronchitis virus (IBV) as important representatives, respectively—all contain M proteins with only *N*-linked sugars, the M proteins of group 2 coronaviruses such as mouse hepatitis virus (MHV) are *O*-glycosylated [13]. An exception is MHV-2, the M protein of which carries both *O*- and *N*-linked sugars [14].

The M protein plays a central role in coronavirus assembly. Together with the minor E protein it is responsible for the assembly of the coronavirus envelope [15–17].

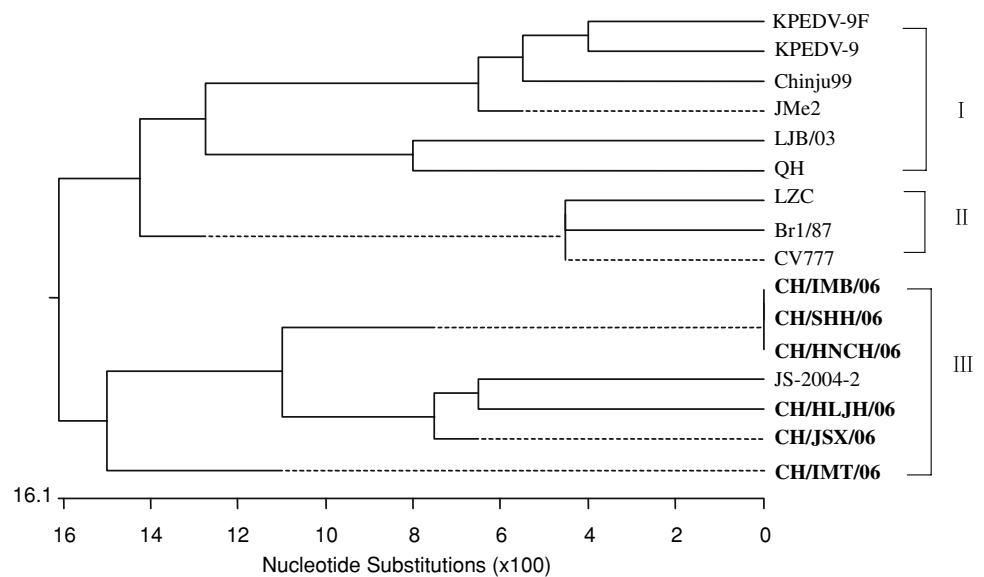
**Table 3** Comparison of the nucleotide and deduced amino acid sequences of the membrane protein genes of Chinese PEDV isolates and PEDV reference strains

Strains	Br1/87	Chinju99	CV777	JMe2	JS-2004-2	KPEDV-9	KPEDV-9F	LJB/03	LZC	QH	CH/HLJH/06	CH/HNCH/06	CH/IMB/06	CH/IMT/06	CH/JSX/06	CH/SHH/06
Br1/87	–	98.7	99.9	99.1	98.4	98.7	98.2	97.8	98.8	98.4	98.8	98.7	98.7	99.0	99.3	98.7
Chinju99	97.3	–	98.8	99.3	97.9	98.8	98.4	97.4	97.8	97.9	98.4	98.2	98.2	98.2	98.5	98.2
CV777	99.6	97.8	–	99.3	98.5	98.8	98.4	97.9	99.0	98.5	99.0	98.8	98.8	99.1	99.4	98.8
JMe2	98.7	98.7	99.1	–	98.4	99.3	98.8	97.8	98.2	98.4	98.8	98.7	98.7	98.7	99.0	98.7
JS-2004-2	98.7	97.8	99.1	99.1	–	97.9	97.5	97.7	97.5	98.2	99.0	98.5	98.5	98.5	98.8	98.5
KPEDV-9	97.3	97.3	97.8	98.7	97.8	–	99.0	97.4	97.8	97.9	98.4	98.2	98.2	98.2	98.5	98.2
KPEDV-9F	97.8	97.8	98.2	99.1	98.2	98.7	–	96.9	97.4	97.5	97.9	97.8	97.8	97.8	98.1	97.8
LJB/03	97.3	96.5	97.8	97.8	96.9	96.5	96.9	–	96.9	98.5	97.2	97.7	97.7	97.4	97.7	97.7
LZC	97.8	96.0	98.2	97.3	97.3	96.0	96.5	96.0	–	97.5	97.9	97.8	97.8	98.1	98.4	97.8
QH	98.7	97.8	99.1	99.1	98.2	97.8	98.2	97.8	97.3	–	97.8	98.2	98.2	97.9	98.2	98.2
CH/HLJH/06	98.7	97.8	99.1	99.1	100.0	97.8	98.2	96.9	97.3	98.2	–	99.0	99.0	99.0	99.6	99.0
CH/HNCH/06	99.1	98.2	99.6	99.6	98.7	98.2	98.7	98.2	97.8	99.6	98.7	–	100.0	98.8	99.4	100.0
CH/IMB/06	99.1	98.2	99.6	99.6	98.7	98.2	98.7	98.2	97.8	99.6	98.7	100.0	–	98.8	99.4	100.0
CH/IMT/06	98.7	96.9	99.1	98.2	98.2	96.9	97.3	96.9	97.3	98.2	98.2	98.7	98.7	–	99.4	98.8
CH/JSX/06	99.6	97.8	100.0	99.1	99.1	97.8	98.2	97.8	98.2	99.1	99.1	99.6	99.6	99.1	–	99.4
CH/SHH/06	99.1	98.2	99.6	99.6	98.7	98.2	98.7	98.2	97.8	99.6	98.7	100	100.0	98.7	99.6	–

Nucleotide identity (%) in upper triangle; Deduced amino acid identity (%) in lower triangle



**Fig. 3** Phylogenetic relationship based on the membrane glycoprotein genes of Chinese PEDV isolates and PEDV reference strains. Phylogenetic tree generated by Jotun Hein Method in MegAlign software is a rooted tree. The Chinese PEDV isolates are in bold



In addition, the M protein directs the incorporation of the S protein [18, 19] and the nucleocapsid [20] into the budding particle.

In the present study, the complete nucleotide and deduced amino acid sequences of M protein genes of six Chinese PEDV isolates were determined and compared with ten PEDV reference strains in order to find genetic diversity among prevailing PEDVs in China. Although PEDV is known to have only one serotype at present, the genome of PEDV is widely to show genetic diversity, according to the analysis results using molecular biological methods [21–23]. The M protein genes of four Chinese strains (LJB/03, JS-2004-2, LJC, and QH) have been sequenced and reported. Although the phylogenetic analysis based on M protein gene of LJB/03 was reported [24], it could not reflect the phylogenetic relationship among the PEDV field strains in China. Although the bi-combined attenuated vaccine against TGEV and PEDV infection is authorized to use in swine farms, PED still occurs in the immunized swine herds in China. So it is necessary to further investigate the prevalence status of PEDV based on the phylogenetic analysis of M protein genes of PEDV field strains in China.

There are no insertion and deletion in the M protein genes of six Chinese PEDV isolates except point mutations. These point mutations may lead to genetic diversity among these isolates. Our findings show that Chinese PEDV isolates are genetically diverse in their M protein genes, both among themselves and as compared with reference strains. The hexamer motif (ATAAAC), which locates at the 5' end of M protein gene, are conserved among the six isolates. They are identical to those found in the corresponding sites in strains CV777, Br1/87, JMe2,

LJC, and QH. The M proteins of six isolates have four hydrophobic domains alternating with short hydrophilic regions in the amino-terminal half of the protein, and the fourth hydrophobic domain may have functions as a cleavable signal peptide. In the center of M proteins of six isolates, directly adjacent to the third hydrophobic domain, are also the stretches of eight amino acids (SWWSFNPE) that are well conserved as mentioned by de Haan [12]. The coronavirus M protein is almost invariably glycosylated in its exposed amino-terminal domain. *N*-linked glycosylation is initiated in the endoplasmic reticulum by the cotranslational linkage of a large oligosaccharide structure to the polypeptide at asparagine residues within the consensus sequence NXS/T (where X is any amino acid). By prediction, five isolates have two *N*-linked glycosylation sites except CH/IMT/06 that has one *N*-link glycosylation site. The M proteins of group 1 and 3 coronaviruses and SARS-CoV all contain with only *N*-linked sugars. The glycosylation status of the MHV M protein was found to influence the ability of the virus to replicate in the liver but not in the brain [25], yet that of PEDV M protein has not been reported.

Sequence homology of M protein genes indicated that six Chinese PEDV isolates were highly homologous to CV777, Br1/87, Chinju99, JMe2, and JS-2004-2 rather than to QH, LJC, LJB/03, KPEDV-9, and KPEDV-9F. Although CH/SHH/06, CH/HNCH/06, CH/IMB/06 were isolated in different provinces, they had identical nucleotide sequences. It showed that the three isolates maybe originated from the same ancestor. CH/HLJH/06 was highly homologous to JS-2004-2, but they were isolated in different provinces at different time. Although CH/IMB/06 and CH/IMT/06 were isolated in the same province, the

sequence homology between them was lower. These results indicated that the Chinese PEDV isolates had genetic diversity with respect to the M protein genes. By phylogenetic analysis, the six Chinese PEDV isolates with JS-2004-2 formed a separate group, which excluded European strains, Korean strains, Japanese strain and three Chinese strains (QH, LZC, LJB/03). The phylogenetic relationship indicated that Chinese PEDV isolates were different from foreign PEDV strains and a new genotype PEDV was prevailing in China. We expect to prevent outbreaks of the PEDV-induced diarrhea more effectively, according to choosing the more efficient PEDV vaccine strain by accurate analysis on genome of prevalent PEDVs in China. Moreover, genetic variation existence in M protein genes of Chinese PEDV isolates and more prevalent PEDVs discovery in China will form the basis for further functional exploration of PEDV.

**Acknowledgment** This work was supported by the National Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, People's Republic of China.

## References

1. J. Oldham, Pig Farming (Oct suppl), 72–73 (1972)
2. M.B. Pensaert, S.G. Yeo, in *Disease of Swine*, ed. by B.E. Straw, J.J. Zimmerman, S. D'Allaire, D.J. Taylor (The Blackwell Publishing Professional, Ames, IA, 2006), pp. 367–372
3. C. Qinghua, N. Xiaoying, Y. Chengyu, *J. Anim. Vet. Sci.* **22**(3), 22 (1992)
4. M.B. Pensaert, P. de Bouck, *Arch. Virol.* **58**, 243–247 (1978)
5. P. de Bouck, M.B. Pensaert, *Am. J. Vet. Res.* **41**, 219–223 (1980)
6. M. Duarte, H. Laude, *J. Gen. Virol.* **75**, 1195–1200 (1994)
7. H.F. Egberink, J. Ederveen, P. Callebaut, M.C. Horzinek, *J. Vet. Res.* **49**, 1320–1324 (1988)
8. F.A. Murphy, E.P.J. Gibbs, M.C. Horzinek, M.J. Studdert, *Veterinary Virology* (Academic Press, San Diego, USA, 1999)
9. S.J. Park, H.J. Moon, J.S. Yang, C.S. Lee, D.S. Song, B.K. Kang, B.K. Park, *Virus Genes* **35**, 321–332 (2007)
10. Y.Z. Chi, H.M. Kwon, H.K. Jeong, J.H. Han, *Korean J. Vet. Res.* **43**, 219–230 (2003)
11. M. Duarte, K. Tobler, A. Bridgen, D. Rasschaert, M. Ackermann, H. Laude, *Virology* **198**(2), 466–476 (1994)
12. C.A.M. de Haan, P.J.M. Rottier, *Adv. Virus Res.* **64**, 165–230 (2005)
13. P.J.M. Rottier, in *The Coronaviridae*, ed. by S.G. Siddell (Plenum Press, New York, 1995), pp. 115–139
14. Y.K. Yamada, M. Yabe, T. Ohtsuki, F. Taguchi, *Virus Res.* **66**, 149–154 (2000)
15. P. Baudoux, L. Besnardeau, C. Carrat, P. Rottier, B. Charley, H. Laude, *Adv. Exp. Med. Biol.* **440**, 377–386 (1998)
16. G.J. Godeke, C.A.M. de Haan, J.W. Rossen, H. Vennema, P.J.M. Rottier, *J. Virol.* **74**, 1566–1571 (2000)
17. H. Vennema, G.J. Godeke, J.W. Rossen, W.F. Voorhout, M.C. Horzinek, D.J. Opstelten, P.J.M. Rottier, *EMBO J.* **15**, 2020–2028 (1996)
18. V.P. Nguyen, B.G. Hogue, *J. Virol.* **71**, 9278–9284 (1997)
19. D.J. Opstelten, M.J. Raamsman, K. Wolfs, M.C. Horzinek, P.J. Rottier, *J. Cell Biol.* **131**, 339–349 (1995)
20. K. Narayanan, A. Maeda, J. Maeda, S. Makino, *J. Virol.* **74**, 8127–8134 (2000)
21. S. Kubota, O. Sasaki, K. Amimoto, N. Okada, T. Kitazima, H. Yasuhara, *J. Vet. Med. Sci.* **61**, 827–830 (1999)
22. A. Bridgen, R. Kocherhans, K. Tobler, A. Carvajal, M. Ackermann, *Adv. Exp. Med. Biol.* **440**, 781–786 (1998)
23. S.G. Yeo, M. Hernandez, P.J. Krell, E.E. Nagy, *Virus Genes* **26**(3), 239–246 (2003)
24. J.H. Fan, Y.J. Li, *Virus Genes* **30**(1), 69–73 (2005)
25. C.A.M. de Haan, M. de Wit, L. Kuo, C. Montalto-Morrison, B. L.Haagmans, S.R. Weiss, P. S.Masters, P.J.M. Rottier, *Virology* **312**, 395–406 (2003)