



OPEN ACCESS

EDITED BY

Tomomi Takano,
School of Veterinary Medicine, Kitasato
University, Japan

REVIEWED BY

Tereza Cristina Cardoso,
Universidade Estadual de São Paulo, Brazil
Long Liu,
Hubei University of Medicine, China

*CORRESPONDENCE

Li Xing
✉ xingli107@gmail.com

†These authors have contributed equally to this work

SPECIALTY SECTION

This article was submitted to
Veterinary Infectious Diseases,
a section of the journal
Frontiers in Veterinary Science

RECEIVED 17 January 2023

ACCEPTED 29 March 2023

PUBLISHED 17 April 2023

CITATION

Wang P-H, Nawal Bahoussi A, Tariq Shah P,
Guo Y-Y, Wu C and Xing L (2023) Genetic
comparison of transmissible gastroenteritis
coronaviruses. *Front. Vet. Sci.* 10:1146648.
doi: 10.3389/fvets.2023.1146648

COPYRIGHT

© 2023 Wang, Nawal Bahoussi, Tariq Shah,
Guo, Wu and Xing. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Genetic comparison of transmissible gastroenteritis coronaviruses

Pei-Hua Wang^{1†}, Amina Nawal Bahoussi^{1†}, Pir Tariq Shah^{1†},
Yan-Yan Guo¹, Changxin Wu^{1,2,3,4} and Li Xing^{1,2,3,4*}

¹Institutes of Biomedical Sciences, Shanxi University, Taiyuan, China, ²Shanxi Provincial Key Laboratory of Medical Molecular Cell Biology, Shanxi University, Taiyuan, China, ³Shanxi Provincial Key Laboratory for Prevention and Treatment of Major Infectious Diseases, Taiyuan, China, ⁴The Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, Shanxi University, Taiyuan, China

Transmissible gastroenteritis virus (TGEV) is a porcine coronavirus that threatens animal health and remains elusive despite years of research efforts. The systematical analysis of all available full-length genomes of TGEVs (a total of 43) and porcine respiratory coronaviruses PRCVs (a total of 7) showed that TGEVs fell into two independent evolutionary phylogenetic clades, GI and GII. Viruses circulating in China (until 2021) clustered with the traditional or attenuated vaccine strains within the same evolutionary clades (GI). In contrast, viruses latterly isolated in the USA fell into GII clade. The viruses circulating in China have a lower similarity with that isolated latterly in the USA all through the viral genome. In addition, at least four potential genomic recombination events were identified, three of which occurred in GI clade and one in GII clade. TGEVs circulating in China are distinct from the viruses latterly isolated in the USA at either genomic nucleotide or antigenic levels. Genomic recombination serves as a factor driving the expansion of TGEV genomic diversity.

KEYWORDS

coronavirus, transmissible gastroenteritis virus (TGEV), porcine respiratory coronavirus (PRCV), phylogeographic network, phylogenetic tree, China, USA, recombination

Introduction

Porcine transmissible gastroenteritis virus (TGEV) is a member of the *Coronaviridae* family that was first identified in the United States of America (USA) in 1946 (1) and was reported subsequently in countries of Europe, Asia, Africa, and America (2). In China, TGEV infection was first detected in Guangdong province in 1956, and since then, sporadic outbreaks have been reported in multiple swine-producing provinces, including Heilongjiang, Jilin, Henan, Gansu, Shanghai, and Guangxi (3–7).

TGEV is a rapidly spreading pathogen that infects swine of all ages and species and causes enteritis with a mortality rate >90%, inversely correlated with the age of pigs (8–10). TGEV has a single-stranded positive-sense RNA genome of ~28.6 kb in length (11) and belongs to the subgenus *Tegacovirus* of the *Alphacoronavirus* genus (12). The viral genome encompasses at least nine open reading frames (ORFs). The 5' two-thirds of the virus genome contains ORF1a and ORF1b, encoding the polyprotein 1a (pp1a) and polyprotein 1b (pp1b), respectively, proteolytically processed by virus-encoded proteases into 16 nonstructural proteins (Nsps 1–16) during the translation (13). The 3' one-third of the virus genome encodes four viral structural proteins, including the spike glycoprotein (S), the envelope protein (E), the membrane protein (M), the nucleocapsid protein (N), and three accessory proteins (NS3A, NS3B, and NS7) (14).

The S glycoprotein of TGEV is a large type 1 membrane protein composed of 1448 amino acids (GenBank: KX499468.1, TGEV AHHF strain) (15) and containing five domains: N-terminal domain (NTD, aa 1–248) with a short peptide signal (aa 1–16) (16), S1 domain (aa 249–670), S2 domain (aa 832–1372) (15), the transmembrane domain (TM, aa 1388–1410), and C-terminal cytoplasmic tail (aa 1411–1448). S glycoprotein of coronaviruses is the main structural protein of the viral envelope, which mediates the entry of virions into the host cell and induces the generation of neutralizing antibodies (17). S glycoprotein of TGEV binds the sialic acid or the porcine aminopeptidase N (pAPN) to ensure the entry of virion into host cells (18). The cysteine-rich motif (CRM) at the TGEV S glycoprotein carboxy-terminus is palmitoylated, which facilitates the assembly of S glycoprotein during the maturation of viral particles (19). Mutational analysis of NTD of S glycoprotein revealed a mild effect on viral virulence (20–22).

A previous study of S glycoprotein of TGEV strain PUR46-MAD (GenBank: M94101.1; Protein ID: AAA47109.1) has identified four major epitope sites within the N-terminal first 543 amino acids in the order of C, B, D, and A (17). Sites A and B contain conformational epitopes that depend on glycosylation (23). The amino acid residues 538K, 591R, and 543G are essential to forming the antigen subsites Aa, Ab and Ac of site A, respectively. Site B also consists of at least three subsites, two of which are overlapped and include tryptophan. The amino acid residues 97W and 144S contribute to the formation of epitopes in site B. However, sites C and D contain glycosylation-independent linear epitopes located at aa 48–52 and aa 373–398, respectively (17, 24).

The non-enteropathogenic porcine respiratory coronavirus (PRCV) is a naturally occurring mutant of TGEV with the deletion of an N terminal segment of S glycoprotein (25). PRCV preferentially infects non-ciliated epithelial cells of the porcine respiratory tract (26).

Vaccination and immunoprophylaxis represent the most effective means to contain and prevent TGEV, in addition to biosecurity measures (27). Since coronaviruses are an ongoing threat to humans, animals and the global economy, this study aimed to review the evolution and molecular characteristics of the full-length genomes of TGEV and PRCV during past decades to help promote the control and surveillance strategies.

TGEV and PRCV fall into two major genogroups GI and GII

The full-length genomes of all available TGEVs, including 19 strains isolated in China between 1973–2021 and 24 strains isolated in the USA between 1952–2014, in addition to 7 PRCV viruses isolated in the USA and UK between 1986 to 2016 (Supplementary Table 1), were retrieved from the NCBI GenBank and aligned with the ClustalW using the MEGA11 software. The maximum likelihood (ML) phylogenetic trees based on the full-length genome sequences or genomic fragments containing different ORFs were inferred using the best-fitting models in the IQ-TREE multicore version 1.6.12 (28) with 1000 bootstraps. As shown in Figure 1A, the complete genome-based phylogenetic tree identified two main clades GI and GII. All viruses isolated in China fell into the GI clade, together with the eight earliest strains

from the USA, including attenuated Purdue P115 (GenBank ID: DQ811788.1), Virulent Purdue (GenBank ID: DQ811789.2), virulent Miller-M6 (GenBank ID: DQ811785.1), attenuated Miller-M60 (GenBank ID: DQ811786.2), TGEV/USA/HB/1988 (GenBank ID: KX900394.1), TGEV/USA/Z/1986 (GenBank ID: KX900393.1), Purdue (GenBank ID: AJ271965.2), and TGEV/Mex/145/2008 (GenBank ID: KX900402.1). The remaining 16 viruses from the USA fell independently into a separate, distinct clade referred to as the GII clade (Figure 1A). To deeply understand the virus evolution, we further constructed phylogenetic trees based on different genomic fragments. In the phylogenetic trees based on ORF1a (Supplementary Figure 1A), and ORF1b (Supplementary Figure 1B), ORF2 (encoding spike glycoprotein of TGEV and PRCV) (Supplementary Figure 1C), and genomic fragment containing ORFs of E, M, and N proteins (Supplementary Figure 1D), all viruses isolated in China and the aforementioned relatively earliest or attenuated vaccine viruses identified in the USA clustered in GI clade, corroborating results of the full-length genome-based phylogenetic analysis (Figure 1). In addition, the phylogenetic tree based on the ORF2 (Supplementary Figure 1C) or genomic fragment containing ORFs of E, M, and N proteins (Supplementary Figure 1D) revealed that the PRCV strains, known as a natural mutant of TGEV with a segmental deletion in the ORF2, clustered independently according to their geographical regions. The two PRCV strains identified in the UK clustered into GI, whereas the five identified in the USA clustered genetically closer to each other into GII, within different sub-lineage (Figure 1, Supplementary Figures 1A, B), demonstrating a complex evolutionary relationship between TGEV and PRCV strains (29).

Before this report, all constructed phylogenetic trees based on the full-length genomes of 30 TGEV isolates (30), 23 isolates (8), or based on the ORF2 of 11 isolates (3), displayed two major clades. Herein, our study: (1) included all available full-length genomic sequences of TGEV and PRCV, (2) showed that the GII clade contains all viruses from the USA, whereas the GI clade encompasses the nine relatively old TGEVs or attenuated vaccine strains from the USA in addition to all TGEVs from China. These findings indicate that the genetic distance of TGEVs from China is relatively far from that of the strains from the USA in the GII clade.

To further determine the genetic relatedness of TGEV genomes, we conducted a genomic similarity analysis using SimPlot software (31) and involving nine representative TGEV isolates from each clade: GI clade (Supplementary Figures 2A, B) and GII clade (Supplementary Figures 2A, C), separately. The Virulent-Purdue TGEV (GenBank ID: DQ811789.2) was included as a query. The nucleotide sequences of China viruses revealed a great similarity (>97%) with the query strain for most parts of the viral genome (Supplementary Figures 2A, B). In contrast, the similarity between viral isolates of the GII clade and the query strain is significantly reduced to <97% for most parts of the viral genome (Supplementary Figures 2A, C). Notably, the sequence similarity plot revealed that the ORF2 region shows the lowest similarity for viral isolates in both GI and GII clades (Supplementary Figures 2B, C, respectively). The genomic similarity plot results are consistent with the phylogenetic trees in that TGEVs isolated in China and USA belong to two distinct evolutionary branches. The low

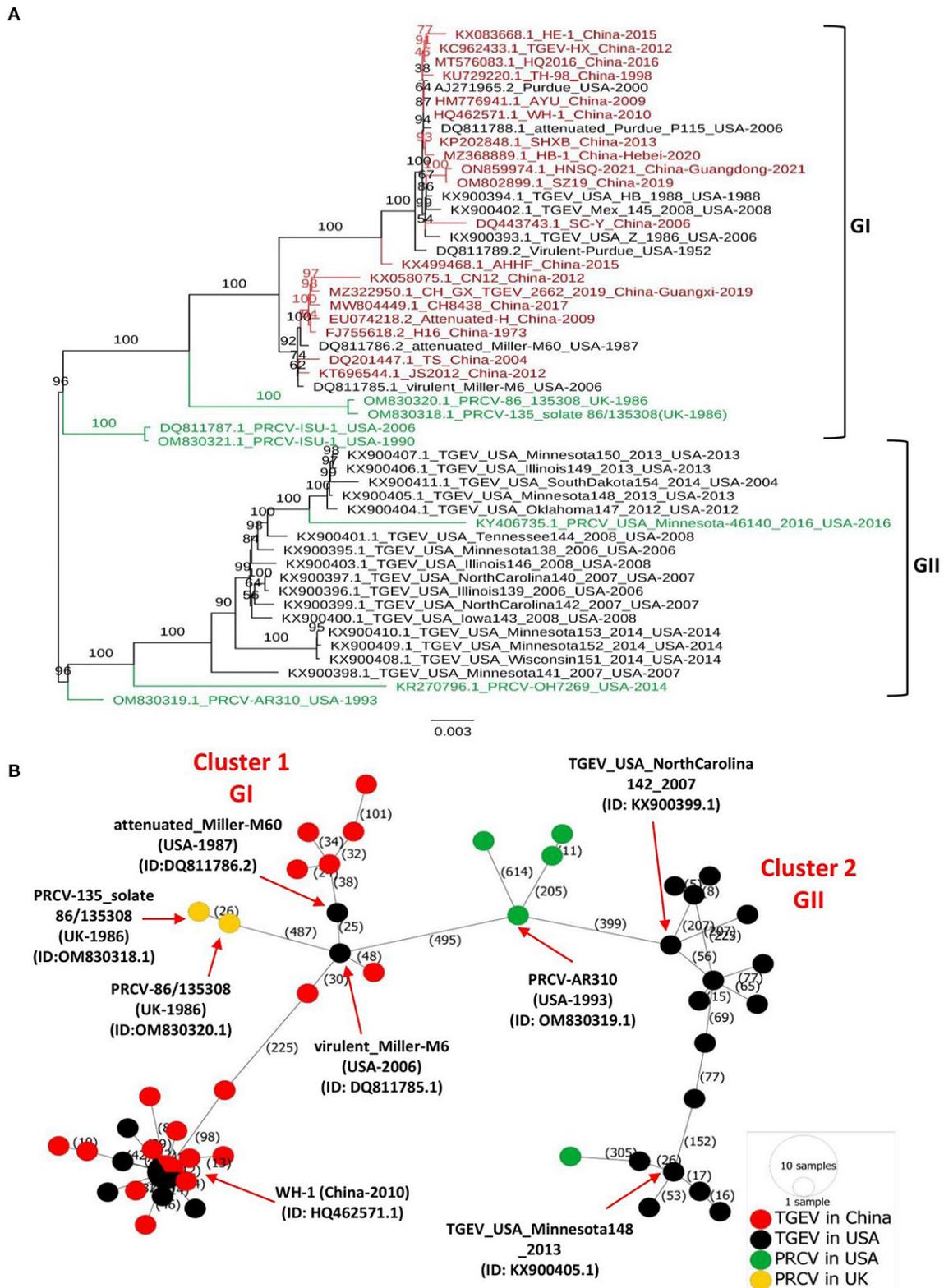


FIGURE 1
 Phylogenetic tree and phylogeographic network of TGEV and PRCV strains. **(A)** Complete genome sequence-based ML tree. All complete genome nucleotide sequences were aligned using the MEGA11 software (37) and edited with the BioEdit v7.2.5. The ML phylogenetic trees were inferred using the IQ-TREE multicore version 1.6.12 (28) with 1000 bootstraps and the best-fitting model TIM + F + I + G4. The numbers on each branch are the bootstrap values (%). The scale bar represents a length corresponding to the 0.003 nucleotide substitutions per site. Virus strains are formatted as GenBank accession number: virus name (country-year of collection). The red indicates TGEV from China, and the black indicates TGEV from the USA. The green color indicates PRCVs. **(B)** Phylogeographic network. The complete genome sequences of TGEV and PRCV were used to infer the Minimum Spanning Network (MSN) implemented by PopArt v1.7 (38). The network included 50 TGEV/PRCV strains, where 24 TGEV strains were from the USA, 19 TGEV strains from China, 5 PRCV strains from the USA and 2 PRCV strains from the UK. The mutation numbers are shown.

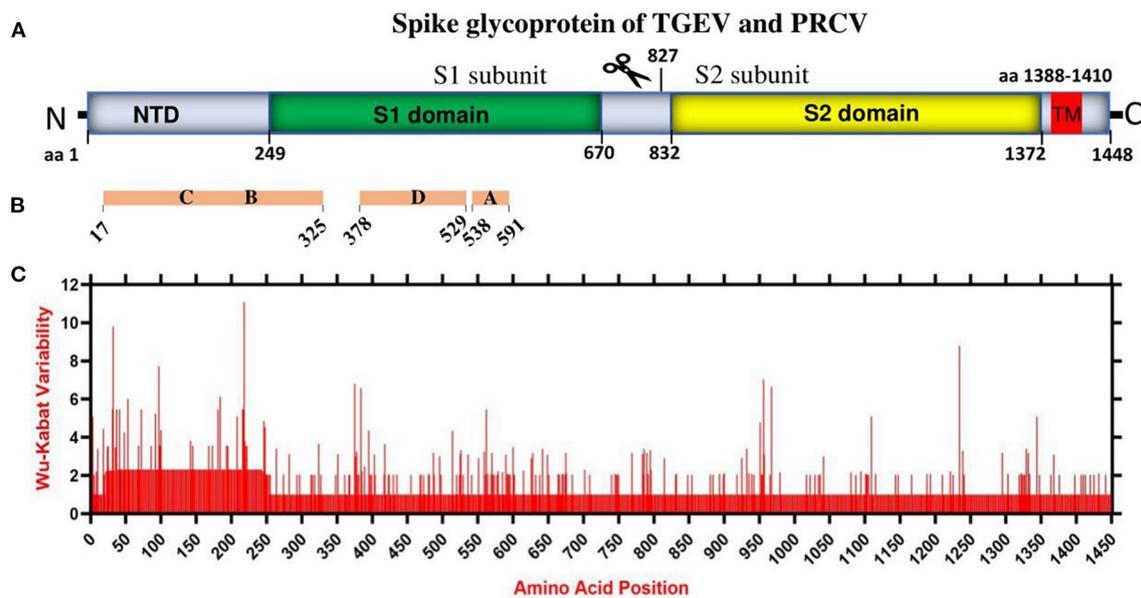


FIGURE 2

The amino acid variation landscape of TGEV and PRCV spike glycoprotein. (A) Domain structure of TGEV S glycoprotein, including NTD (aa 1–248), S1 domain (aa 249–670), S2 domain (aa 832–1,372), TM (aa 1,388–1,410) and C-terminal cytoplasmic tail (aa 1,411–1,448). (B) The brown blocks represent the relative positions of known sites containing neutralizing epitopes in the spike glycoprotein of TGEV (17). (C) Amino acid variation landscape. The amino acid sequences of all available complete spike proteins (a total of 58, [Supplementary Table 1](#)) of TGEV and PRCV were separately retrieved from the NCBI database and were aligned with ClustalW using the MEGA11 software (37). The protein variability was determined using the Wu-Kabat variability coefficient offered by the Protein Variability Server (PVS) (39). The Wu-Kabat variability coefficient describes the susceptibility of an amino acid position to evolutionary replacements and is computed using the following formula: variability coefficient = $N \cdot k/n$, where N is the number of sequences in the alignment, k is the number of different amino acids at a given position and n is the frequency of the most common amino acid at that position. Y-axes represent the Wu-Kabat variability coefficient values, where the estimation limit is set as “1”. Above the limit of “>1” represents variations. X-axes represent the amino acid positions.

similarity of ORF2 region in the genome ([Supplementary Figure 2](#)) indicates that the spike glycoprotein is highly variable.

PRCV-135_solate 86/135308 (GenBank ID: OM830318.1, UK-1986) that were identified to be more distanced within GI clade of the full-length phylogenetic, are also connected to the virulent_Miller-M6 in the Network Cluster 1 ([Figure 1B](#)).

Phylogeographic network of TGEV and PRCV

To map the regional spread and genetic relationships of TGEV and PRCV strains, we performed a phylogenetic network using all available complete genome sequences isolated from 1952 to 2021 ([Supplementary Table 1](#)). The analysis showed that strains within the GI and GII clades cluster into two separate Network Clusters ([Figure 1B](#)). The first Network Cluster (Cluster 1) consists of GI strains, while the second (Cluster 2) comprises GII strains ([Figure 1B](#)). These results support our ML phylogenetic classification of the complete genomes. Further, TGEV and PRCV strains seem to be radiated from strains isolated in the USA. The TGEV WH-1 (GenBank ID: HQ462571.1, China-2010) is identified sharing its genetic ancestor with the USA TGEV strains and connecting most of the GI strains to it through short mutational branches ([Figure 1B](#)). The remaining China TGEV strains are connected to the virulent_Miller-M6 (GenBank ID: DQ811785.1, USA-2006) within the same Network Cluster (Cluster 1). Interestingly, both the PRCV strains isolated in the UK, e.g., PRCV-86_135308 (GenBank ID: OM830320.1, UK-1986) and

Recombination of TGEV

Genomic recombination is a relatively common phenomenon for coronaviruses, which have also been reported to occur in TGEVs. Zhang et al. reported that the TGEV AHHF strain (GenBank: KX499468.1) is a natural recombinant (32). The ORF1a gene of AHHF resulted from the recombination between SC-Y (GenBank: DQ443743.1) and H16 (GenBank: FJ755618.2), while the spike gene of AHHF was produced by recombination between the virulent Purdue (GenBank: DQ811789.2) and the attenuated H strain (GenBank: EU074218.2). TGEV JS2012 (GenBank: KT696544.1) strain is another natural recombinant between Miller M6 (GenBank: DQ811785.1) and Purdue 115 (GenBank: DQ811788.1) (8). Thus, we performed further recombination analysis of the entire TGEV and PRCV genome sequences (from the USA and China) using the seven algorithms of the recombination detection program 4 (RDP, GENECONV, BootScan, MaxChi, Chimera, SiScan, and 3Seq) (33). The analysis identified the occurrence of at least four potential recombination events ([Supplementary Table 2](#)), three of which (Events 1–3) occurred within the GI clade, and one event (Event 4) occurred between

strains in GII clade. The recombination among the USA strains (Event 4) is identified for the first time. In consistence with the previous reports (8, 32), the beginning and ending breakpoints of the identified recombination events are mainly located in ORF1a (Events 1 and 4) and ORF2-containing spike gene (Events 2 and 3) (Supplementary Figure 3). Zhang et al. report analyzed the occurrence of recombination among TGEV AHHF and 14 other strains and identified potential recombination at two locations (ORF1a and ORF2) (32). Consistent with those results, our analysis detected AHHF strain as a recombinant in two events (Events 1 and 2). In Event 2, the AHHF strain was recombined by CN12 and attenuated Purdue P115 in the ORF2, a new possibility of recombination for the AHHF virus. A previous report indicated that TGEV JS2012, a highly pathogenic strain isolated in newborn piglets faces in Jiangsu Province, China, resulted from natural recombination between Miller M6 and Purdue 115 and preserved the genetic integrity and characteristics of virulent strain, despite the recombination breakpoints in the ORF2 (8). Consistent with this report, TGEV JS2012 in our analysis was also identified as a recombinant strain, but between H16 (major parent) and the attenuated Purdue P115 (minor parent) in the ORF2, indicating the existence of alternative recombination. Interestingly, the identified recombination events occurred only between viruses in the same clade (intra-clade). None of the inter-clade recombinations have been found so far.

The landscape of amino acid variation of the spike glycoprotein

Coronaviruses enter the host cells by interaction between the S glycoprotein and the cellular receptor. S glycoprotein represents the chief immunogenic protein that elicits neutralizing antibody production (34) and the main antigen for vaccine research (35, 36). The virulence and antigenicity of TGEV S protein were shown to be sensitive to amino acid (aa) changes (3, 8). The mutation at aa 585 of TGEV HQ2016 engenders a serine-to-alanine change, affecting receptor binding and antigenicity (8). Here, the genome similarity and recombination analysis identified a low similarity of ORF2 (encoding spike glycoprotein) (Supplementary Figure 2) with potential recombination events (Supplementary Figure 3). In addition, the S1 subunit of TGEV spike glycoprotein has been shown to contain the neutralizing epitopes (17) (Figures 2A, B). Thus, we evaluated the amino acid variation patterns of TGEV and PRCV spike proteins by applying the Wu-Kabat variability coefficient provided by the Protein Variability Server (PVS). The method was used to acquire the consensus sequences of spike protein that consists of 1449 aa in TGEV (Figure 2C). The N terminal region, including both NTD and S1 domain, is the most variable, especially the aa positions 1–100, aa 200–250, aa 370–400, and aa 500–600, where the neutralizing epitopes have been identified (17).

Overall, China TGEV isolates are found to keep genetic features similar to that of older or attenuated vaccine strains so far, but distinct from that of the USA strains isolated latterly, which has implications in the monitoring, prevention, and control of TGEV. Furthermore, mutation is the modification of a gene resulted from insertion, deletion, or substitution of a single or multiple base

units. Recombination occurs when the larger genetic fragments from at least two genomes are exchanged. The highly variable regions of spike proteins of TGEV and PRCV (Figure 2) and the recombination events (Supplementary Figure 3) demonstrate that both mutation and recombination are driving the expansion of genetic diversity and transmission of TGEV and PRCV, which need attention for the surveillance and vaccine development.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

P-HW and LX conceived the study. P-HW, PT, and AN performed analysis. Y-YG provided resources in data analysis. P-HW and AN wrote the manuscript. CW and LX carried out administration. AN, PT, and LX revised the manuscript. All authors have read and approved the final manuscript.

Funding

This work was supported by the Programme of Introducing Talents of Discipline to Universities (D21004).

Acknowledgments

We thank all the persons who contributed to the collection and production of TGEV and PRCV genomic sequences in NCBI GenBank.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fvets.2023.1146648/full#supplementary-material>

References

- Doyle LP, Hutchings LM. A transmissible gastroenteritis in pigs. *J Am Vet Med Assoc.* (1946) 108:257–9.
- Pineyro PE, Lozada MI, Alarcon LV, Sanguinetti R, Cappuccio JA, Perez EM, et al. First retrospective studies with etiological confirmation of porcine transmissible gastroenteritis virus infection in Argentina. *BMC Vet Res.* (2018) 14:292. doi: 10.1186/s12917-018-1615-9
- Hu X, Li N, Tian Z, Yin X, Qu L, Qu J. Molecular characterization and phylogenetic analysis of transmissible gastroenteritis virus HX strain isolated from China. *BMC Vet Res.* (2015) 11:72. doi: 10.1186/s12917-015-0387-8
- Li JQ, Cheng J, Lan X, Li XR, Li W, Yin XP, et al. Complete genomic sequence of transmissible gastroenteritis virus TS and 3' end sequence characterization following cell culture. *Virolog. Sin.* (2010) 25:213–24. doi: 10.1007/s12250-010-3108-2
- Wang L, Ge C, Wang D, Li Y, Hu J. The survey of porcine teschoviruses, porcine circovirus and porcine transmissible gastroenteritis virus infecting piglets in clinical specimens in China. *Trop Anim Health Prod.* (2013) 45:1087–91. doi: 10.1007/s11250-012-0329-4
- Weiwai H, Qinghua Y, Liqi Z, Haoifei L, Shanshan Z, Qi G, et al. Complete genomic sequence of the coronavirus transmissible gastroenteritis virus SHXB isolated in China. *Arch Virol.* (2014) 159:2295–302. doi: 10.1007/s00705-014-2080-9
- Yuan D, Yan Z, Li M, Wang Y, Su M, Sun D. Isolation and characterization of a porcine transmissible gastroenteritis coronavirus in Northeast China. *Front Vet Sci.* (2021) 8:611721. doi: 10.3389/fvets.2021.611721
- Guo R, Fan B, Chang X, Zhou J, Zhao Y, Shi D, et al. Characterization and evaluation of the pathogenicity of a natural recombinant transmissible gastroenteritis virus in China. *Virology.* (2020) 545:24–32. doi: 10.1016/j.virol.2020.03.001
- Li X, Li P, Cao L, Bai Y, Chen H, Liu H, et al. Porcine IL-12 plasmid as an adjuvant improves the cellular and humoral immune responses of DNA vaccine targeting transmissible gastroenteritis virus spike gene in a mouse model. *J Vet Med Sci.* (2019) 81:1438–44. doi: 10.1292/jvms.18-0682
- Yuan X, Lin H, Fan H. Efficacy and immunogenicity of recombinant swinepox virus expressing the a epitope of the TGEV S protein. *Vaccine.* (2015) 33:3900–6. doi: 10.1016/j.vaccine.2015.06.057
- Hubert L, Denis R, Bernard D, Murielle G, Jacqueline G, Bernard C. *Molecular Biology of Transmissible Gastroenteritis Virus.* Amsterdam: Elsevier (1990), p. 23. doi: 10.1016/0378-1135(90)90144-K
- Wong ACP, Li X, Lau SKP, Woo PCY. Global epidemiology of bat coronaviruses. *Viruses.* (2019) 11:174. doi: 10.3390/v11020174
- Gorbalenya A, Enjuanes L, Zi Eb Uhr J, Snijder EJ. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* (2006) 117:17–37. doi: 10.1016/j.virusres.2006.01.017
- Penzes Z, González JM, Calvo E, Izeta A, Smerdou C, Méndez A, et al. Complete genome sequence of transmissible gastroenteritis coronavirus PUR46-MAD clone and evolution of the purdue virus cluster. *Virus Genes.* (2001) 23:105–18. doi: 10.1023/A:1011147832586
- Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* (2017) 45:D200–3. doi: 10.1093/nar/gkw1129
- Laude H, Godet M, Bernard S, Gelfi J, Duarte M, Delmas B. Functional domains in the spike protein of transmissible gastroenteritis virus. *Adv Exp Med Biol.* (1995) 380:299–304. doi: 10.1007/978-1-4615-1899-0_48
- Gebauer F, Posthumus WP, Correa I, Suñé C, Smerdou C, Sánchez CM, et al. Residues involved in the antigenic sites of transmissible gastroenteritis coronavirus S glycoprotein. *Virology.* (1991) 183:225–38. doi: 10.1016/0042-6822(91)90135-X
- Ballesteros ML, Sánchez CM, 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) 227:378–88. doi: 10.1006/viro.1996.8344
- Sandra G, Bastian T, Kathrin E, Michael V, Christel S-W. Palmitoylation of the Alphacoronavirus TGEV spike protein S is essential for incorporation into virus-like particles but dispensable for S-M interaction. *Virology.* (2014) 464:397–405. doi: 10.1016/j.virol.2014.07.035
- Sánchez CM, Izeta A, Sánchez-Morgado JM, Alonso S, Sola I, Balasch 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) 73:7607–18. doi: 10.1128/JVI.73.9.7607-7618.1999
- Sánchez CM, Pascual-Iglesias A, Sola I, Zuniga S, Enjuanes L. Minimum determinants of transmissible gastroenteritis virus enteric tropism are located in the N-Terminus of spike protein. *Pathogens.* (2019) 9:2. doi: 10.3390/pathogens9010002
- Wang G, Liang R, Liu Z, Shen Z, Shi J, Shi Y, et al. The N-Terminal domain of spike protein is not the enteric tropism determinant for transmissible gastroenteritis virus in piglets. *Viruses.* (2019) 11:313. doi: 10.3390/v11040313
- Reguera J, Ordoño D, Santiago C, Enjuanes L, Casanovas JM. Antigenic modules in the N-terminal S1 region of the transmissible gastroenteritis virus spike protein. *J Gen Virol.* (2011) 92:1117–26. doi: 10.1099/vir.0.027607-0
- Enjuanes L, Suñé C, Gebauer F, Smerdou C, Camacho A, Antón IM, et al. Antigen selection and presentation to protect against transmissible gastroenteritis coronavirus. *Vet Microbiol.* (1992) 33:249–62. doi: 10.1016/0378-1135(92)90053-V
- Pensaert M, Callebaut P, Vergote J. Isolation of a porcine respiratory, non-enteric coronavirus related to transmissible gastroenteritis. *Vet Q.* (1986) 8:257–61. doi: 10.1080/01652176.1986.9694050
- Peng JY, Punyadarsaniya D, Shin DL, Pavasutthipaisit S, Beineke A, Li G, et al. The cell tropism of porcine respiratory coronavirus for airway epithelial cells is determined by the expression of porcine aminopeptidase N. *Viruses.* (2020) 12:1211. doi: 10.3390/v12111211
- Gerdtts V, Zakhartchouk A. Vaccines for porcine epidemic diarrhea virus and other swine coronaviruses. *Vet Microbiol.* (2017) 206:45–51. doi: 10.1016/j.vetmic.2016.11.029
- Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* (2016) 44:W232–5. doi: 10.1093/nar/gkw256
- Chen F, Knutson TP, Rossow S, Saif LJ, Marthaler DG. Decline of transmissible gastroenteritis virus and its complex evolutionary relationship with porcine respiratory coronavirus in the United States. *Sci Rep.* (2019) 9:3953. doi: 10.1038/s41598-019-40564-z
- Cheng S, Wu H, Chen Z. Evolution of transmissible gastroenteritis virus (TGEV): a codon usage perspective. *Int J Mol Sci.* (2020) 21:7898. doi: 10.3390/ijms21217898
- Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol.* (1999) 73:152–60. doi: 10.1128/JVI.73.1.152-160.1999
- Zhang X, Zhu Y, Zhu X, Shi H, Chen J, Shi D, et al. Identification of a natural recombinant transmissible gastroenteritis virus between Purdue and Miller clusters in China. *Emerg Microbes Infect.* (2017) 6:e74. doi: 10.1038/emi.2017.62
- Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* (2015) 1:vev003. doi: 10.1093/ve/vev003
- Turlewicz-Podbielska H, Pomorska-Mól M. Porcine coronaviruses: overview of the state of the art. *Virolog. Sin.* (2021) 36:833–51. doi: 10.1007/s12250-021-00364-0
- Meng F, Zhao Z, Li G, Suo S, Shi N, Yin J, et al. Bacterial expression of antigenic sites A and D in the spike protein of transmissible gastroenteritis virus and evaluation of their inhibitory effects on viral infection. *Virus Genes.* (2011) 43:335–41. doi: 10.1007/s11262-011-0637-1
- Mou C, Zhu L, Xing X, Lin J, Yang Q. Immune responses induced by recombinant *Bacillus subtilis* expressing the spike protein of transmissible gastroenteritis virus in pigs. *Antiviral Res.* (2016) 131:74–84. doi: 10.1016/j.antiviral.2016.02.003
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* (2011) 7:539. doi: 10.1038/msb.2011.75
- Leigh JW, Bryant D. POPART: full-feature software for haplotype network construction. *Methods Ecol Evol.* (2015) 6:1110–6. doi: 10.1111/2041-210X.12410
- García-Boronat M, Díez-Rivero CM, Reinherz EL, Reche PA. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Res.* (2008) 36:W35–41. doi: 10.1093/nar/gkn211