

Complete genome sequence of a novel S-insertion variant of porcine epidemic diarrhea virus from South Korea

Sunhee Lee¹ · Changhee Lee¹ 

Received: 25 April 2017 / Accepted: 30 May 2017 / Published online: 6 June 2017
© Springer-Verlag GmbH Austria 2017

Abstract Porcine epidemic diarrhea virus (PEDV) is an emerging and re-emerging viral pathogen of pigs that is causing significant economic concerns in the global pork industry. The virulent G2b PEDV subgroup, which was responsible for the 2013–2014 pandemics on the American and Asian continents, hit Jeju Island in 2014, which had been PEDV-free since 2004. This study presents molecular characterization of a novel PEDV variant with a 5-amino acid (aa) insertion (DTHPE) in the spike (S) gene (an S-insertion) that was identified in diarrheic piglets from Jeju Island. The genome of the PEDV strain KOR/KNU-1601/2016 was sequenced and analyzed to characterize the S-insertion variant circulating on Jeju Island. The full-length genome sequence of KNU-1601 is 28,053 nucleotides (nt) in length, which is 24 nt and 15 nt longer than the genome sequences of G1 classical and G2 epidemic field strains, respectively. KNU-1601 shares 99.5–99.6% nucleotide sequence identity at the genome level and 98.0–98.6% amino acid sequence identity at the S gene level with other global G2b PEDV strains. Genetic and phylogenetic analyses indicated that the KNU-1601 variant is most closely related to the G2b field isolates but appears to undergo continuous evolution in the field. These data advance our understanding of the genetic diversity and

evolutionary characteristics of the PEDV field strains circulating in South Korea.

Porcine epidemic diarrhea virus (PEDV), a member of the genus *Alphacoronavirus* in the family *Coronaviridae* of the order *Nidovirales*, is a highly epizootic, deadly enteric virus of pigs and is considered one of the most economically important viral pathogens in countries with intensive swine industries [4]. In South Korea, since the first porcine epidemic diarrhea (PED) epizootic in 1992 [3], the disease has remained rampant, devastating the hog industry. PED epidemics in 2013–2014 rapidly swept across mainland South Korea, and then Jeju Island, and killed hundreds of thousands of piglets in domestic herds [6, 7]. Although the disease first appeared in 1996 on Jeju Island and had since annually plagued the local swine herd until 2003, the province was PEDV-free for the next 10 years. However, in late March 2014, PEDV re-emerged in the Hanlim area and subsequently spread to other districts, causing high mortality in newborn piglets. The re-emergent Jeju PEDV isolates were most genetically and phylogenetically similar to the pandemic genogroup 2b (G2b) strains responsible for the massive 2013–2014 PEDV outbreaks, indicating direct introduction of the virus from mainland South Korea via contaminated sources [6]. Since then, PEDV has plagued Jeju swine farms, leading to financial losses in the provincial pork industry. On one commercial pig farm located in Jeju Province, a PED case was initially confirmed in April 2014. Despite the acquired immune status through vaccination and feedback (intentional exposure) using a G2b-based inactivated vaccine and feces from infected neonatal piglets, respectively, this farm has continued to experience sporadic, mild PEDV-associated diarrheal disease outbreaks. In March 2016, PED-like

Electronic supplementary material The online version of this article (doi:10.1007/s00705-017-3441-y) contains supplementary material, which is available to authorized users.

✉ Changhee Lee
changhee@knu.ac.kr

¹ Animal Virology Laboratory, School of Life Sciences, College of Natural Sciences, BK21 Plus KNU Creative BioResearch Group, Kyungpook National University, Daegu 41566, Republic of Korea

symptoms recurred on the same pig farm, and fecal samples from pigs with acute watery diarrhea were submitted to our laboratory for diagnosis. According to results of RT-PCR using a TGE/PED detection kit (iNtRON Biotechnology, Seongnam, South Korea), all stool samples were found to be positive for PEDV. Subsequently, the complete sequences of the S genes of the KNU-1601 isolates from the same farm were determined by the traditional Sanger method as described previously [5]. Nucleotide (nt) sequencing analysis showed that all of the KNU-1601 strains are genetically nearly identical to each other, with 99.9–100% amino acid (aa) sequence identity, and share 98.0–98.6% aa sequence identity with the recently emerging G2b PEDVs on Jeju Island, mainland South Korea, and in the USA. Furthermore, the KNU-1601 isolate possessed the genetic signature of G2 field strains, consisting of two discontinuous 4-aa and 1-aa insertions at positions 55 and 135 and one 2-aa deletion at positions 160 and 161 within the N-terminal hypervariable region of S compared to the prototype CV777 strain [4, 5; Fig. 1]. These data indicate that the KNU-1601 isolates are

genetically most closely related to the G2b pandemic stains. Intriguingly, the KNU-1601 virus was found to contain a 5-aa (DTHPE) insertion at position 380 in the S1 domain, which is completely absent in the other G1 and G2 isolates for which sequences are available in the GenBank database (Fig. 1). Thus, the S genes of the KNU-1601 group strains are 4,176 nt in length, encoding a 1391-aa protein, which is 24 nt (8 aa) and 15 nt (5 aa) longer than the S genes in the G1 and G2 strains, respectively. To determine the genetic relationship of the novel Korean S-insertion variant to other global PEDV strains, the complete genome of a representative KNU-1601 strain was sequenced and analyzed.

The full-length genomic sequence of KNU-1601 was determined by next-generation sequencing (NGS) technology. Ten overlapping cDNA fragments encompassing the entire genome were generated, pooled in equimolar amounts, and subjected to NGS, using an Ion Torrent Personal Genome Machine (PGM) Sequencer (Life Technologies, Carlsbad, CA) and a 316 v2 sequencing chip (Life Technologies). The KNU-1601 NGS reads were

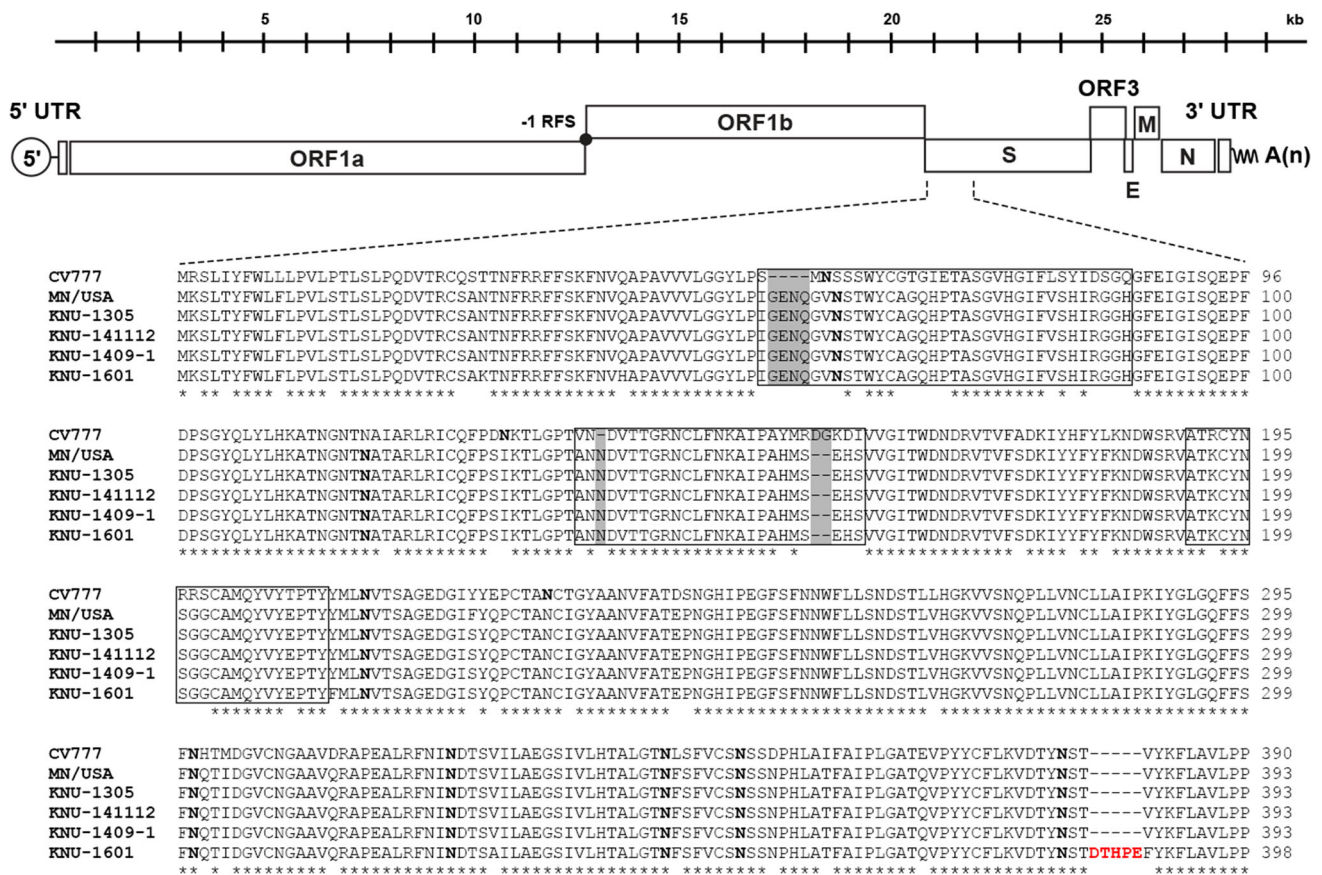


Fig. 1 Alignment of the amino acid sequences of the S proteins of various PEDV strains, including classical G1a, pandemic G2b, and KNU-1601 strains. Potential N-glycosylation sites predicted by the NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) are shown in boldface type. Genetic signatures, consisting of insertions

and deletions (INDELs), in the PEDV epidemic G2b strains are shaded. Amino acids representing potential hypervariable domains (solid boxes) and a unique 5-aa (DTHPE) insertion (red color) are also indicated (color figure online)

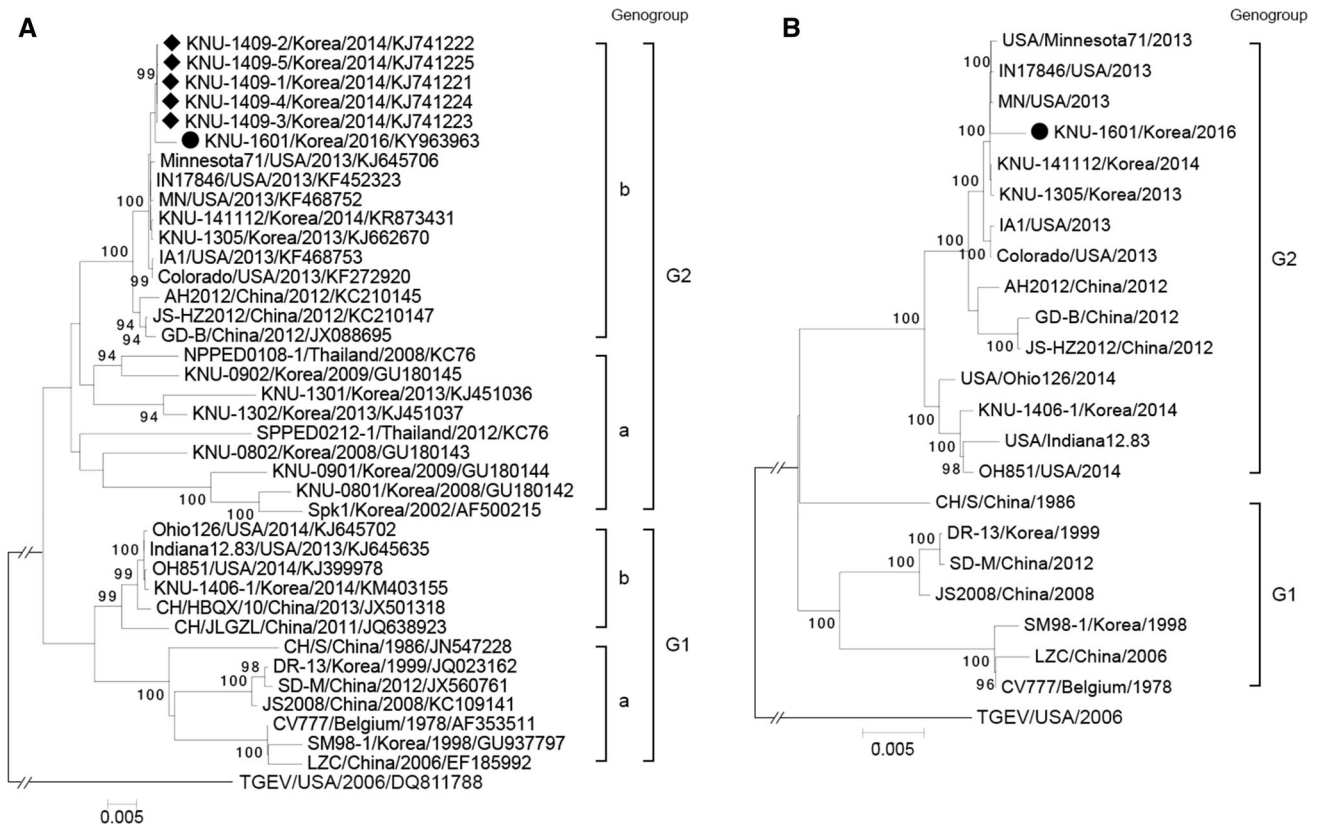


Fig. 2 Phylogenetic analysis based on the nucleotide sequences of the spike genes (A) and full-length genomes (B) of PEDV strains. A region of the spike gene and the complete genome sequence of TGEV were included as the outgroups in A and B, respectively. Multiple sequence alignments were generated with ClustalX, and a phylogenetic tree was constructed from the aligned nucleotide sequences using the neighbor-joining method. Numbers at each branch are

bootstrap values greater than 50% based on 1000 replicates. The names of the strains, countries and dates (year) of isolation, GenBank accession numbers, and genogroups and subgroups proposed in this study are shown. A solid circle indicates the KNU-1601 strain identified in this study; solid diamonds indicate the re-emergent Jeju strains detected in 2014. Scale bars indicate nucleotide substitutions per site

assembled using the complete genome sequence of the PEDV reference strain KOR/KNU-1305/2013 (GenBank accession no. KJ662670). The 5' and 3' ends of the KNU-1601 genome were also determined by rapid amplification of cDNA ends (RACE) as described previously [8]. The KOR/KNU-1601/2016 PEDV sequence data have been deposited in the GenBank database under accession number KY963963. The complete genome sequence of KNU-1601 is 28,053 nt in length, excluding the 3' poly(A) tail. Except for the S-insertion, no additional insertions or deletions were identified in the KNU-1601 genome. The genome of KNU-1601 shared high nucleotide sequence identity (99.5–99.6%) with other complete G2b PEDV genome sequences in GenBank, showing the highest nucleotide sequence identity with the recent epidemic PEDV strain KNU-141112. The number of nt/aa differences and the percent identity shared between KNU-1601 and genogroup representative strains are summarized in Supplementary Table S1. Compared to the complete genome of the Korean prototype G2b strain KNU-1305, the

genome of KNU-1601 had 93 nucleotide differences (99.6% identity). Among these, 90 were located in coding regions, and 45 of these differences were non-synonymous, causing 45 amino acid changes: 22 in ORFs 1a and 1b, 17 in S, two in an accessory ORF3, and four in N. The overlapping ORFs 1a and 1b of KNU-1601 encode 4,117-aa and 2,681-aa polyproteins, respectively, which are proteolytically cleaved into 16 functional nonstructural proteins (nsps). A total of 22 amino acid changes in KNU-1601 were randomly dispersed among 11 nsps; nsp1, nsp2, nsp3, nsp4, nsp5, nsp6, nsp9, nsp12, nsp13, nsp15, and nsp16, which contained one, five, five, two, one, one, two, one, one, two, and one difference(s), respectively, compared to KNU-1305. Subsequent phylogenetic analysis based on the complete S protein clearly delineated the PEDV strains into two distinct genogroup clusters, G1 and G2, which were further divided into subgroups 1a, 1b, 2a, and 2b (Fig. 2A). The KNU-1601 S-insertion strain belongs to subgroup G2b, like recent domestic field isolates, which were clustered most closely with the emergent

US strains that form an adjacent clade within the same subgroup. In addition, the phylogenetic tree based on the genome sequences indicates that the novel S-insertion KNU-1601 variant is grouped within the same cluster as the global epizootic strains (Fig. 2B).

To our knowledge, this is the first report of the complete genome sequence of a novel S-insertion variant. The S glycoprotein of coronaviruses can be functionally divided into two domains, S1 and S2. The former is responsible for binding to a host-specific receptor, while the latter appears to be involved in direct fusion between the viral and cellular membranes [2]. Like other coronavirus S proteins, the PEDV S protein plays a critical role in infection by interacting with the cellular receptor to mediate viral entry and inducing neutralizing antibodies in its natural host [4]. Mutations or insertions/deletions in the S gene have been shown to alter viral pathogenicity and tissue/species tropism [2, 4, 10]. Recent studies have identified novel G2b variants in South Korea, China, and Japan with unique large or small deletions of the S gene when compared to other G2 field PEDV sequences [1, 9, 11]. Interestingly, a Chinese variant with a short deletion at the extreme C-terminal end of the S gene showed reduced virulence in newborn piglets [1], whereas another strain with a novel two-amino-acid deletion at positions 58–59 was highly pathogenic to neonatal pigs [12]. Although we failed to isolate and propagate the new PEDV variant described here in cell culture and were unable to pursue further studies to investigate its biological properties, it is speculated that the pathogenicity of this strain might have been modified. Indeed, the pig farm infected with this novel variant had experienced less-severe clinical signs than those reported for other epidemic G2b strains. The immune status provided by vaccination or intentional exposure may at least contribute to this mild disease outcome. Otherwise, we hypothesize that the S-insertion might cause a conformational change in the S protein, thereby allowing it to evade host immune defenses, including neutralizing antibodies, and ultimately alter viral pathogenicity, leading to persistent infection in the field. Thus, studies using a reverse genetics system are needed to address the specific function of insertions in S in PEDV pathogenesis. Most importantly, further investigations should be conducted to determine how such a PEDV variant was generated and evolved in the field and to strengthen the monitoring of the emergence of novel variants via genetic drift and/or recombination events. Nevertheless, our sequence data provide insights into the genetic diversity and evolution of PEDV field strains in South Korea and suggest that PEDV continues to

undergo evolutionary processes, accumulating the mutations necessary for viral fitness in its natural host.

Acknowledgements This research was supported by the Bio-Industry Technology Development Program through the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (iPET) funded by the Ministry of Agriculture, Food and Rural Affairs (315021–04).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with animals performed by any of the authors.

References

- Fan B, Jiao D, Zhao X, Pang F, Xiao Q, Yu Z, Mao A, Guo R, Yuan W, Zhao P, He K, Li B (2017) Characterization of Chinese porcine epidemic diarrhea virus with novel insertions and deletions in genome. *Sci Rep* 7:44209
- Gallagher TM, Buchmeier MJ (2001) Coronavirus spike proteins in viral entry and pathogenesis. *Virology* 279:371–374
- Kweon CH, Kwon BJ, Jung TS, Kee YJ, Hur DH, Hwang EK, Rhee JC, An SH (1993) Isolation of porcine epidemic diarrhea virus (PEDV) in Korea. *Korean J Vet Res* 33:249–254
- Lee C (2015) Porcine epidemic diarrhea virus: An emerging and re-emerging epizootic swine virus. *Virol J* 12:193
- Lee DK, Park CK, Kim SH, Lee C (2010) Heterogeneity in spike protein genes of porcine epidemic diarrhea viruses isolated in Korea. *Virus Res* 149:175–182
- Lee S, Ko DH, Kwak WK, Lim CH, Moon SU, Lee DS, Lee C (2014) Reemergence of porcine epidemic diarrhea virus on Jeju Island. *Korean J Vet Res* 54:185–188
- Lee S, Lee C (2014) Outbreak-related porcine epidemic diarrhea virus strains similar to US strains, South Korea, 2013. *Emerg Infect Dis* 20:1223–1226
- Lee YN, Lee C (2013) Complete genome sequence of a novel porcine parainfluenza virus 5 isolate in Korea. *Arch Virol* 158:1765–1772
- Park S, Kim S, Song D, Park B (2014) Novel porcine epidemic diarrhea virus variant with large genomic deletion South Korea. *Emerg Infect Dis* 20:2089–2092
- Sato T, Takeyama N, Katsumata A, Tuchiya K, Kodama T, Kusanagi K (2011) Mutations in the spike gene of porcine epidemic diarrhea virus associated with growth adaptation in vitro and attenuation of virulence in vivo. *Virus Genes* 43:72–78
- Van Diep N, Norimine J, Sueyoshi M, Lan NT, Yamaguchi R (2017) Novel porcine epidemic diarrhea virus (PEDV) variants with large deletions in the spike (S) gene coexist with PEDV strains possessing an intact S Gene in domestic pigs in Japan: a new disease situation. *PLoS One* 12:e0170126
- Zhang X, Pan Y, Wang D, Tian X, Song Y, Cao Y (2015) Identification and pathogenicity of a variant porcine epidemic diarrhea virus field strain with reduced virulence. *Virol J* 12:88