Characteristics of Major Structural Protein Coding Gene and Leader-Body Sequence in Subgenomic mRNA of Porcine Reproductive and Respiratory Syndrome Virus Isolated in Japan

Akito SAITO, Toru KANNO[1)], Yosuke MURAKAMI[1)], Masatake MURAMATSU, and Shigeo YAMAGUCHI[1)]*
National Veterinary Assay Laboratory, Kokubunji, Tokyo 185, and [1)]Exotic Diseases Research Division, National Institute of Animal Health, Kodaira, Tokyo 187, Japan
(Received 20 June 1995/Accepted 28 November 1995)

ABSTRACT. Nucleotide sequence, 1713nt in length, of porcine reproductive and respiratory syndrome virus (PRRSV) isolated in Japan was determined. The sequence encompassed 3 overlapping open reading frames (ORFs), ORF5 to ORF7. These ORFs encodes major structural proteins of PRRSV. The deduced amino acid sequence of each ORF showed higher than 87.5% identity with an American isolate, and lower (54.6 to 80.5%) identity with an European isolate. This result supported a previous report about antigenic characteristics of the EDRD-1 strain. Leader-body junction sequence in subgenomic mRNA of the EDRD-1 strain was determined by sequencing cDNA clones of subgenomic RNAs. A common sequence motif of 5 nucleotide, represented by UA(A/G)CC, was identified as the junction sequence. — KEY WORDS: junction sequence, nucleotide sequence, porcine reproductive and respiratory syndrome.
J. Vet. Med. Sci. 58(4): 377–380, 1996

Porcine reproductive and respiratory syndrome (PRRS) was appeared in the United States in 1987 [13]. Thereafter, the syndrome, also referred as mystery swine disease, swine infertility and respiratory syndrome, porcine epidemic abortion and respiratory syndrome, was recognized in many countries [4, 24, 29]. The disease is characterized by severe reproductive failure and respiratory disease in pigs. The causative agent has been isolated in Europe [29] and North America [7] and both isolates were identified as small spherical enveloped virus [3, 5]. Whereas, antigenic diversity was detected between the isolates [23, 30]. Molecular characteristics of both the European (Lelystad virus; LV) and the North American isolates (VR2385, and IAF-exp91 strain) have been reported [18–20]. The viral genome is a positive-stranded, polyadenylated of about 15 Kb, and identified 8 open reading frames (ORFs) [20]. The virus infected cell contains a 3'-coterminal nested set of six subgenomic mRNAs [19, 20, 27]. These mRNAs containes a common leader sequence derived from the 5' non-coding region of the genomic RNA. And a common sequence has been identified at the junction site of leader and body of each subgenomic mRNA [21].

From the similarity of the genomic organization and subgenomic mRNA formation, PRRS virus (PRRSV) is classified to arteriviridae [20, 25] which comprised equine arteritis virus (EAV) [10] and lactate dehydorogenase-elevating virus (LDV) [15] and simian hemorrhagic fever virus.

In Japan, the syndrome recognized from 1989 [28] and isolated the virus in 1994 [22, 26]. One of the Japanese isolates, EDRD-1 strain, has been characterized to have the similar antigenicity to an American isolate but different from European isolate of LV [22].

In this report, we determined genomic and subgenomic mRNA sequences of the Japanese isolate and characterized the 3'-terminal structure protein coding region and leader-body junction sequences.

The virus, EDRD-1 strain, used in this report was isolated in Japan from lung homogenate of a fattening pig with severe chronic respiratory disorder using swine alveolar macrophage cell cultures, and identified as PRRSV serologically [22]. MARC-145 cell line was used for further replication of the isolate. The MARC-145 cell line was established from MA-104 cell line by cell cloning, and identified as highly susceptible cell line to PRRSV replication [14]. After plaque cloned and passed 10th in MARC-145 cell, the virus inoculated onto the MARC-145 cells at a m.o.i. of 0.1. When the cell started to show cytopathic effect, the culture medium and the cells were harvested separately for virus genome and subgenomic RNA extraction. The viral genome was purified by the method as described previously [27] with modification. Briefly, the culture medium was clarified by low-speed centrifugation and the virus was precipitated by ultracentrifugation at 40,000 r.p.m. (Beckman 45Ti rotor) for 2 hr. The virus pellet was resuspended in TNE buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1 mM EDTA), added sodium dodecyl sulfate to make a 1% concentration, extracted with phenol/chloroform, and ethanol precipitated to obtain partially purified genomic RNA. For preparation of subgenomic RNA, the virus infected MARC-145 cells washed with PBS (pH 7.0) were collected with rubber policeman, pelleted by low centrifugation, and the cells were destroyed with 1% TritonX-100 solution. Polyadenylated RNA was purified with Dynabeads oligo-dT (Dynal Co.) according to the instruction manual and used as a subgenomic RNA sample.

To obtain the 3'-terminal genomic or subgenomic cDNA clones, cDNA libraries were constructed by cDNA synthesis kit (Amersham International plc). Briefly, first-strand cDNA was synthesized with oligo (dT) primers by reverse transcription (RT). Second-strand cDNA was synthesized with RNase H and DNA polymerase I, then blunted with T4

* CORRESPONDENCE TO : Dr. YAMAGUCHI, S., Second Research Division, National Institute of Animal Health, Tsukuba, Ibaraki 305, Japan.

DNA polymerase. The blunted cDNA was ligated to *Eco*RI adaptors, phosphorylated with T4 polynucleotide kinase, then ligated with dephosphorylated *Eco*RI-digested lambda phage vector, λgt 10. The resulting recombinant phage vector was introduced in *E. coli* (NM514 strain) by employing *in vitro* packaging method. The library of genomic and subgenomic cDNA had a titer of $10^{7.1}$ p.f.u./m*l* and $10^{6.7}$ p.f.u./m*l* respectively. To identify the PRRSV specific clones in the cDNA library, plaque hybridization was applied. Probe for the hybridization was a digoxigenin-labeled cDNA fragment, which was prepared from LV cDNA clone containing whole ORF7 and 3'-untranslated region. The plaques on the *E. coli* plates were lifted onto nylon membranes. The membranes were then denatured with 0.5 M NaOH and 1.5 M NaCl, neutralized with 0.5 M Tris-HCl pH 7.2 and 1.5 M NaCl, washed with 2 × SSC and dried. Hybridization was performed at 68°C overnight, according to the instruction manual of the kit (Borhringer Mannheim Co.). Hybridization positive plaques were picked and grew to purify DNA, then digested with *Eco*RI to excise the insert cDNA. The *Eco*RI-digested phage DNA were analyzed by southern hybridization with the digoxigenin labeled probe after agarose gel electrophoresis. Excised cDNA fragments were clearly identified in each clone. The largest excised genomic cDNA (1.7 Kb) and several subgenomic cDNAs were subcloned into pBluescript II SK(+) phagemid. The genomic cDNA clone was subjected to unidirectional deletion using exonuclease III and mung bean nuclease, and then double stranded DNA was purified with nucleic acid purification tip (Quiagen Co.). Sequencing have been done with an Automated Laser Fluorescent DNA sequencer from Pharmacia LKB. The AutoCycle sequencing kit (Pharmacia Co.) was used essentially according to the procedure described in the kit. The sequence analysis programs, GENETYX (Software Development Co.) and PCGENE (Intelligenitics, Inc.), were used to assemble and analyze the sequence data.

Nucleotide sequence, 1713nt in length, of the 3'-terminal of EDRD-1 strain genome [excluding the poly-(A) tail] was determined. The nucleotide sequence data reported in this paper have been deposited with the DDBJ database and assigned the accession number D45852. The sequence encompassed 3 overlapping ORFs, with the potential to encode polypeptides with predicted M_rs of 22.2, 19.1 and 13.6 respectively (Table 1), and determined as ORF5, ORF6 and ORF7 from homology with nucleotide sequence of PRRSV [19, 20]. Gene organization such as position of ORFs and number of nucleotides in ORF5 to ORF7 of EDRD-1 strain were extremely identical to that of VR2385 strain of PRRSV. Table 2 shows identities of deduced amino acid sequence of each ORF and nucleotide sequence of 3'-untranslated region between EDRD-1 strain and other PRRSVs, LDV and EAV [11, 12, 18–20]. The deduced amino acid sequence showed higher than 87.5% identity with North American isolates (VR2385 and IAF-exp91 stains), and lower identity, 54.6 to 80.5%, with the LV of the European isolate. This result supports the serological results that the antigenicity of EDRD-1 strain was close to American isolate (46448 strain) but different to European isolate (LV) by immunofluorescence and immunoperoxidase monolayer assay [22]. Even the nucleotide sequence of LV had low identity, the hybridization probe bound to EDRD-1 strain cDNA clones in the condition used. Nucleotide identity of 3'-untranslated region was also high between EDRD-1 strain and North American isolates.

ORF5 and ORF6 protein possessed very hydrophobic domains, which possibly represented membrane-spanning fragments, at the N-terminal and central regions (data not shown), and probably encode viral membrane-associated protein. Three ORFs from 3'-terminal of genome of EAV were supposed to encode major structural protein, ORF5 for glycosilated envelope protein, ORF6 for membrane protein and ORF7 for nucleocapsid protein [9]. Hydrophilicity profile of ORF6 and ORF7 products of EDRD-1 strain showed similar pattern to corresponding products of other PRRSV, LDV and EAV [8, 11, 19, 20, 25]. Whereas ORF5 product of EDRD-1 strain showed more diverse pattern between EAV, but similar to other PRRSV and LDV. These

Table 1. Characteristics of deduced proteins of EDRD-1 strain of PRRSV

ORF	No. of amino acid	M_r (kDa)	Potential N-glycosylation sites
ORF5	200	22.2	3
ORF6	174	19.1	2
ORF7	123	13.6	0

Table 2. Amino acid and nucleotide sequence identities (%) between EDRD-1 strain and other PRRSVs, LDV and EAV

Compared viruses (strains)	Deduced amino acid sequence of EDRD-1 strain			Nucleotide sequence of EDRD-1 strain
	ORF5	ORF6	ORF7	3'UTR[a]
PRRSV (VR2385)	87.5	94.8	94.3	93.4
PRRSV (IAF-exp91)	NE[b]	NE	95.9	93.4
PRRSV (LV)	54.6	80.5	61.1	56.6
LDV (LDV-C)	46.3	49.7	51.2	38.4
EAV (Bucyrus)	24.5	26.3	28.7	30.3

a) 3'-untranslated region.
b) Not examined due to lack of sequence data.

findings suggest each product have a similar topology on the virion and functions. There were 3, 2 and 0 potential N-glycosylation sites in the product of ORF5, ORF6 and ORF7 protein respectively. ORF5 product supposed to encode, envelope protein, and may encode neutralization epitope [2], therefore hydrophilicity profile was more variable than other structural proteins.

To determine the leader-body junction sequence, subgenomic RNA was purified with Dynabeads oligo-dT from EDRD-1 strain infected cell lysate. Two cDNA clones were used for sequencing, and found these clones were identical and had the same nucleotide sequence as genomic cDNA from 3'-terminal poly(A) to 128nt upstream of initiation codon of ORF7. From this result the both clones were identified as subgenomic RNA7. Non-identical sequence of the subgenomic RNA7 to the genomic sequence, upper from 129nt upstream of the initiation codon, was considered as leader sequence. To amplify leader-body junction region, oligonucleotide primers were designed from the subgenomic RNA7 and the genomic sequence as shown in Fig.1. RT-PCR was applied for amplification of the junction region and used for sequence analysis.

RT-PCR products were cloned to pCRII plasmid (Invitrogen, Co.) and 2 to 6 cDNA clones were used to determine sequence of each region. Leader-body junction sequence was determined by alignment of genomic RNA sequence and subgenomic RNA sequence, which were determined with RT-PCR product amplified with primers indicated in Fig. 1. A common sequence motif of 5 nucleotides, represented by UA(A/G)CC, was identified in each genomic and subgenomic sequence. These results suggest the motif is a junction sequence (Table 3). The third sequence of the motif was "A" in subgenomic RNA6 and RNA7, while in subgenomic RNA5 was "G". The junction sequence for ORF5 to ORF7 of EDRD-1 strain share some similarity with those identified in the subgenomic RNAs of LV; (A/U)(C/U)AACC [21], EAV; UCAAC [10] and LDV; (A/C) (A/U)AACC [6]. The junction sequence motif of EDRD-1 strain and their location were completely conserved in VR2385 genomic sequence. Six cDNA clones were used for sequencing of subgenomic RNA6, and four had P1-P3 (sgRNA6)A and the others had P1-P3 (sgRNA6)B sequence. The junction sequence of subgenomic RNA6 had a possibility to be a longer sequence of CCCU(A/U)UAACC. Heterogeneity of junction sequence was also reported in LV and coronaviruses [17, 21]. These heterogeneity were detected only in upstream of the sequence motif and they were identical to body sequence but not leader sequence. Therefore it seemed difficult to explain the subgenomic RNA formation only by the unique leader primed transcription at motif sequence [1, 16]. Further study may be necessary to determine the mechanism of subgenomic RNA formation.

The results of sequence analysis of one of the Japanese isolates revealed origin of the prevalent virus related closely to the American isolate. Further investigation may be necessary to define whether only American type is prevalent in Japan or not.

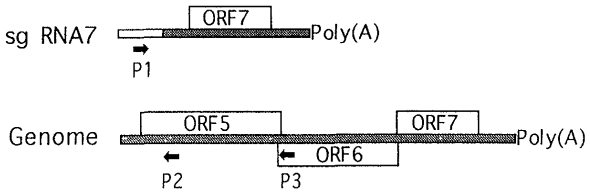

Fig. 1. Location of oligonucleotide primers for amplification of leader-body junction sequence of subgenomic RNAs. Position and direction of primers were indicated by arrow. P1 and P2 primers were used for subgenomic RNA5, and P1 and P3 were for subgenomic RNA6. Each primer designed from EDRD-1 strain sequence was the following. P1 (5'GGTCTGTCCCTAGCACCTTG3'), P2 (5'GACAAAACTCTCCACTGCCC3') and P3 (5'TATCATAATTGGCGTGTAGGTGA3'). Leader and body sequences are indicated by open and shaded box respectively.

Table 3. Sequence of the leader-body junction region of EDRD-1 strain subgenomic and genomic RNAs

RNA	Junction sequence[a]	Position[b]
P1-P2 (sgRNA5)[c]	UCUCCACCCUUAGCCUGUCUUUUUG	45
Genomic ORF5[d]	GCAACUGUUUUAGCCUGUCUUUUUG	45
P1-P3 (sgRNA6)A	CUCCACCCUUUAACCAAAGUUUCAG	23
P1-P3 (sgRNA6)B	CUCCACCCUAUAACCAAAGUUUCAG	23
Genomic ORF6	GCAACCCCUUUAACCAAAGUUUCAG	23
Subgenomic RNA7	CUCCACCCUUUAACCACGCAUUUCU	128
Genomic ORF7	CGGCAAGUGAUAACCACGCAUUUCU	128

a) Common junction sequence motif is indicated by shaded letter.
b) Distance of the first junction sequence motif "U" from each initiation codon of ORF.
c) RNA sequence of the junction region identified between indicated primers and predicted amplified subgenomic RNA is indicated in parenthesis.
d) Genomic RNA sequence of the junction region identified upstream of indicated ORF.

REFERENCES

1. Baker, S. C. and Lai, M. M. C. 1990. *EMBO J.* 9: 4173–4179.
2. Balasuriya, U. B. R., Rossitto, P. V., Demaula, C. D., and MacLachlan, N. J. 1993. *J. Gen.Virol.* 74: 2525–2529.
3. Benfield, D. A., Nelson, E., Collins, J. E., Harris, L., Goyal, S. M., Robison, D., Christianson, W. T., Morrison, R. B., Gorcyca, D., and Chladek, D. 1992. *J. Vet. Diagn. Invest.* 4: 127–133.
4. Bilodeau, R., Dea, S., Martineau, G. P., and Sauvageau, R. 1991. *Vet. Rec.* 129: 102–103.
5. Botner, A., Nielsen, J., and Billehansen, V. 1994. *Vet. Microbiol.* 40: 351–360.
6. Chen, Z. Y., Faaberg, K. S., and Plagemann, P. G. W. 1994. *J. Gen. Virol.* 75: 925–930.
7. Collins, J. E., Benfield, D. A., Christianson, W. T., Harris, L., Hennings, J. C., Shaw, D. P., Goyal, S. M., McCullough, S., Morrison, R. B., Joo, H. S., Gorcyca, D., and Chladek, D. 1992. *J. Vet. Diagn. Invest.* 4: 117–126.
8. Conzelmann, K. K. K., Visser, N., Van Woensel, P., and Thiel, H. J. 1993. *Virology* 193: 329–339.
9. de Vries, A. A. F., Chirnside, E. D., Horzinek, M. C., and Rottier, P. J. M. 1992. *J. Virol.* 66: 6294–6303.
10. de Vries, A. A. F., Chirnside, E. D., Bredenbeek, P. J., Gravestein, L. A., Horzinek, M. C., and Spaan, W. J. M. 1990. *Nucleic Acids Res.* 18: 3241–3247.
11. den Boon, J. A., Snijder, E. J., Chirnside, E. D., Devries, A. A. F., Horzinek, M. C., and Spaan, W. J. M. 1991. *J. Virol.* 65: 2910–2920.
12. Godeny, E. K., Chen, L., Kumar, S. N., Methven, S. L., Koonin, E. V., and Brinton, M. A. 1993. *Virology* 194: 585–596.
13. Keffaber, K. K. 1989. *Am. Assoc. Swine Prac. Newslett.* 1: 1–9.
14. Kim, H. S., Kwang, J., Yoon, I. J., Joo, H. S., and Frey, M. L. 1993. *Arch. Virol.* 133: 477–483.
15. Kuo, L. L., Harty, J. T., Erickson, L., Palmer, G. A., and Plagemann, P. G. W. 1991. *J. Virol.* 65: 5118–5123.
16. Makino, S., Stohlman, S. A., and Lai, M. M. C. 1986. *Proc. Natl. Acad. Sci. U.S.A.* 83: 4202–4208.
17. Makino, S., Soe, L. H., Shieh, C., and Lai, M. M. C. 1988. *J. Virol.* 62: 3870–3873.
18. Mardassi, H., Mounir, S., and Dea, S. 1994. *J. Gen. Virol.* 75: 681–685.
19. Meng, X. J., Paul, P. S., and Halbur, P. G. 1994. *J. Gen. Virol.* 75: 1795–1801.
20. Meulenberg, J. J. M., Hulst, M. M., de Meijer, E. J., Moonen, P. L. J. M., den Besten, A., de Kluyver, E. P., Wensvoort, G., and Moormann, R. J. M. 1993. *Virology* 192: 62–72.
21. Meulenberg, J. J. M., de Meijer, E. J., and Moormann, R. J. M. 1993. *J. Gen. Virol.* 74: 1697–1701.
22. Murakami, Y., Kato, A., Tsuda, T., Morozumi, T., Miura, Y., and Sugimura, T. 1994. *J. Vet. Med. Sci.* 56: 891–894.
23. Nelson, E. A., Christopher-Hennings, J., Drew, T., Wensvoort, G., Collins, J. E., and Benfield, D. A. 1993. *J. Clin. Microbiol.* 31: 3184–3189.
24. Paton, D. J., Brown, I. H., Edwards, S., and Wensvoort, G. 1991. *Vet. Rec.* 128: 617.
25. Plagemann, P. G. W. and Moennig, V. 1992. *Adv. Virus Res.* 41: 99–192.
26. Shimizu, M., Yamada, S., Murakami, Y., Morozumi, T., Kobayashi, H., Mitani, K., Ito, N., Kubo, M., Kimura, K., Kobayashi, M., Miura, Y., Yamamoto, T., and Watanabe, K. 1994. *J. Vet. Med. Sci.* 56: 389–391.
27. Van der Zeijst, B. A. M., Horzinek, M. C., and Moening, V. 1975. *Virology* 68: 418–425.
28. Watanabe, K. 1992. *Proc. Jpn. Pig Soc.* 20: 15–16 (in Japanese).
29. Wensvoort, G., Terpstra, C., Pol, J. M., ter Laak, E. A., Bloemraad, M., de Kluyver, E. P., Kragten, C., van Buiten, L., den Besten, A., Wagenaar, F., Broekhuijsen, J. M., Moonen, P. L. J. M., Zetstra, T., de Boer, E. A., Tibben, H. J., de Jong, M. F., van't Veld, P., Groenland, G. J. R., van Gennep, J. A., Voets, M. T., Verheijden, J. M. H., and Braamskamp, J. 1991. *Vet. Q.* 13: 121–130.
30. Wensvoort, G., de Kluyver, E. P., Luijtze, E. A., den Besten, A., Harris, L., Collins, J. E., Christianson, W. T., and Chladek, D. 1992. *J. Vet. Diagn. Invest.* 4: 134–138.