CHAPTER II

LITERATURE REVIEW

Characteristic of PEDV

Porcine epidemic diarrhea virus is one of the members in family Coronaviridae, order Nidovirales, genus alpha coronavirus (Chen et al., 2011). The particle size of the virus is approximately 130 nm in diameter (95-190 nm) with genomic size of 28-32 kb (Ducatelle et al., 1982). PEDV generally replicates inside the epithelial cell of small intestinal villi and the virus 1-2 particles can be found within an infected cell (Ducatelle et al., 1982). The virion can be detected within infected pig feces with pleomorphic and spherical shapes (Pensaert and Yeo, 2006).

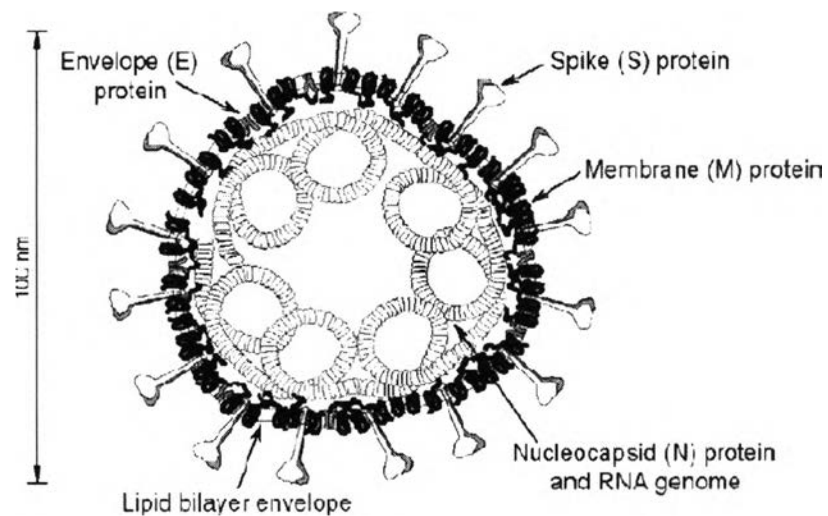

Figure1.The structure of coronavirus (Master, 2006)

Molecular biology of PEDV

PEDV is a member of genus alpha coronavirus (α-CoVs), or group 1 coronavirus of Family Coronaviridae. The other known member of alpha coronavirus such as Feline coronavirus (FCoV), Canine coronavirus (CCoV), Human coronavirus (HCoV-229E and HCoV-NL63) and Transmissible gastroenteritis virus (TGEV) in pig (Gonzalez et al., 2003). TGEV is the same enteropathogenic organism as PEDV that causes severe diarrhea especially in piglets (Song and Park, 2012).

Genus alpha coronavirus consists of two subgroups as subgroups A and B. The subgroups are separated by the number of accessory genes. Subgroup A composes with several short accessory protein encoding genes (ORFs) that are located between S and E genes, including at the 3' end of N gene. The members of subgroup A are TGEV and FCoV. For subgroup B, most of the members have only one accessory protein encoding gene (ORF) that is located between S and E genes, except some strains of Bat coronavirus (Bat CoVs). The members of subgroup B are Human coronavirus 229E (HCoV-229E), Human coronavirus strain NL63 (HCoV-NL63), Bat coronavirus strain HKU2 (Bt-CoV HKU2) and Porcine epidemic diarrhea virus (PEDV) (Gonz'alez et al., 2003). The schematic representation of the genomic organization of alpha coronavirus group has been shown in Figure2.

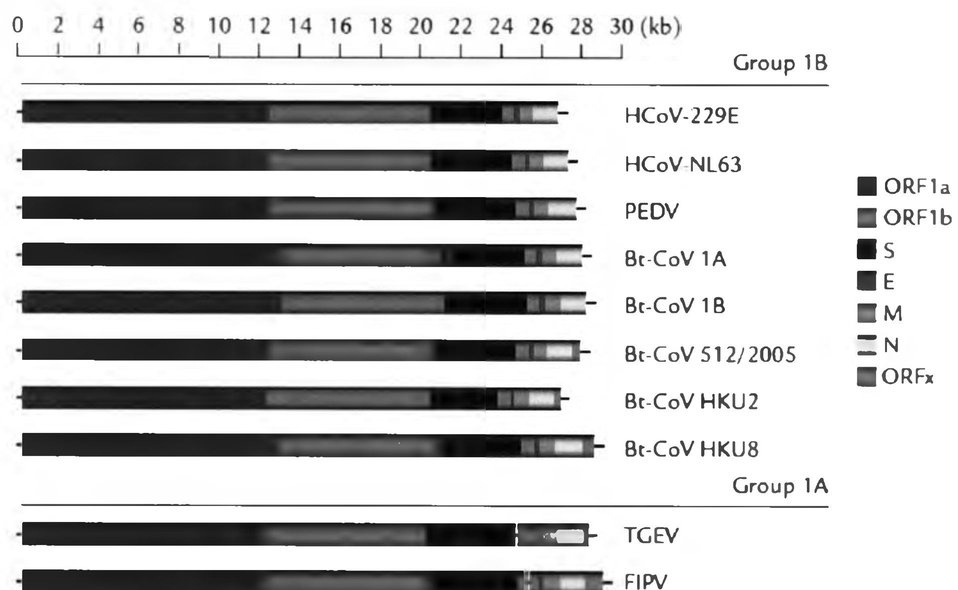

Figure2. The genomic organization of alpha coronavirus group (Dijkman and van der Hoek, 2009) (Human coronavirus 229E (HCoV-229E), Human coronavirus strain NL63 (HCoV-NL63), Porcine epidemic diarrhea virus (PEDV), Bat coronavirus strain 1A (Bt-CoV 1A), Bat coronavirus strain 1B (Bt-CoV 1B), Bat coronavirus strain 512/2005 (Bt-CoV 512/2005), Bat coronavirus strain HKU2 (Bt-CoV HKU2), Bat coronavirus strain HKU8 (Bt-CoV HKU8), Transmissible gastroenteritis virus (TGEV) and Feline coronavirus (FCoV).

หอสมุดกลาง สำนักงานวิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

PEDV is a single positive-strand RNA, with linear genome and non-segmented. The genome contains 6 genes including replicase (Rep), spike (S), open reading frame 3 (ORF3), envelope (E), membrane (M), nucleocapsid (N) gene (Chen et al., 2011), which are 4 structural proteins (S, E, M and N), and 3 non-structural proteins (replicases 1a and 1b, and ORF3). The genome is arranged as 5'- replicase (1a/1b)–S-ORF3–E–M–N-3' (Figure 3) (Song and Park, 2012).

Group	Virus species	Accessory genes (Proteins)*
1	TGEV	[rep] - [*S*] - 3a, 3b - [*E*] - [*M*] - [*N*] - 7
	FIPV	[rep] - [*S*] - 3a, 3b, 3c - [*E*] - [*M*] - [*N*] - 7a, 7b
	HCoV-229E	[rep] - [*S*] - 4a, 4b - [*E*] - [*M*] - [*N*]
	PEDV	[rep] - [*S*] - 3 - [*E*] - [*M*] - [*N*]
	HCoV-NL63	[rep] - [*S*] - 3 - [*E*] - [*M*] - [*N*]
2	MHV	[rep] - 2a, 2b(HE) - [*S*] - 4 - 5a, [*E*] - [*M*] - [*N*], 7b(I)
	BCoV	[rep] - 2a - 2b(HE) - [*S*] - 4a(4.9k), 4b(4.8k) - 5(12.7k) [*E*] - [*M*] - [*N*], 7b(I)
	HCoV-OC43	[rep] - 2a - 2b(HE) - [*S*] - 5(12.9k) - [*E*] - [*M*] - [*N*], 7b(I)
	SARS-CoV	[rep] - [*S*] - 3a, 3b - [*E*] - [*M*] - 6 - 7a, 7b - 8a, 8b - [*N*], 9b(I)
	HCoV-HKU1	[rep] - 2(HE) - [*S*] - 4 - [*E*] - [*M*] - [*N*], 7b(I)
	Bat-SARS-CoV	[rep] - [*S*] - 3 - [*E*] - [*M*] - 6 - 7a, 7b - 8 - [*N*], 9b(I)
3	IBV	[rep] - [*S*] - 3a, 3b, 3c - [*E*] - [*M*] - 5a, 5b - [*N*]

Figure3. Coronavirus accessory genes (Master, 2006)
(Coronavirus genome encodes with replicase protein gene (rep), spike protein gene (S), envelope protein gene (E), membrane protein gene (M), nucleocapsid protein gene (N), haemagglutinin esterase protein gene (HE) and accessory protein gene as open reading frame.

The replicase gene (*rep* gene) or polymerase gene (*pol* gene) which is a highly conserved region. The *rep* gene composes of two overlapping open reading frames (ORFs) as ORF1a and ORF1 b, which encode for two polymerase polyproteins (pp1a and pp1b respectively) with cleavage site (Stephensen et al., 1999; Chen et al., 2011). The cleavage site is separated into 15-16 non-structural proteins (NSPs) by specific protease enzymes (Masters, 2006; Putics et al., 2006; Woo et al., 2010). Some types of non-structural proteins encode enzyme playing an important role in tRNA processing (tRNA splicing) (Woo et al., 2010). Replicase gene is located at 5' of the coronavirus genome, approximate two of the three in length along with genome. The *rep* gene has

the overlapping region between ORF1a and ORF1b that called "slippery sequence" (Kocherhans et al., 2001; Chen et al., 2011) or "intergenic regulatory sequence" or "ribosomal frameshift sites" (Enjuanes et al., 2001; Kocherhans et al., 2001; Chen et al., 2011). The set of nucleotide sequence within this overlapping region consists of seven nucleotides, as UUUAAAC (Enjuanes et al., 2001; Kocherhans et al., 2001; Chen et al., 2011) which are required for the translation of ORF1b (Chen et al., 2011). If there is variation within this seven-nucleotide sequence, the sub-genomic mRNAs will not be produced (Chen et al., 2011). For amino acid sequence of replicase representing the ORF1b is more conserved than ORF1a (Kocherhans et al., 2001).

The amino acid sequence of ORF1a and ORF1b of PEDV is closely related to HCoV-229-E and TGEV. For ORF1a amino acid sequence of HCoV-229E and TGEV shares 59.4% and 52.1% similarity to PEDV amino acid sequence, respectively. For ORF1b amino acid sequence of HCoV-229E and TGEV shows similarity 83.2% and 80.3% to PEDV amino acid sequence, respectively (Kocherhans et al., 2001). In addition, ORF1a amino acid sequence of MHV and IBV shows the sequence identities of 39.5% and 38.7% to PEDV, respectively.

The spike gene (S gene) encoded spike protein or previous term called E2 protein (Lai, 1990). S protein is found on the surface of coronavirus particles that making the viral particle seems to be crown-like appearance under electron microscope (Woo et al., 2010).

S protein separates into sub-domain 1 and sub-domain 2 (Figure 4) (Woo et al., 2010; Shirato et al., 2011). Sub-domain 1 or N-terminal S1 subunit is located between 1-789 amino acid positions (Song and Park, 2012). Sub-domain 2 or C-terminal transmembrane S2 subunit (TM) is located between 790-1,383 amino acid positions (Woo et al., 2010). The S2 subunit is subdivided into three 3 domains as a large ectodomain (with protease cleavage site, putative fusion peptide or F region and two heptad regions or HP regions that are called HR1 and HR2 respectively), a transmembrane domain and an endodomain with cysteine-rich cytoplasmic tail or CT domain (Master, 2006; Woo et al., 2010; Shirato et al., 2011). In addition, S protein can be subdivided into various domains. For instance, a signal domain is located between 1-

18 amino acid position, neutralizing epitopes are located in 4 parts of amino acid position (499-638, 748-755, 764-771 and 1,368-1,374 amino acid), a transmembrane domain is located between 1,334-1,356 amino acid at C-terminal domain and a short cytoplasmic tail with cysteine-rich residues (Lai, 1990; Woo et al., 2010).

S protein plays the important role in receptor binding activity (by S1 subunit) (Master, 2006) and membrane fusion (by S2 subunit) between PEDV and susceptible host cell (Master, 2006). It is also associated with tissue tropism and production of neutralizing antibodies (Enjuanes et al., 2001; Kim et al., 2001; Chang et al., 2002), wherefore S protein could be the target for vaccine development to neutralize coronaviruses (Chang et al., 2002; Song and Park, 2012).

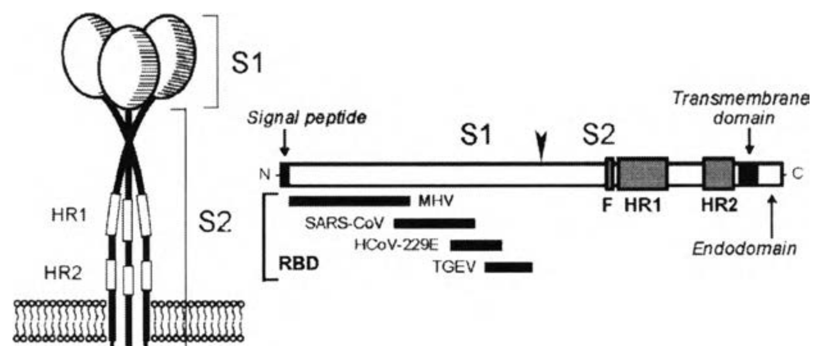

Figure4. Spike protein structure of coronavirus (Master, 2006). S1 subunit located at the N-terminal part of the molecule. S2 portion located next to S1 portion and towards to the right part of C-terminal.

For PEDV neutralization activity, S protein has collagenase-digested fragment (CO-26K) or COE region where contains neutralizing epitope (Chang et al., 2002). This region is highly conserved in PEDV strains (Chang et al., 2002). The CO-26K region of PEDV shared only 28% homology to CO-26K region of TGEV. Therefore the cross neutralized immunity between PEDV and TGEV is not occurred (Chang et al., 2002).

S protein is required for M and E protein co-expression for viral binding to host cell receptor (Lai, 1990), and for promoting the viral budding into intracellular vesicles within infected host cell (Shirato et al., 2011). In addition, S protein associates with viral adaptation *in vitro* including viral attenuation of virulence *in vivo* (Song and Park, 2012).

Therefore, S protein is important target for studying on genetic relationship and diversity between PEDV strains, genetic mutations, PEDV epidemiology and viral function (Song and Park, 2012).

The envelope gene (E gene) encoded envelope protein (E protein). E protein involves in envelope organization (Woo et al., 2010) and serves the neck of the viral particle between the final stages of viral budding process (Master, 2006). E gene is conserved among members in coronavirus groups but it is not often used for phylogenetic studies. The main reason is because of its containing a short sequence (Master, 2006).

The membrane gene (M gene) encoded membrane protein (M protein) or E1 protein (Lai, 1990). M protein is the most common envelope protein that is found on the viral surface (Zhang et al., 2012). M protein is the membrane-spanning glycoprotein with two domains, ectodomain and endodomain (Figure 5) (Master, 2006). The ectodomain composes with three transmembrane segments (tm1, tm2 and tm3) and short amino acid terminal domain (Master, 2006; Zhang et al., 2012). The endodomain with long carboxy-tail is located between 99-226 amino acid positions (Zhang et al., 2012). The amino acid positions at the last 25 residues of tm3 within ectodomain to the upstream of endodomain are the most conserved part in all coronaviruses (Master, 2006).

Figure5. Membrane protein structure (Masters, 2006)

The roles of M protein associate with viral assembly process (S and N protein fusion) (Park et al., 2011; Sato et al., 2011), promote viral budding process (Zhang et al., 2012), induce the production of complement and α-interferon (neutralization antibody) (Masters,2006; Park et al., 2011; Ren et al., 2011; Zhang et al., 2012) and perform and maintain the core of viral structure (Enjuanes et al., 2001). The co-expression between M

and E protein is essential for pseudoparticle formation, which inhibit interferogenic activity as though the complete virions (Park et al., 2011; Song and Park, 2012).

The nucleocapsid gene (N gene) encoded nucleocapsid protein (N protein). N protein is highly conserved in coronaviruses. The nucleocapsid is useful for genotypic comparison between PEDV strains (Li et al., 2012). The identity of PEDV N protein to the other coronaviruses such as MHV, IBV, HCoV-OC4 and BoCV is approximately 12-19% and as compared to FIPV, CCoV, PECV, TGEV and HCoV-229E is 32-37% (Bridgen et al., 1993).

The roles of N protein are required for packaging of viral RNA, transcription process, RNA replication (Schelle et al., 2005), viral core formation along with M protein (Li et al., 2009), and RNA-binding protein (RNA-binding activity) to perform helical nucleocapsid for viral RNA synthesis (Lai, 1990; Schelle et al., 2005; Master, 2006). Moreover, N protein plays the important role for cell-mediated immunity (CMI) production (Song and Park, 2012) and the early stage of PEDV infection gives the high levels of antibody response to N protein (Li et al., 2009).

The accessory gene is different among coronavirus group member (Enjuanes et al., 2001). For PEDV genome has only an accessory gene which is designated as open reading frame 3 (ORF3 gene) (Chen et al., 2011). The ORF3 gene is located between S and E genes (Masters, 2006). It is one of pathogenicity factors (Kweon et al., 1999; Song et al., 2003) and could be a marker for monitoring of nucleotide sequence variation during higher passage adaptation (Song et al., 2003).

Disease distribution and epidemiology

In 1971 in England, the first outbreaks of coronavirus-like were found in several swine herds. The feeder and fattening pigs showed the clinical sign of diarrhea. However, the suckling piglets were not affected. The TGEV infection was elided from these outbreaks because several ages of swine were affected. Then, the name "Epidemic Viral Diarrhea (EVD)" was designated (Pensaert and Yeo, 2006).

Afterward in 1976, the recurring TGEV-like outbreaks were found in all ages of pig including suckling piglets. The name of EVD was divided into two types as EDV type 1 for the cause of diarrhea in weaning pigs and feeder pigs, EVD type 2 for the cause of

diarrhea in all ages of pig including suckling piglets (Pensaert and Yeo, 2006). In 1978, the experiment of isolate designated CV777 was observed. The result indicated CV777 as an enteropathogen to all ages of pig which associated to EVD type 1 and type 2. Thus, the name of "Porcine Epidemic Diarrhea Virus (PEDV)" was renamed and had been used until today (Pensaert, 1984). During 1982-1990, several reports had been published on antibody responses to PEDV infection from many pig herds in Belgium, Switzerland, Germany, France, Netherlands, England, Bulgaria and Taiwan (Pensaert and Yeo, 2006).

The PEDV identification was reported in many European countries with several periods of time (Smid et al., 1993; Martelli et al., 2008). In Belgium, the report was published in 1978. The field trial was conducted by giving the minced piglet intestine with unidentified coronavirus-like particles from natural cases to various ages of experimental pigs. The EM technique was able to detect the PEDV particles within intestinal epithelium and feces of infected pigs (Pensaert and de Bouck, 1978).

In Spain, the first report of PEDV identification was in 1990s (Carvajal et al., 1995), which revealed that 55.9% of total 803 breeding farms had specific antibodies to PEDV (Carvajal et al., 1995). In the Czech Republic, PEDV detection was reported in 1993 (Smid et al., 1993, Song and Park, 2012). In Hungary, the report of PEDV identification was shown in 1995 (Nagy et al., 1996; Martelli et al., 2008). In Italy, repeated PEDV outbreaks were occurred during 2005-2006. The report showed total PEDV outbreaks of 21 times in 2005 and 42 times in 2006, which all ages were affected in the outbreaks, but only piglets died after infection (Martelli et al., 2008). PEDV can continuously circulate in the pig herd after an outbreak if those herds have usually replaced by new susceptible pigs into the population (Pijpers et al., 1993). Nowadays, the situations of acute PEDV outbreak in Europe are rarely found (Pospischil et al., 2002; Martelli et al., 2008).

For North and South American and Australia continents have not been reported of PED outbreak either by viral or antibody based detection (Penseart, 1984; Pospischil et al., 2002). Although the situations of acute PEDV outbreak in Europe were rarely found, but previous serological survey of PED infection had been reported in most

western European countries, excluding in Scandinavian countries and North Ireland (Penseart, 1984; Pospischil et al., 2002; Martelli et al., 2008).

In Asia, PEDV and antibody based detection have been reported in several countries during the 1980s. In Japan, the first outbreak of PEDV was found in 1982 and the major re-emerging has been found during 1993-1994 which showed 30-100% death in piglets (Kubota et al., 1999). In Korea during 1997-1999, the mass investigation of PEDV infection in swine herds found that 48% of pig farms (304/639 farms) were PEDV positive. All infected pigs were found in suckling pigs (Chae et al., 2000). In China, the first outbreak of acute diarrhea in swine herd was found in 1973 which enteropathogenic the infectious agent had specified only TGEV-like particle. Until 1984, PEDV was identified by the fluorescent antibody test and serum neutralization to confirm the causative agent at previous outbreak (Chen et al., 2011). In Vietnam, PEDV outbreak was reported in 2009. The suckling piglets from three provinces in southern part of Vietnam have been severely affected from PEDV infection. The morbidity rate closed to 100% and the mortality rate was between 65-91% (Duy et al., 2011).

In Thailand, the first PEDV outbreak was found in Southern part of Thailand in 1995 (Srinuntapunt et al., 1995) and the isolation of PEDV from Vero cell was successfully cultured and reported in 1997 (Antarasena et al., 1997). The emerging of PEDV outbreaks was found during recent years in several pig producing provinces. The major PEDV outbreaks have re-emerged in Nakornpathom province in December 2007. At this time, the affected pig age was mostly found in suckling period, showing sign of severe diarrhea, dehydration, and milk curd vomiting. The devastating impact on pig production was found up to 100% loss in newborn piglets (Puranaveja et al., 2009).

Noteworthy, the serious consequences of PEDV outbreaks in Asia is most often found in suckling piglets less than 10 days old of age (Pensaert and Yeo, 2006), unlike the outbreaks in western Europe or England that cause of diarrhea in pigs of all ages (Pensaert, 1984; Lee and Yeo, 2003). At the present, PEDV was continuously broken out and also spread out to the other neighbor countries, including Philippines (Morales et al., 2007), Korea, Japan, China, Taiwan, India (Pensaert and Yeo, 2006), Vietnam (Duy et

al., 2011) and Thailand (Puranaveja et al., 2009). PEDV is therefore considered as an endemic pathogen in South East Asia (Pensaert and Yeo, 2006).

The molecular epidemiology of PEDV

The molecular epidemiology of PEDV was studied based on S, M and ORF3 glycoprotein genes (Song and Park, 2012). The genetic relationship among PEDV isolates was described from many countries. In Korea 2007, the study of phylogenetic analysis of PEDV isolates between 2002-2005 based on partial S gene including epitope region, showed similarity to Chinese strain (JS-2004-2) (Park et al., 2007). In contrast, Lee et al (2010) reported that the phylogenetic analysis based on S gene of all Korean PEDV isolates during 2008-2009 remained situating in the same group of Korean parent strain. However, the nucleotide identities of these isolates compared to reference strains CV777, Br1/87, DR13, SM98, DX, JS-2004-2, LJB/03 LZC and PEDV isolated during 2002-2005, showed low homology (Lee et al., 2010).

In 2011, the sequence analysis based on the M gene of 32 Korean PEDV isolated from 2003-2007 revealed different pattern of relationships to Korean parent and vaccine strains. Korean PEDV isolated between 2003 and 2004 were closely resembled to Korean and Chinese field strains, except that some isolates had similarity to vaccine strains (attenuated DR13, KPED-9 and P-5V). Korean PEDV isolated in 2007 had closely related to Chinese field strain and four Thai isolates during 2007-2008 (Park et al., 2011).

In China, CH/S strain is the first PEDV strain isolated in 1986 (Chen et al., 2011). The phylogenetic relationship of Chinese PEDV isolates in 2006 based on M gene showed higher homology to Chinese strain (JS-2004-2) than strains CV777, BR1/87, Chinju99 and JMe2 (Chen et al., 2008). Additionally, Chinese field strains between years 2006-2009 represented the genetic diversity with European strains and vaccine strain (which have been using in China). However, they were closely related to Korean strains (Chen et al., 2011). Other outbreaks were reported in early 2010 in Southern China, the new variant PEDV strain as GD-A was isolated. GD-A genome had a high relatedness to CV777, killed-vaccine strain (based on CV777 strain) and the other foreign strains rather than Chinese field isolates that were isolated in the same year (CH/FJND-3/2011 and BI-2011-1). Moreover, GD-A genome had genetic mutations from the other recent field

isolates, thus GD-A strain was a new variant PEDV strain emerging in China (Fan et al., 2012).

In Vietnam during 2009-2010, PEDV emerging outbreak was occurred in Southern part of Vietnam. The genetic relationship between Vietnamese isolates and reference isolates were analyzed based on partial S and M gene sequences. The phylogenetic analysis of Vietnamese PEDV isolates showed high nucleotide identity to Chinese (JS-2004), Thai isolates (2009) and Korean isolates. So this data indicated the Chinese isolate (JS-2004-2) could be PEDV ancestor which was continuously overspread to the other Asian neighbor countries such as Vietnam, Thailand and Korea (Duy et al., 2011).

In Thailand, PEDV was isolated in 2004 by Thanawongnuwech (unpublished data) but the first report was published in NCBI database in 2008. The latest study of molecular characterization was reported in 2009 that involved with molecular epidemiology of the Thai isolates during 2007-2008 based on partial S and M gene sequences analysis. These result showed Thai isolates had closely related to Chinese field isolates (JS-2004-2 and LJB/03) rather than European and Korean strains (Puranaveja et al., 2009).

Interestingly, all reports of PEDV isolates between 2007-2011 from the Asian outbreaks in Korea, Thailand and Vietnam may have derived from the same Chinese ancestor (Park et al., 2007, Puranaveja et al., 2009, Duy et al., 2011).