Molecular characterization and phylogenetic analysis of transmissible gastroenteritis virus HX strain isolated from China

Xiaoliang Hu Jr1,2, Nannan Li Jr1, Zhige Tian Jr3, Xin Yin Jr2, Liandong Qu4* and Juanjuan Qu*1

Abstract

Background: Porcine transmissible gastroenteritis virus (TGEV) is the major etiological agent of viral enteritis and severe diarrhea in suckling piglets. In China, TGEV has caused great economic losses, but its role in epidemic diarrhea is unclear. This study aims to reveal the etiological role of TGEV in piglet diarrhea via molecular characterization and phylogenetic analysis.

Results: A TGEV-HX strain was isolated from China, and its complete genome was amplified, cloned, and sequenced. Sequence analysis indicated that it was conserved in the 5′- and 3′-non-translated regions, and there were no insertions or deletions in nonstructural genes, such as ORF1a, ORF1b, ORF3a, ORF3b, and ORF7, as well as in genes encoding structural proteins, such as the envelope (E), membrane (M), and nucleoprotein (N) proteins. Furthermore, the phylogenetic analysis indicated that the TGEV-HX strain was more similar to the TGEV Purdue cluster than to the Miller cluster.

Conclusions: The present study described the isolation and genetic characterization of a TGEV-HX strain. The detailed analysis of the genetic variation of TGEVs in China provides essential information for further understanding the evolution of TGEVs.

Keywords: Porcine transmissible gastroenteritis virus, Sequence alignment, Phylogenetic analysis

Background

Transmissible gastroenteritis virus (TGEV) is the etiological agent of transmissible gastroenteritis (TGE), and it can cause viral enteritis and severe diarrhea with high morbidity in pigs of all ages, as well as high mortality in suckling piglets [1]. It occurs at swine-raising farms and results in significant economic losses [2,3].

TGEV is an enveloped virus belonging to the Coronaviridae (CoV) family and the Nidovirales order. It possesses a large 28.5-kb single-stranded, positive-sense RNA genome. About two-thirds of the entire RNA comprises open reading frames (ORFs) 1a and 1b, encoding RNA replicase. The 3′-one-third of the genome comprises genes encoding structural and non-structural proteins [4,5]. The genes of TGEV are arranged in the order of 5′-rep-S-3a-3b-E-M-N-ORF7-3′.

The spike (S) gene of TGEV encodes an approximately 1,450-amino acid protein, with a molecular weight ranging from 128–160 kDa without glycosylation and 150–200 kDa after glycosylation. Functionally, the S glycoprotein is the major target of neutralizing antibodies, and it is also related to host cell tropism [6], interaction with its cellular receptor, pathogenicity, fusion, and hemagglutination activity [7-9]. ORF7 encodes a small hydrophobic protein (HP) during viral replication. The intracellular localization of the HP suggests that it may play an important role in the process of membrane integrity during viral replication and/or virion assembly [10,11].

In this report, we isolated a TGEV-HX strain of TGEV from the feces of piglets in Heilongjiang province in China. To better understand the molecular characteristics of this isolate, its complete genome sequence was
obtained, and a phylogenetic tree was constructed based on the complete ORF sequence of the S gene. The results provide molecular and phylogenetic information for a Chinese isolate of TGEV, which may assist in elucidating the genetic evolution of TGEV in China.

Methods

Ethics statement

Pigs used in this study were approved by the Institutional Animal Care and Use Committee (IACUC) of the Harbin Veterinary Research Institute (HVRI), the Chinese Academy of Agricultural Sciences. No animals were sacrificed specifically for this study. Feces samples were collected at the farm.

Viral isolation and identification

Five fecal samples were collected from piglets with diarrhea in a suburb of Harbin, the capital of Heilongjiang province, P. R. China. As both TGEV and PEDV can cause diarrhea, two pairs of specific primers (TGEV-NF; TGEV-NR; PEDV-NF, PEDV-NR) were employed to identify the kinds of viruses in the above samples (Table 1). PCRs were conducted as below, and the cycling parameters for the PCR included 94°C for 5 min, followed by 30 cycles of 94°C for 0.5 min, 55°C for 0.5 min, and 72°C for 1 min, and a final extension at 72°C for 10 min. Then, PCR-positive viral samples were inoculated into PK-15 cells, which were grown as a monolayer in Dulbecco’s modified Eagle medium (DMEM) (GIBCO, Grand Island, NY, USA) containing 10% fetal calf serum (GIBCO, Grand Island, NY, USA) and 5% CO₂ in air. Viruses were passaged three times and were harvested by three cycles of freezing and thawing. Cellular debris was removed by low speed centrifugation at 3,000 × g (Eppendorf, Hamburg, Germany) for 10 min, and the supernatant was aliquoted and stored at −20°C. Viral titers were determined using the Reed-Muench method [12].

Extraction of viral RNA, reverse transcribed-polymerase chain reaction (RT-PCR) and complete genome sequencing

Viral RNA was extracted from PK-15 cells infected with TGEV-HX using the TRIzol reagent (Invitrogen, Carlsbad, CA, USA). cDNA was generated by adding 4 μl of RNA to the following components: 4 μl of 5× reverse transcription buffer, 4 μl of dNTPs mixture (2.5 mM), 0.5 μl of RNase inhibitor, 5 μl of random primer (50 μM), 0.5 μl (10 U) of AMV reverse transcriptase, and 2 μl of sterile water. The components were gently mixed in an Eppendorf tube and incubated at room temperature for 10 min, then transferred to a water incubator at 42°C for 1 h prior to storage at −20°C. The resulting cDNA was amplified by PCR using LA Taq DNA polymerase (TaKaRa, Tokyo, Japan). Ten pairs of primers were designed based on conserved regions of TGEV strain H165 (Table 1). PCRs were conducted in a total of volume of 50 μl containing 5 μl of 10× buffer, 3 μl of dNTPs mixture (2.5 mM), 8 μl of cDNA, 1 μl of forward primer (10 μM), 1 μl of reverse primer (10 μM), 5 U of LA Taq polymerase (TaKaRa, Tokyo, Japan), and 31 μl of sterile water. The cycling parameters for the PCR included 94°C for 1 min, followed by 30 cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 1 min, and a final extension at 72°C for 10 min. Two primers (P-F, P-R) were employed to confirm the 5′ and 3′ ends of the viral genome by rapid amplification of cDNA ends (RACE) using the RACE cDNA Amplification Kit (Invitrogen, Carlsbad, CA, USA). The PCR products were run on agarose gels, and correctly sized amplicons were observed. Then, the PCR products were purified using the Axygen Gel Extraction Kit (Axygen, USA) and cloned into pMD18-T (TaKaRa, Tokyo, Japan). Three to five independent clones of each TGEV amplicon were sequenced. The DNA was

| Primer  | Sequence 5′-3′ | Position |
|---------|---------------|----------|
| TGEV-NF | TCA1GCAGATGCCAAATTTAAGA | 27,213-27,236 |
| TGEV-NR | TCA1CCTTCTCGTTATGGAATGTT | 27,456-27,479 |
| PEDV-NF | TTTTCAAGGTACTGGCAAAATAATG | 26,382-26,405 |
| PEDV-NR | TTGGAGATCTGGACCTGTTGC | 26,757-26,780 |
| P1      | ATGAGTCCTCAACATTCGAAAG | 315-336 |
| P2      | ATC1AAAACATCCAAAGCACCCT | 4,318-4,339 |
| P3      | AATTCAAGTCCTAAAAACGAT | 4,250-4,271 |
| P4      | GCGTACGATCTCATAAAGACGG | 4,870-4,891 |
| P5      | CATTGTCACCCTTGTTGAAC | 4,820-4,841 |
| P6      | GTAGATGCTAAAAGCTCTACTA | 12,400-12,421 |
| P7      | TCTATGCAGATTTTACTGTTG | 12,300-12,321 |
| P8      | TAAATGATTTAGCTGTTTCC | 15,200-15,221 |
| P9      | AGGCATGTTGATGATGTTG | 15,100-15,121 |
| P10     | AAGCTTGACAAAGCTCT | 18,169-18,187 |
| P11     | CGC1ACTGCTTAAATGCTCT | 18,080-18,101 |
| P12     | ACTAGCTTAAACGTTGCTTG | 20,660-20,681 |
| P13     | ATGAAAACACTATTGGTTG | 20,365-20,383 |
| P14     | TTAATGAGCTGCTTATTTC | 24,690-24,708 |
| P15     | GCTGTGATGATCATAAACG | 24,608-24,629 |
| P16     | TTGGAGGTTATGGGTTGAG | 27,021-27,042 |
| P17     | CTTCTAAATGCGCAACCCAG | 26,910-26,932 |
| P18     | CGACGTACCTGCTTATTGCTTG | 28,056-28,077 |
| P19     | AGAAGGTGACAGCTTATG | 27,963-27,984 |
| P20     | GTATGACTATCATAAAGAG | 28,559-28,580 |
| P-F     | CAAAATGATGATCATAAACG | 139-160 |
| P-R     | ATTTGGCAATGCTAGTATGAA | 28,441-28,464 |
sequenced using an ABI 3730XL Sanger-based Genetic Analyzer (Applied Biosystems, Waltham, MA, USDA).

Electron microscopy
PK-15 cells infected with TGEV were harvested by freezing and thawing three times. One mL of cell culture was centrifuged for 5 min at 800 × g. The supernatant was transferred into a new microfuge tube and centrifuged for 10 min at 13,400 × g. Then, the pellet was negatively stained with 2% phosphotungstic acid and analyzed on a transmission electron microscope (H-7650, Hitachi, Tokyo, Japan) [13].

Sequence alignment and phylogenetic analysis
Sequence data were assembled and analyzed using Clustal X software (1.83) and DNASTAR. To determine the relationship between the TGEV representative isolates and the HX strain, phylogenetic trees based on the S gene were constructed using molecular evolutionary genetics analysis (MEGA) software (version 4.0) using the neighbor-joining (NJ) method. Bootstrap values were estimated for 1,000 replicates. The sequences of the TGEV reference strains were obtained from GenBank, and the details are summarized in Table 2. The sequences obtained in this study were assembled and submitted to GenBank under the accession number KC962433.

Results

Virus culture and electron microscopy analysis
The PCR results confirmed that one of five samples was TGEV-positive, designated TGEV-HX; the other four samples were TGEV-negative and all five samples were PEDV-negative (Figure 1A). After three passages, cytopathic effects (CPE) were found in the PK-15 cells, as evidenced by the cells rounding up and enlarging, the formation of syncytia, and the detachment of cells into

Table 2 Source of transmissible gastroenteritis virus (TGEV) sequences used in the experiment

| Isolate  | Accession no. | Origin | Collection date |
|----------|---------------|--------|-----------------|
| 96-1933  | AF104420      | UK     | 2001            |
| H16      | FJ755618.2    | CHN    | 2010            |
| H165     | EU074218      | CHN    | 2010            |
| TGEV-HX  | KC962433      | CHN    | 2013            |
| ISU-1    | DQ811787      | USA    | 2009            |
| Miller 6 | DQ811785      | USA    | 2009            |
| Miller 60| DQ811786.2    | USA    | 2011            |
| Purdue 115| DQ811788.1   | USA    | 2011            |
| SC-Y     | DQ443743      | CHN    | 2006            |
| TS       | DQ201447      | CHN    | 2006            |
| WH-1     | HQ62571.1     | CHN    | 2011            |

Figure 1 Identification and Isolation of TGEV-HX. (A) Identification of five transmissible gastroenteritis virus (TGEV) samples by PCR. (B) Cytopathic effect (CPE) induced by TGEV HX in the PK-15 cell line. (C) Control (uninfected) PK-15 cells. (D) Electron micrograph of the purified isolate negatively stained with 2% phosphotungstic acid. The scale bar represents 500 nm.
Figure 2 Multiple sequence alignment of S among TGEV strains. (A) A 6-nt deletion in the S gene at nt position 1123–1128 of the HX, SC-Y, and WH-1 strains. (B) Amino acid alignments of deduced S sequences compared with strain HX. (▲) indicates amino acid SBS, (★) indicates amino acids of the HX, SC-Y, and WH-1 strains that are different from those of other transmissible gastroenteritis viruses.
the medium (Figure 1B and C). The median tissue culture infective dose (TCID$_{50}$) of TGEV-HX ($10^{7.25}$/0.1 ml) was measured using the Reed-Muench method. When observed by electron microscopy, the virus displayed a circular shape, and the surface projections were petal-shaped, with a diameter ranging from 100 to 150 nm, which is similar in size to known TGEVs (Figure 1D).

Complete genome sequence of the TGEV-HX strain

The full-length genome sequence of the TGEV-HX strain was deduced by combining the sequences of 10 overlapping cDNA fragments. The genome sequence of the TGEV-HX strain was 28,580 nucleotides (nt) long, including the poly A tail. The 5′ portion of the genome contained a 314-nt non-translated region (NTR), and ORF1a (nt 315–12,368) and ORF1b (12,326–20,368), encoding the viral RNA-dependent RNA replicase. The structural proteins S, E, M, and N were found to be encoded by ORFs S (nt 20,365–24,708), E (nt 25,857–26,105), M (nt 26,116–26,904), and N (nt 26,917–28,065), respectively. The three non-structural protein-coding genes were ORF3a (nt 24,827–25,042), ORF3b (nt 25,136–25,870), and ORF7 (nt 28,071–28,307). The 5′ NTR included a potential short AUG-initiated ORF (nt 114–121), beginning within a Kozak sequences (5′-UCUAUGAA-3′). The 3′ end of the genome contained a 273-nt untranslated sequence and a poly (A) tail. Upstream from the poly (A) tail, there was a 5′-GGAA GAGC-3′ octameric sequence.

The non-structural genes

The replicase genes were composed of ORF1a and ORF1b, which contained a 43-nt common region (nt 12,326–12,368) and a “slippery site” (5′-UUUAAC-3′, nt 12,333–12,339). The ORF1a gene of TGEV-HX was predicted to encode a protein of 4,017 amino acids (aa), while ORF1b was predicted to encode a 2,680-aa protein. Nucleotide sequence analysis indicated that there were no deletions or insertions in the ORF1ab region of the Miller 6 and Purdue TGEV strains.

ORF3a and ORF3b of TGEV-HX were predicted to encode 72-aa and 244-aa proteins, respectively. No deletions or insertions were found in the ORF3a or ORF3b genes of TGEV-HX. The ORF7 gene of TGEV-HX was predicted to encode a 78-aa protein, which contained the common PP1c-binding motif 5′-RVIFLVI-3′ [14]. No deletions or insertions were found in ORF7 of TGEV-HX.

The structural genes

The nucleotide sequence of the S gene of TGEV-HX was 4,344 nt in length, encoding a predicted protein of 1,447 aa. A 6-nt deletion was found in the S gene at nt 1,123–1,128 of the TGEV-HX, SC-Y, and WH-1 strains, which caused the S protein to be two amino acids shorter than those of the Purdue, Miller 6, TS, H16 and H165 strains (Figure 2A). Amino acid 585 of the Purdue, Miller 6, and TS strains was S, while in the TGEV-HX, SC-Y, WH-1, H16, and H165 strains, it was A (Figure 2B). Amino acids 32, 208, 403, 418, 496, 562, 675, 1,109, and 1,234 of TGEV-HX were identical to those of strains SC-Y and WH-1, but different from those of strains TS, H16, and H165.

Sequence analysis revealed no deletions or insertions in the E and N genes of any of the TGEVs. The predicted E, M, and N proteins were 82-, 264-, and 382-aa long, respectively (Table 3).

Homology comparison

To investigate the homology of TGEV-HX to other TGEVs, the nucleotide and predicted amino acid sequences of the nonstructural and structural protein coding genes were compared (Table 4). The results suggested that TGEV-HX showed higher identity to strains SC-Y, WH-1, and Purdue, and less identity to TS, Miller 6, and H165.

Phylogenetic analysis

Phylogenetic analysis of the complete S gene showed that TGEV strains could be divided into two groups (Figure 3). The TGEV-HX strain had a close relationship with the Purdue strain and is more distant evolutionarily from the Miller strains group and strain ISU-1.

Discussion

As an enteropathogenic coronavirus, TGEV is the major cause of viral enteritis and diarrhea in neonatal pigs, resulting in significant economic losses. Currently, TGEV occurs sporadically in parts of Europe, North America, and Asia. The fact that wild and domestic carnivores (foxes, dogs, cats, and possibly minks) seroconvert to
TGEV indicates that they are potential subclinical carriers of TGEV. [15]. In summation, TGEV has become a new challenge for the pig farming industry. As few TGEV genome sequences have been published, little is known about TGEV evolution. The results of this study will provide necessary information for further understanding the evolution of TGEV.

A complete sequence analysis indicated that no deletions or insertions were found in the 5′- and 3′- NTR regions of TGEV-HX, suggesting that its replication and transcription mechanism was not changed. A 6-nt (nt 1,123–1,128) deletion in the S gene was found in the TGEV-HX, SC-Y, and WH-1 strains, but not in the Purdue, Miller 6, TS, H16, and H165 strains. It was showed that this deletion was observed in the attenuated Purdue strains PUR46-C8 and PUR46-MAD, and it was considered to play a role in viral attenuation [16]. Competition studies using monoclonal antibodies led to the prediction of at least four main antigenic sites, designated A, B, C, and D [17,18]. The A and B sites (aa 506–706) have been mapped, and they serve as the major antigenic sites, including the binding site for the viruses host receptor, aminopeptidase N (APN). Single amino acid changes in the S protein can greatly impact its antigenicity [19]. The serine to alanine mutation at amino acid 585 may have a significant influence in receptor binding or neutralizing antibody interactions [18]. Five Chinese strains, TGEV-HX, SC-Y, WH-1, H16, and H165, had and alanine residue at this position, while the Purdue, Miller 6, and TS strains had a serine residue, which suggested that the antigenicity of the S protein of the five Chinese strains may be changed. Furthermore, amino acids 32, 208, 376, 403, 418, 496, 562, 675, 1,109, and 1,234 of TGEV-HX were identical to those of strains SC-Y and WH-1, but different from those of strains TS, H16, and H165.

The nucleotide and amino acid sequence homology analysis of the structural proteins and non-structural proteins indicated that TGEV-HX was highly similar to the SC-Y, Purdue, and WH-1 strains, and had a lower sequence similarity to the Miller 6, TS, H16, and H165 strains. The phylogenetic analysis showed that TGEV-HX was closely related to the SC-Y, WH-1, and Purdue strains, which belonged to the Miller cluster, while the Miller 6, TS, H16, and H165 strains belonged to the Miller cluster, which was consistent with the results of the homology comparison. The data obtained in this study indicated that HX had different ancestors than the early Chinese strain H16, and it might be derived from the same ancestor as the SC-Y, WH-1, and Purdue strains.

Conclusions
The present study provides the complete genome sequence of a TGEV-HX strain from China. By comparing the S gene and protein with those of other TGEV strains, we have gained a further understanding of the genetic structure, diversity, and evolution of the TGEV-HX strain. Our next work is to evaluate the characteristics of mutations in the S gene using a reverse genetic approach in animal experiments.
Transmissible gastroenteritis virus; CoV: Coronaviridae; ORF: Open reading frames; S: Spike; HP: Hydrophobic protein.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
HXL participated in the study design and collection of samples, conducted all laboratory and statistical analyses, participated in the interpretation of analyses, and drafted the manuscript. LNN, YX, and TZG participated in the study, provided oversight of the laboratory analysis, participated in interpretation of analyses, and helped to draft and critically revise the manuscript. QJJ and QLD participated in the study design and interpretation of analyses, and helped to draft and critically revise the manuscript. All authors have read and approved the final manuscript.

Acknowledgment
This work was supported by the National Ministry of Science and Technology of China (no. 2012BAD46B01).

Received: 7 August 2014 Accepted: 5 March 2015
Published online: 21 March 2015

References
1. Enjuanes L, Smerdou C, Castilla J, Anton IM, Torres JM, Sola I, et al., editors. Development of Protection Against Coronavirus Induced Diseases. New York: Springer; 1995. p. 197–211.
2. Enjuanes L, Smerdou C, Castilla J, Anton IM, Torres JM, Sola I, et al. Development of protection against coronavirus induced diseases. Adv Exp Med Biol. 1995;382:197–211.
3. Wesley R, Woods R, Cheung A. Genetic analysis of porcine respiratory coronavirus, an attenuated variant of transmissible gastroenteritis virus. J Virol. 1991;65(6):3369–73.
4. Enjuanes L, Brian D, Cavanagh D, Holmes K, Lai AMWC, Laude H, et al. In: van Regenmortel MHR, Faquet GC, Bishop CM, Carstens EHB, Lemon MK, McGeoch DJ, Maniloff J, Mayo MA, Pringle CR, Wicker B, editors. Virus Taxonomy: Classification and Nomenclature of Viruses. New York: Academic; 2000. p. 835–49.
5. Vaughn EM, Halbur PG, Paul PS. Sequence comparison of porcine respiratory coronavirus isolates reveals heterogeneity in the S, 3, and 3′ genes. J Virol. 1995;69(5):3176–84.
6. Krempf P, Schulze B, Laude H, Henler G. Point mutations in the S protein connect the sialic acid binding activity with the enteropathogenicity of transmissible gastroenteritis coronavirus. J Virol. 1997;71:3285–7.
7. Ballesteros M, Sanchez C, Enjuanes L. Two amino acid changes at the N-Terminus of transmissible gastroenteritis coronavirus spike protein result in the loss of enteric tropism. Virology. 1997;227(2):378–88.
8. Delmas B, Rasschaert D, Godet M, Gelf J, Laude H. Four major antigenic sites of the coronavirus transmissible gastroenteritis virus are located on the amino-terminal half of spike glycoprotein S. J Gen Virol. 1990;71(6):1313–23.
9. Sanchez C, Izeta A, Sanchez M, Alonso S, Sola I, Balach M, et al. Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. J Virol. 1999;73(9):7607–18.
10. Ortega J, Sola I, Almazan F, Ceriani JE, Riquelme C, Balach M, et al. Transmissible gastroenteritis coronavirus gene 7 is not essential but influences in vivo virus replication and virulence. Virology. 2003;308(1):33–22.
11. Tung FY, Abraham S, Sethna M, Hung S, Sethna P, Hogue B, et al. The 9.4Da hydrophobic protein encoded at the 3′ end of the porcine transmissible gastroenteritis coronavirus genome is membrane-associated. Virology. 1992;186(2):675–83.

Submit your next manuscript to BioMed Central and take full advantage of:

• Convenient online submission
• Thorough peer review
• No space constraints or color figure charges
• Immediate publication on acceptance
• Inclusion in PubMed, CAS, Scopus and Google Scholar
• Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit