Archives of Virology (2024) 169:183
https://doi.org/10.1007/s00705-024-06101-8

ORIGINAL ARTICLE

# Molecular and phylogenetic analysis of transmissible gastroenteritis virus strain VET-16, isolated from piglets in Vietnam

Soo Hyun Moon[1] · Gyu-Nam Park[1] · SeEun Choe[1] · Sok Song[1] · Van Phan Le[2] · Yun Sang Cho[1] · Dong-Jun An[1]

Received: 1 March 2024 / Accepted: 17 June 2024 / Published online: 21 August 2024


## Abstract

Porcine transmissible gastroenteritis virus (TGEV) is a major pathogen that causes viral enteritis and severe diarrhea in newborn piglets. TGEV strains have been isolated in the USA, Europe, and China, and their molecular characteristics are well known. However, there have been few reports of molecular analysis of TGEV strains isolated in Southeast Asia. In 2016, we isolated TGEV strain VET-16 from fecal samples collected from piglets in Vietnam and determined its complete genome sequence by Sanger sequencing. We found that, while the full genome of the VET-16 strain was 92.4-99.9% identical to those of other TGEV strains, the ORF3 gene showed very little sequence similarity. Phylogenetic analysis suggested that the VET-16 strain belongs to the Purdue subgroup. Comparison of the predicted amino acid (aa) sequence of the spike protein of strain VET-16 with those of other TGEV strains revealed three aa substitutions (V378L, S379T, and D380N) and a 3-aa insertion (F383_F387insWEK) in antigenic site D of the VET-16 strain. Also, a single aa deletion (ΔF1413) was found in the transmembrane domain of the spike gene of VET-16. Like the ORF3 gene from the TGEV Miller M60 vaccine strain, the VET-16 strain has a large deletion (Δ725 nt) in the ORF3 gene. Previous studies have suggested that these mutations in the spike and ORF3 genes might be associated with a reduction in pathogenicity. The data from this study will facilitate further genetic analysis and research into the evolution of TGEV in pigs in Vietnam.

## Introduction

Porcine transmissible gastroenteritis virus (TGEV) (genus *Alphacoronavirus*, family *Coronaviridae*) is an enveloped virus with a single-stranded positive-sense RNA genome approximately 28.5 kb in length [1]. The genome has nine open reading frames (ORFs) that encode four structural proteins (spike [S], envelope [E], membrane [M], and nucleocapsid [N]) and five non-structural proteins (ORF 1a/1b, ORF 3a/3b, and ORF7) [2]. These genes are arranged in the order 5′-ORF1a-ORF1b-S-ORF3a-ORF3b-E-M-N-ORF7-3′ [1]. TGEV is a pathogen that infects newborn piglets, causing viral diarrhea and enteritis. The mortality rate in piglets less than 2 weeks old is 100% [3, 4]. All pigs are susceptible to infection by TGEV, but piglets under 2 weeks of age are at especially high risk [4]. TGEV was first reported in the USA in 1946 [5], and since then, cases of TGEV infection (or coinfection) have occurred in pork-producing regions of Europe (England [6], Spain [7], and Germany [8]) and Asia (China [9] and Japan [10]), resulting in significant economic losses.

Molecular and phylogenetic analysis of TGEV isolates has led to genotypic classification into two groups: the traditional group (the Purdue and Miller subgroups) and the variant group [11]. The traditional group has been identified in the USA [12, 13], Europe [6, 12], and Asia [5, 14, 15], whereas the variant group has been identified (and is prevalent) mostly in the USA [12]. However, there have been no reports of molecular and phylogenetic analysis of TGEV strains isolated in Southeast Asia since 1982 [16]. Among the TGEV genes, mutations are most common in S and ORF3, and these are strongly associated with virulence and cell tropism [2]. The S1 subunit of the S protein binds to sialic acid moieties and specific receptors on host cells [2, 17]. Four major antigenic sites (A–D) in the S1 subunit of the TGEV have been mapped at its N-terminus. Of

Handling Editor: Sheela Ramamoorthy.

✉ Dong-Jun An
andj67@korea.kr

[1] Viral Disease Division, Animal and Plant Quarantine Agency, Gimcheon 9660, Republic of Korea

[2] Department of Veterinary Microbiology and Infectious Diseases, Faculty of Veterinary Medicine, Vietnam National University of Agriculture, Hanoi 100000, Vietnam

these, antigenic sites A and D are antigenically dominant with respect to neutralization of TGEV *in vitro* [18–20]. The ORF3 gene of TGEV encodes ORF3a and ORF3b, and in many TGEV strains [21–23], as well as other coronaviruses such as porcine respiratory coronavirus (PRCV; a respiratory variant of TGEV) [24, 25], porcine epidemic diarrhea virus [26], and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [27], both 3a and 3b often carry deletions. Previous studies have suggested that ORF3 deletions are associated with viral fitness (which is supported by identification of a naturally occurring truncated ORF3 gene) [27–31] or cell adaptation *in vitro* [13, 32–34].

Because little is known about the molecular characteristics of TGEV strains circulating in Vietnam, the aim of the present study was to perform a detailed analysis of TGEV isolated from piglets in Vietnam. We identified a novel TGEV strain (designated "VET-16") and determined its full genome sequence. Molecular and phylogenetic analysis and additional detailed analysis of the S and ORF3 genes showed that this TGEV strain isolated from piglets in Vietnam has unique molecular features never before identified in other strains of TGEV.

## Materials and methods

### Sample preparation and RNA extraction

In 2016, a survey of 18 farms in Hanoi, Hung Yen, Lao Cai, Tuyen Quang, and Thai Nguyen was conducted to assess the prevalence of piglet diarrhea in northern Vietnam, and TGEV was detected on one farm in Hung Yen, where mild diarrhea symptoms were observed in suckling and weaned piglets. TGEV-positive fecal samples obtained from piglets were diluted 1:2 in phosphate-buffered saline, and RNA was extracted from these samples using a Patho Gene-spin DNA/RNA Extraction Kit (LiliF Diagnostics, South Korea).

### PCR amplification and sequencing

The full-length genome of TGEV was sequenced by one-step reverse transcription polymerase chain reaction (RT-PCR) using universal primers targeting TGEV (Supplementary Table S1) [30] and a HelixCript One-Step RT-PCR Kit [Hot-Taq] [UDG System] (NanoHelix, South Korea). The RT-PCR mixtures (50 µL) comprised 4 µL of RNA template, 2 µL of each F/R primer (10 pmol), 25 µL of 2× Reaction Mix (Hot-Taq containing dUTP), 2 µL of enzyme mix (Hot-Taq containing UDG), and 15 µL of nuclease-free water. The RT-PCR conditions were as follows: UDG activation at 25℃ for 5 min, cDNA synthesis at 55℃ for 50 min, and pre-denaturation at 95℃ for 15 min, followed by 40 cycles of denaturation at 95℃ for 20 s, annealing at 52–55℃ for 40 s, and extension at 72℃ for 1 min 30 s. Post-extension was performed at 72℃ for 5 min. The full genome sequence was divided into 44 fragments, with a 50- to 200-nt overlap between adjacent segments. PCR products were separated by agarose gel electrophoresis and purified by gel extraction. Finally, Sanger sequencing was performed using an Applied Biosystems 3730xl DNA Analyzer.

### Molecular and phylogenetic analysis

To analyze their molecular characteristics using BioEdit software (7.2.5 version), the full genome sequences of 39 TGEV strains obtained from the GenBank database were aligned (Table 1). For bioinformatic analysis, we used the nucleotide Basic Local Alignment Search Tool BLASTn to search the NCBI database. For phylogenetic analysis, the full genome, S, and ORF3 sequences were analyzed using Molecular Evolutionary Genetics Analysis 11 (MEGA 11). All phylogenetic trees based on nucleotide sequences were constructed using the maximum-likelihood method with 1000 replicates in MEGA 11.

### Recombination analysis

Recombination analysis was performed using RDP4 software (which includes RDP, BootScan, and SiScan) to identify likely parental strains and recombination breakpoints, using default settings. The criterion for identifying recombination breakpoints were a $P$-value $< 10^{-6}$ or a recombination score $> 0.6$.

## Results

### Whole-genome sequence of the VET-16 strain

Sanger sequencing of the full-length TGEV VET-16 strain revealed that the genome is 27,867 nucleotides (nt) in length, with a 314-nt 5′ untranslated region (UTR), nine ORFs, including ORF1a (nt 315–12,368), ORF1b (nt 12,332–20,368), S (nt 20,365–24,714), ORF3a (nt 24,833–24,922), ORF3b (nt 24,948–25,151), E (nt 25,138–25,386), M (nt 25,397–26,185), N (nt 26,198–27,346), and ORF7 (nt 27,352–27,588), as well as a 279-nt 3′ UTR with a poly(A) tail.

### Comparison of the VET-16 genome with those of other strains

Sequence comparisons showed that the ORF1a/b, S, E, M, N, and ORF7 genes of VET-16 were 99.0−100% identical to those of the eight Purdue subgroup TGEV strains and 92.4–97.5% identical to those of four variant group

**Table 1** Sequence information for isolated TGEV strains

| No. | Strain | Country | Years | GenBank accession no. |
|---|---|---|---|---|
| 1 | VET-16* | Vietnam | 2016 | PP236916 |
| 2 | HX | China | 2012 | KC962433 |
| 3 | HQ2016 | China | 2016 | MT576083 |
| 4 | AYU | China | 2009 | HM776941 |
| 5 | TH-98 | China | 1998 | KU729220 |
| 6 | SHXB | China | 2013 | KP202848 |
| 7 | Purdue P115 | USA | 2009 | DQ811788 |
| 8 | WH-1 | China | 2011 | HQ462571 |
| 9 | SC-Y | China | 2006 | DQ443743 |
| 10 | 145 | Mexico | 2008 | KX900402 |
| 11 | Virulent Purdue | USA | 1952 | DQ811789 |
| 12 | AHHF | China | 2017 | KX499468 |
| 13 | JS2012 | China | 2012 | KT696544 |
| 14 | TS | China | 2016 | DQ201447 |
| 15 | Miller M6 | USA | 2009 | DQ811785 |
| 16 | SC2021 | China | 2021 | ON858825 |
| 17 | CN12 | China | 2012 | KX058075 |
| 18 | H16 | China | 1973 | FJ755618 |
| 19 | Attenuated H | China | 2009 | EU074218 |
| 20 | Tennessee144 | USA | 2008 | KX900401 |
| 21 | Minnesota148 | USA | 2013 | KX900405 |
| 22 | NorthCarolina140 | USA | 2007 | KX900397 |
| 23 | NorthCarolina142 | USA | 2007 | KX900399 |
| 24 | PUR46-MAD | USA | Unknown | AJ271965 |
| 25 | HE-1 | China | 2015 | KX083668 |
| 26 | Z | USA | 2006 | KX900393 |
| 27 | HB | USA | 1988 | KX900394 |
| 28 | DS01-2022 | China | 2022 | OP805351 |
| 29 | Miller M60 | USA | 2011 | DQ811786 |
| 30 | 96-1933 | U.K | 1996 | AF104420 |
| 31 | FS772/70 | U.K | 1970 | X53128 |
| 32 | TFI | Taiwan | 1983 | Z35758 |
| 33 | NEB72-RT | USA | 1972 | M94099 |
| 34 | DAE | South Korea | 1998 | JQ693050 |
| 35 | TO14 | Japan | Unknown | AF302263 |
| 36 | TOY56 | Japan | 1956 | M94103 |
| 37 | KT2 | South Korea | 2000 | JQ693051 |
| 38 | KT3 | South Korea | 2000 | JQ693052 |
| 39 | 133 | South Korea | 1997 | JQ693049 |

*Isolated in this study.

strains. However, comparative analysis of the ORF3 gene revealed that VET-16 strain showed much less sequence similarity than the other 19 TGEV strains, and this was particularly evident for ORF3a (26.6–27.8% identity). It is noteworthy that the ORF3b gene of the VET-16 strain was most similar (74.6% identity) to that of the Miller M60 strain (Table 2).

Nucleotide BLAST analysis showed that the sequence of the ORF3a gene of VET-16 was very similar to those of previous TGEV and canine coronavirus (CCoV) isolates (Table 3). Indeed, the ORF3a gene of strain VET-16 was 100% identical to that of CCoV, but recombination analysis did not indicate any mixing of VET-16 strain and CCoV sequences in the ORF3a gene.

## Molecular characteristics of the S gene

The S gene (4350 nt) of the VET-16 strain was found to have a 6-nt deletion at nt 1123-1128 (Fig. 1A, B). This deletion has been reported previously in the NEB72-RT, DAE, Purdue P115, WH-1, HX, and HQ2016 strains. This gene also contained four nt substitutions (resulting in three aa changes), a 9-nt (3-aa) insertion in antigenic site D (Fig. 1A, B), and a 3-nt deletion in the TM domain (Fig. 1C, D). The indels in the S gene of VET-16 strain were confirmed by additional Sanger sequencing (Fig. 1B).

Amino acid 72 of the VET-16 strain is asparagine, whereas that in the Virulent Purdue and Purdue P115 strains is aspartic acid. Amino acid 219 of the VET-16 strain is serine, whereas that in the Virulent Purdue, Purdue P115, Miller M6, Miller M60, H16, and attenuated H strains is alanine. Amino acid 585 of the VET-16 strain is alanine, which is the same as that in the Purdue P115, Miller M60, H16, and attenuated H strains. The amino acid residues in the aminopeptidase N (APN) binding site of the VET-16 strain are identical to those in the DAE and NEB72-RT strains (Table 4). The 6-nt deletion of nt 1123-1128 results in the deletion of two aa residues (N375_D376del) from the S protein of the VET-16 strain (Fig. 1E). These deletions have also been observed in the NEB72-RT, DAE, Purdue P115, WH-1, HX, and HQ2016 strains. The VET-16 strain carries three aa substitutions (V378L, S379T, and D380N) and a 3-aa insertion (F383_F387insWEK) in antigenic site D (Fig. 1E). In addition, the VET-16 strain has a single aa deletion (F1413del) in the TM domain of the S protein (Fig. 1E).

## Molecular characteristics of the ORF3 gene

The VET-16 strain has a large deletion (Δ725 nt) in the ORF3 gene (Supplementary Fig. S1), which was confirmed by RT-PCR using specific primers (Large-del-F, 5′-GGA TGCATAGGTTGTTTAG-3′; Large-del-R, 5′-CCACGT ATTGCTATGCTTAC-3′; amplicon size: 1080 bp). Because the start codon of the ORF3a and ORF3b sequence was not deleted, the ORF3 and ORF3b proteins are still expressed, but shortened, resulting in a length of 29 aa and 67 aa, respectively. The length of ORF3b of the VET-16 strain is the same as that of the Miller M60 strain (Fig. 2A). The large deletion was verified by comparing the size of DNA bands visualized on agarose gels using a TGEV-positive control

**Table 2** Nucleotide and amino acid sequence homology between the TGEV VET-16 strain and other TGEV strains

| Group | Strain | ORF1a | ORF1b | S | ORF3a | ORF3b | E | M | N | ORF7 |
|---|---|---|---|---|---|---|---|---|---|---|
| Purdue subgroup (Traditional group) | HX | 99.84/99.78 | 99.94/99.96 | 99.43/98.97 | 27.80/14.08 | 27.76/27.46 | 99.60/98.78 | 99.62/98.85 | 99.91/99.74 | 99.58/100.0 |
| | HQ2016 | 99.90/99.88 | 99.96/99.96 | 99.45/99.03 | 27.80/14.08 | 27.76/27.46 | 99.20/97.56 | 99.62/98.85 | 99.91/99.74 | 99.58/100.0 |
| | TH-98 | 99.87/99.78 | 99.96/99.93 | 99.24/98.83 | 27.80/14.08 | 27.76/27.46 | 99.20/98.78 | 99.87/99.62 | 99.91/99.74 | 99.58/100.0 |
| | HB | 99.93/99.95 | 99.96/99.96 | 99.45/99.17 | 27.80/14.08 | 27.76/27.46 | 99.60/98.78 | 99.87/99.62 | 99.91/99.74 | 99.58/100.0 |
| | Virulent Purdue | 99.83/99.70 | 99.96/99.93 | 99.01/98.28 | 27.35/15.49 | 27.76/27.46 | 99.20/97.56 | 99.62/98.85 | 99.65/99.48 | 99.58/100.0 |
| | Purdue P115 | 99.85/99.85 | 99.94/99.96 | 99.40/98.90 | 27.80/14.08 | 27.76/27.46 | 99.60/98.78 | 99.75/99.24 | 99.83/99.48 | 99.58/100.0 |
| | WH-1 | 99.93/99.90 | 99.98/99.96 | 99.45/99.03 | 27.80/14.08 | 27.76/27.46 | 99.60/98.78 | 99.75/99.24 | 99.91/99.74 | 99.58/100.0 |
| | SC-Y | 99.48/99.23 | 99.83/99.74 | 99.26/98.76 | 27.80/14.08 | 27.76/27.46 | 99.60/98.78 | 99.87/99.62 | 99.83/99.48 | 99.58/100.0 |
| Miller subgroup (Traditional group) | JS2012 | 98.95/99.03 | 99.10/99.63 | 98.14/97.80 | 27.43/12.50 | 27.62/27.05 | 98.39/95.12 | 98.10/97.33 | 98.09/98.17 | 95.78/93.67 |
| | TS | 98.76/98.63 | 98.98/99.52 | 97.87/97.59 | 27.43/12.50 | 27.48/26.64 | 98.39/95.12 | 97.85/96.56 | 98.00/97.91 | 95.78/93.67 |
| | Miller M6 | 98.95/99.05 | 99.07/99.59 | 97.87/97.59 | 27.43/12.50 | 27.62/27.05 | 97.99/93.90 | 98.10/97.33 | 98.09/98.17 | 95.36/93.67 |
| | Miller M60 | 98.90/98.95 | 99.06/99.59 | 97.77/97.45 | 27.88/16.67 | 74.65/55.56 | 99.20/97.56 | 99.62/98.85 | 99.91/99.74 | 99.58/100.0 |
| | H16 | 98.85/98.88 | 99.01/99.55 | 97.71/97.38 | 27.43/12.50 | 27.62/27.05 | 97.59/93.90 | 97.97/96.95 | 98.09/98.17 | 95.78/93.67 |
| | Attenuated H | 98.86/98.90 | 98.99/99.55 | 97.57/97.18 | 27.43/12.50 | 27.62/27.05 | 96.79/91.46 | 97.97/96.95 | 98.00/98.17 | 95.78/93.67 |
| | DS01-2022 | 98.85/98.85 | 98.99/99.55 | 97.64/97.31 | 27.43/12.50 | 27.62/27.05 | 96.79/91.46 | 97.97/96.95 | 98.17/98.17 | 95.78/93.67 |
| Variant group | Tennessee144 | 96.49/97.34 | 97.50/99.22 | 95.04/96.28 | 27.31/16.67 | 26.80/25.41 | 97.19/95.12 | 95.82/96.18 | 95.56/97.38 | 92.41/92.41 |
| | Minnesota148 | 96.31/97.04 | 97.40/99.14 | 94.86/96.21 | 27.11/16.67 | 26.80/25.41 | 97.19/95.12 | 95.69/96.18 | 95.13/97.12 | 92.83/92.41 |
| | NorthCarolina140 | 96.58/97.31 | 97.54/99.22 | 95.02/96.21 | 27.75/16.67 | 26.67/25.00 | 96.79/95.12 | 95.82/95.80 | 95.39/97.12 | 93.25/93.67 |
| | NorthCarolina142 | 96.57/97.29 | 97.61/99.29 | 95.21/96.21 | 27.75/16.67 | 26.80/25.41 | 97.19/95.12 | 95.94/96.18 | 95.39/97.12 | 92.83/92.41 |

Data are presented as percentages.

**Table 3** Nucleotide BLAST results for the ORF3a sequence of TGEV strain VET-16 (searched against the NCBI nucleotides database)

| Description | Query coverage | Percent identity |
|---|---|---|
| Canine coronavirus isolate VuCCoV_191109SY, complete genome | 77% | 100% |
| Canine coronavirus isolate CD44_21052023 non-structural protein 3a (ORF3a), non-structural protein 3b (ORF3b), non-structural protein 3c (ORF3c), envelope protein (E), membrane protein (M), and nucleocapsid protein (N) genes, complete cds | 77% | 100% |
| Canine coronavirus isolate SMU-8, partial genome | 77% | 100% |
| Transmissible gastroenteritis virus strain TGEV/Mex/145/2008, complete genome | 82% | 98.65% |
| Transmissible gastroenteritis virus strain TGEV/USA/Z/1986, complete genome | 82% | 98.65% |
| Feline coronavirus 3c gene for truncated non-structural protein 3c, complete cds, isolate: FCoV/II/JP38/As/58/2015 | 82% | 98.65 % |
| Canine coronavirus isolate HLJ-073, complete genome | 82% | 98.65% |

virus; it was also confirmed by analyzing the signal peaks generated by Sanger sequencing (Fig. 2B).

### Phylogenetic analysis

Phylogenetic trees based on the nucleotide sequences of the complete TGEV, as well as the S and ORF3 genes, revealed that the VET-16 strain belongs to a Purdue subgroup within the traditional TGEV group (Fig. 3). TGEV strain VET-16 is closely related to the strains HB, NEB72-RT, and Purdue P115 but distinct from the TFI strain isolated in Southeast Asia. The phylogenetic tree based on the S gene showed that the VET-16 strain is closely related to the NEB72-RT strain, which was isolated from the respiratory tract of an infected animal. Phylogenetic analysis based on the ORF3 gene indicated that the VET-16 was most closely related to the Purdue P115 vaccine strain. RDP, BootScan, and SiScan analysis showed that the VET-16 strain had a *P*-value $> 10^{-6}$ and a recombination score $< 0.6$, and therefore, no evidence of a recombination event was found.

## Discussion

Analysis of the complete genome sequence of the VET-16 strain revealed that it contains crucial mutations in the spike gene. The amino acids at positions 72, 219, and 585 of the TGEV S protein are considered to be potential determinants of enteric tropism [14, 35]. A previous study suggested that aa substitutions at residue 72 (aspartic acid → asparagine) and residue 219 (alanine → serine) are associated with a loss of gut tropism [35]. Other studies have suggested that a substitution at aa 585 (serine → alanine) is a marker of attenuation [7, 13, 22]. Here, we found that the amino acids at positions 72, 219, and 585 of the S protein of VET-16 isolated from piglets in Vietnam were asparagine, serine, and alanine, respectively. Therefore, we predict that these substitutions may lead to attenuation of the virus due to a loss of intestinal tropism. A previous study showed that the APN binding site (aa 522–744) of the S protein is also associated with tissue tropism and virulence [13, 35]. Interestingly, the APN binding site of the VET-16 strain was found to have the same amino acid sequence as that of the NEB72-RT strain, which has lost intestinal tropism [14]. The S gene of the VET-16 strain contains a 2-aa deletion (N375_D376del) that is also found in the attenuated strains NEB72-RT and Purdue P115. Previous studies have shown that N375_D376del is also present in recombinant TGEV strains that show reduced replication in the enteric tract, which implies a loss of intestinal tropism [14, 36]. This may mean that the VET-16 strain is attenuated, with a reduced growth rate in enteric tissue.

Antigenic site D (aa 378-392) [20] is a neutralization epitope in TGEV [19, 37]. Unlike other TGEV strains, the VET-16 strain contains three aa substitutions (V378L, S379T, and D380N) and a 3-aa insertion (F383_F387insWEK) in antigenic site D. These six aa mutations might change the 3D structure of antigenic site D, which might in turn affect antigenicity and virulence. Mutations in antigenic site D of the VET-16 strain may change the epitope structure of the antigen. It is suspected that changes in the epitope structure may reduce viral pathogenicity.

The deletion of F1413 (resulting in loss of an aromatic amino acid in the TM domain [17], which anchors the S protein to the viral membrane [38]), may reduce the interfacial and hydrophobic properties of the TM peptide. Indeed, the loss of a hydrophobic or aromatic amino acid has been observed to cause a defect of viral fusion in recombinant SARS-CoV and murine coronaviruses *in vitro* [39, 40]. Mutations in the TM domain of the VET-16 strain may act as an attenuation factor by weakening cell-to-cell fusion; however, *in vitro* studies of growth kinetics are required to address this question.

The VET-16 strain also carries notable mutations in the ORF3a/b gene. PRCV, a variant of TGEV that harbors an ORF3a gene deletion, shows a loss of enteric tropism [22, 24]. A previous reverse genetics study demonstrated that deletions in the ORF3a/b gene of recombinant TGEV might be associated with a reduction in virulence and replication in

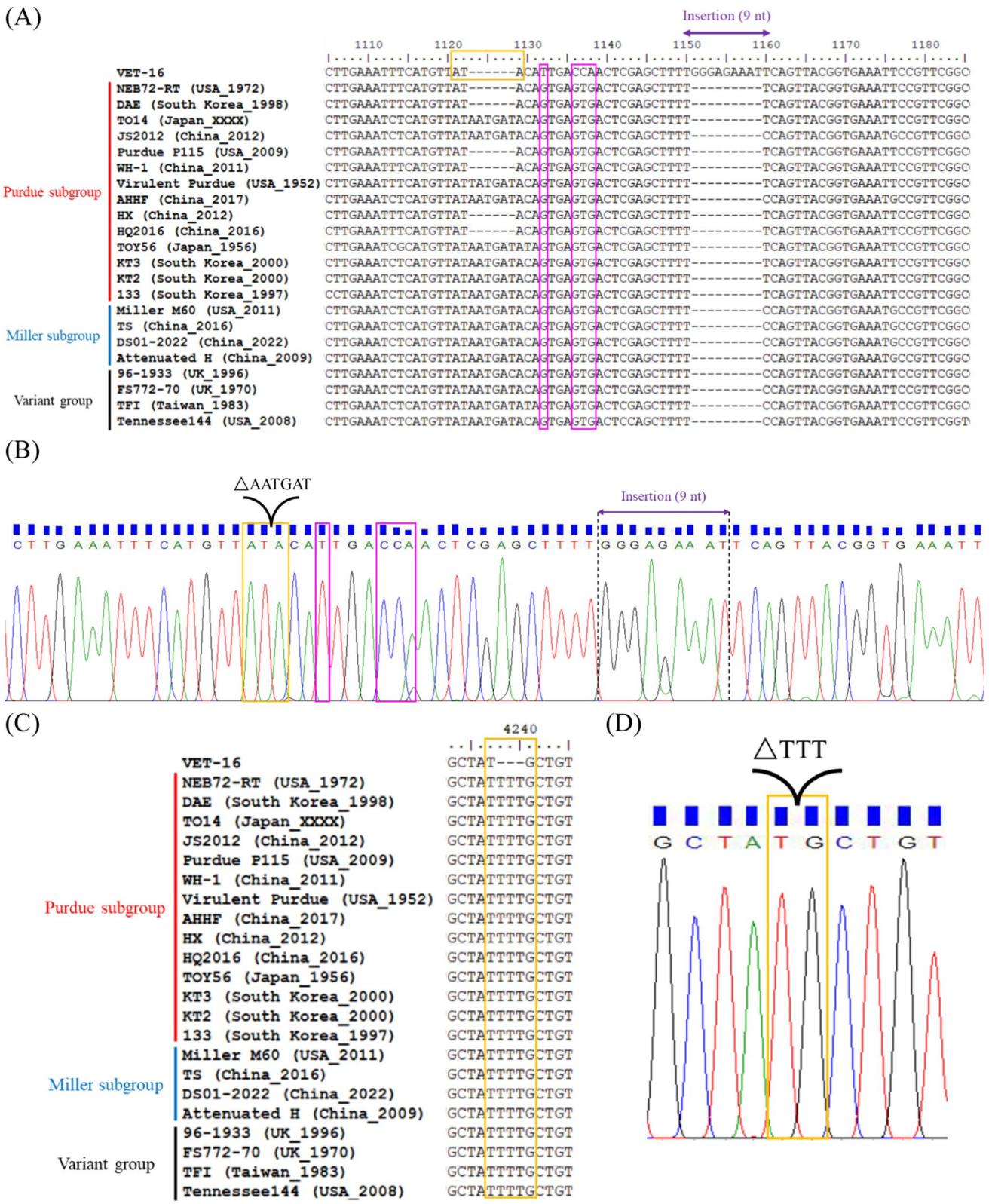


**Fig. 1** Multiple sequence alignment of the spike genes of TGEV strains. (A) Mutation of nucleotide sequences at positions 1122–1151. (B) Sanger sequencing chromatograms confirming a 3-nt deletion, four nt substitutions, and a 9-nt insertion in the VET-16 strain. (C) A 3-nt deletion in the 3' terminal region. (D) Sanger sequencing chromatograms confirming the presence of the 3-nt deletion shown in panel C. (E) Amino acid mutations in antigenic site D and in the transmembrane (TM) domain. Yellow and purple rectangles indicate nucleotide deletions and substitutions, respectively.

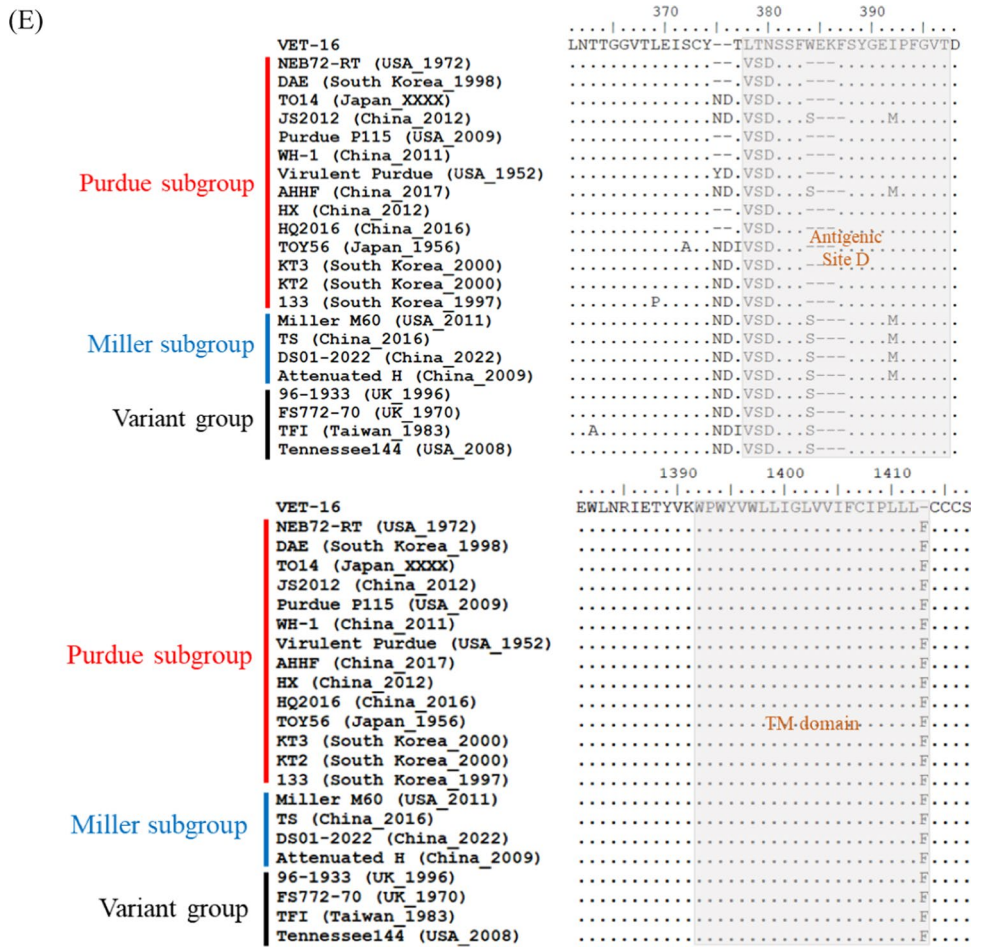


**Fig. 1** (continued)

**Table 4** Comparison of cell tropism-associated amino acid substitutions in the virulent and avirulent strains

| Strain | | Amino acid position† | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **72** | **219** | 525 | 536 | 562 | 564 | **585** | 590 | 615 | 649 | 671 | 675 | 702 |
| Novel strain | VET-16 | **N** | **S** | F | G | H | T | **A** | I | D | E | D | V | K |
| Virulent strain | DAE | * | * | * | * | * | * | * | * | * | * | * | * | * |
| | Virulent Purdue | **D** | **A** | * | D | N | * | **S** | * | * | * | * | L | Q |
| | Miller M6 | * | **A** | * | * | N | * | **S** | V | * | D | * | L | Q |
| | H16 | * | **A** | * | * | D | * | * | V | A | D | * | L | Q |
| Avirulent strain | NEB72-RT | * | * | * | * | * | * | * | * | * | * | * | * | * |
| | Purdue P115 | **D** | **A** | * | * | * | * | * | * | * | * | * | * | Q |
| | Miller M60 | * | **A** | * | * | N | N | * | V | * | D | * | L | Q |
| | Attenuated H | * | **A** | P | * | D | * | * | V | A | D | G | L | Q |

†Amino acid position in TGEV strain Virulent Purdue.

*Amino acid is the same as that in the VET-16 strain.

Aminopeptidase N (APN) binding site (aa 522–744).

Bold underlined letters denote amino residues strongly associated with cell tropism

pigs [36]. In comparison to the virulent Miller M60 strain, the Miller M6 strain harbors a large deletion in the ORF3b gene [13]. Taken together, these data suggest that deletions in ORF3 are associated with viral attenuation. The VET-16 strain contains a large deletion in the ORF3 gene, resulting in truncated ORF3a and ORF3b proteins. This may be

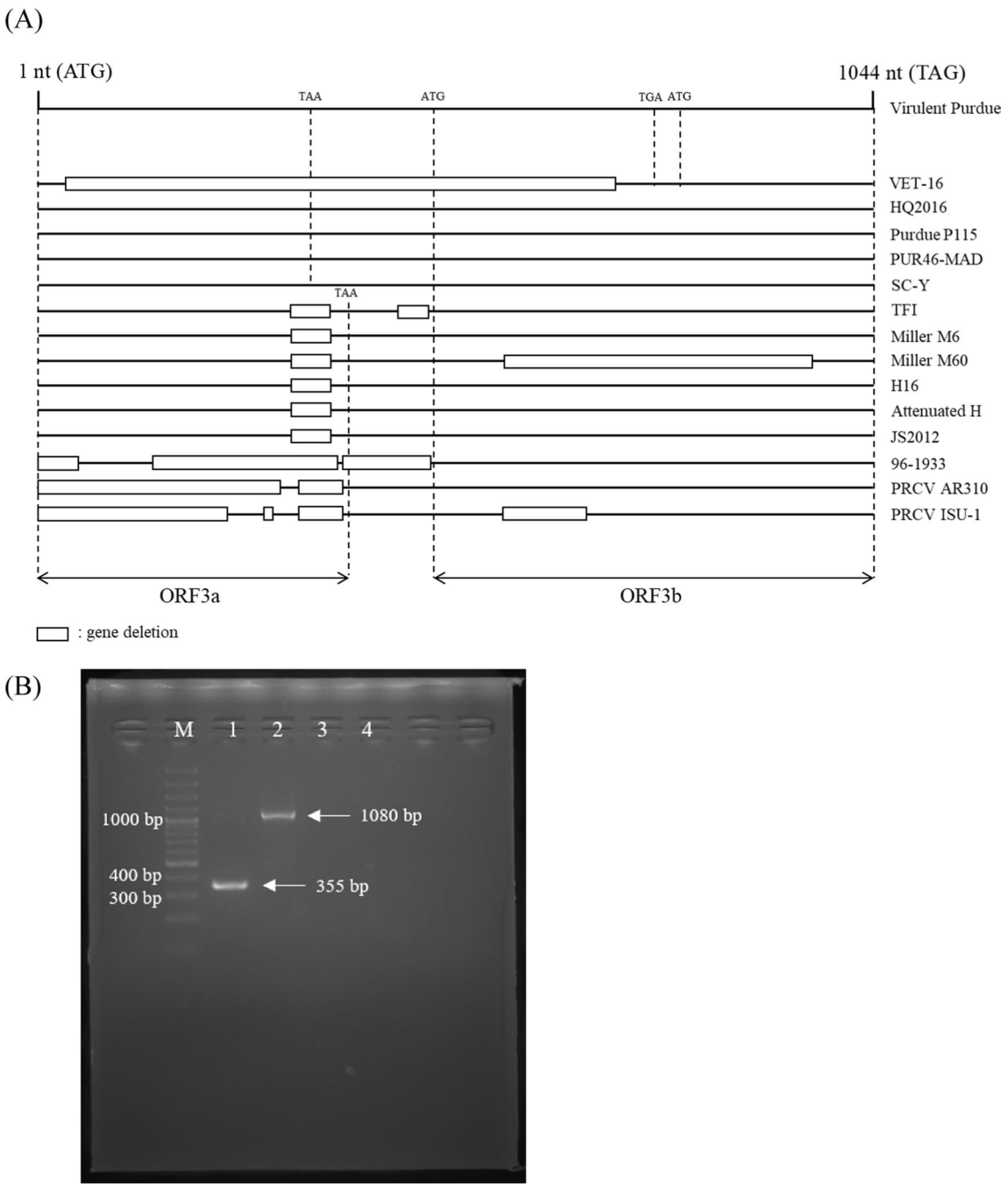


**Fig. 2** Electrophoresis gel of PCR products, confirming the presence of a large deletion in ORF3 (3a and 3b) of the VET-16 strain. (A) Comparison of gene deletions in ORF3 of the VET-16 strain with those in other TGEV strains and two PRCV strains. (B) PCR of the VET-16 strain, with a large deletion in ORF3, and a TGEV-positive control strain with no deletion in ORF3. Lane M, DNA (100 bp) ladder; lane 1, VET-16 strain; lane 2, TGEV, used as a positive control; lane 3, nuclease-free water; lane 4, no-template control

why the VET-16 strain causes only mild diarrhea in piglets. Evaluation of pathogenicity in newborn piglets is required to examine this further.

In general, TGEVs cause severe diarrhea or enteritis in piglets aged less than 2 weeks; however, on the Vietnamese farm where the VET-16 strain was isolated in 2016, infected piglets showed only mild diarrhea symptoms. These mild diarrhea symptoms are likely to be associated with the molecular features of the VET-16 strain identified in this study. If future studies confirm that the VET-16 strain is indeed attenuated, it may be a potential vaccine candidate. To demonstrate an association between the unique

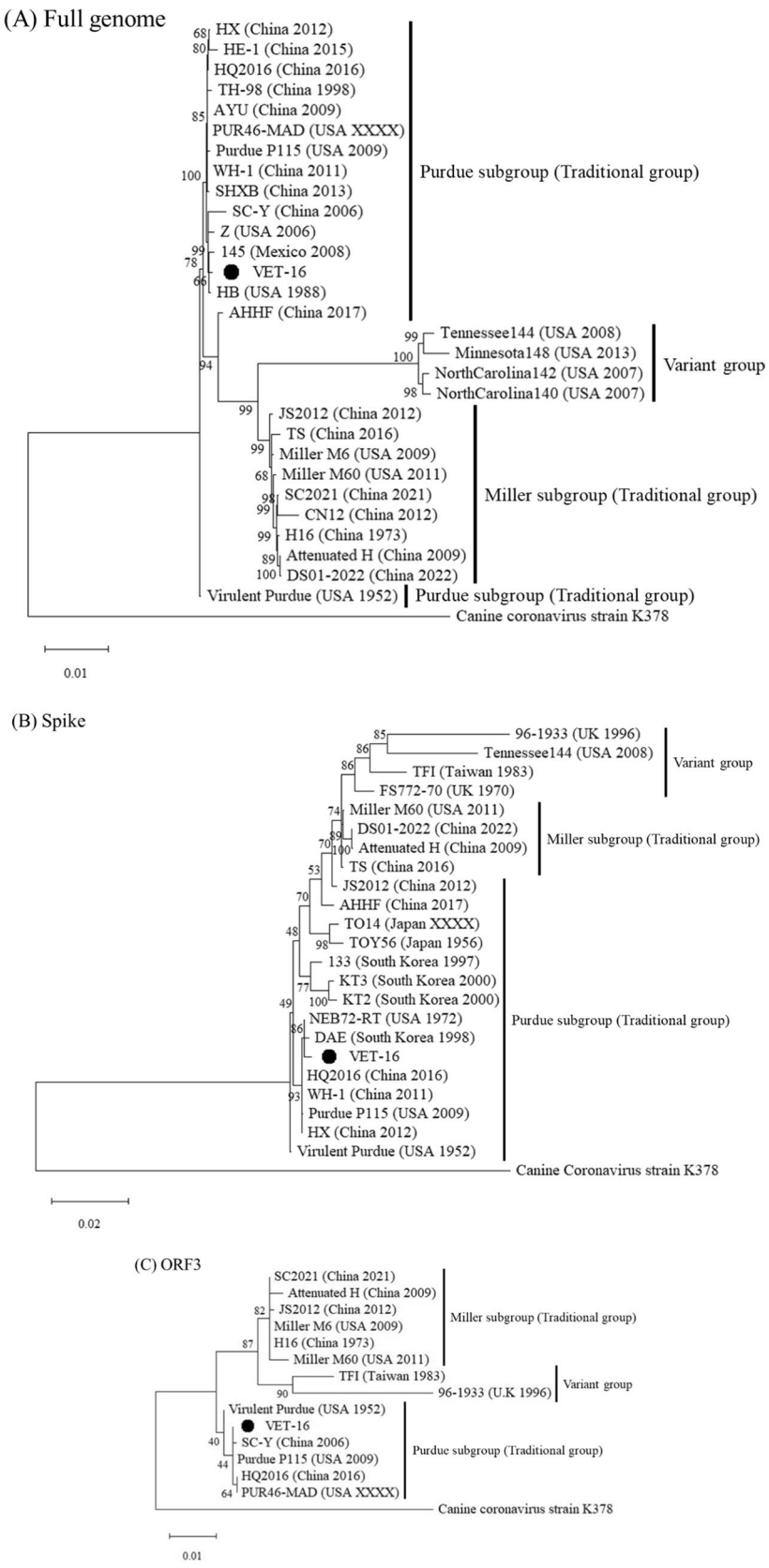


**Fig. 3** Phylogenetic trees based on nucleotide sequences of (A) the full genome, (B) the complete spike gene, and (C) the complete ORF3, built using the maximum-likelihood method with 1000 replicates in MEGA 11. Canine coronavirus strain K378 was used as an outgroup.

characteristics of the VET-16 strain and reduced pathogenicity, clinical signs such as diarrhea should be evaluated in piglets inoculated with the VET-16 strain.

In conclusion, molecular characterization of the VET-16 strain isolated from piglets in Vietnam identified 10 genetic mutations in the S gene and a large deletion in the ORF3 gene. These genetic data suggest that the VET-16 strain may be attenuated and have reduced enteric tropism. Therefore, the VET-16 strain will be a helpful reference for future studies of TGEV evolution.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00705-024-06101-8.

**Acknowledgments** The authors thank Dr. Hyun-Mi Pyo (Baronbio Company, Gimcheon, South Korea) for editing the English text.

**Author contributions** SeEun Choe and Gyu-Nam Park conceived and designed the study. Soo Hyun Moon and Sok Song supervised data collection and analysis. Van Giap Nguyen was responsible for sample collection. Dong-Jun An and Yun Sang Cho checked and finalized the manuscript. All authors read and agreed to the published version of the manuscript.

**Funding** This project was supported by grants (project code no. M-1543083-2023-24-02) from the Animal and Plant Quarantine Agency, Republic of Korea.

**Data availability** The whole genome sequence of the VET-16 strain determined in this study has been deposited in the GenBank database under accession number PP236916.

## Declarations

**Conflict of interest** The authors declare no conflicts of interest.

**Ethical approval** This article does not report any experiments involving human participants or animals.

## References

1. Yuan D, Yan Z, Li M, Wang Y, Su M, Sun D (2021) Isolation and characterization of a porcine transmissible gastroenteritis coronavirus in Northeast China. Front Vet Sci 8:611721
2. Chen Y, Zhang Y, Wang X, Zhou J, Ma L, Li J, Yang L, Ouyang H, Yuan H, Pang D (2023) Transmissible gastroenteritis virus: an update review and perspective. Viruses 15(2):359
3. Enjuanes L, Smerdou C, Castilla J, Antón IM, Torres JM, Sola I, Golvano J, Sánchez JM, Pintado B (1995) Development of protection against coronavirus induced diseases. A review. Adv Exp Med Biol 380:197–211
4. Laude H, Rasschaert D, Delmas B, Godet M, Gelfi J, Charley B (1990) Molecular biology of transmissible gastroenteritis virus. Vet Microbiol 23:147–154
5. Doyle L, Hutchings L (1946) A transmissible gastroenteritis in pigs. J Am Vet Med Assoc 108:257–259
6. McGoldrick A, Lowings JP, Paton DJ (1999) Characterisation of a recent virulent transmissible gastroenteritis virus from Britain with a deleted ORF 3a. Arch Virol 144:763–770
7. Cubero MJ, León L, Contreras A, Astorga R, Lanza I, Garcia A (1993) Transmissible gastroenteritis in pigs in south east Spain: prevalence and factors associated with infection. Vet Rec 132:238–241
8. Akimkin V, Beer M, Blome S, Hanke D, Höper D, Jenckel M, Pohlmann A (2016) New chimeric porcine coronavirus in swine feces, Germany, 2012. Emerg Infect Dis 22:1314–1315
9. Guo R, Fan B, Chang X, Zhou J, Zhao Y, Shi D, Yu Z, He K, Li B (2020) Characterization and evaluation of the pathogenicity of a natural recombinant transmissible gastroenteritis virus in China. Virology 545:24–32
10. Takahashi K, Okada K, Ohshima K (1983) An outbreak of swine diarrhea of a new-type associated with coronavirus-like particles in Japan. Nihon Juigaku Zasshi 45:829–832
11. Cheng S, Wu H, Chen Z (2020) Evolution of transmissible gastroenteritis virus (TGEV): a codon usage perspective. Int J Mol Sci 21:7978
12. Chen F, Knutson TP, Rossow S, Saif LJ, Marthaler DG (2019) Decline of transmissible gastroenteritis virus and its complex evolutionary relationship with porcine respiratory coronavirus in the United States. Sci Rep 9:3953
13. Zhang X, Hasoksuz M, Spiro D, Halpin R, Wang S, Stollar S, Janies D, Hadya N, Tang Y, Ghedin E, Saif L (2007) Complete genomic sequences, a key residue in the spike protein and deletions in nonstructural protein 3b of US strains of the virulent and attenuated coronaviruses, transmissible gastroenteritis virus and porcine respiratory coronavirus. Virology 358:424–435
14. Sánchez CM, Gebauer F, Suñé C, Mendez A, Dopazo J, Enjuanes L (1992) Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. Virology 190:92–105
15. Zhang X, Zhu Y, Zhu X, Shi H, Chen J, Shi D, Yuan J, Cao L, Liu J, Dong H, Jing Z, Zhang J, Wang X, Feng L (2017) Identification of a natural recombinant transmissible gastroenteritis virus between Purdue and Miller clusters in China. Emerg Microbes Infect 6:e74
16. Chen CM, Pocock DH, Britton P (1993) Genomic organisation of a virulent Taiwanese strain of transmissible gastroenteritis virus. Adv Exp Med Biol 342:23–28
17. Sanchez CM, Pascual-Iglesias A, Sola I, Zuñiga S, Enjuanes L (2019) Minimum determinants of transmissible gastroenteritis virus enteric tropism are located in the N-terminus of spike protein. Pathogens 9(1):2
18. Delmas B, Rasschaert D, Godet M, Gelfi J, Laude H (1990) Four major antigenic sites of the coronavirus transmissible gastroenteritis virus are located on the amino-terminal half of spike glycoprotein S. J Gen Virol 71(Pt 6):1313–1323
19. Enjuanes L, Suñé C, Gebauer F, Smerdou C, Camacho A, Antón IM, González S, Talamillo A, Méndez A, Ballesteros ML et al (1992) Antigen selection and presentation to protect against transmissible gastroenteritis coronavirus. Vet Microbiol 33:249–262
20. Gebauer F, Posthumus WP, Correa I, Suñé C, Smerdou C, Sánchez CM, Lenstra JA, Meloen RH, Enjuanes L (1991) Residues involved in the antigenic sites of transmissible gastroenteritis coronavirus S glycoprotein. Virology 183:225–238
21. Kim L, Hayes J, Lewis P, Parwani AV, Chang KO, Saif LJ (2000) Molecular characterization and pathogenesis of transmissible gastroenteritis coronavirus (TGEV) and porcine respiratory coronavirus (PRCV) field isolates co-circulating in a swine herd. Arch Virol 145:1133–1147
22. Li JQ, Cheng J, Lan X, Li XR, Li W, Yin XP, Li BY, Yang B, Li ZY, Zhang Y, Liu JX (2010) Complete genomic sequence of transmissible gastroenteritis virus TS and 3' end sequence characterization following cell culture. Virol Sin 25:213–224
23. Zhang X, Zhu Y, Zhu X, Chen J, Shi H, Shi D, Dong H, Feng L (2017) ORF3a deletion in field strains of porcine-transmissible gastroenteritis virus in China: a hint of association with porcine respiratory coronavirus. Transbound Emerg Dis 64:698–702

24. Keep S, Carr BV, Lean FZX, Fones A, Newman J, Dowgier G, Freimanis G, Vatzia E, Polo N, Everest H, Webb I, McNee A, Paudyal B, Thakur N, Nunez A, MacLoughlin R, Maier H, Hammond J, Bailey D, Waters R, Charleston B, Tuthill T, Britton P, Bickerton E, Tchilian E (2022) Porcine respiratory coronavirus as a model for acute respiratory coronavirus disease. Front Immunol 13:867707
25. Page KW, Mawditt KL, Britton P (1991) Sequence comparison of the 5' end of mRNA 3 from transmissible gastroenteritis virus and porcine respiratory coronavirus. J Gen Virol 72(Pt 3):579–587
26. Lu Y, Huang W, Zhong L, Qin Y, Liu X, Yang C, Wang R, Su X, Du C, Mi X, Wang H, He Y, Zhao W, Chen Y, Wei Z, Ouyang K (2021) Comparative characterization and pathogenicity of a Novel Porcine Epidemic Diarrhea Virus (PEDV) with a naturally occurring truncated ORF3 gene coinfected with PEDVs possessing an intact ORF3 gene in piglets. Viruses 13(8):1562
27. Lam JY, Yuen CK, Ip JD, Wong WM, To KK, Yuen KY, Kok KH (2020) Loss of orf3b in the circulating SARS-CoV-2 strains. Emerg Microbes Infect 9:2685–2696
28. Lednicky JA, Cherabuddi K, Tagliamonte MS, Elbadry MA, Subramaniam K, Waltzek TB, Morris JG Jr (2021) In-frame 12-nucleotide deletion within open reading frame 3a in a SARS-CoV-2 strain isolated from a patient hospitalized with COVID-19. Microbiol Resour Announc 10(8):e00137-e221
29. Lu Y, Su X, Du C, Mo L, Ke P, Wang R, Zhong L, Yang C, Chen Y, Wei Z, Huang W, Liao Y, Ouyang K (2020) Genetic diversity of porcine epidemic diarrhea virus with a naturally occurring truncated ORF3 gene found in Guangxi. China. Front Vet Sci 7:435
30. Xu L, Dai H-b, Luo Z-p, Zhu L, Zhao J, Lee F-q, Liu Z-y, Nie M-c, Wang X-t, Zhou Y-c, Xu Z-w (2023) Characterization and evaluation of the pathogenicity of a natural gene-deleted transmissible gastroenteritis virus in China. Transbound Emerg Dis 2023:2652850
31. Zhang YH, Li HX, Chen XM, Zhang LH, Zhao YY, Luo AF, Yang YR, Zheng LL, Chen HY (2022) Genetic characteristics and pathogenicity of a novel porcine epidemic diarrhea virus with a naturally occurring truncated ORF3 gene. Viruses 14(3):487
32. Beall A, Yount B, Lin CM, Hou Y, Wang Q, Saif L, Baric R (2016) Characterization of a pathogenic full-length cDNA clone and transmission model for porcine epidemic diarrhea virus strain PC22A. MBio 7:e01451-e1415
33. Liu Y, Zhang X, Liu J, Xia H, Zou J, Muruato AE, Periasamy S, Kurhade C, Plante JA, Bopp NE, Kalveram B, Bukreyev A, Ren P, Wang T, Menachery VD, Plante KS, Xie X, Weaver SC, Shi PY (2022) A live-attenuated SARS-CoV-2 vaccine candidate with accessory protein deletions. Nat Commun 13:4337
34. Wang K, Lu W, Chen J, Xie S, Shi H, Hsu H, Yu W, Xu K, Bian C, Fischer WB, Schwarz W, Feng L, Sun B (2012) PEDV ORF3 encodes an ion channel protein and regulates virus production. FEBS Lett 586:384–391
35. Ballesteros ML, Sánchez CM, Enjuanes L (1997) Two amino acid changes at the N-terminus of transmissible gastroenteritis coronavirus spike protein result in the loss of enteric tropism. Virology 227:378–388
36. Sola I, Alonso S, Zúñiga S, Balasch M, Plana-Durán J, Enjuanes L (2003) Engineering the transmissible gastroenteritis virus genome as an expression vector inducing lactogenic immunity. J Virol 77:4357–4369
37. Jiménez G, Correa I, Melgosa MP, Bullido MJ, Enjuanes L (1986) Critical epitopes in transmissible gastroenteritis virus neutralization. J Virol 60:131–139
38. Huang Y, Yang C, Xu X-f, Xu W, Liu S-w (2020) Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacol Sin 41:1141–1149
39. Bos EC, Heijnen L, Spaan WJ (1995) Site directed mutagenesis of the murine coronavirus spike protein. Effects on fusion. Adv Exp Med Biol 380:283–286
40. Corver J, Broer R, van Kasteren P, Spaan W (2009) Mutagenesis of the transmembrane domain of the SARS coronavirus spike glycoprotein: refinement of the requirements for SARS coronavirus cell entry. Virol J 6:230