Comparison of the 5′ Leader Sequences of North American Isolates of Reference and Field Strains of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV)

CORRINA TAN,1* LEO CHANG,2 SHUO SHEN,1 DING XIANG LIU1 & JIMMY KWANG1†

1Institute of Molecular Agrobiology, 1 Research Link, National University of Singapore, Singapore 117604
2Besta Pte Ltd, Block 31, #07–30, Commonwealth Lane, Singapore 149553
E-mail: jimmykwang@ima.org.sg

Accepted September 7, 2000

Abstract. The 5′ leader is documented to be an important regulatory element in many (+) ssRNA virus genome. To understand the significance of the 5′ leader RNA of PRRSV, we determined the complete leader sequences of fifteen different North American strains of PRRSV and predicted their secondary structures. Viruses analysed included three reference strains and nine field strains originating from different geographic locations. To further examine the leader region, one of the field strains was adapted to grow in tissue culture, and three clones were isolated. We also predicted the secondary structures of two European strains based on their published sequences. The predicted RNA secondary structures of the leader sequences suggested the existence of three conserved domains formed by the 5′ region of the leader among the North American strains, two of which were conserved in the European strains. A variable structural domain was predicted from the 3′ region of the leader sequences of the North American strains, where all tissue culture-adapted isolates were characterized by a stem-loop while field isolates were characterized by an internal bulge within the stem-loop.

Key words: 5′ leader region, 5′ non-coding region, porcine reproductive and respiratory syndrome virus, 5′ RACE, nucleotide sequence, RNA secondary structure prediction

Introduction

Porcine reproductive and respiratory syndrome virus (PRRSV) is an important pathogen that causes reproductive failure in breeding swine and respiratory tract illness in piglets (1), one of the most economically significant diseases of swine herds worldwide. PRRSV is a member of the family Arterividae in the order Nidovirales, together with lactate dehydrogenase-elevating virus (LDV), equine arteritis virus (EAV), and simian hemorrhagic fever virus (SHFV). It is a small, enveloped virus with a positive-sense single-stranded RNA genome. The genome is 5′ capped and 3′ polyadenylated and contains two large replicase open reading frames (ORF1a and 1b), and several smaller ORFs corresponding to the structural genes (2,3).

In infected cells, arteriviruses produce a nested set of six to eight major coterminous subgenomic mRNAs (sgmRNAs), each thought to express only the relative 5′ terminal ORF. These sgmRNAs all have leader sequence derived from the genome joined at specific leader-body junction sites at their 5′ ends (2–6). The 5′ leader had been documented to be an important regulatory element in many viruses. For the closely related coronavirus, the leader regulates viral genome replication, sgmRNA transcription, sgmRNA translation and viral persistence in vitro (7–10). Furthermore, leader sequences of other viruses, such as picornaviruses, echoviruses and Venezualan equine encephalitis virus contain attenuation/virulence determinants (11–14).
To initially examine the significance of the leader region in sgRNA transcription and translation, we report the complete leader sequences and their predicted RNA secondary structures for different North American isolates of PRRSV. Among these isolates were three reference strains and nine field strains originating from different geographic locations. In addition, included in this study were three clones of a field strain adapted for growth in tissue culture. Results demonstrate three conserved structural domains formed by the 5' region of the leader and a variable structural domain formed by the 3' region of the leader.

Materials and Methods

Viruses and Cells

The PRRSV strains used in this study are listed in Table 1. Reference strains VR2332 and Schering-Plough, and field strains NVSL, IOWA and 12068-96 were obtained from Dr. Fernando A. Osorio, Department of Veterinary and Biomedical Sciences, University of Nebraska-Lincoln, U.S.A. Sub-strains JK100-1, JK100-2 and JK100-3 were tissue culture-adapted populations exhibiting distinct CPE morphology generated from field strain JK100. At passage 25, strain JK100 virus suspension was serially diluted to obtain one subclone per mixture, and after propagation for an additional 5 passages, RNA was extracted from each subpopulation. All viruses were passaged in MARC-145 cells, a subclone of the monkey kidney cell line MA-104 (15). MARC-145 cells were maintained in complete DMEM medium (GIBCO BRL, U.S.A.) supplemented with 10% new-born calf serum and infected at low multiplicity of infection (0.1 pfu/cell) as previously described (15).

Partial Purification of Viruses and RNA Extraction

Virus-infected cells were frozen and thawed three times and cell debris was removed by centrifugation at 5000 rpm for 10 min (Beckman JA-25.50) at 4°C. Virus was harvested from supernatant by ultracentrifugation through a 20% sucrose cushion (in TNE buffer) at 150,000 g (Beckman SW28) for 2 h at 4°C. The pellet was suspended in TNC buffer (20 mM Tris-Cl, pH7.4, 100 mM NaCl, 2 mM CaCl2) and stored at −80°C. Viral genomic RNA was extracted using the TRIzol reagent (Life Technologies, U.S.A.) according to supplier’s instructions.

5' Rapid Amplification of cDNA Ends (5' RACE)

The 5' end sequence of viral genomic RNA was amplified using the 5'/3' RACE kit (Boehringer Mannheim, Germany) according to manufacturer’s protocol with some modification. Briefly, first strand cDNA was synthesized using AMV reverse transcriptase with viral-specific reverse primer 1 (5'-atggtgtcagtagcttg-3') based on a consensus sequence of VR2332 (4,5) and 16244B (16). Upon purification of cDNA, either a poly(A)-tail or poly(C)-tail was added with terminal transferase. The respective tailed cDNA was then amplified with a mix of Taq and Pwo DNA polymerases (Boehringer Mannheim, Germany) with viral-specific reverse primer 2 (5'-acacgtgcaccgatacg-3') based on the consensus sequence of VR2332 and 16244B, and either a oligo-dT anchor primer (5'-ggccacgcgtcgtagctggg-3') or oligo-dC anchor primer (5'-ggccacgcgtgctagctggg-3') (where v̂ is non tailed nucleotide) depending on the tailed cDNA used. PCR cycling conditions used were 94°C for 5 min, followed by 30 cycles of 94°C for 30 sec, 55°C for 30 sec and 72°C for 1 min, and finally 72°C for 7 min. Purified PCR products were sequenced. Each sequence was confirmed with PCR products derived from four separately generated cDNAs to prevent identification of minor quasispecies.

Primer Extension

Primer extension (17) was performed on viral genomic RNA extracted from strains VR2332, BI and SP. The viral RNA was hybridized at 65°C for 1.5 hr to reverse primer 2 which was end-labeled with (γ-32P)-ATP (10,000 cpm) and purified with the Nuctrap Probe Kit (Stratagene, U.S.A.). The labeled primer was extended with the Expand reverse transcriptase (Boehringer Mannheim, Germany) at 55°C for 1 hr prior to treatment with DNase-free RNase. Primer extension products were analyzed in
sequencing gels together with sequenced samples of 5' RACE products to determine the 5'-terminal nucleotide.

**DNA Sequencing and Computer Analyzes**

PCR amplified products were automatically sequenced using viral-specific reverse primer 2 and forward primer 3 (5'-ggcaggtgtttgctctat-3'), based on 16244B sequence with the DYE Terminator Cycle Sequencing kit (Perkin Elmer, U.S.A.) (18). Nucleotide chromatogram traces were analyzed with the Sequencher version 3.0 program. Sequence alignments were conducted using the Clustal programme included in the LASERGENE package (DNASTAR, U.S.A.). RNA secondary structures were predicted with the computer program Mfold at 37°C (19). The GenBank accession numbers of the field strains used in this study are listed in Table 1.

**Results**

**Determination and Analysis of Primary Sequences**

The 5' RACE products were used to determine sequence of the 5' leader regions. The 5' RACE was performed using both poly(A)-tailed and poly(C)-tailed cDNAs. This helped identify whether the 5' terminal nucleotide of the RNA was a uracil or guanine, in which case the 3' end of the first-strand cDNA would terminate with either an adenine or a cytosine, respectively. This method identified an adenine as the extreme 5' terminal nucleotide as was confirmed by the primer extension analyses (results not shown).

A multiple alignment of the PRRSV leader sequences is shown in Fig. 1. Our sequence data of the VR2332 leader differed from that reported by (5) in that the additional uracil at the extreme 5' end was
not observed in either our primer extension and modified 5' RACE analyses. The leader regions of the different isolates were either 189 or 190 nucleotides in length. Pair-wise comparisons of the leader sequences identified high degree of identity (between 86.8% and 99.5%) among them, indicating that the field isolates originating from different geographic locations represented distinct subspecies of the North American genotype. Detailed examination revealed that nucleotide substitutions within these leaders are common in RNA viruses. For instance, G/A substitutions at positions 31 and 74, and U/C substitutions at positions 103, 121 and 130 occurred with regularity. The additional nucleotide at position 128 observed in some PRRSV leaders in all cases represented a guanine. The hexanucleotide 5'-UUACC-3' defining the leader-body junction sequence was conserved. Two AUG codons, at positions 1 and 23, with in-frame opal (UGA) terminator codons at positions 100 and 53, respectively, were conserved among the reference and field strains. However, the sequence context around the two PRRSV intra-leader ORFs does not conform to the Kozak consensus for a functional initiation codon, with that of the presumptive AUG at the start codon of ORF1a emerges as the favored sequence. This suggest the possibility that ribosome can bypass the non-functional AUGs and initiate translation at the authentic initiator codon.

**Computer Prediction and Analysis of the Potential RNA Secondary Structures**

RNA secondary structures based on the PRRSV primary leader sequences were predicted using the Mfold programme. Previous studies with the Mfold programme have demonstrated that well-determined stable structures domains (i.e. features for which few competing alternative structures form in sub-optimum foldings) are predicted more reliably than those that are poorly determined. Leader RNA structures were therefore predicted with well-determined domains formed in optimum folding identified by the Mfold energy dot plot analysis that evaluate the number of possible base-pairing for each nucleotide.

The putative folding structures based on the PRRSV leader sequences are shown in Fig. 2 and Fig. 3. Three of the predicted stem-loop structures, designated 1, 2 and 3, in the 5' half of the leader region from nucleotides 2–110 were formed in all isolates. Stem-loops 1, 2 and 3 were formed by nucleotides 2–45, 46–70 and 78–106, respectively. The 3' half of the leader sequence, however formed a fourth variable domain, which for different PRRSV isolates had one of three alternative optimal foldings, designated 4a, 4b and 4c, formed by nucleotides 113–184.

All of the leader structural interactions were formed either by the Watson-Crick type base-pairing, or G:U wobble pairs, and almost all nucleotide substitutions maintained these two types of interactions. Nucleotide mutations in the leader sequences were selected to maintain the general form of the secondary structures. The majority of the nucleotide substitutions occurred in the loop regions and did not affect base-pairing. Most sequence changes in stems did not disturb the base-pairing scheme. Dual nucleotide changes, such that alteration of the primary sequence at two residues in a stem were compensatory to retain a structure, were observed for instance at positions 123 and 138 (T : A to C : G), and 16 and 29 (U : G to G : C) for strain c653, and positions 85 and 98 (A : T to G : C) for strain 11. Nucleotides in bulge area were catered for base changes to maintain stem stability, for instance, substitution of cytosine by uracil at position 68 of strain JK101 selected for the adjacent adenine within the bulge (results not shown).

**Variable Structures Differentiate Field Isolates From Cell Culture-Adapted Isolates**

Field isolates exhibited different virus titres in vitro, with different numbers of passages required for cell culture-adaptation. For example, isolate 12068-96 generated extensive CPE and high virus titres at passage 8, while JK100 remained very poorly adapted and generated minimal CPE and low virus titres even at passage 10. Field isolates that had been adapted to the tissue culture system were NVSL (10^6 pfu/ml), IOWA (10^5 pfu/ml) and 12068-96 (10^5 pfu/ml), while JK100 (10^4 pfu/ml), JK101 (10^2 pfu/ml), 11 (10^2 pfu/ml) and the Taiwanese strains (all three isolates exhibited at 10^4 pfu/ml virus titres) remained poorly-adapted at the time RNA was isolated.

PRRSV leader sequence predicted a variable domain (Fig. 3). Domain 4a was unique among field strains NVSL, IOWA, 12068-96 and reference strains VR2332 (10^6 pfu/ml), BI (10^7 pfu/ml), SP (10^7 pfu/ml). Domain 4b was unique among field strains JK100, JK101 and 11. Domain 4c was unique among the Taiwanese field strains (p319, c635, p653). These
Fig. 1. Alignment of the full length PRRSV leader nucleotide sequences against North American reference strain VR2332. Strains used for alignment (description in Table 1) were shown on the left while nucleotide positions were shown to the right of the alignment. Nucleotides identical to VR2332 were indicated by dots. Substitutions were indicated by the nucleotide. Deletion was indicated by dashes. Asterisks identified the leader-body junction.
Fig. 2. Predicted RNA structures of the North American reference and field strains leaders. Conformation ‘a’ is represented by predicted structure of VR2332. Conformation ‘b’ is represented by predicted structure of JK101. Conformation ‘c’ is represented by predicted structure of p319. All three conformations (a, b, c) contain structurally conserved domains 1, 2 and 3 formed by nucleotides 2–45, 46–70 and 78–106 respectively. Conformations a, b and c represents the structurally variable domains 4a, 4b and 4c respectively (boxed), formed by nucleotides 113–184. Domain 4a was unique among VR2332, BI, SP, NVSL, IOWA, 12068-96, JK100-1, JK100-2 and JK100-3. Domain 4b was unique among JK100, JK101 and 11. Domain 4c was unique among the Taiwanese strain p319, c635 and p653. Internal loop that differentiate the tissue culture adapted virus from the field strain were indicated by arrow. The minimum free energy (dG) for these conformations ranged from $-62.21$ to $-72.31$. 

Tan et al.
Fig. 3. (a) Predicted RNA structure of the leader sequence of field strain JK100 and its tissue culture-adapted clones (JK100-1, JK100-2 and JK100-3). The structural switch at domain 4 (boxed) suggested that the loss of the internal bulge as indicated by arrow corresponds with cell culture-adaptation. (b) Reference strains VR2332, BI, SP, and tissue culture-adapted strains NVSL, IOWA and 12068-96 were characterized by domain 4a which lack an internal bulge within the stem-loop. Poorly tissue culture-adapted strains JK101, 11, p319, p653 and c635 were characterized by domain 4b and 4c which possessed an internal bulge within the stem-loop.
results suggest that domain 4 may play some role in differentiating the field isolates that were poorly cell culture-adapted from the cell culture-adapted isolates, with domain 4a characterizing the cell culture-adapted isolates, while domain 4b and 4c characterizing those that were poorly tissue culture-adapted.

**Absence of the Internal Bulge Within Variable Domain 4 Indicates Cell Culture-Adaptation**

To examine whether structural changes in domain 4 occur upon cell culture-adaptation, the leader sequence of field strain JK100 was analyzed prior to and after being adapted in vitro. After repeated passages, strain JK100 exhibited 10⁸ pfu/ml virus titre and three cell culture-adapted clones exhibiting unique CPE morphology were isolated. The putative leader structure of JK100 and its clones maintained domain 1, 2 and 3, but not domain 4 (Fig. 3). JK100 exhibited a domain 4b structure and its tissue culture-adapted clones all exhibited a domain 4a structure, despite the different CPE morphologies. The structural switch from domain 4b (JK100) to domain 4a (JK100-1, JK100-2, JK100-3) suggests that the loss of the internal bulge within the stem-loop sub-domain of domain 4 corresponds with cell culture-adaptation. This is consistent with the other PRRSV leaders, which all field isolates poorly adapted to tissue culture either exhibited domain 4b or 4c, both of which possess a stem-loop sub-domain with an internal bulge. All cell culture-adapted and reference strains exhibit domain 4a, which lacks an internal bulge within the stem-loop sub-domain (Fig. 3).

**Discussion**

In this study, the complete leader primary sequences of fifteen PRRSV strains were determined. For each strain, RNA secondary structure was predicted from the sequence data. The viruses analyses included ATCC VR2332, commercially available vaccine strains SP and BI, nine field isolates originating from different geographic locations, and three tissue culture sub-strains, each of which exhibited a distinct CPE morphology to the progenitor field strain JK100.

Pair-wise alignment of the leader sequences of the field strains to those of the European strains Lelystad (23) and 111/92 (5), demonstrated that, in contrast to the high conservation to the North American reference sequences, field strains leaders were approximately 30 nucleotides shorter and only 33% identical to those of the European strains. Several groups using or having used the modified live vaccine observed the emergence of new PRRSV sub-species in pigs, with virulent strains arising in some instances upon gaining entry into susceptible animals and spreading to non vaccinated herds (24–28). The North American prototype-based vaccines (SP and BI) have been most commonly used for PRRSV vaccination to date and sequence conservation suggests that the field strains from various geographic regions used in this study represent different sub-species of the North American genotype.

Notwithstanding the distant leader sequences between the North American and European PRRSV, the leader-body junction were observed to be conserved among these two PRRSV types. The leader-body junction sequence, also known as the transcription-regulating sequence (TRS), was similar to the TRS which precedes every subgenomic mRNA (4,5), and to all arterivirus consensus sequences described to date (29–31). This conserved sequence was required for the base-pairing interaction between the leader TRS and the body TRS in a site specific manner (32), which (6) demonstrated to be essential for mRNA transcription.

Besides the well-documented function of the TRS in mRNA transcription, the function of most length of the PRRSV leader sequence remains incomprehensive in comparison to the progress made with other viruses. In this study, RNA secondary structure predictions revealed that the sequence divergence in the 5′ region of the leaders of the different PRRSV isolates was constrained to maintain the overall stability of three stem-loop structures (domains 1, 2 and 3). Putative RNA structures of European Lelystad and 111/92 PRRSV leaders revealed that two of these North American stem-loop structures (domains 2 and 3) are maintained in the European stains (results not shown). Structural conservation indicates that most non-conforming or non-compensatory mutational changes will not be tolerated and suggests functional roles for the 5′ region of the leaders of PRRSV will be maintained during evolution. The significance of secondary structures in viral leader sequences for initiation of mRNA translation has been well documented (11,13,33,34). In addition, many of these studies show that variations in leader structural elements essential for translation initiation result in
structures predicted in the 5′ leader sequence, in contrast to the conserved secondary leader (domain 4). Variability in this portion of the leader sequence, in contrast to the conserved secondary structures predicted in the 5′ region, suggests a lesser role for this region of the leader in PRRSV regulation. Nevertheless, such RNA structural variation within the 5′ region of the leaders of PRRSV characterizes the cell culture-adapted isolates from the field isolates that are poorly tissue culture-adapted. The stability of domain 4 is altered, in that the internal bulge within the stem-loop sub-domain is lost upon tissue culture-adaptation. Structural changes in the 5′ region of the leader may be a prerequisite for PRRSV quasispecies best adapted to tissue culture, although we cannot discount the involvement of other sequence difference.

Acknowledgments

We thank Dr. Fernando A. Osorio of the University of Nebraska, Department of Veterinary and Biomedical Sciences for providing the different strains of PRRSV.

References

1. Rossow K.D., Vet Pathol 35, 1–20, 1998.
2. de Vries A.A.F., Horzinek M.C., Rottier P.J.M., and de Groot R.J., Seminars in Virology 8, 33–47, 1997.
3. Snijder E.J. and Meulenberg J.J.M., J. General Virology 79, 961–979, 1998.
4. Nelsen C.J., Murtaugh M.P., and Faaberg K.S., J Virology 73, 270–280, 1999.
5. Oleksiewucz M.B., Bettnaer A., Nielsen J., and Storgaard T., Arch Virology 144, 981–987, 1999.
6. Marle G.V., Dobbe J.C., Gultyaev A.P., Luytjes W., Spaan W.J.M., and Snijder E.J., Proceedings of the National Academy of Sciences USA 96, 12056–12061, 1999.
7. Chen W. and Baric R.S., J Virology 69, 7529–7540, 1995.
8. Hofmann M.A., Senanayake S.D., and Brain D.A., Proceedings of the National Academy of Sciences USA 90, 11733–11737, 1993.
9. Tahara S.M., Dietlin T.A., Bergmann C.C., Nelson G.W., Kyuwa S., Anthony R.P., and Stohsman S.A., Virology 202, 621–630, 1994.
10. Zhang X., Liao C.L., and Liu M.M., J Virology 68, 4738–4746, 1994.
11. Hallen A.A., Stewart S.R., and Semler B.L., J Virology 70, 1467–1474, 1996.
12. Kinney R.M., Chang G.J., Tsuchiya K.R., Sneider J.M., Roehrig J.T., Woodward T.M., and Trent D.W., J Virology 67, 1269–1277, 1993.
13. Macadam A.J., Stone D.M., Almond J.W., and Minor P.D., Trends in Microbiology 2, 449–454, 1994.
14. Romero J.R. and Rothbart H.A., J Virology 69, 1370–1375, 1995.
15. Kim H.S., Kwang J., Yoon I.J., Joo H.S., and Freg M.L., Arch Virology 135, 477–483, 1993.
16. Allende R., Lewis T.L., Lu Z., Rock D.L., Kutsch G.F., Ali A., Doster A.R., and Osorio F.A., J General Virology 80, 307–315, 1999.
17. Sambrook J., Fritsch E.F., and Maniatis T., Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbour, New York, 1989.
18. Sanger F., Nicklen S., and Coulson A.R., Proceedings of the National Academy of Sciences USA 74, 5463–5467, 1977.
19. Zuker M., Mathews D.H., and Turner D.H., J Molecular Biology 288, 911–940, 1998.
20. Domingo E., Martinez-Salas E., Sobrino F., de la Torre J.C., Portela A., Ortin J., Lopez-Galindez C., Perez-Brena P., Villanueva N., and Najera R., Gene 40, 1–8, 1985.
21. Kozak M., Gene 234, 187–208, 1999.
22. Zuker M. and Jacobson A.B., Nucleic Acids Research 23, 2791–2798, 1995.
23. Meulenberg J.J., Bos-de Ruijter J.N., van de Graaf R., Wensvoort G., and Moormann R.J., J Virology 72, 380–387, 1998.
24. Andreyev V.G., Wesley R.D., Mengeling W.L., Vorwald A.C., and Lager K.M., Arch Virology 142, 993–1001, 1997.
25. Bettnaer A., Strandbygaard B., Sørensen K.J., Have P., Madsen K.G., Smeegaard M.E., and Alexandersen S., Vet Rec 141, 497–499, 1997.
26. Madsen K.G., Hansen C.M., Madsen E.S., Strandbygaard B., Bettnaer A., and Sørensen K.J., Arch Virol 143, 1683–1700, 1998.
27. Rossow K.D., Shivers J.L., Yeske P.E., Polson D.D., Rowland R.R., Lawson S.W., Murtaugh M.P., Nelson E.A., and Collins J.E., Vet Rec 144, 444–448, 1999.
28. Rowland R.R., Steffen M., Ackerman T., and Benfield D.A., Virology 259, 262–266, 1999.
29. Chen Z., Kuo L., Rowland R.R.R., Even C., Faaberg K.S., and Plagemann P.G.W., J General Virology 74, 643–660, 1993.
30. den Boon J.A., Kleijnen M.F., Spaan W.J.M., and Snijder E.J., J Virology 70, 4291–4298, 1996.
31. Godeny E.K., de Vries A.A.F., Wang X.C., Smith S.L., and de Groot R.J., J Virology 72, 862–867, 1998.
32. Lai M.M., Adv. Exp. Med. Biol. 380, 463–471, 1995.
33. Duke G.M., Hofmann M.A., and Palmenberg A.C., J Virology 66, 1602–1609, 1992.
34. Fukushima S., Katayama K., Kurilahara C., Ishiyama N., Hoshino F.B., Ando T., and Oya A., Biochem Biophys Res Commun 199, 425–432, 1994.
35. Shen S., Kwang J., Liu W., and Liu D.X., Arch Virology 145, 871–883, 2000.