Genetic Analysis of Porcine Respiratory Coronavirus, an Attenuated Variant of Transmissible Gastroenteritis Virus

R. D. WESLEY,* R. D. WOODS, AND A. K. CHEUNG
National Animal Disease Center, U.S. Department of Agriculture, Agricultural Research Service, P.O. Box 70, Ames, Iowa 50010

Received 7 December 1990/Accepted 8 March 1991

The genome and transcriptional pattern of a newly identified respiratory variant of transmissible gastroenteritis virus were analyzed and compared with those of classical enterotropic transmissible gastroenteritis virus. The transcriptional patterns of the two viruses indicated that differences occurred in RNAs 1 and 2(S) and that RNA 3 was absent in the porcine respiratory coronavirus (PRCV) variant. The smaller RNA 2(S) of PRCV was due to a 681-nucleotide (nt) deletion after base 62 of the PRCV peplomer or spike (S) gene. The PRCV S gene still retained information for the 16-amino-acid signal peptide and the first 6 amino acid residues at the N terminus of the mature S protein, but the adjacent 227 residues were deleted. Two additional deletions (3 and 5 nt) were detected in the PRCV genome downstream of the S gene. The 3-nt deletion occurred in a noncoding region; however, the 5-nt deletion shortened the potential open reading frame A polypeptide from 72 to 53 amino acid residues. Significantly, a C-to-T substitution was detected in the last base position of the transcription recognition sequence upstream of open reading frame A, which rendered RNA 3 nondetectable in PRCV-infected cell cultures.

Transmissible gastroenteritis virus (TGEV), like other coronaviruses, is a pleomorphic enveloped virus that contains a large, positive-sense, single-stranded RNA genome (21). TGEV replicates via a leader RNA-primed mechanism that generates a nested set of subgenomic mRNAs sharing common 3′ polyadenylated termini and extending for different lengths in the 5′ direction (8). A common RNA leader sequence primes transcription initiation and thus is present at the 5′ end of the full-length genomic RNA and of each subgenomic mRNA. This leader sequence for TGEV is about 90 nucleotides (nt) long (12, 19). During transcription, the newly synthesized leader RNA oligonucleotide dissociates from the full-length negative-sense template and hybridizes with an intergenic recognition sequence that is immediately upstream of each open reading frame (ORF). Current data indicated that for TGEV an octameric sequence, TAC TAAC, is the recognition sequence upstream of each large ORF (24). Transcription proceeds by extension of the leader RNA primer sequence at the 3′ end. This leader RNA-primed transcription mechanism for synthesis of a nested set of mRNAs is a feature unique to coronaviruses. Negativesense subgenomic RNAs are produced during replication of TGEV and bovine coronavirus, and these function as an alternative pathway to generate subgenomic mRNAs (6, 20).

For TGEV, at least eight intracellular mRNAs are synthesized during virus replication (20, 24). Although each mRNA, except mRNA 8, is polycistronic in coding capacity, only the unique region of each mRNA is translationally active. Thus, each mRNA expresses only a single polypeptide. The three major structural proteins, the peplomer or spike (S) glycoprotein, the integral membrane (M) glycoprotein, and the nucleocapsid (N) protein, are expressed from mRNAs 2(S), 6(M), and 7(N), respectively. Full-length mRNA 1 may encode one or two replicases, and a major 14-kDa intracellular protein of unknown function is expressed by mRNA 8 (4, 25). Nucleotide sequence analysis revealed that the ORFs of mRNAs 3, 4, and 5 are capable of encoding polypeptides of 7.9, 27.7, and 9.3 kDa, respectively (1, 20, 24). The functions of these proteins have not been determined; however, mRNA 3 and, possibly, mRNA 4 are not necessary for virus replication (26). There is evidence to suggest that the polypeptides expressed by mRNAs 3 and 4 function in the pathogenesis of TGEV (26).

TGEV causes a fatal diarrheal disease in neonatal piglets because it selectively infects and destroys the small-intestinal enterocytes required for nutrient absorption and fluid regulation (11). Additionally, TGEV replicates in porcine respiratory tract tissues but this does not result in primary respiratory disease (7). A spontaneously occurring variant of TGEV, designated porcine respiratory coronavirus (PRCV), was identified and isolated in Belgium in 1983 to 1984 (14). Recently, an independent variant, PRCV-Ind/89, was isolated from pigs in the United States (27). These PRCV isolates are antigenically very similar to TGEV and replicate more extensively than TGEV in the respiratory tracts of young and adult pigs without causing clinical disease. However, PRCV differs from TGEV in that it undergoes only limited replication in an unidentified submucosal cell type of the small intestine (2). Thus, PRCV is an attenuated variant of TGEV because it does not replicate in the enterocytes that line the small intestine. In this report, we compare the genetic structures and transcriptional patterns of respiratory variant PRCV-Ind/89 and virulent enterotropic virus TGEV-PP3.

Both TGEV and PRCV are cytopathic for swine testicular (ST) cells and were plaque purified on this cell line (10). The Miller strain of TGEV was the sixth pig passage of the virus as an intestinal homogenate. This virus was designated PP3 after three rounds of plaque purification. The respiratory variant of TGEV was obtained from a nasal swab of an infected pig (27). The respiratory virus used in these studies (PRCV-Ind/89) was obtained after two rounds of plaque purification in ST cells.

Intracellular RNAs. To determine the transcriptional patterns of TGEV-PP3 and PRCV-Ind/89, intracellular RNA
was isolated from confluent ST cells infected at a multiplicity of infection of 0.02 PFU per cell. At 17 h postinfection, when a few cells had rounded, total intracellular RNA was prepared by guanidinium isothiocyanate extraction and pelleting of the RNA through a CsCl cushion (26). RNA was denatured with glyoxal and dimethyl sulfoxide and separated by electrophoresis in a 1% agarose gel (9). After electrophoresis, the RNAs were transferred to Gene Screen nylon membranes (Dupont) in 10× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate, pH 7.0) and cross-linked with UV light. Hybridization was carried out as described previously (24) but at 60°C by using as the probe a 32P-end-labeled oligonucleotide, 5'-CAGCATGGAGGAAGACGAGCATC TCG-3' (HP-1), specific for the 3' end of the TGEV genome (Fig. 1A). The blots were washed at 60°C with three changes of 2× SSC at room temperature, followed by a 2× SSC wash at 60°C for 1 h. Dried filters were exposed to Kodak XAR-2 film at −70°C with an intensifying screen for 6 to 17 h.

The intracellular RNA profiles of PP3 and PRCV-Ind/89 are shown in Fig. 1B, and a schematic diagram showing the relative positions of the PP3 intracellular RNAs is illustrated in Fig. 1A. For both PP3 and PRCV, RNAs 1 to 8 comigrated and were present in similar concentrations. However, there are differences in the RNA 1, 2(S), and 3 profiles of these viruses. In PRCV-infected cells, RNA 3 was not detectable, RNA 2(S) migrated as a smaller 7.5-kb species instead of the 8.2-kb RNA species of PP3, and RNA 1 migrated slightly faster than the corresponding RNA 1 species of PP3.

Immunoprecipitation of intracellular S proteins indicated that the smaller 7.5-kb RNA 2(S) of PRCV encoded a PRCV S protein that migrated faster by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) than the PP3 S protein (data not shown). It is not known whether the smaller 7.5-kb RNA 2(S) of PRCV is sufficient to account for the more rapid migration of the PRCV RNA 1 species or whether additional deletions might be present in the PRCV genome upstream of the S gene.

**Mapping of deletions in the PRCV S gene and ORF A.** Nuclease S1 protection experiments were carried out to locate the deleted coding sequences in the S gene and sequences downstream of the S gene. The probes used in these experiments were produced from PP3 genomic RNA, and their positions relative to the S gene are shown in Fig. 2A. Plasmid RP1 and subclone Hpa contain sequences at the 5' end of the S gene. In addition, RP1 also contains 930 nt of the Pol-encoding gene. Subclone XE contains 2,179 nt from the middle of the S gene, and plasmid F180 contains the remaining 3' S gene sequences plus downstream sequences of ORF A and the 5' half of ORF B.

For plasmid RP1, 32P-labeled runoff transcripts were hybridized to unlabeled PRCV total intracellular RNA. After digestion with nuclease S1, two protected bands of 2,600 and 1,000 nt were observed (Fig. 2B). This indicated that a region of pRP1 was not colinear with the PRCV RNA. Since RP1 is 4,254 nt long, approximately 650 bases were not protected. Experiments with subclone XE showed that the middle portions of the PRCV and PP3 S genes were colinear. Protection experiments with the Hpa runoff transcript generated a single band of 800 nt. These results indicate that PRCV contains a deletion of 650 to 700 nt near the 5' end of the S gene.

Nuclease S1 protection results obtained with runoff transcripts of clone F180 indicated that the PP3 and PRCV sequences were also not completely colinear in this region of the genome. The major bands protected were 1,100 and 500 nt, although three minor bands were also present (Fig. 2B). This shows that changes in the PRCV genome have occurred either in the 3' region of the S gene or in sequences downstream of the S gene. By nucleotide sequence analysis
VOL.

protected with runoff transcripts generated from plasmids P180 and TGEV plasmids other than PRCV S gene. The PRCV sequence.

1991

Fig. 2. (A) Schematic diagram showing the S gene, the 3' end of the Pol gene, ORFs A and B, and the relative positions and sizes of plasmids F180 and RPI and subclones Hpa and XE. (B) S1 nuclease fragments protected with runoff transcripts generated with F180, RPI, pHpa, and pXE were analyzed on a 1% agarose gel after glyoxal-dimethyl sulfoxide treatment and heat denaturation. The sizes of the 32P-labeled input probes are shown in the left margins, and those of the protected bands are indicated on the right. The input lanes represent the probe used in that experiment, while the other lanes represent the S1 nuclease experimental results with ST, PP3, or PRCV RNA, as indicated.

(see below), it was shown that the changes occurred downstream of the S gene.

Nucleotide sequence analysis of the S gene deletion. To localize the large PRCV S gene deletion accurately, cDNA clones were prepared from total PRCV-Ind/89 intracellular RNA by using oligonucleotides 5'-GCTAGGGACTGGC-3' (no. 15), which is complementary to bases 917 to 930 of the TGEV-PP3 S gene (23). These cDNAs were cloned into the EcoRI site of lambda ZAP (Stratagene) and excised by superinfection with helper phage to produce Bluescript plasmids containing cDNA inserts. Nucleotide sequencing was carried out on double-stranded plasmids (5) by the dyeoxy-chain termination method (18).

The first 200 nt at the 5' end of the PRCV S gene are shown in Fig. 3. Sixty-two nt into the PRCV S gene, a 681-nt deletion occurred compared with the PP3 S gene. Thus, PRCV retained the sequences encoding the 16-amino-acid signal peptide (17). The 6 amino acid residues at the N-terminal end of the mature S protein were also retained; however, the next 227 amino acids were deleted. This accounts for a 30- to 35-kDa reduction in the size of the PRCV S protein.

Genetic basis for the lack of PRCV mRNA 3. To confirm the absence of mRNA 3 in PRCV and determine the nature of the gaps indicated by the S1 protection experiments with the F180 runoff transcript, PRCV cDNA clones were obtained and sequenced. Specific primers, oligonucleotides 5'-GCT TACAAGCATAGGG-3' (no. 3A) and 5'-ATGACCATTG CATG-3' (no. 3), were used to generate these cDNAs. Oligonucleotide 3A is complementary to a region within ORF B, and oligonucleotide 3 is complementary to a region of ORF C (Fig. 1).

The 500 nt immediately downstream of the PRCV S gene are shown in Fig. 4. The ATG initiation codons of ORFs A and B are located at positions 103 and 378, respectively. A noncoding region of 102 nt is present between the S gene and ORF A. A second noncoding region of 116 nt, from bases 262 to 377, exists between ORFs A and B. The transcription recognition sequence TTCTAAAC for ORF B (mRNA 4) is present in both PP3 and PRCV. However, upstream of ORF A at position 79 (Fig. 4), a C-to-T substitution in PRCV, which altered the transcription recognition sequence, apparently accounted for the lack of a detectable mRNA 3. Two deletions were present in this region of the PRCV genome, a 3-nt deletion at position 19 in the noncoding region between the S gene and ORF A and a 5-nt deletion at position 263. The second deletion introduced a termination code that shortened the ORF A product from 72 to 53 amino acid residues. These deletions produced nuclease S1-sensitive sites that yielded the 1,100- and 500-nt fragments in experiments with the F180 runoff transcripts. Incomplete S1 digestion at the deletion sites in these experiments accounted for the three minor bands that were observed.

Here we show that in the peplomer gene and in the downstream ORF A region of the TGEV genome, deletions and a significant point mutation that are associated with production of a pneumotropic TGEV variant, designated PRCV-Ind/89, have occurred. Similar but not identical genetic changes have occurred in the European PRCV (15). In both instances, the PRCVs have lost the capacity to infect and destroy the absorptive epithelial cells lining the small intestine and thus, in contrast to TGEV, no longer cause serious diarrheal disease in young pigs. Although PRCV has retained or perhaps acquired additional specificity for lung tissues and nasal epithelial cells, in the absence of secondary infection by other opportunistic organisms, it does not appear to cause serious respiratory disease. We have shown that a deletion of 681 nt has occurred near the 5' end of the PRCV S gene that encodes the more variable N-terminal

![Diagram](http://jvi.asm.org/)

FIG. 2. (A) Schematic diagram showing the S gene, the 3' end of the Pol gene, ORFs A and B, and the relative positions and sizes of plasmids F180 and RPI and subclones Hpa and XE. (B) S1 nuclease fragments protected with runoff transcripts generated with F180, RPI, pHpa, and pXE were analyzed on a 1% agarose gel after glyoxal-dimethyl sulfoxide treatment and heat denaturation. The sizes of the 32P-labeled input probes are shown in the left margins, and those of the protected bands are indicated on the right. The input lanes represent the probe used in that experiment, while the other lanes represent the S1 nuclease experimental results with ST, PP3, or PRCV RNA, as indicated.

The first 200 nt at the 5' end of the PRCV S gene are shown in Fig. 3. Sixty-two nt into the PRCV S gene, a 681-nt deletion occurred compared with the PP3 S gene. Thus, PRCV retained the sequences encoding the 16-amino-acid signal peptide (17). The 6 amino acid residues at the N-terminal end of the mature S protein were also retained; however, the next 227 amino acids were deleted. This accounts for a 30- to 35-kDa reduction in the size of the PRCV S protein.

Genetic basis for the lack of PRCV mRNA 3. To confirm the absence of mRNA 3 in PRCV and determine the nature of the gaps indicated by the S1 protection experiments with the F180 runoff transcript, PRCV cDNA clones were obtained and sequenced. Specific primers, oligonucleotides 5'-GCT TACAAGCATAGGG-3' (no. 3A) and 5'-ATGACCATTG CATG-3' (no. 3), were used to generate these cDNAs. Oligonucleotide 3A is complementary to a region within ORF B, and oligonucleotide 3 is complementary to a region of ORF C (Fig. 1).

The 500 nt immediately downstream of the PRCV S gene are shown in Fig. 4. The ATG initiation codons of ORFs A and B are located at positions 103 and 378, respectively. A noncoding region of 102 nt is present between the S gene and ORF A. A second noncoding region of 116 nt, from bases 262 to 377, exists between ORFs A and B. The transcription recognition sequence TTCTAAAC for ORF B (mRNA 4) is present in both PP3 and PRCV. However, upstream of ORF A at position 79 (Fig. 4), a C-to-T substitution in PRCV, which altered the transcription recognition sequence, apparently accounted for the lack of a detectable mRNA 3. Two deletions were present in this region of the PRCV genome, a 3-nt deletion at position 19 in the noncoding region between the S gene and ORF A and a 5-nt deletion at position 263. The second deletion introduced a termination code that shortened the ORF A product from 72 to 53 amino acid residues. These deletions produced nuclease S1-sensitive sites that yielded the 1,100- and 500-nt fragments in experiments with the F180 runoff transcripts. Incomplete S1 digestion at the deletion sites in these experiments accounted for the three minor bands that were observed.

Here we show that in the peplomer gene and in the downstream ORF A region of the TGEV genome, deletions and a significant point mutation that are associated with production of a pneumotropic TGEV variant, designated PRCV-Ind/89, have occurred. Similar but not identical genetic changes have occurred in the European PRCV (15). In both instances, the PRCVs have lost the capacity to infect and destroy the absorptive epithelial cells lining the small intestine and thus, in contrast to TGEV, no longer cause serious diarrheal disease in young pigs. Although PRCV has retained or perhaps acquired additional specificity for lung tissues and nasal epithelial cells, in the absence of secondary infection by other opportunistic organisms, it does not appear to cause serious respiratory disease. We have shown that a deletion of 681 nt has occurred near the 5' end of the PRCV S gene that encodes the more variable N-terminal

![Diagram](http://jvi.asm.org/)

FIG. 3. Comparison of cDNA sequences of PRCV and the PP3-Miller strain of TGEV. [S] indicates the methionine initiation code of the S gene. The region that encodes the cleaved signal peptide is indicated by the bar. The filled triangle between bases 62 and 63 indicates the location of a 681-nt insertion in the TGEV S gene. Other single-base differences between PP3 and PRCV are depicted below the primary PRCV sequence.
globular portion of the peplomer protein. This portion of the protein projects from the virus lipid envelope and is thought to contain sites for virus attachment to cells (22). Large deletions in the 5′ half of the peplomer gene of another coronavirus, mouse hepatitis virus, also have been observed (3, 13). These deletions have either occurred spontaneously or been selected as neutralization-resistant mutants that replicated in the presence of neutralizing monoclonal antibody and produced virus variants with reduced neurovirulence. Deletions could occur during coronavirus replication if the mRNA transcript dissociated from the negative-sense template and reinitiated at a downstream site on either the same or another negative-stranded template.

The genomic RNA of a European PRCV isolate (RM 4) has been cloned and studied (15). The 3′ sequence of 7,519 nt has been determined. There are general similarities between the genomes of the U.S. and European PRCVs that distinguish these viruses from enterotropic TGEV strains. The U.S. and European PRCVs have large deletions that map to the 5′ end of the S gene, and both have genetic changes that eliminate mRNA 3 (the unique region of this transcript encodes ORF A). However, the specific differences in the genomes of the two viruses suggest that the more recent U.S. isolate arose independently of the European isolate. The deletion near the 5′ end of the RM 4 S gene is slightly smaller (672 nt), and two deletion regions have virtually eliminated ORF A of RM 4. One deletion removed the upstream transcription recognition sequence and the ATG initiation codon, while a second deletion (31 bases) occurred in the body sequence of ORF A.

It has been our experience that PRCV requires little or no adaptation to grow on cultured ST cells, suggesting that the PRCV peplomer, like the TGEV peplomer, retains the capacity to bind to the ST cell receptor. However, PRCV has lost the ability to replicate in swine intestinal epithelial cells in vivo. One possible explanation for the reduced enteric tropism of PRCV is that the 227 amino acid residues that are deleted near the N terminus of the PRCV-Ind/89 peplomer protein are critical for binding to the cell receptor of swine enterocytes in vivo. Another TGEV variant, the small-plaque virus, is also avirulent for baby pigs and lacks the ability to replicate in enteric enterocytes. In small-plaque virus, however, the S gene region is colinear with the S gene of the PP3 virus by S1 protection assays (26). Small-plaque virus and the PRCVs share a common feature in that they lack ORF A or the mRNA that encodes ORF A. For small-plaque virus, there is a large 462-nt deletion that has eliminated ORF A and the 5′ portion of ORF B. For PRCV-Ind/89, ORF A is reduced to only 53 amino acid residues but, in addition, a C-to-T substitution in the transcription recognition sequence has rendered mRNA 3 nondetectable, presumably eliminating the potential ORF A protein product. As described above, ORF A has been eliminated in PRCV RM 4 by deletions. In addition, mRNA 3 of another avirulent, high-passage TGEV-Miller strain virus (Miller 60) is not detectable by Northern (RNA) blot analysis (unpublished data).

We cannot ascertain whether the deletion in the S gene, the lack of mRNA 3, or perhaps other genetic differences have altered the tropism of PRCV-Ind/89 for enteric enterocytes. Current data suggest that ORF A is an important factor involved in TGEV pathogenesis, since all of the attenuated TGEVs examined to date have mutations that eliminate the hypothetical ORF A polypeptide.

In a typical TGEV RNA pattern (Fig. 1B), RNA 3 is more abundant than RNA 4 and both transcripts are preceded by the transcription recognition sequence ^GGTAAAC (24). For PRCV-Ind/89, a C-to-T substitution has changed the recognition sequence from AACTAAAC to AACTAAAT, which subsequently reduced mRNA 3 to a nondetectable level. Similar C-to-T substitutions have occurred in the recognition sequences preceding ORF B of TGEV-Purde and the FS772-70 strain of TGEV (1, 16). In these viruses, the corresponding mRNA 4 transcript is either absent or present at a reduced level. Thus, the terminal cytosine of the consensus recognition sequence may function as an important determinant for initiation of TGEV transcription.

We thank David Michael for technical assistance and Linda Hornung for typing the manuscript.

REFERENCES


