大韓獸醫學會誌(2007) 第47卷 第2號
Korean J Vet Res(2007) 47(2) : 169~174

Analysis for nucleotide sequence of the small membrane (sM) protein gene of porcine epidemic diarrhea virus Chinju99 isolated in Korea

Sang-Geon Yeo*, Changhee Lee[1]

College of Veterinary Medicine
[1]College of Natural Sciences, Kyungpook National University, Daegu 702-701, Korea
(Accepted: May 31, 2007)

Abstract : To provide information for the molecular pathogenesis and antigenic structures of Korean isolates of porcine epidemic diarrhea virus (PEDV), the small membrane (sM) protein gene of Chinju99 strain, which was previously isolated from piglets suffering from severe diarrhea was characterized and further analyzed with other PEDV strains. The sM gene of Chinju99 generated by reverse transcription and polymerase chain reaction had a single open reading frame with 231 bases consisting of 24.2% adenine, 18.6% cytosine, 18.1% guanine and 39.0% thymine nucleotides. Nucleotide sequence of the gene revealed 97.8% homology to those of Belgian strain CV777 and British strain Br1/87, and 97.0% to Chinese strain LZC. The gene encoded a protein with 76 amino acids, and putative amino acid sequence of the gene revealed 98.7% homology to those of CV777 and Br1/87, and 96.1% to LZC. The amino acids of Chinju99 sM gene consisted of mostly hydrophobic residues, and there were one potential N-myristylation site and one potential threonine (T)-linked phosphorylation site recognized. Also, there was a transmembrane region with 46 amino acids, and Chinju99 was more close to CV777 and Br1/87 than to LZC in phylogenetic analysis on the sM amino acid sequences.

Key words : amino acids, nucleotides, PEDV, sM gene

Introduction

Porcine epidemic diarrhea virus (PEDV) is a member of the genus *Coronavirus* in the family *Coronaviridae,* and it causes severe diarrhea and dehydration in suckling piglets, which leads to a high mortality up to 90% in 1-2 weeks old piglets [12, 13, 16]. The genome of PEDV is a positive-sense, single-stranded RNA with 28,033 nucleotides and encodes genetic information for virion proteins of important functions [8, 12].

Among the proteins, spike glycoprotein (S, M_r 180-200 KDa) plays an important role in the induction of the immune response by attaching the virion to cellular receptors and subsequent penetrating into the cells. Nucleocapsid protein (N, M_r 57-58 KDa) is the phosphoprotein that binds to viral genomic RNA to form the helical nucleocapid. Membrane gylcoprotein (M, M_r 27-32 KDa) is the matrix of the viral envelope and mediates immune reaction and course of the infection. Small membrane protein (sM, M_r 9 KDa), which together with M, is essential for virion assembly [3, 5, 12, 13].

Porcine epidemic diarrhea has been one of the enteric diseases frequently occurred in pig population in Korea since the PEDV was first reported there in 1993 [9]. In 2003, a strain of PEDV was isolated from piglets suffering from severe diarrhea and it was named as Chinju99 after its biological and physicochemical properties were first determined in Korea [10]. Lee and Yeo [11] molecularly cloned 1,326 nucleotides encoding open reading frame (ORF) of entire Chinju99 N gene, and there were 96.4-96.5% homology of the nucleotide sequences between the Chinju99, a Belgian strain

This research was supported by Kyungpook National University Research Fund, 2004.
Nucleotide sequence data reported is available in the GenBank database under accession No. EF612793.
*Corresponding author: Sang-Geon Yeo
College of Veterinary Medicine, Kyungpook National University, Daegu 702-701, Korea
[Tel: +82-53-950-5968, Fax: +82-53-950-5955, E-mail: sgyeo@knu.ac.kr]

CV777 and a British strain Br1/87. Yeo *et al*. [17] reported that the ORF of entire Chinju99 S gene consisted of 4,152 nucleotides and its nucleotide sequence was 94.5% homologous with those of the CV777 and Br1/87. Baquilod and Yeo [1] also reported nucleotide sequence of entire Chinju99 M gene with 681 bases and it was 97-99% homologous with those of CV777, Br1/87, a Korean strain KPEDV-9, a Japanese strain JMe2, and two Chinese strains JS2004-2 and LJB-03.

Therefore, nucleotide sequences of the genes for major structural proteins of the Chinju99, such as S, N and M have been unraveled, except for the sM. To our knowledge, studies on the nucleotide sequences of the sM gene of any Korean PEDV isolates have not been reported.

The sM protein, together with M protein constructs viral matrix that is essential for virion assembly, and its physicochemical properties become important information if the protein is served as an antigen for certain diagnostic assay or research on the molecular pathogenesis. Among physicochemical properties of a protein, conformational complexity gained by posttranslational modification such as glycosylation, phosphorylation, methylation, acylation and myristylation is important in expressing biological activity and proper conformation of the protein [12]. Hydrophobicity of the protein is a factor affecting accuracy in the enzyme-linked immunosorbent assay (ELISA). Hydrophobic protein usually elicits nonspecific background reactions while it is used as the antigen in the ELISA for the detection of antibodies [2].

In the present study, therefore, the nucleotide and putative amino acid sequences of the Chinju99 sM gene were determined and further analyzed with those of foreign PEDV strains to provide molecular information on the gene and its protein of the Korean PEDV isolates.

Materials and Methods

Virus

PEDV Chinju99 strain, which was previously isolated from intestines of piglets suffering from severe diarrhea in Chinju, Korea was used in the present study [10]. The virus was cultured in monolayer of Vero cells grown in minimal essential medium (Invitrogen, USA) 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 [7].

Extraction of virus genomic RNA

The medium was removed at the early stage of rounding degeneration and syncytium formation in the virus-infected Vero cells. Genomic RNA of the virus was extracted from the cells by RNeasy minikit (QIAGEN, USA). The RNA was dissolved in diethyl pyrocarbonate-treated distilled water and stored at −70°C.

Primers for cDNA synthesis

To generate cDNA for the sM gene of Chinju99, sense primer, 5'-CAACTAGACGAGTATGCTAC-3' and antisense primer, 5'-CAACGGGAATAGAACCGTTA-3' were designed and aligned to the nucleotide sequence of sM gene of CV777 (accession No. Z24733) available from the GenBank database of National Center for Biotechnology Information (NCBI, USA). A DNA with 276 bases, including entire sM gene can be amplified by these primers.

Synthesis of cDNA for sM gene

To synthesize the first-strand cDNA, reverse transcription (RT) was done for the RNA using Superscript III reverse transcriptase kit (Invitrogen, USA). A 10 µl of viral RNA was mixed with 2 µl of 100 pM antisense primer, 4 µl of 5X first-stand buffer, 1 µl of 10 mM dNTP mixture, 1 µl of 0.1 M DTT, 1 µl of RNase inhibitor, 1 µl of reverse transcriptase (200 U/µl). The reaction mixture was incubated at 45°C for 50 min and heated at 70°C for 15 min to stop the reaction. The double-stranded cDNA for the sM gene was synthesized by polymerase chain reaction (PCR) using a reagent kit (Takara, Japan). For the PCR, a 2 µl of the first-strand cDNA template was added to 5 µl of 10X PCR buffer, 2 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 (Biometra, Germany) following the program of pre-denature at 94°C for 2 min and 30 cycles of 1 min at 94°C, 1 min at 45°C and 1 min at 72°C, and a final extension at 72°C for 7 min. The PCR product was run on 1% agarose-gel in TBE buffer (0.089 M tris-borate, 0.089 M boric acid, 0.002 M EDTA, pH 8.0) by electrophoresis.

Sequencing of sM nucleotides

Direct nucleotide sequencing reaction was done by PCR with reagent mixture of 1 µl of template cDNA for the sM gene, 1 µl of each sense primer and antisense primer, 2 µl of dye reagent of fluorochrome-labeled ddNTPs, DNA polymerase and buffer, and 5 µl of distilled water. The PCR was carried out in a thermocycler (Perkin-Elmer, USA) following 30 cycles of 10 sec at 96°C, 5 sec at 50°C and 4 min at 60°C. The sequencing products were run on the gel by automated DNA analyzer (ABI Prism 3100 DNA analyzer).

Analysis of nucleotide and amino acid sequences

Nucleotide and deduced amino acid sequences of Chinju99 sM gene were analyzed and compared by computer program ClustalW (version 1.82) using data available from CV777 and Br1/87 (accession No. Z24733) [4], and a Chinese PEDV strain LZC (accession No. EF185992) registered in NCBI. The protein chemistry of Chinju99 sM amino acids was analyzed using protein statistic programs PEPSTATS (Pasteur Institute, France), PredictProtein (European Molecular Biology Laboratory, Germany), ProtScale (Swiss Institute of Bioinformatics, Swiss) and GeneBee (Russian Foundation for Basic Research, Russia).

Results

A DNA with 280 bases in approximate containing entire Chinju99 sM gene was generated by RT-PCR using primers specific to sM gene of PEDV CV777 (Fig. 1).

The nucleotide sequence of Chinju99 sM gene consisted of 231 bases in an ORF containing 56 adenine (24.2%), 43 cytosine (18.6%), 42 guanine (18.1%) and 90 thymine (39.0%) nucleotides, and a

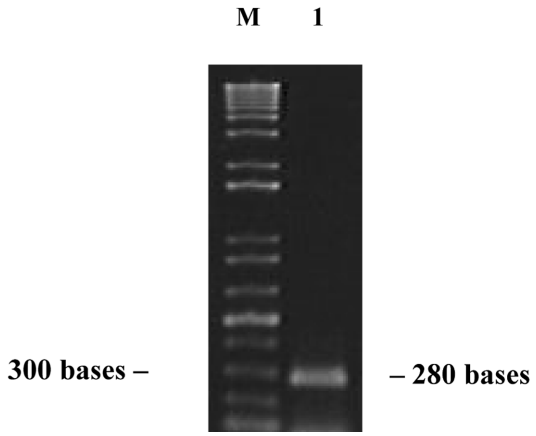

Fig. 1. A DNA synthesized from PEDV Chinju99 sM gene by RT-PCR: M, 1 kb DNA ladder marker (Invitrogen); 1, a DNA with 280 bases in approximate containing Chinju99 sM gene.

```
Chinju99  ATGCTACAAT TAGTGAATGA TAATGGTCTA GTAGTTAATG TTATACTTTG GCTTTTCGTA  60
CV777     ---------- ---------- ---------- ---------- ---------- ----------  60
Br1/87    ---------- ---------- ---------- ---------- ---------- ----------  60
LZC       ---------- ---------- ---------- ---------- ---------- ----------  60

Chinju99  CTTTTTTTTC TGCTTATTAT AAGCATTACC TTCGTCCAAT TGGTTAATCT GTGCTTTACT  120
CV777     --C-----C- ---------- ---------- ---------- ---------- ------C---  120
Br1/87    --C-----C- ---------- ---------- ---------- ---------- ------C---  120
LZC       --C-----C- ---------- ---------- ---------- ---------- ------C---  120

Chinju99  TGTCACCGGT TGTGTAATAG CGCAGTTTAC ATACCTATAG GGCGCCTGTA TAGAGTTTAT  180
CV777     ---------- ---------- ---------T -C-------- ---------- ----------  180
Br1/87    ---------- ---------- ---------T -C-------- ---------- ----------  180
LZC       ---------- ---------- ---------T -C-------- ---------- ----------  180

Chinju99  AAGTCTTACA TGCGAATTGA CCCCCTCCCC AGTACTGTTA TTGACGTATA A  231
CV777     ---------- ---------- C-C------- ---------- ---------- -  231
Br1/87    ---------- ---------- C-C------- ---------- ---------- -  231
LZC       ---------- ---------- G-G------- ---------- ---------- -  231
```

Fig. 2. Nucleotide sequence of Chinju99 sM gene and comparison of the sequence to those of other PEDV strains: both start and stop codons were expressed in bold letters; only the nucleotides of other strains mismatching the Chinju99 sequence were shown.

```
Chinju99  MLQLVNDNGL VVNVILWLFV LFFLLIISIT FVQLVNLCFT CHRLCNSAVY IPIGRLYRVY  60
CV777     ---------- ---------- ---------- ---------- ---------- T---------  60
Br1/87    ---------- ---------- ---------- ---------- ---------- T---------  60
LZC       ---------- ---------- ---------- ---------- ---------- T---------  60

Chinju99  KSYMRIDPLP STVIDV *  76
CV777     ------DP-- ------ *  76
Br1/87    ------DP-- ------ *  76
LZC       ------ER-- ------ *  76
```

Fig. 3. Putative amino acid sequence of Chinju99 sM gene and comparison of the sequence to those of other PEDV strains: one potential N-myristylation site was underlined; one potential threonine (T)-linked phosphorylation site by casein kinase II was italicized; a stretch of amino acids in shadowed area was the helical transmembrane region; *, translation termination.

Table 1. Homology of nucleotide and amino acid sequences of Chinju99 sM gene to those of other PEDV strains

% Homology of nucleotides (amino acids)			
Strains	CV777	Br1/87	LZC
Chinju99	97.8 (98.7)	97.8 (98.7)	97.0 (96.1)

GC content was 36.8%. The nucleotide sequence of Chinju99 sM gene had 5 mismatches from those of CV777 and Br1/87 at nucleotide position 63, 69, 117, 150 and 152, and 7 mismatches from LZC at nucleotide position 63, 69, 117, 150, 152, 201 and 203. Meanwhile, CV777 and Br1/87 were 100% homologous in the sM nucleotide sequence, and LZC had 2 mismatches of the sM nucleotide sequence from those of CV777 and Br1/87 at nucleotide position 201 and 203 (Fig. 2). The Chinju99 sM gene revealed 97.8% homology of nucleotide sequence to CV777 and Br1/87, and 97.0% to LZC (Table 1).

The protein encoded by Chinju99 sM gene consisted of 76 amino acids and its molecular weight was calculated as approximately 8821.7 Da. Amino acid sequence deduced from Chinju99 sM gene had 1 amino acid mismatch from those of CV777 and Brl/87 at amino acid position 51, and 3 mismatches from LZC at the position 51, 67 and 68. LZC had 2 amino acid mismatches from those of CV777 and Br1/87 at the position 67 and 68. The Chinju99 sM protein had one potential N-myristylation site starting with glycine (G) at the amino acid position 9. Also, there was one potential threonine (T)-linked phosphorylation site by casein kinase II at the position 72, and an internal stretch with 46 (60.5%) amino acids running from the position 12 to 57 was determined as the helical transmembrane region (Fig. 3). The amino acids of

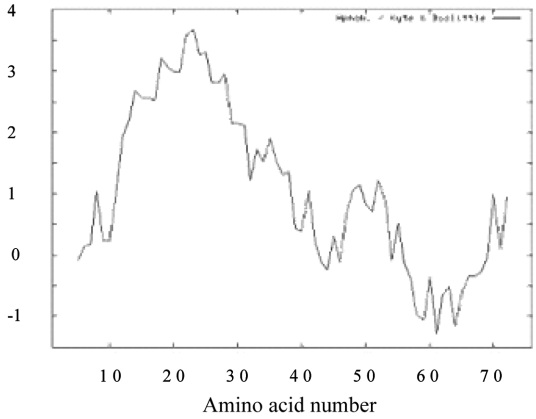

Fig. 4. Hydrophobicity predicted in Chinju99 sM amino acids determined by the Kyte-Doolittle scale: the most hydrophobic residue with score of 3.667 was found at number 23.

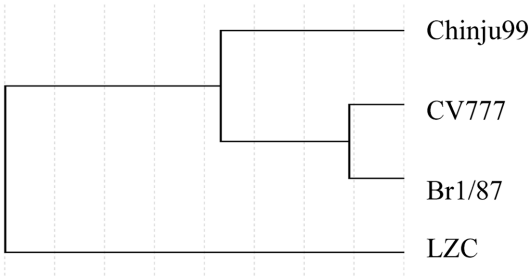

Fig. 5. Phylogenetic relationship of amino acid sequence of Chinju99 sM gene with those of other PEDV strains.

Chinju99 sM gene showed 98.7% sequence homology to CV777 and Br1/87, and 96.1% to LZC (Table 1).

The Chinju99 sM protein consisted of mostly hydrophobic amino acids, with average hydrophobicity score of 1.074 and the highest score of 3.667 at position 23 (Fig. 4). In phylogenetic tree analysis on the amino acid sequence of sM gene, Chinju99 was more close to CV777 and Br1/87 than to LZC (Fig. 5).

Discussion

The nucleotide sequence of entire Chinju99 sM gene was established in the present study, as it had a single ORF with 231 bases consisting 24.2% adenine, 18.6% cytosine, 18.1% guanine and 39.0% thymine nucleotides, and the GC content was somewhat low as 36.8%.

While sM nucleotide sequence of Chinju99 was aligned and compared to those of CV777, Br1/87 and LZC, it was recognized that the sM nucleotides were well conserved among these strains with sequence homology of 97.0-97.8%, although Chinju99 sM gene had 5-7 nucleotide mismatches from these strains. Duarte *et al*. [4] reported that the sM gene was strictly conserved between CV777 and Br1/87. On the other hand, Rasschaert *et al*. [15] and Raabe and Siddell [14] aligned sM nucleotide sequence of PEDV with those of other coronaviruses of the antigenic group I in which PEDV is included and reported that the sM nucleotide sequence of PEDV revealed a 54% homology for that of human coronavirus (HCV) 229E and a 29% for that of transmissible gastroenteritis virus. Therefore, they suggested that PEDV is more close to HCV 229E among antigenic group I coronaviruses. Meanwhile, M gene also was found well conserved among Chinju99, CV777 and Br1/87 as the Chinju99 M gene showed 98% nucleotide homology to these strains [1]. Also, Chinju99 S gene had 94.5% nucleotide homology to CV777 and Br1/87 [17], and Chinju99 N gene had 96.4-96.5% nucleotide homology to these strains [11].

Chinju99 sM gene encoded a protein with 76 amino acids and its M_r was estimated as 8821.7 Da, which was homologous to the sM protein (9 kDa) of other coronaviruses [6, 12]. The protein had only 1 amino acid mismatch from those of CV777 and Br1/87, and homology of the amino acid sequence to these strains was 98.7%. This mismatching amino acid was brought out by substitution of the nucleotide 'C' in ACA codon at nucleotide position 151-153 in CV777 and Br1/87 to 'T' of ATA codon at the same position in Chinju99, which altered the relevant amino acid from threonine (T) to isoleucine (I). Therefore, it was recognized that although CTC, TTC, TTC and TAT codons, which ended respectively with remaining 4 mismatched nucleotides at position 63, 69, 117 and 150 in CV777 and Br1/87 changed to CTT, TTT, TTT and TAC in Chinju99, all of these underwent substitution of the third nucleotide of the codon which was silent in phenotypic expression [12]. Therefore, the CTT, TTT, TTT and TAC codons were still translated into leucine (L), phenylalanine (F), phenylalanine (F) and tyrosine (Y), respectively, as those translated from CTC, TTC, TTC and TAT codons in CV777 and Br1/87.

Chinju99 sM protein had one of each potential N-myristylation site and T-linked phosphorylation site. These signals for the posttranslational modification will enable the protein to be more complex structure and have a predicted molecular mass for biological function [12]. The protein also had an internal transmembrane region spanning amino acid position 12 to 57 that is relevant to 60.5% of the total amino acids. Amino acids in the membrane-associated protein of viruses usually maintain hydrophobic properties [5]. This feature was also supported with the Chinju99 sM protein as it contained mostly hydrophobic amino acids with average hydrophobicity score of 1.074, especially those located in the predicted transmembrane region. Duarte *et al*. [4] also reported that sM protein of CV777 and Br1/87 had a long hydrophobic stretch as well as found in TGEV and HCV 229E. Likewise, Baquilod and Yeo [1] reported that Chinju99 M protein also consisted of mostly hydrophobic amino acids.

As regards phylogenetic relationship of sM amino acids, Chinju99 was more close to CV777 and Br1/87 than to LZC. In the aspect of relationship to other coronaviruses, Duarte *et al*. [4] reported that sM amino acid sequences of CV777 and Br1/87 showed a high identity with HCV 229E than with TGEV in phylogenetic alignment.

To our knowledge, this is the first published report on the nucleotide sequence of the sM gene of Korean PEDV isolates, and structural features of the Chinju99 sM protein determined in this study can be the feasible information for the molecular pathogenesis and antigenic structure of PEDV isolates.

Conclusion

The nucleotide and deduced amino acid sequences of the sM gene of PEDV Chinju99 were determined and further analyzed with those of other PEDV strains, and the results were as follows.

The Chinju99 sM gene had a single ORF with 231 bases consisting of 24.2% adenine, 18.6% cytosine, 18.1% guanine and 39.0% thymine nucleotides, which encoded a protein with 76 amino acids. The nucleotide

sequence of the gene revealed 97.8% homology to those of CV777 and Br1/87, and 97.0% to LZC. Also, the amino acid sequence deduced from the gene revealed 98.7% homology to those of CV777 and Br1/87, and 96.1% to LZC.

The Chinju99 sM protein had one of each potential N-myristylation site and potential threonine (T)-linked phosphorylation site at amino acid position 9 and 72, respectively. Also, there was a transmembrane region with 46 amino acids from the position 12 to 57. The protein consisted of mostly hydrophobic amino acids with average hydrophobicity score of 1.074, and it was more close to CV777 and Br1/87 than to LZC in phylogenetic analysis.

References

1. **Baquilod GSV**, **Yeo SG**. Analysis for nucleotide sequence of the membrane protein gene of porcine epidemic diarrhea virus Chinju99. Korean J Vet Res 2006, **46**, 355-361.
2. **Cho HJ**, **Masri SA**, **Deregt D**, **Yeo SG**, **Thomas EJG**. Sensitivity and specificity of an enzyme-linked immunosorbent assay for the detection of bovine viral diarrhea virus antibody in cattle. Can J Vet Res 1991, **55**, 56-59.
3. **Duarte M**, **Laude H**. Sequence of the spike protein of the porcine epidemic diarrhoea virus. J Gen Virol 1994, **75**, 1195-1200.
4. **Duarte M**, **Tobler K**, **Bridgen A**, **Rasschaert D**, **Ackermann M**, **Laude H.** Sequence analysis of the porcine epidemic diarrhea virus genome between the nucleocapsid and spike protein genes reveals a polymorphic ORF. Virology 1994, **198**, 466-476.
5. **Fields BN**, **Knipe DM**, **Howley PM**, **Chanock RM**, **Melnick JL**, **Monath TP**, **Roizman B**, **Straus SE**. Fields Virology. 3rd ed. pp. 62-72, 1075-1093, Lippincott-Raven, Philadelphia, 1996.
6. **Godet M**, **L'Haridon R**, **Vautherot JF**, **Laude H**. TGEV coronavirus ORF4 encodes a membrane protein that is incorporated into virions. Virology 1992, **188**, 666-675.
7. **Hofmann M**, **Wyler R**. Propagation of the virus of porcine epidemic diarrhea in cell culture. J Clin Microbiol 1988, **26**, 2235-2239.
8. **Kocherhans R**, **Bridgen A**, **Ackermann M**, **Tobler K**. Completion of the porcine epidemic diarrhoea coronavirus (PEDV) genome sequence. Virus Genes 2001, **23**, 137-144.
9. **Kweon CH**, **Kwon BJ**, **Jung TS**, **Kee YJ**, **Hur DH**, **Hwang EK**, **Rhee JC**, **An SH**. Isolation of porcine epidemic diarrhea virus (PEDV) in Korea. Korean J Vet Res 1993, **33**, 249-254.
10. **Lee HK**, **Yeo SG**. Biological and physicochemical properties of porcine epidemic diarrhea virus Chinju99 strain isolated in Korea. J Vet Clin 2003, **20**, 150-154.
11. **Lee HK**, **Yeo SG**. Cloning and sequence analysis of the nucleocapsid gene of porcine epidemic diarrhea virus Chinju99. Virus Genes 2003, **26**, 207-212.
12. **Murphy FA**, **Gibbs EPJ**, **Horzinek MC**, **Studdert MJ**. Veterinary Virology. 3rd ed. pp. 62-63, 73, 228, 495-501, Academic Press, San Diego, 1999.
13. **Pensaert MB**, **Yeo SG**. Porcine epidemic diarrhea. In: Straw BE, Zimmerman JJ, D'Allaire S, Taylor DJ (eds.). Diseases of Swine. 9th ed. pp. 367-372, Blackwell Publishing, Ames, 2006.
14. **Raabe T**, **Siddell S**. Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA 5 unique regions. Nucleic Acids Res 1989, **17**, 6387.
15. **Rasschaert D**, **Gelfi J**, **Laude H**. Enteric coronavirus TGEV: partial sequence of the genomic RNA, its organization and expression. Biochimie 1987, **69**, 591-600.
16. **Timoney J**, **Gillespie JH**, **Scott FW**, **Barlough JE**. Hagan and Bruner's Microbiology and Infectious Diseases of Domestic Animals. pp. 897-898, Comstock Publishing, Ithaca, 1988.
17. **Yeo SG, Hernandez M, Krell PJ, Nagy É.** Cloning and sequence analysis of the spike gene of porcine epidemic diarrhea virus Chinju99. Virus Genes 2003, **26**, 239-246.