Virus Genes 26:2, 207–212, 2003
© 2003 Kluwer Academic Publishers. Manufactured in The Netherlands.

Cloning and Sequence Analysis of the Nucleocapsid Gene of Porcine Epidemic Diarrhea Virus Chinju99*

HEE-KYUNG LEE[‡] & SANG-GEON YEO[†]

College of Veterinary Medicine, Institute of Animal Medicine, Gyeongsang National University, Chinju 660-701, Republic of Korea

Received January 22, 2003; Accepted February 4, 2003

Abstract. The nucleocapsid (N) gene of the porcine epidemic diarrhea virus (PEDV) Chinju99 which was previously isolated in Chinju, Korea was cloned and sequenced to establish the information for the development of genetically engineered diagnostic reagents. Also, sequences of the nucleotides and deduced amino acids of the Chinju99 N gene were analyzed by alignment with those of CV777 and Br1/87. The nucleotide sequence encoding the entire N gene open reading frame (ORF) of Chinju99 was 1326 bases long and encoded a protein of 441 amino acids with predicted M_r of 49 kDa. It consisted of 405 adenine (30.5%), 293 cytosine (22.1%), 334 guanines (25.2%) and 294 thymines (22.2%) residues. The Chinju99 N ORF nucleotide sequence was 96.5% and 96.4% homologous with that of the CV777 and Br1/87, respectively. The Chinju99 N protein revealed 96.8% amino acid identity with that of Br1/87 and CV777, respectively. The amino acid sequence contained seven potential sites for threonine (T)- or serine (S)-linked phosphorylation by each protein kinase C and casein kinase II.

Key words: nucleotide sequence, N gene, PEDV

Introduction

Porcine epidemic diarrhea virus (PEDV) causes an acute infection in piglets of 1–2 weeks old, and the disease is characterized by severe enteritis and diarrhea, leading to death with mortality up to 90% [1,2]. PEDV is a member of the genus coronavirus of the family *Coronaviridae* [3]. The genome consists of a single molecule of positive-sense, single-stranded RNA, 27–32 kb in size, which is transcribed into a nested set of several 3′-coterminal subgenomic mRNAs for the production of structural and non-structural proteins [3,4].

Among structural proteins of the virion, spike (S) glycoprotein (180–220 kDa) plays an important role in the attachment of the virion on the host's receptors and penetration into the intestinal villous cells by fusion. The S glycoprotein also induces the production of neutralizing antibodies in the host [5–7], and therefore, is an important substance for the immunity against PEDV. On the other hand, nucleocapsid (N) protein (55–58 kDa) is known as a basic phosphoprotein associated with the genome [1,3,8,9], which can be the target for the accurate and early diagnosis of PEDV infection by molecular techniques. Cloning and nucleotide sequencing have been done on these genes of CV777 and Br1/87 strains [5,10]. The gene products can be the feasible alternative to develop genetically engineered vaccines and diagnostic reagents.

Since isolation of PEDV in Korea was first reported in 1993 [11], the virus has been one of the major causes for the death of suckling piglets in pig

*Nucleotide sequence data reported is available in the GenBank database under the accession no. AF237764.

[†]Author for all correspondence.
E-mail: sgyeo@nongae.gsnu.ac.kr
[‡]Present address: Department of Microbiology, College of Medicine, Ulsan University, Seoul 138-736, Republic of Korea.

farming. Park et al. [12] cloned a DNA of 750 bases from N gene of the viral RNA in swine feces, but no further studies on the viral isolation and gene cloning have been reported. In the development of genetically engineered proteins for diagnostic reagents against PEDV, molecular characterization of the N gene is rudimental that still need further elucidation. PEDV infections occur frequently in Korea, and developmental efforts should be geared toward rapid diagnosis and control of the disease. To our knowledge, nucleotide sequences of the full-length N gene of Korean PEDV isolates have not been reported.

In the present study, a DNA clone was constructed for the full-length N gene open reading frame (ORF) of PEDV isolated in Chinju, Korea. The complete sequences of nucleotides and deduced amino acids of the N gene were determined, and further analyzed with those of other PEDVs for the information in the production of genetically engineered diagnostic reagents.

Materials and Methods

Virus

A strain of PEDV, Chinju99 which was previously isolated from the intestinal tissues of piglets suffering from severe diarrhea by Virology Laboratory of Gyeongsang National University College of Veterinary Medicine, Chinju, Korea (data not shown), was used. The virus was propagated in monolayer of Vero cells grown in minimal essential medium (MEM) containing streptomycin (100 μg/ml), penicillin (100 U/ml) and trypsin (10 μg/ml) in a 5% CO_2 incubator at 37°C following the methods of Hofmann and Wyler [13].

Extraction of Viral RNA

When syncytial formation appeared in the Vero cells after propagation of the virus, the wasted MEM was removed. The cells were washed with PBS (pH 7.2) and lysed by Trizol® reagent (Invitrogen, USA) at 2 ml per tissue culture flask (25 cm^2), and homogenized by passing the cell lysate several times through a pipette. Viral RNA was extracted from the homogenate following the manufacturer's suggestions and dissolved in diethyl pyrocarbonate-treated distilled water.

Primers Used for cDNA Synthesis

A pair of sense and antisense primers was designed and aligned based on nucleotide sequences of the N gene of CV777 and Br1/87 [10,14] from the GenBank data base (National Center for Biotechnology Information, USA). The sense primer NF1 (5′CCGAGTGCGGTTCTCACAGAT3′) and antisense primer NR1 (5′CATAGCCAGGATAAGCCGGTC3′) were used to generate cDNA for the N gene of Chinju99 and relative position of the primers are shown in Fig. 2.

cDNA Synthesis for the N Gene

Synthesis of the first-strand cDNA for the N gene was carried out by reverse transcription (RT) using Superscript II® reverse transcriptase reagent kit (Invitrogen) following manufacturer's suggestions. The viral RNA was mixed with 1 μl of 100 pM of the antisense primer, 4 μl of 5X first-strand buffer, 1 μl of 10 mM dNTP mixture, 2 μl of 0.1 M DTT, 1 μl of RNase inhibitor (40 U/μl), 1 μl of reverse transcriptase (200 U/μl) and brought to 20 μl with distilled water. The reaction mixture was incubated for 50 min at 42°C, and the reaction was stopped by heat for 15 min at 70°C. To degrade RNA template, the reaction mixture was treated with RNase H (1 U) for 20 min at 37°C.

The ds-cDNA for the N gene was synthesized by polymerase chain reaction (PCR) using a reagent kit

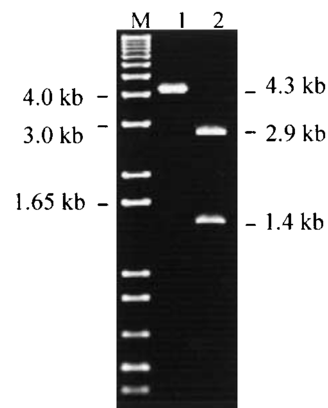

Fig. 1. Construction of recombinant DNA clone for full-length N gene of PEDV Chinju99: lane 1, Recombinant DNA (4.3 kb) between N DNA (1.4 kb) and pTZ19R plasmid DNA (2.9 kb) observed by digestion with *Sst*I; lane 2, N DNA (1.4 kb) was identified from pTZ19R plasmid DNA (2.9 kb) after digestion of the recombinant DNA with *Hin*dIII and *Sst*I; M, 1 kb plus ladder DNA marker (Invitrogen).

(Perkin-Elmer, USA). A 10 µl portion of the first-strand cDNA template was added to 5 µl of 10X PCR buffer, 4 µl of 25 mM $MgCl_2$,1 µl of 10 mM dNTP mixture, 1 µl of each 100 pM sense and antisense primers, 1 µl of *Taq* DNA polymerase (5 U/µl) and brought to 50 µl with distilled water. The PCR was carried out in a thermocycler (Perkin-Elmer) following the program of 2 min at 94°C and 30 cycles of 1 min at 94°C, 1 min at 55°C and 1 min at 72°C, and a final extension at 72°C for 5 min. The PCR products were resolved by electrophoresis in 1% agarose gel.

Cloning of cDNA

Following the routine methods in gene cloning [15], the PCR-generated N gene ds-cDNAs were blunt-ended with Klenow enzyme (2 U) and 1 µl of 0.5 mM dNTPs (Invitrogen) in 20 µl reaction volume and cloned into the *Sma*I site of pTZ19R plasmid DNA by ligation using T4 DNA ligase (1 U) (Invitrogen). The recombinant plasmid DNAs were transformed into competent *Escherichia coli* DH5α cells by heat shock for 45 s at 42°C. After adding SOC medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM $MgSO_4$, 20 mM glucose), the tube was shaken for 1 h at 220 rpm, 37°C. The transformed cells were plated onto Luria Bertani (LB) agar (Invitrogen) containing ampicillin (50 µg/ml), X-gal (40 µg/ml) and isopropylthio-B-galactoside (20 µg/ml) (Invitrogen) and incubated overnight at 37°C. Transformed colonies were cultured in LB broth with ampicillin (50 µg/ml) by shaking at 220 rpm, overnight, at 37°C, and were subjected to DNA extraction by alkaline-lysis, restriction enzyme digestion and electrophoresis in 1% agarose gel for the identification of recombinant DNA clones.

Nucleotide Sequencing

Nucleotide sequencing was done for the N gene-recombinant DNA clones using Dye Terminator Cycle Sequencing kit (Perkin-Elmer) by the automatic sequencer (ABI prism 377, Advanced Biotechnologies, USA).

Analysis on Sequences of Nucleotides and Amino Acids

The sequences of nucleotides and deduced amino acids were analyzed by ClustalW, version 1.82 using data available from GenBank and the European Molecular Biology Laboratory (EMBL). N gene nucleotide and amino acid sequences of Chinju99 were compared with CV777 and Br1/87 [10] (EMBL accession No. Z14976). The protein chemistry of Chinju99 amino acids was analyzed using protein statistic programs PEPSTATS (Pasteur Institute, France) and PredictProtein (EMBL).

Results

Cloning of N Gene

In the synthesis of ds-cDNA of the Chinju99 N gene, a DNA fragment of 1.4 kb in approximate was amplified by RT-PCR using primers specific to N gene of PEDV. The DNA was cloned into pTZ19R vector DNA (Fig. 1) and subjected to sequencing.

Analysis of N Gene Nucleotide and Amino Acid Sequences

The nucleotide sequence encoding the entire Chinju99 N gene was 1326 bases in length and contained a single ORF. The gene had 46 and 48 nucleotide mismatches compared to CV777 and Br1/87, respectively (Fig. 2). It consisted of 405 adenine (30.5%), 293 cytosine (22.1%), 334 guanine (25.2%) and 294 thymine (22.2%) nucleotides, and a GC content of 47.3%. The gene showed 96.5% and 96.4% nucleotide sequence homology to that of CV777 and Br1/87, respectively.

The Chinju99 N gene encoded a protein of 441 amino acids with predicted M_r of 49 kDa. There were seven potential threonine (T)- or serine (S)-linked phosphorylation sites by each protein kinase C and casein kinase II recognized in the protein. The Chinju99 N protein had 14 amino acid mismatches compared to those of CV777 and Br1/87 (Fig. 3) and showed 96.8% amino acid sequence identity with these strains.

Discussion

Bridgen et al. [10] previously cloned a gene of 1326 nucleotides in a single large ORF capable of encoding a 441 amino acid protein of 49 kDa from PEDV CV777 and Br1/87, which were very similar

```
                                                   NF1
                                        -59 CCGAGTGCGGTTCTCACAGAT -39 ------------

Chinju99   1  ATGGCTTCTGTCAGCTTTCAGGATCGTGGCCGCAAACGGGTGCCATTATCCCTCTATGCCCCTCTTAGGGTTACTAATGATAAACCCCTTTCTAAGGTAC   100
CV777      1                                            T                            C G                    100
Br1/87     1                                            T                            C G                    100

Chinju99 101  TTGCAAACAACGCTGTACCCACTAACAAGGGGAATAAGGACCAGCAAATTGGGTACTGGAATGAGCAAATTCGCTGGCGTATGCGCCGTGGTGAGCGAAT   200
CV777    101                                                        C                                       200
Br1/87   101                                                        T                                       200

Chinju99 201  TGAACAACCTTCAAATTGGCATTTCTACTACCTCGGAACAGGACCTCACGGCGACCTCCGTTATAGGACTCGTACTGAGGGTGTTTTCTGGGTCGCTAAA   300
CV777    201         C                                                                         T            300
Br1/87   201         C                                                                         T            300

Chinju99 301  GAAGGCGCAAAAACTGAACCCACTAACCTGGGTGTCAGAAAGGCGTCTGAAAAGCCAATCATTCCAAAATTCTCTCAACAGCTCCCCAGTGTAGTTGAAA   400
CV777    301        G        TT                                                                     G       400
Br1/87   301        G        TT                                                                     G       400

Chinju99 401  TTGTTGAACCTAACACACCTCCTGCTTCACGTGTAAATTCGCGTAGCAGGAGTCGTGGTAATGGCAACAACAGGTCCAGATCTCCGAGTAACAACAGAGG   500
CV777    401                      C              C      T   T     A                                         500
Br1/87   401                      C      CG      C      T   T     A                                         500

Chinju99 501  CAACAACCAGTCCCGTGGTAATTCACAGAATCGTGGAAATAACCAGGGTCGTGGAGCTTCTCAGAACAGAGGAGGCAATAATAATAACAATAACAAGTCT   600
CV777    501   T                                                                                            600
Br1/87   501   T                                                                                            600

Chinju99 601  CGTAACCAGTCCAAGAGCAGGAACCAGTCAAATGACCGTGGTGGTGTGACATCACGCGATGACCTGGTGGCTGCTGTCAAGGATGCCCTTAAATCTTTGG   700
CV777    601          T A                 A        T            A                                           700
Br1/87   601          T A                 A        T            A                                           700

Chinju99 701  GTATTGGAGAAAATCCTGACAGGCATAAGCAACAGCACAAGCCTCAGCAGGAAAAGTCTGACAACAGCGGGAAAAATACACCTAAGAAGAACAAATCCAG   800
CV777    701                           G  A              C                                                  800
Br1/87   701                           G  A              C                                                  800

Chinju99 801  GGCCACTTCGAAGGAACGTGACCTTAAAGACATCCCAGAGTGGAGGAGAATTCCCAAGGGCGAAAATAGCGTACGAGCTTGCTTCGGACCCAGGGGGGGC   900
CV777    801                 C                                  GC         A                                900
Br1/87   801                 C                                  GC         A                                900

Chinju99 901  TTCAAAAATTTTGGAGATGCGGAATTTGTCGAAAAAGGTGTTGATGCGCCAGGCTATGCTCAGATCGCCAGTTTAGCACCAAATGTTGCAGCATTGCTCT  1000
CV777    901       C                       T                                                               1000
Br1/87   901       C                       T                                                               1000

Chinju99 1001 TTGGTGGTAATGTGGCTGTTCGCGAGCTAGCGGACTCTTACGAGATTATATATAACTATAAAATGACTGTGCCAAAGTCTGATCCAAATGTTGAGCTTCT  1100
CV777    1001                T                 C C                                                         1100
Br1/87   1001                T                 C C                                                         1100

Chinju99 1101 TGTTGCACAGGTGGATGCATTTAAAACTGGGAATGCAAAACCCCAGAGAAAGAAGGAAAAGAAGAACAAGCGTGAAACCATTCAGCAGCAGAATGAAGAG  1200
CV777    1101   T                    T                         CG T    C                                   1200
Br1/87   1101   T                    T                         CG T    C                                   1200

Chinju99 1201 GCCATCTACGATGATGTGGGCGTGCCATCTGATGTGACCCATGCCAATCTTGAGTGGGACACAGCTGTTGATGGTGGTGATACGGCCGTTGAAATTATCA  1300
CV777    1201                T C                 G A                                                       1300
Br1/87   1201                T C                 G A                                                       1300

Chinju99 1301 ACGAGATTTTTGACACAGGAAATTAA  1326 --- 1336 GACCGGCTTATCCTGGCTATG 1356
CV777    1301       C C T               1326                NR1
Br1/87   1301       C C T               1326
```

Fig. 2. Nucleotide sequence of Chinju99 N gene and comparison of the sequence with CV777 (EMBL accession No. Z14976) and Br1/87 [10]: Initiation codon ATG and termination codon TAA were underlined; only the nucleotides of CV777 and Br1/87 which mismatched the Chinju99 sequence were included; regions corresponding to the two primers used for cloning were underlined and labeled as NF1 and NR1.

in both length and sequence to coronavirus N proteins, and therefore represented it as the PEDV N gene. In the present study, the N gene of the PEDV Chinju99 was cloned and sequencing was done for the cDNA clones. The resulting sequence data showed a single ORF of 1326 nucleotides encoding a protein of 441 amino acids with M_r of 49 kDa predicted by PEPSTATS program. Chinju99 N gene

```
Chinju99   1  MASVSFQDRGRKRVPLSLYAPLRVTNDKPLSKVLANNAVPTNKGNKDQQIGYWNEQIRWRMRRGERIEQPSNWHFYYLGTGPHGDLRYRTRTEGVFWVAK 100
CV777      1                                                          T                                           100
Br1/87     1                                                                                                      100

Chinju99 101  EGAKTEPTNLGVRKASEKPIIPKFSQQLPSVVEIVEPNTPPASRVNSRSRSRGNGNNRSRSPSNNRGNNQSRGNSQNRGNNQGRGASQNRGGNNNNNNKS 200
CV777    101                                            A                                                         200
Br1/87   101                                            A   T                                                     200

Chinju99 201  RNQSKSRNQSNDRGGVTSRDLVAAVKDALKSLGIGENPDRHKQQHKPQQEKSDNSGKNTPKKNKSRATSKERDLKDIPEWRRIPKGENSVRACFGPRGG 300
CV777    201      NN                                       Q  K                                          A         300
Br1/87   201      NN                                       Q  K                                          A         300

Chinju99 301  FKNFGDAEFVEKGVDAPGYAQIASLAPNVAALLFGGNVAVRELADSYEIIYNYKMTVPKSDPNVELLVAQVDAFKTGNAKPQRKKEKKNKRETIQQQNEE 400
CV777    301                  S                                T                  S           L             L  H   400
Br1/87   301                  S                                T                  S           L             L  H   400

Chinju99 401  AIYDDVGVPSDVTHANLEWDTAVDGGDTAVEIINEIFDTGN*                                                           441
CV777    401         A                                                                                            441
Br1/87   401         A                                                                                            441
```

Fig. 3. Putative amino acid sequence of Chinju99 N gene and comparison of the sequence with CV777 (EMBL accession No. Z14976) and Br1/87 [10]: Only the amino acids of CV777 and Br1/87 which mismatched the Chinju99 sequence were included; *, translation termination; seven potential threonine (T)- or serine (S)-linked phosphorylation sites by protein kinase C were underlined; seven potential T- or S-linked phosphorylation sites by casein kinase II were denoted in italic.

also had 96.8% amino acid sequence identity with that of CV777 and Br1/87 [10], although there were 14 amino acid mismatches recognized. Therefore, the Chinju99 N protein revealed the same features for the nucleotide and putative amino acid sequences in the CV777 and Br1/87, although PEDV N protein is known to possess M_r of 55–58 kDa by polyacrylamid gel electrophoresis [8,9].

The PEDV N protein is known as a phosphorylated, structural protein associated with viral genome [1,3,8,9], which appears abundantly in virus-infected cells [9]. Therefore, the appearance of N protein can be a clue to the replication of PEDV and used for the early and accurate diagnosis so far as the virus replicates in the infected cells. The Chinju99 N protein had each seven potential T- or S-linked phosphorylation sites by protein kinase C or casein kinase II, respectively. Similarly, the CV777 and Br1/87 [10] contained six serine (S) residues as possible phosphorylation sites by these enzymes, although some of the S-linked phosphorylation sites were different with those of the Chinju99.

In conclusion, the full-length nucleotide sequence in the coding region of N gene of PEDV Chinju99 was determined in the present study. Trials were done to analyze the nucleotide and putative amino acid sequences of the Chinju99 N gene comparing to those of other PEDVs. However, we could elucidate molecular properties of the N gene by mere comparison to those of CV777 and Br1/87, because the full-length nucleotides of the PEDV N gene have been determined only in these strains. Nevertheless, it was recognized that Chinju99 N gene has the minor differences in the structural features of putative protein compared to those of CV777 and Br1/87. This can be the feasible information for the development of genetically engineered N protein for the rapid and accurate diagnosis of PEDV infections in Korea. Moreover, the genetic information gained from the Chinju99 N gene can be used for diagnostic work such as PCR and nucleic acid hybridization. To our knowledge, this is the first published report on the full-length nucleotides and molecular characteristics of the N gene of Korean PEDV isolates.

Acknowledgements

This study was supported by a grant (No. 981-0613-065-2) from the Korea Science and Engineering Foundation (KOSEF), Ministry of Science, Korea.

References

1. Pensaert M.B., in Straw B.E., D'Allaire S., Mengeling W.L., and Taylor D.I. (eds), *Diseases of Swine*. The Iowa University Press, Ames, 1999, pp. 179–185.
2. Timoney J.F., Gillespie J.H., Scott F.W., and Barlough J.E., *Hagan and Bruner's Microbiology and Infectious Diseases of*

Domestic Animals. Comstock Publishing Associates, Ithaca, 1988, pp. 897–898.

3. Murphy F.A., Gibbs E.P.J., Horzinek M.C., and Studdert M.J., *Veterinary Virology*. Academic Press, San Diego, 1999, pp. 496–501.
4. Holmes K.V., and Lai M.M.C., in Field B.N., Knipe D.M., Howley P.M., Chanock R.M., Melnick J.L., Monath T.P., Roizman B., and Straus S.E. (eds), *Fields Virology*. Lippincott-Raven Publishers, Philadelpia, 1996, pp. 1075–1093.
5. Duarte M., and Laude H., J Gen Virol *75*, 1195–1200, 1994.
6. Spaan W., Cavanagh D., and Horziner M.C., J Gen Virol *69*, 2939–2952, 1988.
7. Utiger A., Tobler K., Bridgen A., and Ackermann M., Virus Genes *10*, 137–148, 1995.
8. Egberink H.F., Ederveen J., Callebaut P., and Horzinek M.C., Am J Vet Res *49*, 1320–1324, 1988.
9. Knuchel M., Ackermann M., Muller H., and Kihm U., Vet Microbiol *32*, 117–134, 1992.
10. Bridgen A., Duarte M., Tobler K., Laude H., and Ackermann M., J Gen Virol *74*, 1795–1804, 1993.
11. Kwon C.H., Kwon B.J., Jung T.S., Kee Y.J., Hur D.H., Hwang E.K., Rhee J.C., and An S.H., Korean J Vet Res *33*, 249–254, 1993.
12. Park J.Y., Kim C.J., Shin K.S., Kim W.Y., Kang S.Y., Park Y.H., Han H.J., and Park Y.H., Korean J Vet Res *35*, 515–530, 1995.
13. Hofmann M., and Wyler R., J Clin Microbiol *26*, 2235–2239, 1988.
14. Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M., and Laude H., Virol *198*, 466–476, 1994.
15. Ausubel F., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A., and Struhl K., *Short Potocols in Molecular Biology*. John Wiley & Sons, New York, 1992.