A provisional complete genome-based genotyping system for rotavirus species C from terrestrial mammals

Tohru Suzuki¹,* and Ayako Hasebe²

Abstract
Rotaviruses (RVCs) have been detected in pigs, humans, cows, ferrets and dogs. Despite their zoonotic impact and pathogenicity, the genetic characterization of RVCs is incomplete, unlike rotaviruses A (RVAs), whose genetics are well studied. Several studies reported partial and complete genomic sequences for multiple porcine and canine RVCs. We aimed to establish a complete genome-based genotyping system for RVCs, by analysing complete genome data from 22 porcine RVCs identified in Japan from 2002 to 2010, along with those from multiple human, bovine, porcine and ferret RVCs reported in several previous studies. Comparative sequence analyses among RVCs from various host species demonstrated that porcine RVCs had a high level of genetic diversity. In addition, phylogenetic analyses of all 11 RNA segments indicated that porcine RVCs could be classified into multiple genotypes. Notably, the VP4 dendrogram divided bovine RVCs into multiple genotypes. Consequently, the provisional genotype classification for RVCs from terrestrial mammals revealed the existence of genotypes 18G, 21P, 13I, 4R, 6C, 6M, 9A, 8N, 6T, 5E and 4H for the genes VP7, VP4, VP6, VP1, VP2, VP3, NSP1, NSP2, NSP3, NSP4 and NSP5, respectively, based on the cut-off values as defined by the Rotavirus Classification Working Group. The system established in this study deepens our understanding of RVC evolution and facilitates the discovery of genetic events (gene reassortment and interspecies transmission) among RVCs.

INTRODUCTION

Rotaviruses, members of the family Reoviridae, are a major causative agent of gastroenteritis in humans and many animal species worldwide [1]. Rotaviruses are currently divided into nine species (A to I) based on their distinct antigenicity, and on the sequence diversity of the inner capsid protein, VP6 [2, 3]. Rotavirus C (RVC) was first isolated as a causative agent of diarrhoea in piglets in 1980 [4, 5]. Subsequently, RVCs have been detected in humans, cows, ferrets and dogs [6–10]. Human RVCs have often been detected in sporadic cases and outbreaks of diarrhoea in various age groups, although these have mostly been detected in children under 3 years of age [11–15]. RVC infections in humans have been reported in many countries in North and South America, Asia, Africa, Europe and Oceania [16–27]. In addition to human RVCs, bovine RVCs have also been identified in sporadic cases and outbreaks of diarrhoea, predominantly in adult dairy cows [8, 28–31]. Moreover, a high prevalence of antibodies for RVCs was detected in cattle (47–56%) in Japan and the United States (USA) [32].

Aside from human and bovine RVCs, several studies have found an association between porcine RVCs and gastroenteritis in suckling, weaning and post-weaning pigs [33–40]. In addition, a sero-epidemiological study using enzyme-linked immunosorbent assay (ELISA) demonstrated a high prevalence of RVC antibodies in pigs (93–97%) in Japan and the USA [32]. Moreover, molecular surveys using reverse-transcription PCR (RT-PCR) have been carried out in the USA, Canada, Italy, Japan, Korea and the Czech Republic [36–38, 41–45]. Although the detection rates of porcine RVCs were found to be different between the ELISA and RT-PCR methods, these data indicate that porcine RVCs are widely distributed and transmitted in many swine-producing countries.

The rotavirus genome is composed of 11 double-stranded RNA (dsRNA) segments encoding six structural proteins (VP1–VP4, VP6 and VP7) and five or six non-structural proteins (NSP1–NSP5/6). The structural proteins form infectious triple-layered particles surrounding the genomic dsRNA, while NSPs are primarily involved in dsRNA replication and transcription, cellular pathogenesis and...
maturation of virions [46]. The two outer capsid proteins, VP7 and VP4, elicit independently neutralizing antibody responses and form the basis for the G and P genotypes [1]. The dual genotyping system that differentiates these segments encoding VP7 and VP4 has been commonly used for RV species.

Recently, a complete genome-based genotyping system comprising classification of all 11 dsRNA segments, in place of the dual genotyping system, was proposed by the Rotavirus Classification Working Group (RCWG) [47–49]. Complete genome analysis facilitates a molecular understanding of pathogen evolution. It is also a powerful tool for monitoring gene reassortment and interspecies transmission among pathogens. Following this proposal by the RCWG, it became difficult to establish a complete genome-based genotyping system for RVCs, in contrast to RVAs, because available sequence data were limited to several RVC strains from a few host species [28, 30, 36, 41, 50–54]. More recently, molecular studies on RVCs by several research groups have determined complete sequences of a partial or entire RNA segments in several bovine RVCs from Japan; a number of porcine RVCs from the USA, Japan, Korea and the Czech Republic; and one canine RVC from Hungary [31, 37–40, 44, 45, 55, 56].

Although the above studies also constructed a genotyping system for RVCs on specific RNA segments, a systematic genotype classification incorporating all 11 genomic RNA segments has not yet been established. The aim of the present study was to establish a complete genome-based genotyping system for RVCs, especially those from terrestrial mammals, by appending further complete genomic sequence data from 22 porcine RVCs collected in Japan to sequence data from other RVCs available in GenBank, and by calculating the cut-off values for genotype classification on each RNA segment based on the definition recommended by the RCWG. This system would facilitate the molecular epidemiology and evolutionary study of RVCs.

RESULTS AND DISCUSSION

The VP1 ORF nucleotide sequences of 21 porcine RVC strains other than the 87-G2 strain, were 3246 nucleotides (nt) in length, owing to a lack of 27 nt in the 5′-terminal ends of the VP1 segments as compared to those of human, bovine and canine RVC strains (Table S1, available in the online Supplementary Material). The VP2 ORF nucleotide sequences from 20 porcine RVC strains, excluding the 87-G2 and 93-Z4 strains, were 2655 nt in length, like those of human RVC strains but unlike their counterparts from bovine and canine RVC strains (2646 nt). The VP3 products of 20 porcine RVC strains, excluding the 86-H5 and 87-G2 strains, contained one ORF (2082 nt in length), the same as the ORFs of human and canine RVC strains but different from the ORFs of bovine RVC strains (2088 nt). The VP4, VP6 and VP7 ORF nucleotide sequences from 22 porcine RVC strains are summarized in our previous study [39, 40]. The NSP1 ORF nucleotide sequences from 21 porcine RVC strains, excluding the 87-I4 strain, were 1173 nt in length, because of an absence of 9–15 nt in the 5′-terminal ends of the NSP1 genes as compared to homologues from human, bovine and canine RVC strains. The ORF nucleotide sequences of the NSP2 and NSP3 products from 22 porcine RVC strains were 939 and 1209 nt in length, like those of human, bovine and canine RVC strains. The NSP4 products of 22 porcine RVC strains contained one ORF (453 nt in length), the same as the ORFs of the human and canine RVC strains but different from those of the bovine RVC strains (441 nt) The NSP5 ORF nucleotide sequences from 21 porcine RVC strains, excluding the 87-G2 strain, were 633 nt in length, like the canine RVC ORF, but unlike the human (639 nt) and bovine (630 nt) RVC ORFs.

The percentage identities of ORF nucleotide sequences on the 10 remaining genes, except for the VP3 gene, among the 22 porcine RVC strains exhibited considerable variability, to a greater extent than that observed among human and bovine RVC strains. In contrast, considerable diversity in the VP3 ORF nucleotide sequences was found among the 20 porcine RVC strains (79.7–99.7 % homology), similar to that among human RVC strains (83.9–99.3 % homology) as described previously [54]. Furthermore, comparison of the ORF nucleotide sequences on all 11 genes among human, bovine, canine and porcine RVCs, including the additional 22 porcine RVCs, revealed that porcine RVCs had a high level of diversity, which was clearly different from those of RVCs from other host species (Table S2).

The VP1 amino acid sequences from the 21 porcine RVC strains contain a consensus motif of RNA polymerase that is highly conserved among RNA viruses [57, 58] (data not shown). These findings thus suggest that the VP1 products from these strains function as RNA-dependent RNA polymerases, as is the case for other rotaviruses. Comparative sequence analysis of VP3 ORF nucleotide sequences among human, bovine, and porcine RVC strains, plus the additional 20 porcine RVC strains, demonstrated the presence of the KX(D/N)GNNH and TAMD motifs, which represent a potential active site of guanylyltransferase and a possible casein kinase II phosphorylation site, were conserved among all RVC strains [54, 59]. In addition, comparative sequence analysis of the NSP1 ORF nucleotide sequences among human, bovine, canine and porcine RVCs, including the additional 21 porcine RVCs, revealed that the zinc-binding motif, which is known to be involved in evasion of the innate immune response to viral infection through binding to interferon regulatory factor 3 (IRF3), was conserved among all RVCs [60]. Moreover, comparison of the NSP2 ORF nucleotide sequences among human, bovine, canine and porcine RVCs, plus the additional 20 porcine RVCs, demonstrated that the histidine triad motif (HφφφφHφφφ), which is important for NTPase activity, was conserved among all RVCs, as is in other RV species [61, 62]. Therefore, these results indicate that the histidine triad motif is essential for the viral replication process of rotaviruses.
The contemporary genotype classification for the VP4 and VP7 segments was re-analysed using all RVC strains, i.e. combining the 22 porcine RVC strains with several human RVCs from multiple countries; several bovine RVCs from Japan; a large number of porcine RVCs from the USA, Japan, Korea and the Czech Republic; and one canine RVC from Hungary [30, 31, 37, 38, 40, 41, 44, 45, 54–56]. Cut-off values for the division of VP4 and VP7 genotypes, which were calculated from the frequency distribution of pairwise sequence identities, were set to 85 % at the nucleotide level. Our analysis based on these cut-off values and the results from previous reports revealed the presence of 21 (P1–P21) and 18 (G1–G18) genotypes for VP4 and VP7 segments, respectively (Fig. 1a, b). Porcine RVCs, in contrast to other RVCs, which were grouped into one genotype, were divided into multiple genotypes. Moreover, bovine RVCs were clearly divided into two genotypes reflecting the differences in the VP4 ORF nucleotide sequences described previously [31]. The present phylogenetic analysis of the VP6 segment was carried out by re-analysing the ORF nucleotide sequences from multiple human, bovine, porcine and canine RVCs reported in previous studies, plus the 22 porcine RVCs [30, 31, 36, 38, 39, 44, 45, 52–56]. Based on the cut-off value of 87 % at the nucleotide level, our analysis indicated the presence of 13 genotypes (I1–I13); this result adds a canine and a ferret genotype to the multiple genotypes reported previously [45] (Fig. 1c). Aside from the VP4, VP6 and VP7 genes, phylogenetic analyses of the remaining genes were performed using a data set including the 22 porcine RVC strains, and several human, bovine, porcine and canine RVC strains [30, 31, 44, 45, 54–56]. The cut-off value for the division of these genes was calculated from the frequency distribution of pairwise nucleotide sequence identities based on the definition proposed by the RCWG. The genotype classification for the VP1 gene based on the cut-off value of 84 % revealed the presence of four genotypes (R1–R4) independent of the host species (Fig. 1d). The present analysis of RVC VP2 according to the cut-off value of 85 % demonstrated the presence of two new porcine genotypes (C5 and C6), in addition to the prototype porcine (C1) including the Cowden strain, human (C2), bovine (C3) and canine (C4) genotypes (Fig. 1e). Based on the cut-off value of 85 %, the provisional genetic classification for the VP3 gene revealed the presence of another genotype (M6), in addition to the five genotypes (M1–M5) reported in a previous study [56] (Fig. 1f). Moreover, the dendrogram prepared in this study indicates that human RVCs belonging to the M3 cluster might have been derived from porcine RVCs through gene reassortment, as described in a previous study [54]. The present analysis of RVC NSP1 gene based on the cut-off value of 84 % revealed the presence of three new porcine genotypes (A7–A9), in addition to the six genotypes (A1–A6) described in a previous report [56] (Fig. 1g). The RVC NSP2 analysis based on the cut-off value of 87 % revealed the presence of two additional porcine genotypes (N7 and N8), in addition to the six genotypes reported in a previous study [56] (Fig. 1h). The analysis of RVC NSP3 in the present study, based on the cut-off value of 85 %, demonstrated the presence of six genotypes (T1–T6), which contained one genotype additional to the five reported in a previous study [56] (Fig. 1i). Based on the cut-off value of 81 %, the provisional genetic classification for the NSP4 gene revealed the presence of five genotypes (E1–E5), comprising one additional canine genotype to the four described previously [45] (Fig. 1j). The present analysis of RVC NSP5 based on the cut-off value of 80 % revealed the presence of four NSP5 genotypes (H1–H4) for each host species, which agreed with the result of a previous report [45] (Fig. 1k).

In conclusion, we established a tentative complete genome-based genotyping system for RVCs from multiple human, bovine, porcine, ferret and canine sources. The genotype classification for genes VP7, VP4, VP6, VP1, VP2, VP3, NSP1, NSP2, NSP3, NSP4 and NSP5 revealed the existence of 18, 21, 13, 4, 6, 6, 9, 8, 6, 5 and 4 genotypes, according to nucleotide identity cut-off values of 85, 85, 87, 84, 85, 84, 87, 85, 81 and 80 %, respectively. Nevertheless our system presented here remains incomplete, unlike that for RVA, because available sequence data were limited to RVCs from several terrestrial mammals. However, in future a better complete genome-based genotyping system for RVCs, comparable to that of RVAs, will be constructed by adding complete genomic sequence data from new host species to the sequence data presented in the present study.

**METHODS**

**Samples**

The origins of samples used in this study were described in our previous reports [39, 40] (Table S1). The viral RNA was extracted from a 10 % faecal suspension in minimum essential medium, using a QIAamp viral RNA mini kit (Qiagen, Venlo, Limburg, Netherlands), according to the manufacturer’s instructions. Full-length nucleotide sequences of individual genes from porcine RVCs were amplified by RT-PCR using a set of primers as described elsewhere [30]. RT-PCR was carried out using a PrimeScript High Fidelity RT-PCR Kit (Takara Bio, Inc., Shiga, Japan) with a random 6-mer primer under the following conditions: 45 °C for 30 min and 94 °C for 2 min; 35 cycles of 98 °C for 10 s, 55 °C for 15 s and 68 °C for 2 min; and a final extension step (68 °C, 7 min). The PCR products were cloned into the pCR2.1 TOPO vector (Thermo Fisher Scientific, Carlsbad, CA, USA), and several clones were obtained. Thereafter, the clones were sequenced using a BigDye Terminator v3.1 Cycles Sequencing Kit on an automated ABI Prism 3130 Genetic Analyzer (Thermo Fisher Scientific, Carlsbad, CA, USA).

**Sequence and phylogenetic analyses**

The nucleotide sequences determined in this study were submitted to the DNA Data Bank of Japan (DDBJ), and are retrievable from GenBank. The accession numbers for each RNA segment of each porcine RVC strain are summarized in Table S1. The sequence data were aligned using the Clustal W...
Fig. 1. Phylogenetic trees (a–k) based on the ORF nucleotide sequences of individual genes of RVCs. The dendrograms were constructed using the maximum-likelihood method, as implemented by MEGA6 software. The number next to each node represents the percentage of bootstrap support (of 1000 replicates) for the clusters on the right node. Bootstrap values <70% are not shown. The dotted lines represent the division of genotypes based on cut-off values at the nucleotide level. Genotypes are specified on the right. Symbols (filled circles) above strains indicate the 22 porcine RVCs identified in this study. GenBank accession numbers are also shown below the strains.
Fig. 1. (cont.)
Fig. 1. (cont.)
Fig. 1. (cont.)

(f)
Fig. 1. (cont.)

Fig. 1. (cont.)

Fig. 1. (cont.)

(i)
method in MEGA6 software [63]. Genetic distances were calculated using the Kimura two-parameter correction at the nucleotide level [47]. Phylogenetic analyses were conducted using the maximum-likelihood method with the general time reversible nucleotide substitution model and 1000 bootstrap replicates. Classification of all RNA segments of RVCs, including 22 porcine RVCs, was performed using cut-off values calculated based on the definition recommended by the RCWG [47, 48]. Briefly, the cut-off value for genetic classification of individual genes was estimated as the percentage separating the intra-genotype identities (the nucleotide identities among strains belonging to the same genotype) and the inter-genotype identities (the nucleotide identities among strains belonging to different genotypes).
Funding information
The authors received no specific grant from any funding agency.

Acknowledgements
We thank A. Miyazaki, H. Tsunemitsu and S. Ohashi from the National Institute of Animal Health, NARO, for providing advice and discussion.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References


---

**Five reasons to publish your next article with a Microbiology Society journal**

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.