**Original Article**
J Vet Sci 2018, 19(1), 35-43 · https://doi.org/10.4142/jvs.2018.19.1.35

# Analysis of structure-function relationship in porcine rotavirus A enterotoxin gene

Sharad Saurabh[1], Shubhankar Sircar[1], Jobin J. Kattoor[1], Souvik Ghosh[2], Nobumichi Kobayashi[3], Krisztian Banyai[4], Obli R. VinodhKumar[1], Ujjwal K. De[1], Nihar R. Sahoo[1], Kuldeep Dhama[1], Yashpal S. Malik[1,*]

[1]*Indian Council of Agricultural Research-Indian Veterinary Research Institute, Uttar Pradesh 243122, India*
[2]*Department of Biomedical Sciences, One Health Center for Zoonoses and Tropical Veterinary Medicine, Ross University School of Veterinary Medicine, P.O. Box 334, Basseterre, Federation of Saint Kitts and Nevis*
[3]*Sapporo Medical University School of Medicine, Sapporo 060-8556, Japan*
[4]*Institute for Veterinary Medical Research, Centre for Agricultural Research, Hungarian Academy of Sciences, Budapest 1143, Hungary*

Rotavirus (RV)-infected piglets are presumed to be latent sources of heterologous RV infection in humans and other animals. In RVs, non-structural protein 4 (NSP4) is the major virulence factor with pleiotropic properties. In this study, we analyzed the *nsp4* gene from porcine RVs isolated from diarrheic and non-diarrheic cases at different levels of protein folding to explore correlations to diarrhea-inducing capabilities and evolution of *nsp4* in the porcine population. Full-length *nsp4* genes were amplified, cloned, sequenced, and then analyzed for antigenic epitopes, RotaC classification, homology, genetic relationship, modeling of NSP4 protein, and prediction of post-translational modification. RV presence was observed in both diarrheic and non-diarrheic piglets. All *nsp4* genes possessed the E1 genotype. Comparison of primary, secondary, and tertiary structure and the prediction of post-translational modifications of NSP4 from diarrheic and non-diarrheic piglets revealed no apparent differences. Sequence analysis indicated that *nsp4* genes have a multi-phyletic evolutionary origin and exhibit species independent genetic diversity. The results emphasize the evolution of the E9 *nsp4* genotype from the E1 genotype and suggest that the diarrhea-inducing capability of porcine RVs may not be exclusively linked to its enterotoxin gene.

**Keywords:** enterotoxins, *nsp4* gene evolution, porcine, rotavirus, viral nonstructural proteins

## Introduction

Rotavirus (RV)-induced gastroenteritis is one of the most important public health concerns for humans and animals around the world. Consequences of gastroenteritis are ruinous, mainly in resource-poor countries, and can lead to colossal financial losses in animal health and production industries [2,13]. Among the nine rotavirus groups (RVA–RVI) identified to date, RVA is the most frequent cause of diarrhea among mammalian and avian species [28,29]. RVA is also important due to its high prevalence and pathogenicity in humans as well as in pigs [33], and on several occasions RV-infected piglets served as a prospective source for heterologous infection in humans and other animal species [11,18,40]. There are reports of porcine RVs from India [5,12,19,20,25], and in many parts of India, pig farming mainly sustains the livelihood of poor farmers in general and tribal farmers in particular.

The RV genome consists of 11 segments of double-stranded RNA that encode six structural (VP1 to VP4, VP6, and VP7) and five/six non-structural (NSP1 to NSP5/NSP6) proteins. Genome segment 10 encodes the NSP4 protein (the first viral enterotoxin/viroporin), a multifunctional glycoprotein of 175 amino acid residues that has an important role in viral pathogenesis and morphogenesis [45]. Its cytoplasmic tail portion (45th to 175th amino acids) includes biological properties [38], and the portion from 135th to 175th amino acid residues exhibits a high degree of sequence variation [16]. The interspecies variable domain (ISVD; spanning the 135th to 146th amino acid residues), a flexible region located at the extreme C-terminus, is critical for RV virulence and its diarrhea-inducing capabilities; mutations at the 135th and 139th amino acid residues result in loss of viroporin function [44]. Extensive primary structure analyses of NSP4 sequences, however, did not detect any correlation between amino acid substitution and the symptomatic/

Received 4 May 2016, Revised 29 Oct. 2016, Accepted 23 Nov. 2016
*Corresponding author: Tel: +91-5812302777; Fax: +91-5812301757; E-mail: malikyps@ivri.res.in
pISSN 1229-845X
eISSN 1976-555X

asymptomatic phenotypes of the virus [35].

The gene encoding NSP4 proteins has been classified into 15 genotypes (E1 to E15) based on 85% identity cut-off values [26,32]. When genetic sequences are paralleled, close identity is often revealed between human and animal RVs [41]. The evolution of porcine RVs, especially through the exchange of genomic segments among different host-specific RVs can lead to emergence of novel porcine strains having the capability to infect humans [30]. Reports are available that confirm the occurrence of re-assortment events between human and porcine RV strains and the emergence of RVs strains with rare genotypes [27]. In this study, we sought to determine whether there were correlations between enterotoxins and their diarrhea-inducing capabilities at different folding levels of the NSP4 protein of porcine RVA. We also aimed to elucidate the genetic relationship among genes encoding enterotoxins, for which the NSP4 sequences of RVA strains from diarrheic and non-diarrheic porcine viral specimens together with genes from humans were analyzed.

## Materials and Methods

### Viral specimens

A total of 250 fecal samples from diarrheic (n = 85) and non-diarrheic piglets (n = 165) between 1 and 12 weeks of age were collected from March 2013 to June 2015 from an organized pig farm in the Bareilly district, Uttar Pradesh, India. The samples were processed as described previously [23] and were stored at −20°C.

### Primer designing, RT-PCR, and RVA diagnosis

Nucleotide sequences of VP6 genes of porcine origin were downloaded, aligned by using the ClustalW program and analyzed for the presence of conserved motif(s). Degenerate primers were designed based on the identified conserved motif. Viral RNA extraction followed by cDNA synthesis was done as described in our earlier report [24]. Diagnostic PCR was done by using primers based on the VP6 gene, *viz*., RVA-D-F 5′ TTTGATCACTAAYTATTCACC3′ and RVA-D-R 5′ GGTCACATCCTCTCACTA3′ for specific amplification of a 226 bp size product.

### Characterization of *nsp4* genes and their products

Genes encoding full-length NSP4 protein (742 bp) were amplified from 9 porcine RVA (PoRVA)-positive samples representing 5 diarrheic (PoRVA-30, PoRVA-32, PoRVA-34, PoRVA-36, PoRVA-C3) and 4 non-diarrheic viral specimens (PoRVA-172, PoRVA-173, PoRVA-174, PoRVA-175) by using the reported primer pairs [23]. Amplified genes were cloned in pDrive cloning vector (Qiagen, Germany) and sequenced with the M13 universal primer pair. Open reading frames (ORFs) and untranslated regions were identified through EditSeq tool implemented in DNASTAR software (DNASTAR, USA). The GC contents were calculated by using EditSeq tool. The ORFs of individual genes were theoretically translated and molecular weights/isoelectric points were calculated. Antigenic epitopes in the pre-defined ISVD region, based on the Jameson-Wolf antigenic index, were predicted with the help of the Protean tool implemented in DNASTAR software and using various predictive algorithms.

### Molecular modeling of NSP4 protein and prediction of post-translational modification sites

Secondary and tertiary structure predictions were carried out by using the SOPMA secondary structure prediction method [14] and the I-TASSER server [42], respectively. Putative sites of post-translational modification (PTM), *viz*., phosphorylation, glycosylation, and SUMOylation were predicted by using the NetPhos 2.0 [4], NetOGlyc 4.0 [37], NetNGlyc 1.0 (Center for Biological Sequence Analysis, USA), and GPS-SUMO 2.0 [46] servers, respectively, on the ExPASy portal. The threshold/cut-off values for each prediction were set as described previously [8].

### RotaC classification, homology, and evolutionary analysis

Nucleotide and amino acid sequences were aligned by applying the Clustal W method using MegAlign software ver. 5.0 (DNASTAR, USA), and percentage identities were calculated. Nucleotide sequences of *nsp4* genes that had > 95% similarity during megablast analysis by the basic local alignment search tool (National Center for Biotechnology Information, USA) were retrieved from the NCBI database and subjected to phylogenetic analysis with the help of Molecular Evolutionary Genetic Analysis (MEGA; ver. 6.05) [39]. Nucleotide sequences were aligned by using the default pairwise and multiple alignment parameters of Clustal W in MEGA and applied for selection of the best substitution model for further analysis using FindModel tool [34]. Gaps were positioned to minimize nucleotide mismatches and treated as missing data in all analyses. The maximum likelihood statistical method was applied to construct a phylogenetic tree with the selected model. Tree robustness was tested by using the bootstrap method based on 1,000 replicates. Further, the RotaC2.0 web-based tool for classification of RVA was used to determine the genotype of the NSP4 sequences [22]. To elucidate the genetic diversity among *nsp4* genes, a haplotype network was inferred by using the SplitsTree4 program (ver. 4.13.1) with minimum spanning and parsimony splits networks [15]. The genealogy of the *nsp4* genes was estimated by using the TCS network in PopART (Population Analysis with Reticulate Trees) software ver. 1.7 [9]. To analyze the selection pressure, Tajima's D test was performed on DNA Sequence Polymorphism (DnaSP 5.10.01) [21].

## Results

### RVA diagnosis

The RVA diagnostic PCR yielded a 226 bp amplicon in 166 viral specimens, confirming the presence of RVA in the porcine population on the sampled farm (data not shown). Of the 166 RVA-positive specimens, 5 diarrheic and 4 non-diarrheic clinical samples were randomly selected for amplification of the full length *nsp4* gene. Amplified genes (742 bp) were sequenced and submitted to GenBank (National Center for Biotechnology Information, USA), *i.e.*, PoRVA-C3 (KJ650568), Por-30 (KP868550), Por-32 (KP868551), Por-34 (KP868552), Por-36 (KP868553), Por-172 (KT695800), Por-173 (KT695801), Por-174 (KT695802), and Por-175 (KT695803). GenBank accession numbers of other RVA strains/isolates of human (HuRVA, n = 9) and porcine (n = 12) origin are: WH-a (JN650611), LL3354725 (KC139788), LLP48 (KJ126820), Mani-362-07 (GQ240627), NIV9893 (FJ685615), RMC321 (AF541921), RMC-A7 (AY601542), RMC-G60 (AY601543), 42-1 (KF500194), 156-1 (KF500205), 174-1 (KF500216), C-1 (KF500227), F8P4-A (JN974786), NMTL (JF781167), P343 (AB972865), 82 (JN974788), CMP034 (DQ534017), 07-ire (FJ492833), RV0104 (KC164673), OSU-C5111 (PoRVA reference strain, E1 genotype, KJ450851), Wa (HuRVA reference strain, E1 genotype, AF200225), and KUN (HuRVA reference strain, E2 genotype, D88829).

### Primary structure of nsp4 gene and its product

The full-length cDNA encoding viral enterotoxins ranged from 701 to 751 bp in length with constant ORFs of 528 bp. One of the *nsp4* genes (KJ650568) cloned in the study was partial because we could only clone cDNA of 599 bp. UTRs (5′ and 3′) exhibited variation in length, and they ranged from 34–41 bp and 162–182 bp, respectively. The GC content of full-length *nsp4* were 37.8% to 41.74%. Despite the gene variations, the NSP4 enterotoxins comprised of 175 amino acids residues with ~20 kDa molecular weight and slightly alkaline isolelectric point values, except in a few genes. Comparison of the ISVD region (131–141 amino acid residues) in enterotoxins from diarrheic and non-diarrheic porcine specimens revealed no apparent differences at the 135th and 138th positions (Fig. 1). An absence of apparent differences was also evident when antigenic indices of diarrheic and non-diarrheic enterotoxins were compared (Table 1). However, it is noteworthy that, compared to non-diarrheic strains, only one diarrheic strain (PoRVA-C3) exhibited a very high antigenic index (Table 1).

### Molecular modeling of NSP4 protein

Predicted secondary structures of porcine enterotoxins were made up mainly of alpha helices. The alpha helices, extended strands, beta turns, and random coils ranged from 52.08% to 59.64%, 16.00% to 19.43%, 4.57% to 6.86%, and 17.47% to 24.57%, respectively. There were slight differences observed in NSP4 proteins from diarrheic and non-diarrheic specimens. The proportion of alpha helices (58.29%–59.64%) and beta turns (5.14%–6.29%) were slightly higher in non-diarrheic enterotoxins with concomitant lower content of extended strands and random coils compared to those of diarrheic viroporin (52.08%–57.71%, alpha helices and 4.57%–5.56%, beta turns). The tertiary structure prediction of viral enterotoxins showed that the fold consists of five helices, random coils, and a few extended strands (data not shown). Folds resemble the

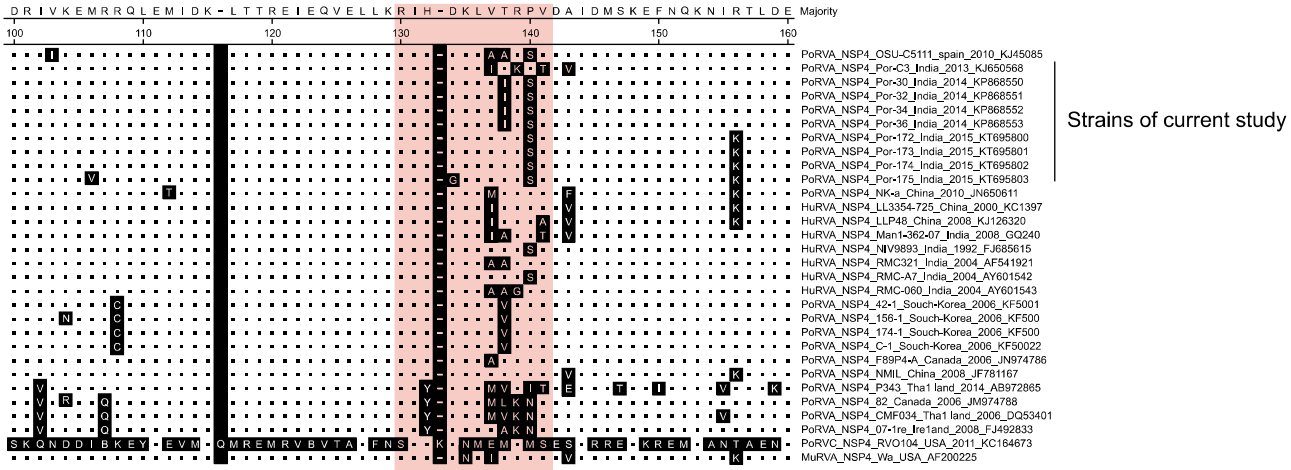


**Fig. 1.** Homology of the interspecies variable domains (ISVD) in NSP4 enterotoxins from porcine and human isolates. The red area indicates the ISVD where no apparent changes in the amino acid residues were observed. The present study included porcine NSP4s isolated from diarrheic (PoRVA-C3, 30, 32, 34, and 36) and non-diarrheic viral specimens (PoRVA-172, 173, 174, and 175). Porcine reference strain OSU-C5111 (E1 genotype, KJ450851) was included for better representation of homology.

**Table 1.** Comparison of antigenic indices of strains used in this study with a reference strain

| Position | Antigenic index | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | OSU-C5111* | PoRV A-C3 | PoRV A-30 | PoRV A-32 | PoRV A-34 | PoRV A-36 | PoRV A-172 | PoRV A-173 | PoRV A-174 | PoRV A-175 |
| 131 | 0.9 | 2.76 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.6 |
| 132 | 0.75 | 2.82 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.3 |
| 133 | 0.45 | 2.78 | −0.3 | −0.3 | −0.3 | −0.3 | −0.15 | −0.15 | −0.15 | −0.45 |
| 134 | 0.75 | 2.74 | 0.3 | 0.3 | 0.3 | 0.3 | 0.9 | 0.9 | 0.9 | 0.45 |
| 135 | 0.75 | 2.7 | 0.6 | 0.6 | 0.6 | 0.6 | 0.9 | 0.9 | 0.9 | 0.6 |
| 136 | 0.3 | 3 | −0.3 | −0.3 | −0.3 | −0.3 | 0.45 | 0.45 | 0.45 | 0.45 |
| 137 | 0.3 | 2 | −0.3 | −0.3 | −0.3 | −0.3 | 0.45 | 0.45 | 0.45 | 0.45 |
| 138 | 0.3 | 1.7 | 0.3 | 0.3 | 0.3 | 0.3 | 0.6 | 0.6 | 0.6 | 0.6 |
| 139 | 0.3 | 1.55 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 |
| 140 | 0.6 | 0.95 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
| 141 | 0.3 | −0.1 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 |

*Reference strain for porcine rotavirus A.

helical bundle-like structure formed by alpha helices together with strands and turns. The ISVD region is comprised mainly of coils followed by helices. The predicted structures showed high similarity to the pentameric structures of RV NSP4 (PDBs 3miwA, 2o1jA, and 2o1kA).

### Putative PTM sites in porcine RVA NSP4

Porcine RVA NSP4 enterotoxins contained 6–11 putative phosphorylation sites (data not shown). There were slight differences in the number of sites when porcine enterotoxins from diarrheic and non-diarrheic specimens were compared. Non-diarrheic NSP4 possessed 11 sites, whereas those from diarrheic specimens had only 9 sites. When analyzed for glycosylation sites, enterotoxins were predicted to have 2 N-glycosylation sites at the 8th and 18th amino acid residues. The O-glycosylation sites were predicted in only two porcine strains, *viz*., PoRVA-C3 (at 130th amino acid residue) and Por-RV0104 (two sites at 63rd and 109th). All NSP4 proteins exhibited one site (at 92nd amino acid) for type-II SUMOylation, with the exception of one non-diarrheic sample (PoRVA-173, KT695801) that possessed two peptides with potential type-II SUMOylation at the 92nd and 163rd amino acid residue (data not shown).

### RotaC classification, homology, and evolutionary analysis

All *nsp4* genes cloned in the study belonged to the E1 genotype as evidenced by their RotaC classification. The *nsp4* genes isolated from diarrheic specimens exhibited more than 99% homologies among them, except for PoRVA-C3 that showed a lower percentage of identity (92.4%) at the nucleotide level. When the deduced amino acid sequences (primary structure of protein) were compared, there was no difference detected among the diarrheic specimens. However, strain PoRVA-C3 shared only 86% homology with other diarrheic specimens at the amino acid level. The *nsp4* genes from non-diarrheic specimens also showed nearly 100% homology among them at the nucleotide and amino acid level. When diarrheic and non-diarrheic enterotoxins were compared, they exhibited > 95% and > 96% homologies, respectively, at the nucleotide and amino acid level (data not shown). Nevertheless, comparison of non-diarrheic enterotoxins with PoRVA-C3 revealed 86%–92% and 80.9%–86.9% homology at the nucleotide and amino acid level, respectively. The *nsp4* genes from the present study shared 76.4%–93.4% and 86.2%–97.5% homologies with selected *nsp4* genes of porcine and human origin, respectively.

In our phylogenetic analysis, RVA isolates with E2 genotype were robustly separated from those with the E1 and E9 genotypes. PoRVA isolates with the E9 genotype formed a separate subclade within the major cluster group of isolates with the E1 genotype (panel A in Fig. 2). The major cluster was further divided into two subclades representing isolates of both human and porcine origin. All *nsp4* genes isolated in the study were grouped into one cluster, with the exception of PoRVA-C3 (panel A in Fig. 2). The above observations were also valid for a phylogenetic tree developed with amino acid sequences, except that the number of subclades within the major cluster increased to four (panel B in Fig. 2). The minimum spanning network generated through the SplitsTree4 program is presented in panel A in Fig. 3. The PoRVA isolate and a human reference strain, *i.e*., "KUN" (accession No. D88829), with the E2 genotype exhibited significant diversity and remained outside of the large cluster containing isolates with E1 and E9 genotypes of porcine and human origin. Diversity among

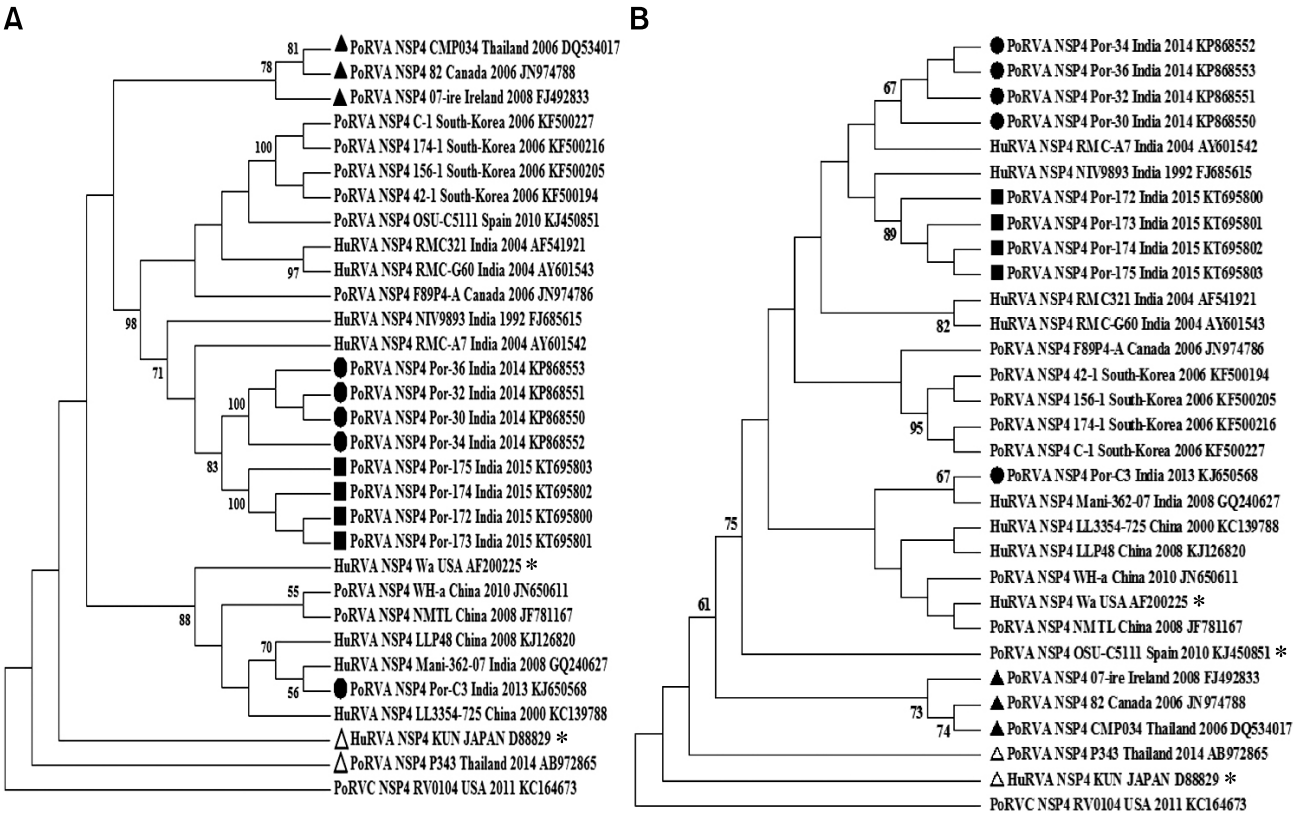


**Fig. 2.** Phylogenetic analysis of nucleotide (A) and amino acid sequences (B) of porcine rotavirus group A (RVA) enterotoxins. RVA isolates with E2 and E9 genotypes were represented as empty (Δ) and solid (▲) triangles, respectively. Isolates with the E2 genotype were robustly separated from those with the E1 and E9 genotypes. Porcine RVA isolates with the E9 genotype were relatively close to those with the E1 genotype and constituted a separate sub-clade suggesting multiphyletic origin of members with E1/E9 and E2 genotypes. All PoRVA isolates identified in the study were clustered together except PoRVA-C3. Round dots (●) represent the diarrheic isolates, while square dots (■) show non-diarrheic PoRVA isolates. Reference prototype porcine strain of OSU-C5111 (E1 genotype, KJ450851), human reference strains Wa (E1 genotype, AF200225) and KUN (HuRVA reference strain, E2 genotype, D88829) are represented by asterisks (*).

members of the E1 and E9 genotypes is represented in a parsimony splits network (panel B in Fig. 3). PoRVA isolates with the E9 genotype exhibited a notable distance from those with the E1 genotype, and members of both genotypes constituted a separate cluster. The genealogy of *nsp4* genes from porcine and human RVA isolates of E1 and E9 genotypes, estimated by using the TCS network (panel C in Fig. 3), supported phylogenetic and genetic diversity analysis observations and elucidated the occurrence of several nucleotide substitution events over time. In the current data set, the TCS network identified PoRVA-30 (accession No. KP868550) as an ancestral haplotype (panel C in Fig. 3). The total dataset contained 180 segregation sites as indicated by the results of the Tajima's D test ($D = -0.34698$, $p > 0.1$).

## Discussion

The non-structural protein 4 of RV is considered a key factor in symptomatic viral infection. Amino acid residues at specific positions in the ISVD or flexible regions were reported to exhibit high variation in a species-specific manner, and mutations in this region (at 135th and 138th amino acid residues within the ISVD) lead to altered virulence in porcine strains [44]. Follow-up studies in humans, felines, and murines, however, were unable to establish a sequence/motif-function relationship in NSP4 proteins by analyzing virulent and attenuated pairs [1,7,31]. Moreover, it has been suggested that the coordinated interaction of a few or all viral proteins results in RV virulence [10]. Furthermore, huge primary structure divergence in the C-terminal region in enterotoxins has been shown to have a significant role in the conformational variation leading to altered biological properties [16,35,44]. Thus, specific tertiary interactions of certain amino acid residues at the ISVD or the flexible C-terminus seem to be important for the retention of diarrhea-inducing capabilities in different NSP4s. In the light of above facts, we undertook a systematic

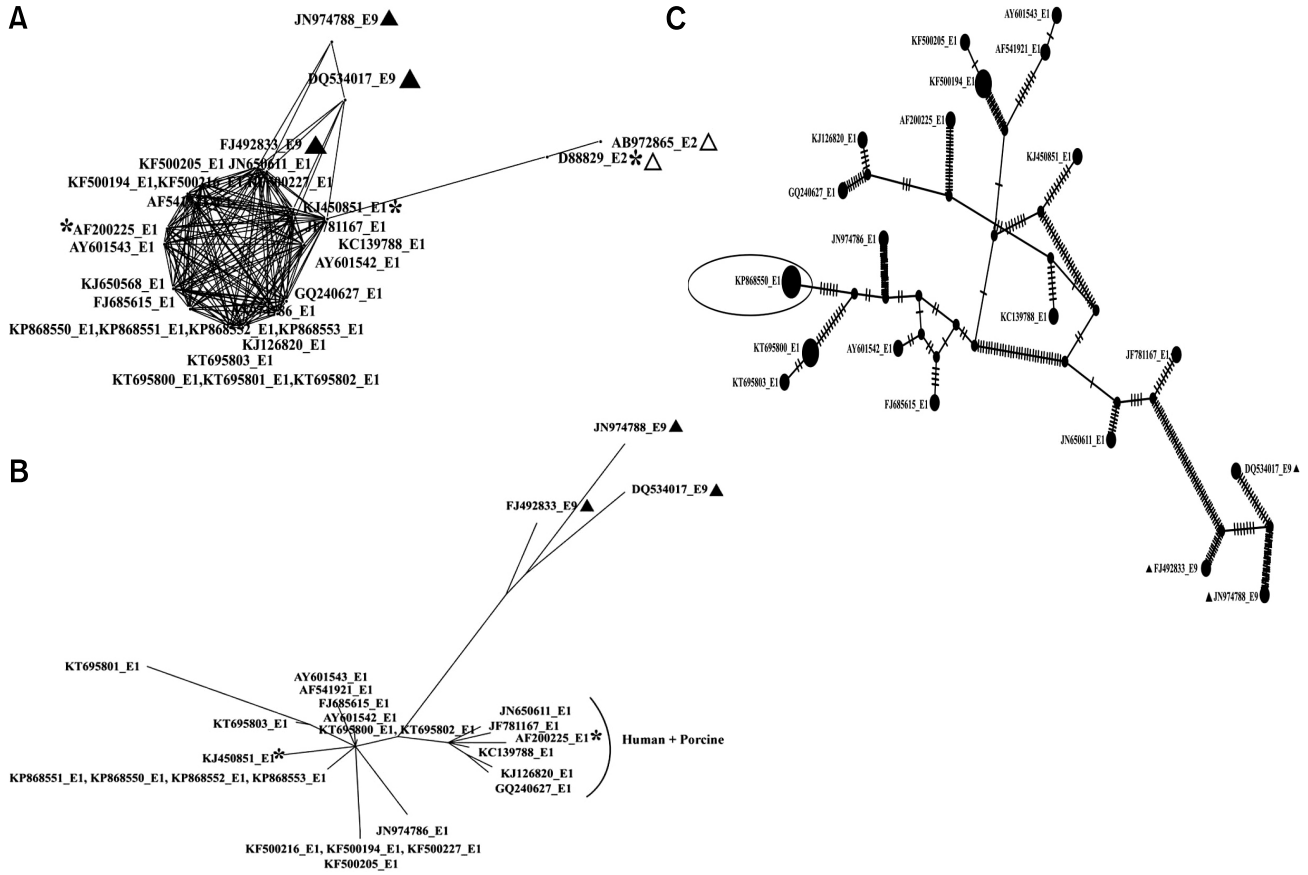


**Fig. 3.** Haplotype networks of *nsp4* genes with different genotypes of human and porcine origin. (A) Minimum spanning network showing closeness of E1 and E9 (▲) genotypes in reference to the E2 (Δ) genotype. (B) Parsimony splits network exhibiting diversity among members of the E1 and E9 genotypes. Grouping of human and porcine isolates of Indian and Chinese origin suggest possible re-assortment events among them. Black asterisks represent reference RVA strains. (C) The genealogy of *nsp4* genes from porcine and human RVA isolates of E1 and E9 genotypes. The TCS network showing 27 haplotypes connected parsimoniously. Individual discs indicate haplotypes with the size of the disc proportional to the number of accessions in the haplotype. RVA strains with E1 and E9 genotypes exhibited distance from each other. The PoRVA-30 strain (encircled in eclipse) isolated in the study was identified as the basal (ancestral) haplotype. Hatch marks represent nucleotide changes during evolution.

analysis of NSP4 proteins from porcine RVA strains at different levels of protein folding and correlated the role of any PTM in the enterotoxin function/virulence with an *in silico* analysis.

In this study, viral specimens were collected over a period of 2 years. Presence of RVA was observed in both symptomatic and asymptomatic porcine viral specimens encompassing all age groups. Prior to 2013, there was no reort of RVA infection in the porcine population of Uttar Pradesh, India [25], but our study indicated a high prevalence of RVA in piglets of this region. In addition, animals that were apparently healthy harbored the virus, thus indicating the sensitivity of the PCR performed with our primer set. Despite the observed variation in the length of cDNAs encoding viral enterotoxins, the length of ORFs and proteins were found to be conserved. The theoretical MW of the translated product was calculated to be ~20 kDa, and it was observed to undergo co-translational glycosylation to 29 kDa and eventually processed to 28 kDa protein [36]. Virulence is usually associated with glycosylation of certain virus proteins, *e.g.*, enterotoxins. N-linked glycosylation is shown to be crucial for NSP4-mediated cytotoxicity in RV viremia [43]. In our analysis, we predicted two conserved N-glycosylation sites in NSP4 sequences. Interestingly, porcine NSP4 sequences cloned from asymptomatic viral specimens were also predicted to possess two N-glycosylation sites, as observed in symptomatic specimens. Additionally, O-glycosylation sites were predicted in two sequences isolated from symptomatic specimens. It is noteworthy that there is no experimental evidence reported to date for O-linked glycosylation in rotaviral enterotoxins; rather some researchers have concluded that recombinant NSP4 did not enter the Golgi apparatus [3]. Our predictions need experimental validation, and it would be interesting to explore the possible effect(s) of

this type of glycosylation in some porcine NSP4s, if any. Phosphorylation of a protein may amend its biological activity and specific protein-protein interactions. Our prediction for phosphorylation in NSP4 suggested many putative sites and is in agreement with a previous analysis [8]. However, we could not find experimental evidence that supports phosphorylation of NSP4 *in vivo*. Further experimental support is required to investigate whether RV enterotoxins are phosphorylated and, if yes, the biological significance of this kind of modification. Likewise, SUMOylation offers functional flexibility to proteins and is involved in different biological processes such as protein localization and functions. SUMOylation in RV proteins, *viz*., VP1, VP2, NSP2, VP6, and NSP5 has been shown to positively regulate replication and viral protein production [6]. Our results suggest that there could be SUMOylation in NSP4 proteins. Prediction of PTMs in rotaviral enterotoxins revealed different types of modifications of which only glycosylation has been experimentally validated previously. *In silico* phosphorylation and SUMOylation in NSP4 have also been previously; therefore, a concerted effort to investigate these modifications may elucidate RV pathogenesis. This information might also be useful in designing and development of treatment stratagems.

Comparison of the primary structure of the ISVD region from diarrheic and non-diarrheic specimens did not correlate with the presence of certain amino acids at specific positions, as was reported previously [44]. We did not observe apparent changes in the ISVD region, *e.g*., serine was present instead of proline at the 138th position in porcine isolates with both clinical signs (both symptomatic and asymptomatic specimens). Likewise, valine instead of alanine was present at the 135th position. Comparison of antigenic indices also exhibited no observable difference, except in one diarrheic strain (PoRVA-C3). The high antigenic index of PoRVA-C3 may be attributed to its different primary structure and, eventually, to different tertiary interactions. These observations were concomitant with previous conclusions that the primary structure of NSP4 is unrelated to its biological functions [1,7,31]. We, therefore, looked into folding of viral enterotoxins and compared the predicted secondary structures. When secondary structures of enterotoxins from diarrheic and non-diarrheic specimens were compared, we did not detect significant differences between them. The absence of significant differences was also true for the three-dimensional structure predictions for enterotoxins. NSP4 proteins from both symptomatic and asymptomatic specimens exhibited highest similarity with a single template only (PDB, 3miwA). This may be due to similar fold structures of NSP4 proteins irrespective of clinical conditions. However, it might also be due to non-availability of a crystal structure of enterotoxins isolated from asymptomatic viral specimens. Whatever the reason, we were unable to differentiate diarrheic NSP4 from non-diarrheic NSP4 at the three-dimensional folding level. Further, we hypothesized whether PTMs have a role in diarrhea-inducing capabilities. We did observe slight changes in the number of phosphorylation sites when non-diarrheic and diarrheic porcine viral specimens isolated in this study were compared. Those differences, however, were not evident in other porcine sequences present in the public database (National Center for Biotechnology Information, USA). Likewise, no significant difference was observed in the SUMOylation sites in NSP4 sequences.

Our phylogenetic analysis indicated that genetic diversity among *nsp4* genes is not species specific as all members of the E1 genotype from both porcine and human origins clustered together. The PoRVA isolates with the E9 genotype were found relatively close to those of the E1 genotype, and together they constituted a separate cluster suggesting a multiphyletic origin of members with E1/E9 and E2 genotypes. Our observations support those in a previous report in which human and porcine isolates of E1 and E9 genotypes grouped together. Moreover, a HuRVA isolate with an E2 genotype clustered independently [41], as was observed in our study. The PoRVA isolate with the E2 genotype is believed to be an example of a bovine to porcine interspecies transmission event [17]; therefore, it formed a separate cluster in the phylogenetic analysis results. Closeness of the E1 and E9 genotype was further evident in genetic diversity study; however, despite this closeness, the E9 genotype exhibited sufficient diversity to constitute a separate cluster in the parsimony splits network. The E9 genotype is only represented in porcine RVA isolates in contrast to E1 that is present in isolates of both human and porcine origin. It is possible that a pool of RVA isolates with the E9 genotype circulating in porcine population might have evolved from those with the E1 genotype. Future studies should focus on comparison of biological functions of porcine NSP4s from both genotypes to elucidate whether this closeness has any role at the protein level. In addition, the results of the genetic diversity analysis supported those of the phylogenetic analysis as 2 PoRVA and 2 HuRVA isolates from China were grouped with the HuRVA isolate (GQ240627) from Manipur, India. Other Indian HuRVA isolates either constituted independent clusters or grouped together. Among the PoRVA isolates, those from Korea formed one cluster, while Indian isolates identified in this study constituted two separate clusters; one grouping *nsp4* genes from diarrheic specimens and the other from non-diarrheic specimens. Further, one PoRVA isolate from a non-diarrheic specimen (accession No. KT651801) constituted an independent cluster, revealing further diversity among *nsp4* genes from non-diarrheic specimens. One PoRVA isolate from Canada (JN974786) remain independent from others in the study, indicating circulation of a different RV strain in that region. Association of human and porcine RV strains from different geographical locations, as observed in this study, indicate the possibility of re-assortment events between species, as has been reported previously [27,30]. Observations in genetic diversity

and genealogy studies among genes encoding enterotoxins were correlated and clearly indicated evolutionary distances of the RVA isolates with different genotypes. This is due to the occurrence of several nucleotide substitution events during evolution as shown in TCS network; however, these mutations were random as the Tajima's D test results were not statistically significant.

This study was intended to draw putative correlation if any, between RV enterotoxins and its diarrhea-inducing capabilities at the protein level; however, we did not detect any evidence of such a correlation. Our preliminary observation based on *in silico* analysis of deduced amino acid sequences and structural prediction indicated that the diarrhea-inducing capabilities and/or pathogenesis of porcine RVA may not be linked to its enterotoxin. There might be involvement of certain receptor(s) inside the host and the host's immune status for development of clinical signs after infection. Additionally, our results suggested that the porcine E9 genotype might have evolved from the E1 genotype. Further studies should be undertaken to investigate possible interspecies transmission events between Indian RV populations of porcine and human origin.

## Acknowledgments

The authors of the manuscript thank and acknowledge their Institute (Indian Veterinary Research Institute, Izatnagar), and the corresponding author (YS Malik) thanks the Agricultural Education Division, Indian Council of Agricultural Research (ICAR), Ministry of Agriculture & Farmers Welfare, New Delhi, for awarding the ICAR National Fellowship.

## Conflict of Interest

The authors declare no conflicts of interest.

## References

1. **Angel J, Tang B, Feng N, Greenberg HB, Bass D.** Studies of the role for NSP4 in the pathogenesis of homologous murine rotavirus diarrhea. J Infect Dis 1998, **177**, 455-458.
2. **Attoui H, Becnel J, Belaganahalli S, Bergoin M, Brussaard C, Chappell J, Ciarlet M, del Vas M, Dermody T, Dormitzer P.** Family *Reoviridae*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds.). Virus Taxonomy. Classification and Nomenclature of Viruses. pp. 541-637, Academic Press, Waltham, 2012.
3. **Berkova Z, Crawford SE, Trugnan G, Yoshimori T, Morris AP, Estes MK.** Rotavirus NSP4 induces a novel vesicular compartment regulated by calcium and associated with viroplasms. J Virol 2006, **80**, 6061-6071.
4. **Blom N, Gammeltoft S, Brunak S.** Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol 1999, **294**, 1351-1362.
5. **Bora DP, Barman NN, Bhattacharyya DK, Dutta LJ.** Faecal excretion of rotavirus by naturally infected pregnant sows in organized pig farms–a potent source of infection to newborns. Indian J Anim Sci 2011, **81**, 575-577.
6. **Campagna M, Marcos-Villar L, Arnoldi F, de la Cruz-Herrera CF, Gallego P, González-Santamaría J, González D, Lopitz-Otsoa F, Rodriguez MS, Burrone OR, Rivas C.** Rotavirus viroplasm proteins interact with the cellular SUMOylation system: implications for viroplasm-like structure formation. J Virol 2013, **87**, 807-817.
7. **Chang KO, Kim YJ, Saif LJ.** Comparisons of nucleotide and deduced amino acid sequences of NSP4 genes of virulent and attenuated pairs of group A and C rotaviruses. Virus Genes 1999, **18**, 229-233.
8. **Chattopadhyay S, Bagchi P, Dutta D, Mukherjee A, Kobayashi N, Chawlasarkar M.** Computational identification of posttranslational modification sites and functional families reveal possible moonlighting role of rotaviral proteins. Bioinformation 2010, **4**, 448-451.
9. **Clement M, Posada D, Crandall KA.** TCS: a computer program to estimate gene genealogies. Mol Ecol 2000, **9**, 1657-1659.
10. **Deepa R, Rao CD, Suguna K.** Structure of the extended diarrhea-inducing domain of rotavirus enterotoxigenic protein NSP4. Arch Virol 2007, **152**, 847-859.
11. **Dhama K, Chauhan RS, Mahendran M, Malik SVS.** Rotavirus diarrhea in bovines and other domestic animals. Vet Res Commun 2009, **33**, 1-23.
12. **Dubal ZB, Bhilegaonkar KN, Barbuddhe SB, Kolhe RP, Kaur S, Rawat S, Nambiar P, Karunakaran M.** Prevalence and genotypic (G and P) determination of porcine group A rotaviruses from different regions of India. Trop Anim Health Prod 2013, **45**, 609-615.
13. **Estes MK, Kapikian AZ.** Rotaviruses. In: Knipe DM, Howley PM, Griffin DE, Martin MA, Lamb RA, Roizman B, Straus SE (eds.). Fields Virology. 5 ed. Vol. 2. pp. 1917–1974, Lippincott Williams & Wilkins, Philadelphia, 2007.
14. **Geourjon C, Deléage G.** SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 1995, **11**, 681-684.
15. **Huson DH, Bryant D.** Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 2006, **23**, 254-267.
16. **Jagannath M, Vethanayagam RR, Reddy BS, Raman S, Rao CD.** Characterization of human symptomatic rotavirus isolates MP409 and MP480 having 'long' RNA electropherotype and subgroup I specificity, highly related to the P6[1],G8 type bovine rotavirus A5, from Mysore, India. Arch Virol 2000, **145**, 1339-1357.
17. **Komoto S, Pongsuwanna Y, Ide T, Wakuda M, Guntapong R, Dennis FE, Haga K, Fujii Y, Katayama K, Taniguchi K.** Whole genomic analysis of porcine G10P[5] rotavirus strain P343 provides evidence for bovine-to-porcine interspecies transmission. Vet Microbiol 2014, **174**, 577-583.
18. **Kuga K, Miyazaki A, Suzuki T, Takagi M, Hattori N, Katsuda K, Mase M, Sugiyama M, Tsunemitsu H.** Genetic diversity and classification of the outer capsid glycoprotein VP7 of porcine group B rotaviruses. Arch Virol 2009, **154**, 1785-1795.
19. **Kumar M, Bhilegaonkar KN, Agarwal RK.** Prevalence and

characterization of rotavirus from faecal samples of children and animals. Indian J Anim Sci 2011, **81**, 993-999.
20. **Kusumakar AL, Savita, Malik YPS, Minakshi, Prasad G.** Occurrence of mammalian group A rotavirus in swine population of central India. Indian J Anim Sci 2008, **78**, 475-477.
21. **Librado P, Rozas J.** DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 2009, **25**, 1451-1452.
22. **Maes P, Matthijnssens J, Rahman M, Van Ranst M.** RotaC: a web-based tool for the complete genome classification of group A rotaviruses. BMC Microbiol 2009, **9**, 238.
23. **Malik YS, Kumar N, Sharma K, Ghosh S, Bányai K, Balasubramanian G, Kobayashi N, Matthijnssens J.** Molecular analysis of non structural rotavirus group A enterotoxin gene of bovine origin from India. Infect Genet Evol 2014, **25**, 20-27.
24. **Malik YS, Kumar N, Sharma K, Haq AA, Kumar A, Prasad M.** Sequence and phylogenetic analysis of bovine rotavirus isolates (G6 genotypes) from India. Adv Anim Vet Sci 2013, **1**, 41-43.
25. **Malik YS, Kumar N, Sharma K, Sircar S, Dhama K, Bora DP, Dutta TK, Prasad M, Tiwari AK.** Rotavirus diarrhea in piglets: a review on epidemiology, genetic diversity and zoonotic risks. Indian J Anim Sci 2014, **84**, 1035-1042.
26. **Matthijnssens J, Ciarlet M, McDonald SM, Attoui H, Bányai K, Brister JR, Buesa J, Esona MD, Estes MK, Gentsch JR, Iturriza-Gómara M, Johne R, Kirkwood CD, Martella V, Mertens PPC, Nakagomi O, Parreño V, Rahman M, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Patton JT, Desselberger U, Van Ranst M.** Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). Arch Virol 2011, **156**, 1397-1413.
27. **Matthijnssens J, Rahman M, Ciarlet M, Zeller M, Heylen E, Nakagomi T, Uchida R, Hassan Z, Azim T, Nakagomi O, Van Ranst M.** Reassortment of human rotavirus gene segments into G11 rotavirus strains. Emerg Infect Dis 2010, **16**, 625-630.
28. **Médici KC, Barry AF, Alfieri AF, Alfieri AA.** Porcine rotavirus groups A, B, and C identified by polymerase chain reaction in a fecal sample collection with inconclusive results by polyacrylamide gel electrophoresis. J Swine Health Prod 2011, **19**, 146-150.
29. **Miyazaki A, Kuga K, Suzuki T, Kohmoto M, Katsuda K, Tsunemitsu H.** Annual changes in predominant genotypes of rotavirus A detected in the feces of pigs in various developmental stages raised on a conventional farm. Vet Microbiol 2013, **163**, 162-166.
30. **Mukherjee A, Ghosh S, Bagchi P, Dutta D, Chattopadhyay S, Kobayashi N, Chawla-Sarkar M.** Full genomic analyses of human rotavirus G4P[4], G4P[6], G9P[19] and G10P[6] strains from North-eastern India: evidence for interspecies transmission and complex reassortment events. Clin Microbiol Infect 2011, **17**, 1343-1346.
31. **Oka T, Nakagomi T, Nakagomi O.** A Lack of consistent amino acid substitutions in NSP4 between rotaviruses derived from diarrheal and asymptomatically-infected kittens. Microbiol Immunol 2001, **45**, 173-177.
32. **Papp H, Al-Mutairi L, Chehadeh W, Farkas S, Lengyel G, Jakab F, Martella V, Szűcs G, Bányai K.** Novel NSP4 genotype in a camel G10P[15] rotavirus strain. Acta Microbiol Immunol Hung 2012, **59**, 411-421.
33. **Papp H, László B, Jakab F, Ganesh B, De Grazia S, Matthijnssens J, Ciarlet M, Martella V, Bányai K.** Review of group A rotavirus strains reported in swine and cattle. Vet Microbiol 2013, **165**, 190-199.
34. **Posada D, Crandall KA.** MODELTEST: testing the model of DNA substitution. Bioinformatics 1998, **14**, 817-818.
35. **Rajasekaran D, Sastri NP, Marathahalli JR, Indi SS, Pamidimukkala K, Suguna K, Rao CD.** The flexible C terminus of the rotavirus non-structural protein NSP4 is an important determinant of its biological properties. J Gen Virol 2008, **89**, 1485-1496.
36. **Srivastava S, Jain A.** Rotavirus nonstructural protein 4 (NSP4)-viral enterotoxin with multiple roles in pathogenesis of diarrhoea in children. J Appl Pharm Sci 2015, **5**, 146-153.
37. **Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KTBG, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, Wandall HH, Levery SB, Clausen H.** Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J 2013, **32**, 1478-1488.
38. **Suguna K, Rao CD.** Rotavirus nonstructural proteins: a structural perspective. Curr Sci 2010, **98**, 352-359.
39. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 2013, **30**, 2725-2729.
40. **Tate JE, Chitambar S, Esposito DH, Sarkar R, Gladstone B, Ramani S, Raghava MV, Sowmyanarayanan TV, Gandhe S, Arora R, Rarashar UD, Kang G.** Disease and economic burden of rotavirus diarrhoea in India. Vaccine 2009, **27S**, F18-24.
41. **Theuns S, Heylen E, Zeller M, Roukaerts IDM, Desmarets LMB, Van Ranst M, Nauwynck HJ, Matthijnssens J.** Complete genome characterization of recent and ancient Belgian pig group A rotaviruses and assessment of their evolutionary relationship with human rotaviruses. J Virol 2015, **89**, 1043-1057.
42. **Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y.** The I-TASSER Suite: protein structure and function prediction. Nat Methods 2015, **12**, 7-8.
43. **Yang W, McCrae MA.** The molecular biology of rotaviruses X: intercellular dissemination of rotavirus NSP4 requires glycosylation and is mediated by direct cell-cell contact through cytoplasmic extrusions. Arch Virol 2012, **157**, 305-314.
44. **Zhang M, Zeng CQY, Dong Y, Ball JM, Saif LJ, Morris AP, Estes MK.** Mutations in rotavirus nonstructural glycoprotein NSP4 are associated with altered virus virulence. J Virol 1998, **72**, 3666-3672.
45. **Zhang M, Zeng CQY, Morris AP, Estes MK.** A functional NSP4 enterotoxin peptide secreted from rotavirus-infected cells. J Virol 2000, **74**, 11663-11670.
46. **Zhao Q, Xie Y, Zheng Y, Jiang S, Liu W, Mu W, Liu Z, Zhao Y, Xue Y, Ren J.** GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Res 2014, **42**, W325-330.