Evolution and homologous recombination of the hemagglutinin–esterase gene sequences from porcine torovirus

Yingying Cong · Dante S. Zarlenga · Juergen A. Richt · Xin Wang · Yang Wang · Siqingaowa Suo · Jingfei Wang · Yudong Ren · Xiaofeng Ren

Abstract The objective of the present study was to gain new insights into the evolution, homologous recombination, and selection pressures imposed on the porcine torovirus (PToV), by examining the changes in the hemagglutinin–esterase (HE) gene. The most recent common ancestor of PToV was estimated to have emerged 62 years ago based upon HE gene sequence data obtained from PToV isolates originating from Spain, South Korea, Netherlands, Hungary, and Italy and using the HE gene of Bovine torovirus isolates Niigata1 (AB661456) and Niigata3 (AB661458) as outgroups. The HE gene sequence data segregated all the PToV isolates into two well-supported monophyletic groups; however, various isolates from Spain, Italy, and South Korea did not segregate geographically suggesting very recent translocation of the viruses to these localities. Evidence of recombination was observed between two South Korean isolates that partitioned into two distinct subclades. Data further suggest that most of the nucleotides in the HE gene are under negative selection; however, changes within codon 237 showed an evidence of positive selection.

Y. Cong · X. Wang · S. Suo · Y. Ren · X. Ren (✉)
College of Veterinary Medicine, Northeast Agricultural University, 59 Mucai Street, Xiangfang District, Harbin 150030, China
e-mail: renxf@neau.edu.cn

D. S. Zarlenga
Animal Parasitic Diseases Laboratory, Agricultural Research Service, Beltsville, MD, USA

J. A. Richt
Department of Diagnostic Medicine/Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, KS, USA

Y. Wang
College of Life Science, Northeast Agricultural University, 59 Mucai Street, Xiangfang District, Harbin 150030, China

J. Wang
Centre for Animal Infectious Disease Diagnosis and Technical Services and State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin 150030, China

Introduction

Toroviruses are enveloped, positive-stranded RNA viruses that, together with coronaviruses, belong to the *Coronaviridae* family of the Nidovirales order and cause enteric diseases in animals and humans [1]. Toroviruses identified, thus far have been grouped on the basis of host tropism by the International Committee on Taxonomy of Viruses (ICTV) [2], i.e., bovine (BToV), porcine (PToV), human (HToV), and equine (EToV) toroviruses. The latter virus is considered the prototypical torovirus that was isolated from a horse with diarrhea in Berne, Switzerland, in 1972 [3]. A morphologically related virus was later found on a cattle farm in Breda, Iowa, in the 1980s [4]. Over the years, there have been several reports describing the presence of torovirus-like particles in the feces of humans and swine [5, 6]; however, the first formal identification and characterization of a torovirus from swine did not materialize until 1998 [7].

Toroviruses resemble coronaviruses in their genome organization, gene order, and replication strategy, but differ

in genome size, host range, and virion architecture. In addition, the morphology of toroviruses is unique in that they appear as spherical, elongated, or kidney-shaped particles with tubular nucleocapsids surrounded by a membrane carrying large spikes [8, 9]. Research on toroviruses has been limited because they are very difficult to propagate in cell culture. The one exception has been the equine isolate from Berne, Switzerland (BEV; 3). Also, in 2004, a new BToV isolate obtained from the intestinal contents of a calf with diarrhea was successfully propagated in cell culture using the human HRT-18 cell line [10].

Toroviruses contain a single, positive-strand RNA molecule approximately 30 kb in length. Two large and overlapping open reading frames, ORF1a and ORF1b, that encode the RNA polymerase which is important for replication and transcription of the genome, can be found in the 5′ two-thirds of the genome. The four additional ORFs 2–5, which encode the spike (S), membrane (M), hemagglutinin–esterase (HE), and nucleocapsid (N) structural proteins, are present in the remaining one-third of the genome [11–16]. Sequence data for PToV strains are limited to the S, M, HE, and N genes where complete HE gene sequences so far have only been documented from Spanish, Italian, and South Korean isolates which were verified by electron microscopy or by diagnostic RT-PCR.

Although, recent reports demonstrate that recombination has occurred in coronaviruses [17, 18], only a few studies have systematically investigated recombination events among toroviruses [19, 20]. Therefore, the objective of the present study was to gain new insights into the evolution, homologous recombination, and selection pressures imposed on the PToV by focusing on sequence data available for the HE gene.

Materials and methods

Sequence data

Complete PToV and BToV HE gene sequences were retrieved from GenBank[TM]. All the sequences were aligned using Clustal W [21] and BioEdit [22] then visually confirmed. The information regarding the strain name, isolation time, and place of origin is summarized in Table 1.

Rates of nucleotide substitution and virus ancestry

The nucleotide substitution rate was determined using the Model Test V3.7 [23], assuming the HKY85 model for nucleotide substitution, and incorporating a gamma distribution for "among-site" rate variation with six rate categories. The evolutionary rates and ages of the most recent common ancestors (TMRCA) for PToV were estimated using the Bayesian Markov Chain Monte Carlo (MCMC) method within the BEAST (V 1.6.2) software package [24]; rates were estimated using both the relaxed clock (uncorrelated lognormal) and the SRD06 nucleotide substitution models [25]. These analyses were run using the Hasegawa–Kishino–Yano DNA substitution model where the sequences were first partitioned into codons followed by 100 million generations using MCMC and subsampling each 10,000 generations. The MCMC chains were run to achieve the convergence using the Tracer program with uncertainty in parameter estimates reflected in the values of the 95 % highest probability density (HPD). The best tree was computed using TreeAnnotator (BEAST V1.6.2) with the first 10 % removed as burn-in. Trees were visualized and edited with FigTree (BEAST V 1.6.2).

Recombination analysis

Based on multiple alignments, complete HE gene sequences were used for recombination analysis with the Recombination Detection Program 3 (RDP3) [26]. To identify potential recombination events and putative breakpoints, the programs RDP3, GENECONV, BOOT-SCAN, MaxChi, CHIMAERA, SISCAN, PhylPro, LARD, and 3seq in RDP (Beta 4.13) were used. In addition, similarity plots and scanning for potential recombination events were performed using SimPlot V3.5.1 [27]. Genetic algorithm recombination detection (GARD) methods [28] were used to refine the recombination analysis and estimate breakpoint locations. Putative breakpoints were verified by both the Shimodaira–Hasegawa test and manually by checking phylogenetic trees for nonrecombinant and recombinant segments through MEGA V5.0 [29]. Statistical significance was set at $p < 0.05$. All analyses were conducted in triplicate.

Selection pressure analysis

Evidence for non-neutral selection was examined by calculating the ratio of nonsynonymous (dN) to synonymous (dS) substitutions using Maximum likelihood (ML) phylogenetic reconstruction and the general reversible nucleotide substitution model available through the Datamonkey web server [30]. To detect non-neutral selection, single-likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL), internal fixed-effects likelihood (IFEL), and random-effects likelihood (REL) within the HYPHY software package [31] were used as implemented in Datamonkey. Significance levels were set at $p = 0.1$, $p = 0.1$, $p = 0.1$, and Bayes factor = 50, and used to estimate the rates of dN and dS within each codon. The values dN/dS > 1, dN/dS = 1, and dN/dS < 1 were used to define the

Table 1 PToV HE gene sequences used in this study

Strain	GeneBank accession no.	Location	Years
BRES	FJ232070.1	Italy	2002
Markelo	AJ575363.1	The Netherlands	1995
P4	AJ575364.1	Italy	1990
P9	AJ575365.1	Italy	1996
P10	AJ575366.1	Italy	1999
P78	AJ575367.1	Hungary	2000
12_11	GU299773.1	Spain	2005
13_11	GU299774.1	Spain	2005
14_7	GU299775.1	Spain	2005
52_7	GU299776.1	Spain	2005
52_11	GU299777.1	Spain	2005
07-55-4	GU187335.1	South Korea	2007
07-55-5	GU187334.1	South Korea	2007
07-56-11	GU187333.1	South Korea	2007
07-56-14	GU187332.1	South Korea	2007
07-56-22	GU187331.1	South Korea	2007
07-56-23	GU187330.1	South Korea	2007
07-109-11	GU181255.1	South Korea	2007
07-109-12	GU181254.1	South Korea	2007
07-109-13	GU181253.1	South Korea	2007
07-109-14	GU181252.1	South Korea	2007
07-109-15	GU181251.1	South Korea	2007

positive selection (adaptive molecular evolution), neutral mutations, and negative selection (purifying selection), respectively.

RNA secondary structure

The RNA secondary structure of the PToV 07-56-14 HE gene was predicted using RNA structure V4.6 [32]. A maximum of 750 suboptimal structures with up to 20 % difference in free energy from the lowest free energy were generated. The window size is set to 0 to allow the predictions of structures with subtle variations.

Results

Phylogenetic analysis and rate of change in the HE gene

The Maximum Clade Credibility (MCC) tree was constructed based on an alignment (Supplementary Fig. 1) of currently available HE gene sequences of PToV, and the BToV isolates Niigata3 (GenBank accession number AB661458) and Niigata1 (GenBank accession number AB661456). Due to the possibility that the recombined region represents a "hotspot" within the HE gene, this section was removed prior to further analysis (Fig. 1a).

This also permitted direct comparison to the ML trees used to analyze the recombined sequences. The complete MCC tree incorporating both recombined and nonrecombined regions is provided in Supplementary Fig. 2. As shown in Fig. 1a, all the PToV HE sequences partitioned into two distinct clades where the most recent common ancestor in this study is estimated to have emerged 62 years ago. Both clades contained isolates from Spain, Italy, and South Korea. Upon eliminating 07-56-14 from the analysis (Fig. 1b), no demonstrable changes were observed in the architecture or evolutionary rates. Of note, when the entire sequence was analyzed (Supplementary Fig. 2), the overall architecture of the resulting tree as well as the estimated divergence times were congruent with those generated when the recombined regions were removed (Fig. 1a), with the exception of the placement of the mutated 07-56-14 which aligned with virions in clade I.

Based on the Bayesian MCMC approach and assuming a relaxed molecular clock (uncorrelated lognormal), the mean nucleotide substitution rate for the entire sequence set was estimated to be 2.18×10^{-2} substitutions per site per 4.97 years (95 % HPD, 5.66×10^{-3}–7.12×10^{-2}).

Recombination in HE gene sequences

In order to evaluate the presence of recombination within PToV HE gene sequences, we used a suite of

Fig. 1 Maximum clade credibility tree for 22 PToV HE gene sequences **a** The MCC tree was constructed using all the currently available HE gene sequences of PToV (Table 1) and the BToV isolates Niigata1 and Niigata3 as outgroups. The tree was generated wherein the putative recombination region i.e., position 226-470 of 07-56-14, was removed from all the sequences. Estimated TMRCAs of these lineages are shown with their 95 % HPD rates in *parentheses*. All the PToV sequences partitioned into two distinct clades [clade I (*Red*)-upper; clade II (*Blue*)-lower]. **b** The MCC tree was constructed based on all currently available HE gene sequences of PToV (Table 1) and the BToV isolates Niigata1 and Niigata3. The tree was generated wherein the putative recombined virus (07-56-14) was removed from the analysis. Estimated TMRCAs of these lineages are shown with their 95 % HPD rates in parentheses. All PToV sequences partitioned into two distinct clades [group I (*Red*)-upper, group II (*Blue*)-lower] (Color figure online)

recombination detection methods including RDP, GEN-ECONV, BootScan, MaxChi, Chimaera, and 3Seq. After comparing 22 different HE gene sequences, one putative recombination event was detected with a high degree of confidence (*p*-values ranging from 3.47×10^{-2} to 4.76×10^{-4}) as judged by all the six detection methods (Supplementary Table 1).

Using MCC, it was determined that the HE gene of the South Korean isolate 07-56-14 (clade 2, Fig. 1a)

partitioned within the same major clade as 07-109-13 and 07-56-23, the putative parent virions. There was a region of the sequence i.e., position 226-470b, that displayed disproportionately lower levels of similarity between 07-56-14 and 07-56-23 relative to the other regions. Using Sim-Plot (Fig. 2a), the South Korean isolate 07-56-14 exhibited the greatest similarity to 07-56-23 at positions spanning <220b and >470b. Likewise, in the area spanning ~220–470b, peak similarity was observed between 07-56-14 and 07-109-13. In the bootscan plot (Fig. 2b), the positions corresponding to 226b and 470b were defined as crossover points in 07-56-14. GARD analyses (Fig. 2c) further corroborated these putative crossover points by suggesting six possible breakpoints between bases 226-470 in 07-56-14, with an average support exceeding 0.70.

Additional confirmation of a recombination event among the relevant PToV strains was obtained from ML trees using MEGA 5 as presented in Fig. 3a, b. The analysis in Fig. 3a was based on the putative nonrecombinant regions (positions 1-225b + 471-1290b); the analysis in Fig. 3b was based on the proposed recombinant region (position: 226-470b). In Fig. 3a, 07-56-14 clustered with 07-56-23 and Markelo within a larger group equivalent to clade 1 in Fig. 1a. This contrasts with the MCC tree wherein 07-56-14 clustered with clade 2 virions. In Fig. 3b, 07-56-14 clustered within a large clade that included 07-109-13 and the other South Korean isolates (07-109-11/12/14/15). These data are consistent with a recombination event between 07-56-23 and 07-109-13-like viruses which in turn gave rise to the HE gene of 07-56-14.

Selection pressure on the HE gene

The ratio of dN to dS substitutions per nucleotide site was estimated from an alignment of 22 complete HE gene sequences; the putative recombinant PToV 07-56-14 was excluded from the analysis. The predominant selective pressure acting on the PToV was negative, reflected in a dN/dS value of <1 (Table 2). Using FEL (*p* = 0.1), 99 % of the codons were under negative selection pressure. Only one site, represented by codon 237 was identified as under positive selection (FEL *p* value < 1.0; SLAC *p*-value < 0.3).

RNA secondary structure of PToV isolate 07-56-14 HE gene

In order to evaluate if RNA secondary structure played a role in the recombination event, we modeled the structure of the entire 07-56-14 HE RNA (Supplementary Fig. 3). The putative region of recombination, which is believed to have occurred at positions 226 and 470, are depicted in Fig. 4a, b, respectively. One breakpoint (470b) resides in a

Fig. 2 Recombination analysis of isolates 07-56-14, 07-56-23, and 07-109-13 **a** Simplot-based nucleotide similarity comparison of the complete sequences of 07-56-14 with those of representative PToVs from both clades using a window size of 200 nt and a step size of 20 nt. The *x-axis* indicates nucleotide positions along the alignment and the *y-axis* denotes the level of similarity where 1.0 implies 100 % identity. **b** Bootscan plots showing the likelihood that sequence 07-56-14 is derived from a recombination event between sequences 07-56-23 and 07-109-13 based upon the pairwise distance and using PToVs from both clades. Data was modeled with a window size 200, step size 20, and 1,000 Bootstrap replicates. The *y-axis* indicates the percentage of bootstrap values that support the clustering of 07-56-14 with each parental strain. Boot strap values over 70 % are considered significant. **c** Putative breakpoint locations in isolate 07-56-14. Breakpoints defining possible locations for recombination between 07-56-23 and 07-109-13 were derived from GARD analysis. Peak heights are proportional to probability
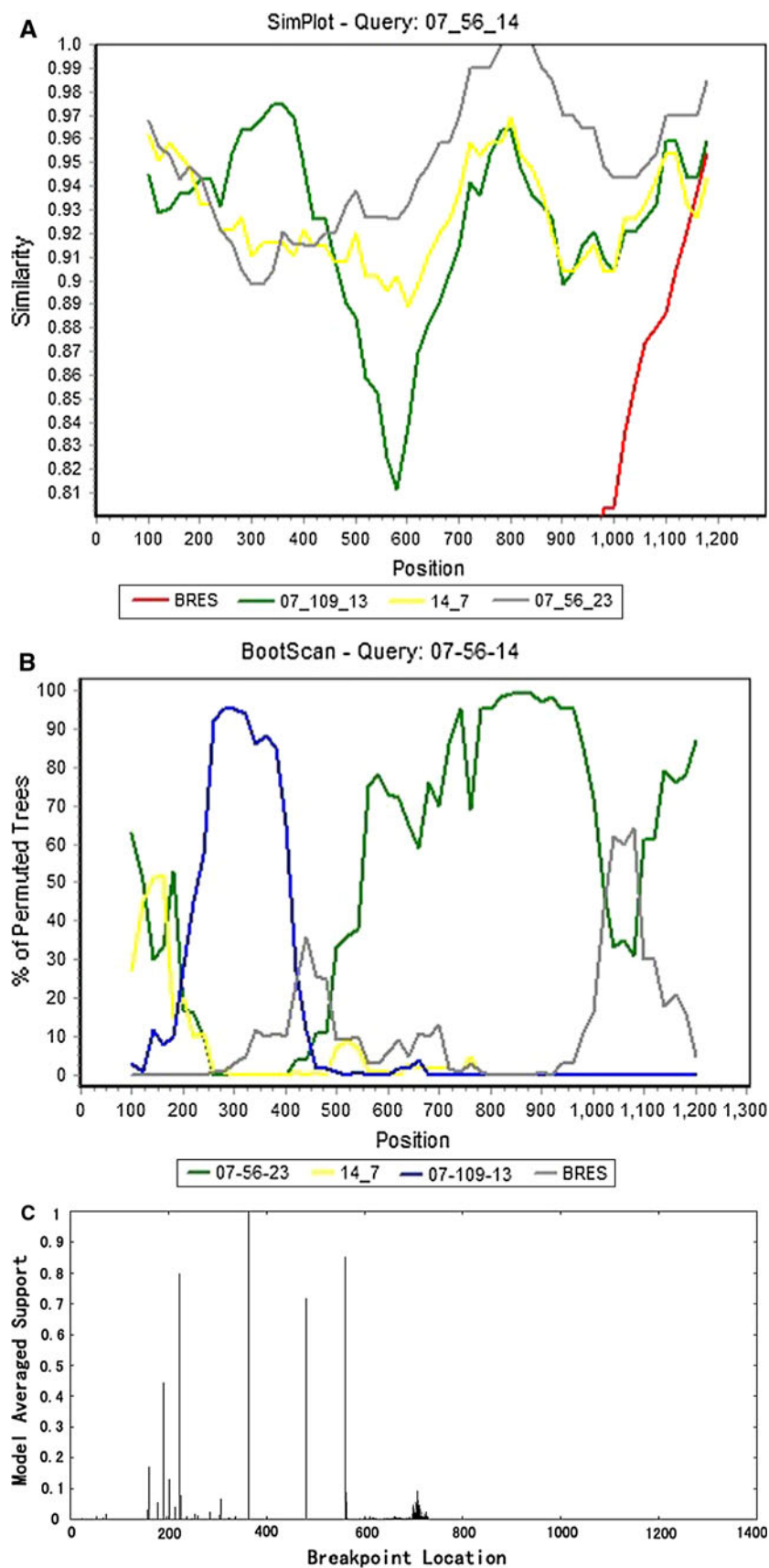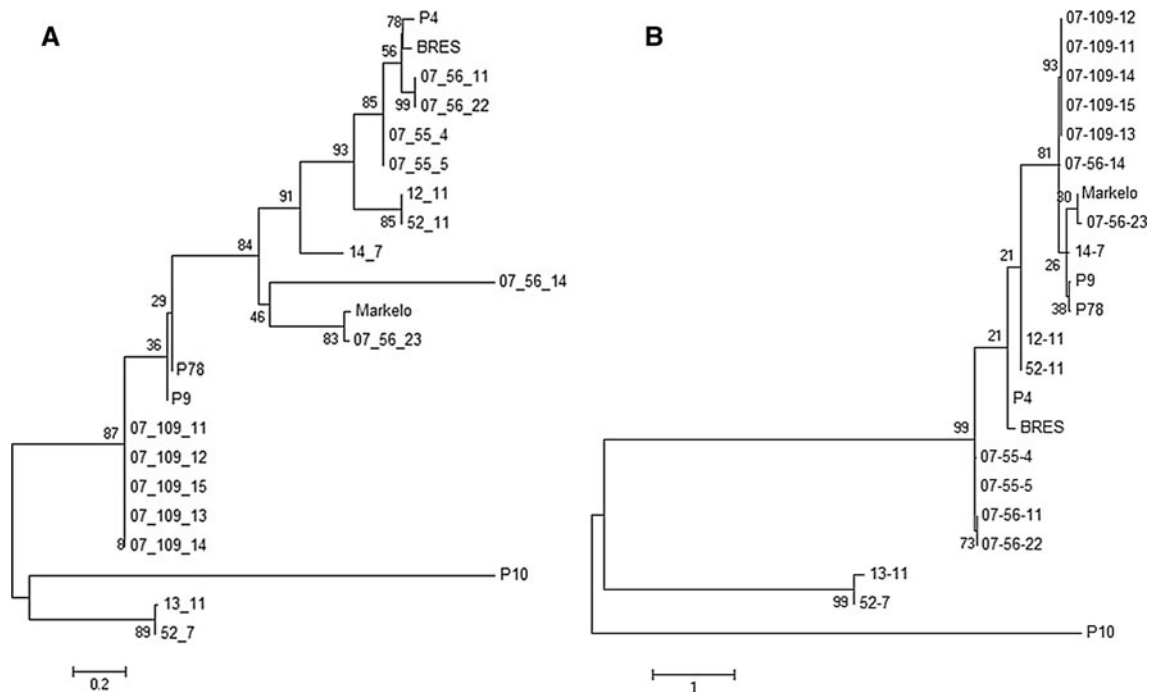
Fig. 3 Maximum likelihood trees for recombinant (**a**) and nonrecombinant (**b**) regions of isolate 07-56-14. **a** Evaluation of the nonrecombined sequences (positions 1-225 and 471-1290) giving rise to 07-56-14. A consensus ML tree (1,000 replicates, Hasegawa–Kishino–Yano model) was constructed using the nonrecombinant regions (positions 1-225 and 471-1290) homologous to 07-56-23. **b** Evaluation of the recombined sequences (position 226-470) giving rise to 07-56-14. A consensus ML tree (1,000 replicates, Hasegawa–Kishino–Yano model) was constructed using the putative recombinant region (position 226–470) homologous to 07-109-13

Table 2 Evidence for positive and negative selection using four detection methods

	SLAC	FEL	REL	IFEL
Positive selection	0	1	0	0
Negative selection	16	139	63	124
Overall dN/dS	0.193	NS	0.594	NS

single-stranded region and both occur within the part of the molecule that is AU rich.

Discussion

Toroviruses have been described as pathogens causing enteric diseases and diarrhea in animals and humans [33, 34]. Epidemiological information on PToV derived from studies in different countries [35–37], indicates that these toroviruses are circulating in pig populations worldwide with a high prevalence. Detailed analyses of the toroviruses described in the present study revealed that the most recent common ancestor of the HE genes of PToV and BToV emerged approximately 62 years ago. In addition, our results also suggest that the most recent common ancestor of clades I and II PToV date back 56 years.

The first recorded isolation and identification of a torovirus was EToV, which occurred in 1972 [3]. However,

based on the analysis of the HE-derived MCC tree, the ancestor of PToV emerged earlier than the first identification of EToV. Closely resembling the HE fusion protein (rHE) of influenza C virus (ICV), the coronavirus and torovirus HEs display acetylesterase activities specific for N-acetyl-9-O-acetylneuraminic acid [38, 39]. The ICV rHE serves as a receptor-binding and receptor-destroying protein, and it has been argued that the HEs of toroviruses and coronaviruses have a similar function [40]. Although, toroviruses and coronaviruses are related, they apparently acquired their HE genes through independent heterologous RNA recombination events [41]. Meanwhile, one of the most remarkable examples of homologous recombination can be observed in the torovirus HE gene, where Smits et al. [42] showed two distinct recombinant events.

Both recombination and selection can negatively impact the accuracy of phylogenetic inference and molecular dating. In the event that bases 226-470 may represent a "hotspot" for recombination, we removed them from the
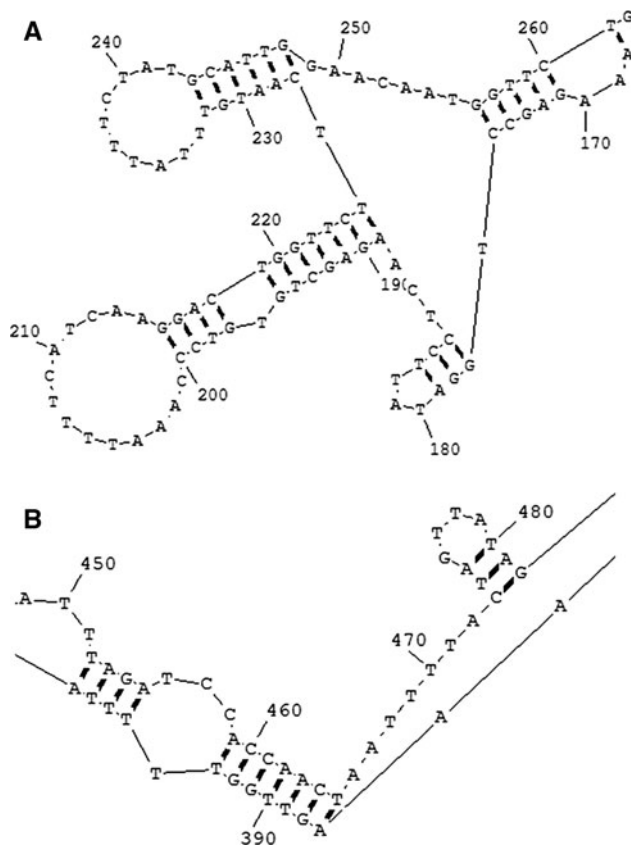
Fig. 4 RNA secondary structure of PToV isolate 07-56-14 HE gene. **a** The predicted RNA secondary structure of 07-56-14 HE gene proximal to base 226. **b** The predicted RNA secondary structure of 07-56-14 HE gene proximal to base 470

analysis prior to evaluating the phylogenetic inference and molecular dating (Fig. 1a). In addition, we performed phylogenetic analysis and molecular dating with the putative recombinant isolate 07-56-14 removed from the analysis (Fig. 1b). Results showed only minor changes in molecular dating (±2 %) at all the nodes. With respect to the phylogenetic inference, the profiles comparing Fig. 1a (recombinant region removed) and Fig. 1b (recombinant 07-56-14 removed) were identical except for the clustering of 14-7 with 52-11 and 12-11 in clade 1, and minor variations in the clustering of the 07-109 series of isolates. Changes in the clustering of 14-7, 52-11, and 12-11 along with the recombinant 07-56-14 also occurred when the entire sequences were analyzed. As such we concluded that the recombination did not demonstrably affect the phylogenetic architecture or molecular dating of our isolates, but the recombination event caused ambiguity in the placement of 07-56-14. This may result from examining HE genes from extant viruses and not having available other informative isolates that span earlier periods in the evolution of clade 2 virions.

Herein, we increased the sensitivity of recombination analysis by utilizing all the GenBank™ available PToV HE sequences in our analyses. The recombination regions in the HE gene visualized by Simplot were consistent with the corresponding breakpoint analyses conducted by GARD. The RDP program further supported a recombination event in the 17-56-14 isolate approximately 244b in length and close to the N terminus of the HE gene. Upon closer scrutiny of the chimeric sequences within the 17-56-14 PToV HE gene, we concluded that these sequences were more consistent with the genomic reorganization of torovirus field variants than with the artifacts derived from PCR amplification [20, 42]. The PToV HE protein appears to provide a selective advantage for the virus at least during a natural infection.

To enhance the predictive character of our analysis, we used multiple programs to evaluate the presence of recombination within the PToV HE gene, including Simplot and RDP, within which reside RDP3, GENECONV, BOOTSCAN, MaxChi, CHIMAERA, SISCAN, PhylPro, LARD, and 3seq. RDP is best suited for scanning sequences in pairs to detect the recombination among a group of aligned DNA sequences [43]. Phylogenetic signaling was not used because of problems identifying the recombinants [26, 44]. Also, most of the programs used scan for putative recombination breakpoints predicated upon the sequence information available which is not possible with phylogenetic signaling [27, 45]. As with most programs of this nature, the quality of the sequence alignment is critical for accurately predicting breakpoints. In general, these programs traverse the query sequence and compare it to a panel of reference sequences. Simplot has the added ability to ignore the gaps in the alignment when generating the plot and to identify the exact similarity value at any point in the sequence. Simplot is not routinely applied to large-scale recombination studies; however, it can accurately identify the position of reorganization, and therefore, breakpoints.

In like manner, we used two methods to construct evolutionary trees, ML, and MCC that apply different parameters in arriving at a final dendrogram. ML is a character-based algorithm that generates trees with specified branch lengths that have the highest conditional probability of recreating the initial dataset. MCC is predicated upon Bayesian phylogenetic inference wherein it tallies the number of times each clade appears in sampled posterior trees to arrive at a tree with MCC [24]. The MCC tree is not a consensus tree but one of the trees actually present in the sample produced by BEAST with the highest product of all the posterior probabilities for all the clades in the tree. MCC identifies a single "point estimate" tree that is central to the distribution of all the trees. The MCMC method infers ancestral states while accommodating uncertainty about the phylogenetic tree and model

parameters [46], allows for easy interpretation of results, and incorporates prior information while providing computational advantages [46]. Topological differences can and do occur between the methodologies predominantly when the sequences are less similar [47] as exemplified when comparing Figs. 1a and 3a, both of which were generated with the recombined sequences removed. As noted, several nodes in Fig. 3a (ML tree) are neither well-supported nor are the relative branch lengths when compared to Fig. 1a (MCC tree). This was not unexpected given the recombinant nature of this sequence. Lack of congruence between trees depicting virus phylogenies can be particularly problematic given the rapid rate of mutation among virions.

RNA viruses are genetically very flexible. During the replication of their genomes, nucleotide substitutions occur at high frequencies presumably to allow rapid adaptations to various selective pressures. However, the balance of the data strongly supports the conclusion that most nucleotides in the HE gene of PToVs are under negative selection with positive selection being observed only at codon 237.

In this study, putative recombinant PToV strain 07-56-14 and its major parent 07-56-23 were identified in the South Korean farm A; its minor parent 07-109-13, was from the South Korean farm C. With the data available, it is difficult to explain why strain 07-56-14 and its major parent 07-56-23 formed a monophyletic clade with isolates derived from Spain, Italy, the Netherlands, and Hungary, including others from South Korea; however, the absence of geographical influence was observed in all the generated trees. Given the paucity of information related to the transmission dynamics of toroviruses, it is possible that anthropogenic factors and/or swine husbandry practices in this region of the world caused the recent dissemination of isolates via the translocation of the host species and/or by-products. Although, swine domestication is believed to have originated 5000–7000 BC in the Middle East, Eastern Mediterranean, or in Southeast Asia, it was not until the last 400–500 years that the host movement due to human intervention took foothold. Even more recently, the raising of pigs in confinement has provided both positive and negative impact on the transmission of swine diseases. Whereas heavy confinement has rendered academic, swine sylvatic pathogens such as parasites, viral pathogens can quickly spread among animals reared under strict confinement and among localities as the animals are transported between production and/or finishing facilities. As such, viral population structure and transmission patterns can be heavily impacted by the patterns of host translocation. The identification of well-defined clades containing geographically distinct isolates coincides temporally with increased transportation and confinement rearing of pigs and with the increase in exportation and importation of meat products worldwide.

Upon examining the predicted RNA secondary structure of the 07-56-14 HE RNA, we found that the position 470 resides within a lengthy single-stranded region and equally important, both putative crossover sites occur within an AU rich portion of the sequence. Prior studies have noted that the recombination in viruses has a predilection for single-stranded regions of the RNA [48], which tend to be less stable than those rich in GC bases. Further, the entire sequence is approximately 64 % in AU residues and within the recombined region, the %AU increases to nearly 69 % which would tend to further destabilize the structure and enhance replicase switching during RNA strand replication [49, 50]. It should be noted that PToV is a positive-strand virus and evidence has been advanced demonstrating that homologous recombination may be more frequent in positive-stranded viruses than in negative-stranded viruses [51]. Roossinck [52] suggested that some plant and animal viruses might have common origins and that plant viruses frequently use recombination and re-assortment as driving forces in evolution. In the present study, we inferred that the recombination and re-assortment do indeed play roles in PToV HE gene evolution.

In summary, our analysis revealed substantial sequence divergence among the PToV HE genes from geographical isolates and showed that this gene is under negative selection. In addition, we provided evidence for RNA recombination in one PToV isolate from South Korea presumably with other isolates from South Korea and conclude that this was likely a recent event. Similar recombination events have been observed in a close relative, the coronavirus [53]. Furthermore, the rapid and continual changes within the torovirus genome require that the persistent surveillance and updating of the phylogeny be performed to better understand the evolution of this pathogen and ascertain what role anthropogenic effects play in dissemination of the pathogen.

References

1. M. Koopmans, M.C. Horzinek, Virus Res. **43**, 233 (1994)
2. Index of Viruses—Coronaviridae, in ICTVdB—The Universal Virus Database, version 4, ed. by C. Büchen-Osmond (Columbia University, New York, 2006)

3. M. Weiss, F. Steck, M.C. Horzinek, J. Gen. Virol. **64**, 1849 (1983)
4. G.N. Woode, D.E. Reed, P.L. Runnels, M.A. Herrig, H.T. Hill, Vet. Microbiol. **7**, 221 (1982)
5. P.J. Durham, L.E. Hassard, G.R. Norman, R.L. Yemen, Can. Vet. J. **30**, 876 (1989)
6. M.L. Penrith, G.H. Gerdes, J.S. Afr, Vet. Assoc. **63**, 102 (1992)
7. A. Kroneman, L.A. Cornelissen, M.C. Horzinek, R.J. de Groot, H.F. Egberink, J. Virol. **72**, 3507 (1998)
8. A.F. de Vries, M.C. Horzinek, P.J.M. Rottier, R.J. de Groot, Semin. Virol. **8**, 33 (1997)
9. E.J. Snijder, M.C. Horzinek, J. Gen. Virol. **74**, 2305 (1993)
10. M. Kuwabara, K. Wada, Y. Maeda, A. Miyazaki, H. Tsunemitsu, Clin. Vaccine Immunol. **14**, 998 (2007)
11. L.A. Cornelissen, C.M. Wierda, F.J. van der Meer, A.A. Herrewegh, M.C. Horzinek, H.F. Egberink, R.J. de Groot, J. Virol. **71**, 5277 (1997)
12. J.A. Den Boon, E.J. Snijder, J.K. Locker, M.C. Horzinek, P.J. Rottier, Virology **182**, 655 (1991)
13. M. Koopmans, J. Ederveen, G.N. Woode, M.C. Horzinek, Am. J. Vet. Res. **47**, 1896 (1986)
14. E.J. Snijder, J.A. den Boon, W.J. Spaan, G.M. Verjans, M.C. Horzinek, J. Gen. Virol. **70**, 3363 (1989)
15. E.J. Snijder, M.C. Horzinek, W.J. Spaan, J. Virol. **64**, 331 (1990)
16. E.J. Snijder, J.A. Den Boon, W.J. Spaan, M. Weiss, M.C. Horzinek, Virology **178**, 355 (1990)
17. A.L. Hughes, Arch. Virol. **156**, 1823 (2011)
18. S.K. Lau, P. Lee, A.K. Tsang, C.C. Yip, H. Tse, R.A. Lee, L.Y. So, Y.L. Lau, K.H. Chan, P.C. Woo, K.Y. Yuen, J. Virol. **85**, 11325 (2011)
19. E.J. Snijder, J.A. den Boon, M.C. Horzinek, W.J. Spaan, Virology **180**, 448 (1991)
20. G.X. Luo, J. Taylor, J. Virol. **64**, 4321 (1990)
21. M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Bioinformatics **23**, 2947 (2007)
22. T.A. Hall, Nucleic Acids Symp. Ser. **41**, 95 (1999)
23. D. Posada, K.A. Crandall, Bioinformatics **14**, 817 (1998)
24. A.J. Drummond, A. Rambaut, BMC. Evol. Biol. **7**, 214 (2007)
25. M.A. Suchard, R.E. Weiss, J.S. Sinsheimer, Mol. Biol. Evol. **18**, 1001 (2001)
26. D.P. Martin, P. Lemey, M. Lott, V. Moulton, D. Posada, P. Lefeuvre, Bioinformatics **26**, 2462 (2010)
27. Z. Liang, A.S. Kumar, M.S. Jones, N.J. Knowles, H.L. Lipton, J. Virol. **82**, 11545 (2008)
28. S.L. Kosakovsky Pond, D. Posada, M.B. Gravenor, C.H. Woelk, S.D. Frost, Bioinformatics **22**, 3096 (2006)
29. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, Mol. Biol. Evol. **28**, 2731 (2011)
30. S.L. Pond, S.D. Frost, Bioinformatics **21**, 2531 (2005)
31. S.L. Kosakovsky Pond, S.D. Frost, Mol. Biol. Evol. **22**, 1208 (2005)
32. D.H. Mathews, J. Sabina, M. Zuker, D.H. Turner, J. Mol. Biol. **288**, 911 (1999)
33. T. Ito, S. Katayama, N. Okada, K. Masubuchi, S. Fukuyama, M. Shimizu, J. Clin. Microbiol. **48**, 1795 (2010)
34. A.M. Maestre, A. Garzón, D. Rodríguez, PLoS ONE **6**, e20972 (2011)
35. J. Pignatelli, M. Jimenez, J. Luque, M.T. Rejas, A. Lavazza, D. Rodriguez, Virus Res. **143**, 33 (2009)
36. D.J. Shin, S.I. Park, Y.J. Jeong, M. Hosmillo, H.H. Kim, H.J. Kim, H.J. Kwon, M.I. Kang, S.J. Park, K.O. Cho, Arch. Virol. **155**, 417 (2010)
37. J. Pignatelli, L. Grau-Roma, M. Jiménez, J. Segalés, D. Rodríguez, Vet. Microbiol. **146**, 260 (2010)
38. P.B. Rosenthal, X. Zhang, F. Formanowski, W. Fitz, C.H. Wong, H. Meier-Ewert, J.J. Skehel, D.C. Wiley, Nature **396**, 92 (1998)
39. S.L. Smits, G.J. Gerwig, A.L. van Vliet, A. Lissenberg, P. Briza, J.P. Kamerling, R. Vlasak, R.J. de Groot, J. Biol. Chem. **280**, 6933 (2005)
40. M.A. Langereis, Q. Zeng, G.J. Gerwig, B. Frey, M. von Itzstein, J.P. Kamerling, R.J. de Groot, E.G. Huizinga, Proc. Natl. Acad. Sci. USA. **106**, 15897 (2009)
41. R.J. de Groot, Glycoconj. J. **23**, 59 (2006)
42. S.L. Smits, A. Lavazza, K. Matiz, M.C. Horzinek, M.P. Koopmans, R.J. de Groot, J. Virol. **77**, 9567 (2003)
43. D. Martin, E. Rybicki, Bioinformatics **16**, 562 (2000)
44. J. Cui, M. Gao, X.F. Ren, Infect. Genet. Evol. **11**, 1957 (2011)
45. S.L. Kavita, C.B. Robert, S.P. Ramesh, G. Deepak, S.K. Smita, G.N. Nicole, I. Roxann, W.S. Haynes, C.R. Stuart, J. Virol. **73**, 152 (1999)
46. J.P. Huelsenbeck, F. Ronquist, Bioinformatics **17**, 754 (2001)
47. B.G. Hall, Mol. Biol. Evol. **22**, 792 (2005)
48. R.C. Olsthoorn, J. van Duin, J. Virol. **70**, 729 (1996)
49. P.D. Nags, J.J. Bujarski, J. Virol. **71**, 3799 (1997)
50. N. Shapka, P.D. Nagy, J. Virol. **78**, 2288 (2004)
51. R.C. Olsthoorn, A. Bruyere, A. Dzianott, J.J. Bujarski, J. Virol. **76**, 12654 (2002)
52. M.J. Roossinck, Annu. Rev. Phytopathol. **35**, 191 (1997)
53. M.J. Stanhope, J.R. Brown, H. Amrine-Madsen, Infect. Genet. Evol. **4**, 15 (2004)