# MOLECULAR DISSECTION OF PORCINE REPRODUCTIVE AND RESPIRATORY VIRUS PUTATIVE NONSTRUCTURAL PROTEIN 2

Kay S. Faaberg, Jun Han, and Yue Wang*

## 1. INTRODUCTION

Porcine reproductive and respiratory syndrome virus (PRRSV) belongs to the family *Arteriviridae* in the order *Nidovirales*.[1] The virus is now known to consist of two different genotypes based on the finding that the prototypes viruses, Lelystad virus (European-like, Type 1[2]) and VR-2332 (North American-like, Type 2[3]), display only approximately 60% nucleotide identity. PRRSV has since been shown to consist of multiple virus isolates that vary within each genotype as much as 20% in nucleotide composition[4,5] (Faaberg, unpublished data). Within this genetic backdrop, seemingly novel field isolates of PRRSV suddenly appeared in the southwestern region of the State of Minnesota, U.S.A., in 2002. Phylogenic analysis, based on the ORF5 gene (encoding the viral attachment protein) of 916 unique PRRSV isolates, revealed that these isolates were most similar to those found in Canada in the early 1990s[6,7] (Faaberg, unpublished data). The PRRSV isolates were determined to have the restriction fragment length polymorphism pattern designated 1-8-4[8] and thus were named MN184 isolates. In order to examine the MN184 isolates more closely, we determined the full-length nucleotide sequences of two field isolates differing in apparent virulence. MN184 isolate comparison revealed that differences existed throughout the genome, most notably in nonstructural protein (Nsp) 2, for which no function has been assigned, except by comparison with the genome of equine arteritis virus.[9–11] The comparison of the MN184 field isolates to the prototypic strain VR-2332,[3] and to the first Type 1 strain seen in the United States, EuroPRRSV,[5] was then completed.

*University of Minnesota, Saint Paul, Minnesota 55108.

## 2. METHODS

### 2.1. PRRSV Strains

MN184 field isolates A (moderate clinical signs) and B (severe clinical signs) were obtained from Kurt Rossow, D.V.M., Ph.D., at the Minnesota Veterinary Diagnostic Laboratory (MVDL) after a single round of PRRSV amplification on freshly isolated porcine alveolar macrophages (PAM). Strains VR-2332 (U87392)[3,12] and EuroPRRSV (AY366525)[5] have been described previously.

### 2.2. Determination of Complete Genomes of MN184A and MN184B Isolates

Viral RNA was purified from infected PAM supernatant using the QIAamp viral RNA kit (Qiagen). RNA was converted to DNA using random hexamers and sequence-specific forward and reverse primers by One-Step RT-PCR (Qiagen). 5'- and 3'- rapid amplification of cDNA ends (RACE) was performed using 5' and 3'-Full Race Core Set (TaKaRa Bio Inc.) on purified viral RNA. PCR primers were derived from several sources, including those described for amplification of strain VR-2332,[3,12] strain JA142 (AY424271)[5] and newly generated MN184 sequence. The detailed primer set used to delineate the MN184 genome will be described elsewhere. The individual nucleotide sequences were assembled using the SeqMan II program in the Lasergene software suite (Version 6; DNASTAR, Inc.). A minimum of three-fold sequence coverage of each genome was obtained.

### 2.3. Genome Analysis

The complete nucleotide sequences for all four PRRSV genomes were analyzed using the Genetics Computer Group Wisconsin Package (GCG, Version 10.3-UNIX; Accelrys, Inc.). The genomes of MN184A and MN184B have been deposited in GenBank.

## 3. RESULTS

### 3.1. Complete Nucleotide Sequences of MN184 Isolates

The MN184 isolates were amplified only by a single round of growth on freshly isolated PAM, the host cell, in order to identify as much variation in the nucleotide sequences as possible. This variation, as suggested by nucleotide degeneracy at individual and discrete nucleotides during sequence analysis of individual PCR product tracefiles, was considerable. Notably, isolate MN184B, which produced severe clinical signs in the field, exhibited much more variation than isolate MN184A.

The complete genomes were found to be identical in length (15,019 bases excluding the polyA tail) and the shortest genome identified to date, including the Type 1 strain EuroPRRSV which possesses only 15,047 bases. However, the isolates were more closely related to Type 2 strains, the shortest of which had been the Chinese strain HB-2

(15,398 bases)[13]. The two MN184 isolates were 97.8% identical in nucleotide sequence, and genome-wide possessed only 326 nucleotide differences which included the degenerate sites mentioned above (Table 1).

### 3.2. Comparison with Prototypic PRRSV Strains VR-2332 and EuroPRRSV

The two MN184 isolates displayed considerable genetic distance from both strain EuroPRRSV (~57%) and strain VR-2332 (~85%). These differences were seen throughout the genome but were greatest in the regions encoding putative Nsp2 (94.5% identity) and the viral attachment protein (open reading frame 5; 97.7% identity). The identified Nsp2 region of ORF1a (the replicase polyprotein) contained the majority of the nucleotide degeneracy seen when analyzing the complete genomes and also included a large deletion when compared with strain VR-2332. We chose this PRRSV genomic region for further bioinformatic analysis.

### 3.3. The Genomic Region Encoding Nsp2 Contains Several Putative Domains

The putative Nsp2 region, originally identified as spanning amino acids (a.a.) 384-1363,[14] was projected to include a.a. 384-1578 of the strain VR-2332 ORF1a protein through genetic analysis of both *Coronaviridae* and *Arteriviridae*.[15] Nsp2 has been previously shown to be the key region of length difference between Type 1 and Type 2 isolates[3] and also revealed by several investigators to vary extensively between North American-like Type 2 isolates[3,13,14,16] as well as between North American Type 1 isolates.[5,17] Comparative sequence analysis with other Nidoviruses has identified a cysteine protease domain near the N-terminal end.[9] The amino acid makeup of this region is over 10% proline, and contains many PxxP motifs, the signature binding motif of Src homology 3 (SH3) domains, which suggests that Nsp2 may be involved in signal transduction mechanisms. For Type 1 strains, there is a leucine zipper motif (Lx6Lx6Lx6L) near C-terminal end.[5] Several B-cell epitopes were identified for Nsp2 using bacteriophage display.[18]

Further bioinformatic analysis of Nsp2 of strain VR-2332 using SignalP 3.0 (www.cbs.dtu.dk/services/SignalP/)[19] and Interproscan (www.ebi.ac-.uk/InterProScan/)[20] revealed the existence of several other domains.

**Table 1.** Nucleotide similarity and divergence of the four PRRSV genomes.

| | | Percent Nucleotide Similarity | | | |
|---|---|---|---|---|---|
| | | VR-2332 | MN184A | MN184B | EuroPRRSV |
| Divergence | VR-2332 | - | 84.5 | 84.7 | 56.6 |
| | 184A | 14.9 | - | 97.8 | 57.6 |
| | 184B | 13.9 | 1.3 | - | 57.4 |
| | EuroPRRSV | 61.0 | 61.6 | 61.0 | - |

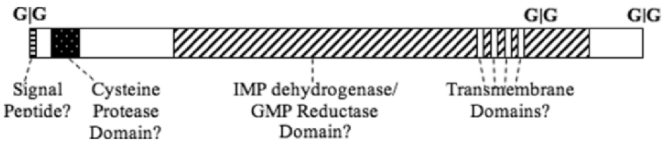


**Figure 1.** Putative Nsp2 protein is predicted to contain several domains. The first G|G represents the putative site of Nsp1β/Nsp2 cleavage at nt 383|384 and the predicted Nsp2 cleavage site at 1363 (second G|G[14]) or at a.a. 1578 (third G|G[15]). A signal peptide at a.a. 22, an equine arterivirus Nsp2-type cysteine protease domain (a.a. 45-152),[9] an IMP dehydrogenase/GMP reductase domain (a.a. 284-1092), and four transmembrane domains located at a.a. 875-895, 910-930, 959-979, 988-1008 were identified.

## 4. CONCLUSIONS

The identification of the MN184 isolates in Minnesota 10 years after similar strains were identified in Quebec, Canada, was unexpected. The MVDL has collected over 4400 ORF5 field isolate sequences from the swine producing regions of United States and Canada, generated in the course of diagnostic analysis, suggesting that these novel isolates appeared suddenly and with no known direct origin. The complete genome sequence determination and analysis of two of these isolates, MN184A of reported moderate virulence and MN184B of reported severe virulence, was undertaken to attempt to determine regions on the genome coding for virulence factors. Two genetic regions were revealed by sequence comparison and found to code for Nsp2 (94.5% identity) and ORF5 (97.7% identity), but the number of identified differences noted genome wide (326 nucleotides) precluded immediate identification of such virulence determinants. Nucleotide degeneracy in Nsp2 may be a factor in predicting virulence, but this must be addressed in separate publication.

The complete genomes of MN184A and MN184B consist of 15019 bases and thus identify the smallest PRRSV isolates to date. These Type 2 isolates were found to diverge from strain VR-2332 nucleotide sequence by approximately 15%. To begin to analyze the region found to account for the most sequence divergence between the MN184 isolates (Nsp2), we submitted the Nsp2 region of strain VR-2332 for bioinformatic analysis. The Nsp2 protein had been previously characterized to have over 10% proline, several PxxP motifs, a leucine zipper motif[5] and several B-cell epitopes.[18] However, further analysis identified a potential signal sequence, an unusual IMP dehydrogenase/GMP reductase domain and four transmembrane domains. We are now poised to begin molecular exploration of this extremely variable region of the PRRSV replicase.

## 5. REFERENCES

1. Cavanagh, D., 1997, Nidovirales: a new order comprising Coronaviridae and Arteriviridae, *Arch. Virol*. **142**:629-633.
2. Meulenberg, J, J., Hulst, M. M., de Meijer, E. J., Moonen, P. L., den Besten, A., de Kluyver, E. P., Wensvoort, G., and Moormann, R. J., 1993, Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV, *Virology* **192**:62-72.
3. Nelsen, C. J., Murtaugh, M. P., and Faaberg, K. S., 1999, Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents, *J. Virol.* **73**:270-280.

4. Kapur, V., Elam, M. R., Pawlovich, T. M., Murtaugh M. P., 1996, Genetic variation in porcine reproductive and respiratory syndrome virus isolates in the midwestern United States, *J. Gen. Virol.* **77**:1271-1276.
5. Ropp, S. L., Wees, C. E., Fang, Y., Nelson, E. A., Rossow, K. D., Bien, M., Arndt, B., Preszler, S., Steen, P., Christopher-Hennings, J., Collins, J. E., Benfield, D. A., and Faaberg, K. S., 2004, Characterization of emerging European-like porcine reproductive and respiratory syndrome virus isolates in the United States, *J. Virol.* **78**:3684-3703.
6. Mardassi, H., Athanassious, R., Mounir, S., Dea, S., 1994, Porcine reproductive and respiratory syndrome virus: morphological, biochemical and serological characteristics of Quebec isolates associated with acute and chronic outbreaks of porcine reproductive and respiratory syndrome, *Can. J. Vet. Res.* **58**:55-64.
7. Murtaugh, M. P., Yuan, S., Faaberg, K. S., 2001, Appearance of novel PRRSV isolates by recombination in the natural environment, *Adv. Exp. Med. Biol.* **494**:31-36.
8. Wesley, R. D., Mengeling, W. L., Lager, K. M., Clouser, D. F., Landgraf, J. G., Frey, M. L., 1998, Differentiation of a porcine reproductive and respiratory syndrome virus vaccine strain from North American field strains by restriction fragment length polymorphism analysis of ORF 5, *J. Vet. Diagn. Invest.* **10**:140-144.
9. Snijder, E. J., Wassenaar, A. L., Spaan, W. J., and Gorbalenya, A. E., 1995, The arterivirus Nsp2 protease. An unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases, *J. Biol. Chem.* **270**:16671-16676.
10. Pedersen, K. W., van der Meer, Y., Roos, N., and Snijder, E. J., 1999, Open reading frame 1a-encoded subunits of the arterivirus replicase induce endoplasmic reticulum-derived double-membrane vesicles which carry the viral replication complex, *J. Virol.* **73**:2016-2026.
11. Snijder, E. J., van Tol, H., Roos, N., and Pedersen, K. W., 2001, Non-structural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex, *J. Gen. Virol.* **82**:985-994.
12. Yuan, S., Mickelson, D. M., Murtaugh, M. P., and Faaberg, K. S., 2001, Complete genome comparison of porcine reproductive and respiratory syndrome virus parental and attenuated strains, *Virus Res.* **74**:99-110.
13. Gao, Z. Q., Guo, X. and Yang, H. C., 2004, Genomic characterization of two Chinese isolates of porcine respiratory and reproductive syndrome virus, *Arch. Virol.* **149**:1341-1351.
14. Allende, R., Lewis, T. L., Lu, Z., Rock, D. L., Kutish, G. F., Ali, A., Doster, A. R., and Osorio, F. A., 1999, North American and European porcine reproductive and respiratory syndrome viruses differ in non-structural protein coding regions, *J. Gen. Virol.* **80**:307-315.
15. Ziebuhr, J., Snijder, E. J., and Gorbalenya, A. E., 2000, Virus-encoded proteinases and proteolytic processing in the *Nidovirales, J. Gen. Virol.* **81**:853-879.
16. Shen, S., Kwang, J., Liu, W., and Liu, D. X., 2000, Determination of the complete nucleotide sequence of a vaccine strain of porcine reproductive and respiratory syndrome virus and identification of the Nsp2 gene with a unique insertion, *Arch. Virol.* **145**:871-883.
17. Fang, Y., Kim, D. Y., Ropp, S., Steen, P., Christopher-Hennings, J., Nelson, E. A., and Rowland, R. R., 2004, Heterogeneity in Nsp2 of European-like porcine reproductive and respiratory syndrome viruses isolated in the United States, *Virus Res.* **100**:229-235.
18. Oleksiewicz, M. B., Botner, A., Toft, P., Normann, P., and Storgaard, T., 2001, Epitope mapping porcine reproductive and respiratory syndrome virus by phage display: the nsp2 fragment of the replicase polyprotein contains a cluster of B-cell epitopes, *J. Virol.* **75**:3277-3290.
19. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S., 2004, Improved prediction of signal peptides: SignalP 3.0, *J. Mol. Biol.* **340**:783-795.
20. Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M. D., Durbin, R., Falquet, L., Fleischmann, W., Gouzy, J., Hermjakob, H., Hulo, N., Jonassen, I., Kahn, D., Kanapin, A., Karavidopoulou, Y., Lopez, R., Marx, B., Mulder, N. J., Oinn, T. M., Pagni, M., Servant, F., Sigrist, C. J., and Zdobnov, E. M., 2001, The InterPro database, an integrated documentation resource for protein families, domains and functional sites, *Nucleic Acids Res.*, **29**:37-34.