Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

VIROLOGY 217, 495–507 (1996)
ARTICLE NO. 0144

Molecular Characterization of Transmissible Gastroenteritis Coronavirus Defective Interfering Genomes: Packaging and Heterogeneity

ANA MÉNDEZ, CRISTIAN SMERDOU, ANDER IZETA, FÁTIMA GEBAUER, and LUIS ENJUANES[1]

Centro Nacional de Biotecnología, CSIC, Department of Molecular and Cell Biology, Campus Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain

Received September 12, 1995; accepted January 9, 1996

Three transmissible gastroenteritis virus (TGEV) defective RNAs were selected by serial undiluted passage of the PUR46 strain in ST cells. These RNAs of 22, 10.6, and 9.7 kb (DI-A, DI-B, and DI-C, respectively) were detected at passage 30, remained stable upon further passage in cell culture, and significantly interfered with helper mRNA synthesis. RNA analysis from purified virions showed that the three defective RNAs were efficiently packaged. Virions of different densities containing either full-length or defective RNAs were sorted in sucrose gradients, indicating that defective and full-length genomes were independently encapsidated. DI-B and DI-C RNAs were amplified by the reverse transcription-polymerase chain reaction, cloned, and sequenced. DI-B and DI-C genomes are formed by three and four discontinuous regions of the wild-type genome, respectively. DI-C contains 2144 nucleotides (nt) from the 5′-end of the genome, two fragments of 4540 and 2531 nt mostly from gene 1b, and 493 nt from the 3′ end of the genome. DI-B and DI-C RNAs include sequences with the pseudoknot motif and encoding the polymerase, metal ion binding, and helicase motifs. DI-B RNA has a structure closely related to DI-C RNA with two main differences: it maintains the entire ORF 1b and shows heterogeneity in the size of the 3′ end deletion. This heterogeneity maps at the beginning of the S gene, where other natural TGEV recombination events have been observed, suggesting that either a process of template switching occurs with high frequency at this point or that the derived genomes have a selective advantage.

INTRODUCTION

Transmissible gastroenteritis virus (TGEV) is a member of the *Coronaviridae* family (Cavanagh *et al.,* 1993, 1990; Siddell *et al.,* 1983) which causes enteric and respiratory infections in swine, leading to 100% mortality in neonatal animals (Cavanagh *et al.,* 1990; Enjuanes and Van der Zeijst, 1995; Saif and Wesley, 1992). It is an enveloped virus with a positive-stranded polyadenylated RNA genome of 28.5 kb that undergoes replication entirely within the cytoplasm of infected cells (Eleouet *et al.,* 1995; Lai, 1990). Coronaviruses initiate infection by translating the genomic RNA into the RNA-dependent RNA polymerase. The genome is then replicated into a minus sense RNA, which serves as template for both synthesis of the genomic RNAs and transcription of the six to seven subgenomic mRNAs. The mRNAs form a 3′ coterminal nested set; that is, the sequence of each mRNA is contained entirely within the next larger mRNA. Although mRNAs are polycistronic, only the 5′-most open reading frame is translated. Coronavirus subgenomic mRNAs all include a 5′ leader sequence. However, on coronavirus genomic RNA this leader is present only once at the 5′ end, implying that synthesis of the subgenomic mRNAs involves fusion of noncontiguous sequences by discontinuous transcription (Jeong and Makino, 1994; Lai *et al.,* 1983; Spaan *et al.,* 1983; Zhang and Lai, 1994). The first mechanism proposed to explain the discontinuous transcription (the leader-priming hypothesis) postulates that a trans-acting leader RNA, transcribed from the 3′ end of a negative copy of the genomic RNA, binds to intergenic promoter sequences on negative-stranded genomic templates and is extended to produce the subgenomic RNAs (for reviews see Lai, 1990 and Van der Most and Spaan, 1995). In addition to the subgenomic mRNAs, minus-stranded copies of mRNAs were identified first in TGEV (Sethna *et al.,* 1989) and then in bovine coronavirus (Hofmann *et al.,* 1990). These negative-stranded RNAs contained an antileader sequence (Sethna *et al.,* 1991) and have been shown to function in mRNA synthesis through replicative intermediates (RIs) (Sawicki and Sawicki, 1990). The precise role of the negative-stranded RNAs in the replicative cycle of coronaviruses has not been fully clarified. It has been proposed that negative-subgenomic RNAs serve as templates for the amplification of subgenomic mRNAs (Brian *et al.,* 1994; Schaad and Baric, 1994) and also to be end products (Jeong and Makino, 1992). The hypothesis of the intergenic promoter sequences acting as transcription attenuator sequences (Sawicki and Sawicki, 1990) has led to two other transcription models. According to one of these models, a nested set of subgenomic negative strands is synthesized first. In this process the intergenic

[1] To whom correspondence and reprint requests should be addressed. Fax 341- 585 45 06. E-mail: Lenjuanes@Samba.CNB.UAM.ES.

0042-6822/96 $18.00

promoter sequences act as transcription termination sequences. These RNAs then serve as templates for leader-primed transcription (Sawicki and Sawicki, 1990). Alternatively, negative-stranded RNAs produced from genomic RNA terminate at internal intergenic promoter sequences, detach, and hybridize to the 3′ end of the leader on the genomic RNA. Then, an extension of the negative-stranded RNA will add a negative-sense leader sequence (for a review see Van der Most and Spaan, 1995).

During transcription and replication the RNA replicase may pause and jump from one virus RNA template to another or from one segment of a template to another, carrying the incomplete nascent strand with it, which may then be elongated at the new template resumption site. The discontinuous and nonprocessive nature of transcription may, in this way, give rise to defective RNAs (D-RNAs). D-RNAs are deletion mutants maintaining the *cis*-signals required for replication, but dependent on other viral functions to be supplied *in trans* by a helper virus. D-RNAs have been isolated for many animal viruses (Baric *et al.,* 1983; Barrett and Dimmock, 1986; Makino *et al.,* 1988a, 1984; Roux *et al.,* 1991; Spaan *et al.,* 1983; Van der Most *et al.,* 1991).

The large size of coronavirus genomes (26 to 31 kb) has been to date a constraint on the assembly of full-length, infectious cDNA clones that would greatly facilitate the molecular analysis of these viruses. An alternative approach is the cloning of subgenomic D-RNAs. Several cDNAs have been generated from coronavirus D-RNAs, including mouse hepatitis virus (MHV), infectious bronchitis virus (IBV), and bovine coronavirus (BCV) (Chang *et al.,* 1994; Makino *et al.,* 1988a; Penzes *et al.,* 1994). cDNA constructs of small MHV D-RNAs have been successfully used to define *cis*-acting signals required for replication, transcription, and encapsidation of the virus (Fosmire *et al.,* 1992; Joo and Makino, 1995; Kim and Makino, 1995; Lin and Lai, 1993; Makino *et al.,* 1991, 1988b; Van der Most *et al.,* 1991, 1994), as well as to introduce site-specific mutations into the 3′-end of the genome by homologous RNA recombination (Masters *et al.,* 1994; Peng *et al.,* 1995; Van der Most *et al.,* 1992) and to engineer chimeric viruses expressing heterologous genes (Liao *et al.,* 1995; Lin and Lai, 1993). D-RNAs may also interfere with wild-type (*wt*) virus replication, a characteristic potentially relevant in protection against viral disease (Barrett and Dimmock, 1986).

This article provides the first report of TGEV D-RNAs, which are replicated and packaged into virus particles. These RNAs were shown to be defective and interfering. Two of these D-RNAs (DI-B and DI-C of 10.6 and 9.7 kb, respectively) were amplified by reverse transcription-polymerase chain reaction (RT-PCR), and their primary structure was determined. A dominant homogeneous population of DI-C molecules was identified, while DI-B was composed of a population of D-RNAs with high heterogeneity at the 3′ end.

MATERIALS AND METHODS

Cells and viruses

Swine testis (ST) (McClurkin and Norman, 1966) and PM cells (Nunn and Johnson, 1979) were grown in DMEM-medium supplemented with 8% fetal calf serum. Intestinal porcine epithelial cells (IPEC-1), provided by Helen M. Berschneider, were grown in DMEM/F12 (1:1) medium, supplemented with 5 μg/ml insulin, 5 μg/ml transferrin, 5 ng/ml selenium, 5 μg/ml epidermal growth factor, 0.01 *M* L-glutamine, antibiotics, and 5% fetal calf serum (FCS) (Berschneider and Powell, 1992). The plaque-cloned PUR46-MAD, PUR46-SW37-CC4, and HOL87 strains of TGEV were grown, purified, and titrated as described before (Sánchez *et al.,* 1990).

Orthophosphate metabolic labeling

Virus from the indicated passage number was used to infect ST-cells grown in 60-mm-diameter plates with a m.o.i. of 10 PFU/cell. Cells were incubated at 37° for 1 hr, the inoculum was removed and the cells were washed with phosphate minus medium containing 2% dialyzed FCS, and covered with 3 ml of the same medium. Plates were incubated at 37° for 3 hr, actinomycin D was added to a final concentration of 2.5 μg/ml, and cells were incubated for 3 additional hr. Then, [^{32}P]H_3PO_4 was added to a final concentration of 100 μCi/ml and replicating RNAs were labeled for 3 hr at 37°. At 10 hr postinfection (hr p.i.) cytoplasmic RNA was extracted as indicated below, and gel electrophoresis was conducted after denaturation with 2.2 *M* formaldehyde (Sambrook *et al.,* 1989).

Cytoplasmic RNA extraction

Cytoplasmic extracts from ST infected cells were prepared by lysing 4×10^6 cells in 200 ml of TSM buffer (0.15 *M* NaCl, 0.01 *M* Tris–hydrochloride, pH 7.6, 5 m*M* $MgCl_2$, 0.2% Nonidet P-40), and pelleting nuclei by centrifugation at 13,000 *g* for 30 sec. RNA was isolated by the addition of 200 ml urea–SDS lysis buffer (1.5% SDS, 15 m*M* EDTA, 0.24 *M* NaCl, 0.04 *M* Tris–hydrochloride, pH 7.6, 8 *M* urea), vigorous vortexing, and phenol-chloroform extraction. RNA extraction from cells grown in roller bottles (8×10^7 cells per bottle) was performed. A standard yield of 400 μg of total cytoplasmic RNA per roller bottle was obtained. Cytoplasmic RNA containing poly(A^+) RNA was selected with paramagnetic oligo (dT) particles using the Poly(A)Ttract mRNA isolation system (Promega) according to manufacturer instructions.

RNA isolation from purified virions

To obtain RNA stocks for RT-PCR amplification, genomic RNA was extracted from purified virus. ST cells grown in 12 roller bottles, each with a growth surface of 500 cm^2, were infected with a m.o.i. of 5 PFU/cell using TGEV passage 1 or DI containing virus from passage

41. Medium containing virus was harvested at 22 hr p.i. (cytopathic effect 80%) and clarified by centrifugation in a Sorvall GSA rotor for 20 min at 6,000 rpm. Virions were pelleted through a 5-ml 15% w/v sucrose cushion by centrifugation at 25,000 rpm in a Kontron TST28.18 rotor for 2 hr. To clear virus from remaining sucrose, viral pellets were resuspended in TEN buffer (10 m*M* Tris–HCl, pH 7.4, 1 m*M* EDTA, 1 *M* NaCl) and sedimented by centrifugation under the same conditions. Viral pellets were resuspended in 500 μl of NTE (10 m*M* Tris–HCl, pH 7.4, 1 m*M* EDTA, 100 m*M* NaCl) buffer, SDS was added to a final concentration of 2%, and the mixture was digested with 180 ng of proteinase K (Boehringer Mannheim) for 30 min at 37°. RNA was extracted twice with phenol-chloroform and precipitated with ethanol. A standard yield of 8 μg of viral RNA per roller bottle was obtained.

RNA analysis by Northern blot

Northern hybridizations were performed as described (Sambrook *et al.,* 1989), using a leader specific probe (5′-CGAGTTGGTGTCCGAAGACAAAATCT-3′) complementary to nt 66 to 91 of TGEV genome or the indicated oligonucleotides. The probes were 5′ ^{32}P labeled using [γ^{32}P]ATP and T4 polynucleotide kinase.

Growth of DI genomes after infection at low m.o.i.

To determine if the replication of subgenomic RNAs was dependent on coinfection with helper virus, PUR46-MAD passage 41 containing DI genomes was used to infect ST cells at m.o.i.: 2, 1, 0.5, 0.25, 0.12, 0.06, and 0.01 PFU/cell, and intracellular RNAs were extracted at 9 hr p.i. and studied by Northern blot analysis as described above.

Fractionation of virus particles with full-length or DI genomes

ST cells were grown in 12 roller bottles. Cells were infected at m.o.i. 5 PFU/ml with passage 41 of PUR46-MAD TGEV. Virus was harvested at 22 hr p.i., clarified, and divided into three aliquots: two of them were underlaid with 31 and 15% (w/v) sucrose cushions prepared in TEN with 0.2% (v/v) Tween 20, respectively, and the third one was concentrated by centrifugation through a 5-ml 15% (w/v) sucrose and then placed on the top of a 15–42% linear sucrose gradient. Samples were centrifuged in a Kontron TST28.18 rotor at 25,000 rpm for 2 hr. The sucrose gradient in the third tube was fractionated into 14 aliquots of 1.5 ml. Virus from each fraction was independently pelleted in the same centrifugation conditions. RNA from each sample was extracted as indicated above. RNA from virus pelleted through the sucrose cushions (31 and 15% w/v) and from gradient fractions with densities of 1.20 or 1.15 g/ml was loaded in a denaturing gel and studied by Northern blot analysis using a leader-specific probe.

RT-PCR amplification of DI RNAs

DI-B and DI-C RNAs from purified virions or Poly(A+)-selected cytoplasmic RNA isolated from passage 41 of TGEV-infected cells were amplified by RT-PCR. Four separate fragments (*a, b, c,* and *d*) were obtained for each DI RNA using primers 1 to 8 (Table 1). Primer 1 was deduced from the leader sequence (Page *et al.,* 1990; Sethna *et al.,* 1991). Primer 2 was deduced from the 5′ end of the FIPV sequence kindly provided by R. de Groot (Utrecht University). Other ORF 1 primers were selected from RNA regions showing sequence homology between MHV, IBV, HCV-229E, and FIPV (Boursnell *et al.,* 1987; Herold *et al.,* 1993; Lee *et al.,* 1991; Pachuk *et al.,* 1989). The presence of sequences complementary to the primers in *wt* TGEV RNA or DI RNAs was tested by Northern blot analysis, and those primers which recognized TGEV genomic and defective RNAs were selected for RT-PCR amplifications. First strand cDNA synthesis of fragments *a, b, c,* and *d* (see Results) was primed with oligonucleotides 2, 4, 6, and 8, respectively, in the presence of 100 ng of template RNA, 20 pmol oligonucleotide, 1× PCR buffer (20 m*M* Tris–HCl, pH 8.4, 50 m*M* KCl, 2.5 m*M* $MgCl_2$, 0.1 mg/ml BSA), 500 μM dNTP mix, 10 m*M* DTT, 12 U HPRI (Boehringer Mannheim), and 12 U AMV RT (Seikagaku) in a final volume of 20 μl. Samples were incubated at 42° for 60 min, and 10 μl of the resulting cDNAs were amplified by PCR by adding 10 pmol of the second primer (oligonucleotides 1, 3, 5, and 7, respectively), 4 μl 10× PCR buffer, 2.5 U *Taq* Polymerase (Perkin–Elmer), and water to a final volume of 50 μl, in a 9600 Perkin–Elmer thermocycler. The sizes of the obtained fragments were as follows: fragment *a,* 1.9 kb (DI-B and DI-C); fragment *b,* 2.8 kb (DI-B and DI-C); fragment *c,* 4.6 kb (DI-B) and 3.5 kb (DI-C); fragment *d,* 1.9 kb (DI-B) and 2.1 kb (DI-C).

Cloning and sequencing of DI-B and DI-C

Fragments *a, b, c,* and *d* were purified using Geneclean II kit (BIO 101, Inc) and cloned into Bluescript SK$^-$ (Stratagene). Fragments *a* and *d* were cloned into the *Eco*RV site using standard cloning procedures (Sambrook *et al.,* 1989). Fragment *b* was digested with *Xba*I and *Hin*dIII (the last enzyme cut the cDNA at an internal position of the fragment), and the resulting 2.84-kb fragment was cloned . The 3.6-kb fragment *c* was cleaved into two fragments of 0.95 and 2.6 kb by restriction endonuclease *Pst*I. Cohesive ends were then generated into these fragments by further digestion with *Eco*RI, as targets for this enzyme were included in the primers used for fragment *c* amplification, and the resulting *Eco*RI–*Pst*I fragments were cloned. The various clones were sequenced using Sequenase 2.0 kit (USB). Sequence data were compiled using the UWGCG (University of Wisconsin, Genetic Computer Group) sequence analysis software package. To detect possible mutations intro-

duced by reverse transcriptase or *Taq* polymerase, at least two clones from independent RT-PCR reactions were sequenced for each fragment. Nucleotide positions that differed between two clones were directly sequenced from the PCR product, using the fmol DNA Sequencing System (Promega), according to the manufacturer instructions.

The 5′ end of the DI-RNAs was determined by sequencing RNA from purified virions as previously described (Fichot and Girard, 1990). The primer 5′-TAATCAACGCTTGTCCTC-3′ hybridizing to nucleotides 127 to 147 from the 5′ end was used at a concentration of 0.28×10^{-6} *M;* dGTP, dCTP, and dTTP were used at concentrations of 0.5×10^{-3} *M;* [α-^{35}S]dATP was used at a concentration of 0.5×10^{-4} *M* (300 μCi/ml); and reverse transcriptase (Moloney murine leukemia virus; Perkin–Elmer) was used at a concentration of 2×10^4 U/ml. Reaction mixtures were incubated at 37° for 60 min before analysis of the products by electrophoresis on 6% polyacrylamide sequencing gels. To determine the position of the last nucleotide a primer extension analysis was performed (Sambrook *et al.,* 1989) using the same primer employed in direct RNA sequencing.

To determine the sequence of the 3′ end of DI-C and DI-B, cDNAs containing the complete 3′ end of DI-C and DI-B were cloned by RT-PCR using RNA extracted from DI containing stocks of purified TGEV. An oligo dT primer and an oligo with the virus sequence from nucleotide −316 to −296, in relation to the 5′ end of the S gene was used (Eleouet *et al.,* 1995). The RT-PCR fragment derived from DI-C had 0.8 kb as expected (results not shown) and those derived from DI-B had a size ranging from 0.6 to 0.7 kb and were cloned into a pGEM-T vector (Promega). The sequences of four clones were determined using primers hybridizing in the pGEM-T vector.

Amplification, cloning, and sequencing of the junction site between regions III and IV of DI-B and DI-C genomes

RNA from passage 41 purified virus was amplified by RT-PCR with primers 8 and 9 (Table 1). The resulting fragments of approximately 1.4 kb were cloned into pGEM-T plasmid (Promega) according to manufacturer instructions and were sequenced using oligonucleotide 5′TCTGTACAAGAGTAGACAGC 3′ complementary to positions 205–224 from the 3′-end of the TGEV genome, with both the Sequenase kit (USB) and the fmol DNA Sequencing System (Promega).

RESULTS

Generation of TGEV defective RNAs by passage at high m.o.i.

The PUR46-MAD strain of TGEV was passaged undiluted 46 times in ST cells (m.o.i. between 50 and 100 PFU/cell). After infection, cells were incubated at 37° for 21–24 hr. At that time around 80% of the cells showed cytopathic effect. Supernatant was then harvested and divided into two aliquots: one of them was stored at −70° and the other one was used for the subsequent infection.

RNA from cells infected with virus of increasing passage number was metabolically labeled and analyzed by electrophoresis in a denaturing gel (Fig. 1). During early passages, only *wt* genomic and subgenomic RNAs were detected. However, three new RNA species of 22, 10.6, and 9.7 (DIs A, B, and C, respectively) were observed at passage 30. These subgenomic RNAs persisted in high amounts during at least 15 subsequent passages in ST cells, significantly interfering with helper genomic and subgenomic RNA synthesis (Fig. 1, lanes corresponding to passage levels 30 to 45). These data indicate that at least one of the three new RNA species is an interfering RNA.

In addition to the serial passage of TGEV PUR46 strain, 35 undiluted virus passages were also performed in ST cells with the HOL87 strain of PRCV (Sánchez *et al.,* 1990) and PUR46-*mar*1C.C12, a monoclonal antibody (MAb) escaping mutant of TGEV (Gebauer *et al.,* 1991). PUR46-SW37-CC4 was passaged in IPEC-1 cells 10 times. Virus RNA from infected cells was metabolically labeled and analyzed in all cases (data not shown). No defective RNAs were detected in any case.

RNAs A, B, and C are defective RNAs

To determine whether subgenomic RNAs A, B, and C had the standard structure of defective RNAs, i.e., if they had the 5′ and 3′-ends of the *wt* genome and internal deletions to justify their smaller size as compared with the *wt* genome Northern hybridization analysis was performed using ^{32}P-labeled oligonucleotides complementary to the leader and the 3′-UTR sequences. Both oligonucleotides hybridized to all parental mRNAs and also detected DI-A, DI-B, and DI-C RNAs (data not shown), indicating that these are in fact viral RNAs with internal deletions. As a first approach to study which genomic sequences were present in the DIs, total RNA from ST-infected cells was also probed with oligonucleotides mapping in S, M, and N structural protein genes. None of them detected the D-RNAs (data not shown), indicating that some deletions mapped to structural genes.

To analyze whether the defective RNAs were able to replicate independently of the *wt* virus, ST cells were infected with virus from passage 41, containing all DI RNAs, at low m.o.i. (1, 0.5, 0.25, 0.125, 0.06, and 0.01 PFU/cell), and the intracellular RNAs studied at 9 hr p.i. The results (not shown) indicated that the three subgenomic RNAs were lost at m.o.i. = 0.1. In addition, when the three DIs were passaged from ST cells into other porcine cell lines, either PM (results not shown) or IPEC cells, all three DI RNAs were recovered after the first passage,

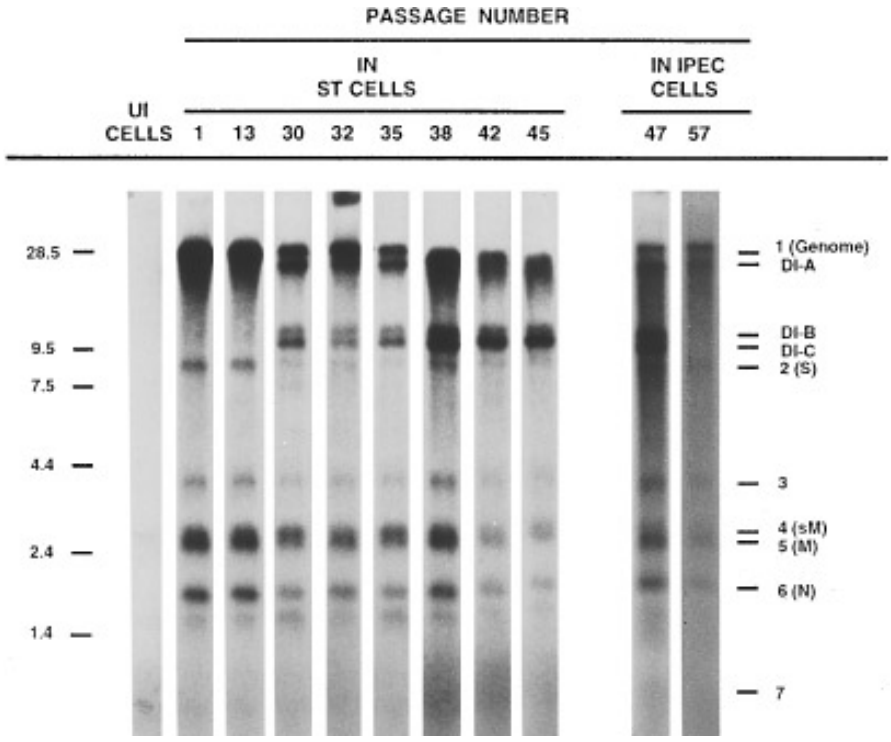

FIG. 1. Electrophoretic analysis of metabolically labeled TGEV RNAs. Cytoplasmic RNA from ST and IPEC cells infected with virus at different passage numbers was extracted and resolved in a denaturing agarose gel after [^{32}P]orthophosphate metabolic labeling. Similar RNA analysis with TGEV passaged in PM cells gave the same results as in IPEC cells (results not shown). Passage numbers are indicated on the top of each lane. Bars on the left indicate molecular weight markers (genomic TGEV RNA and GibcoBRL markers) expressed in kb. Bars on the right indicate the various TGEV mRNAs species and DI RNAs. UI, uninfected.

but only DI-A persisted after 10 more passages (Fig. 1, lanes corresponding to passage levels 47 and 57).

Packaging of DI RNAs

To study whether defective RNAs were encapsidated, virions from passage 1 and 41 were partially purified by centrifugation through a 15% w/v sucrose cushion. RNA from pelleted virions was extracted and studied by Northern Blot analysis (Fig. 2A). Defective RNAs A, B, and C were detected in addition to full-length RNA in virions from passage 41, indicating that all these RNAs were efficiently packaged.

To determine whether defective RNAs were copackaged into virions together with full-length genome or independently encapsidated, virions from passage 41 were purified by centrifugation through sucrose cushions and through a continuous sucrose gradient. RNA extracted from purified virions was analyzed by Northern blot using a leader-specific oligonucleotide (Fig. 2B). After sedimentation through a 31% w/v (1.19 g/ml) sucrose cushion only the *wt* genome was detected in pelleted virions (Fig. 2B, lane b). However, when the sucrose cushion density was decreased to 15% w/v (1.11 g/ml) both the full-length and the defective RNAs were detected (Fig. 2B, lane c). An almost complete separation of *wt* and defective virions was achieved in continuous sucrose gradients (15–42%). Lower fractions (density around 1.20 g/ml) were enriched in standard virions, while defective virions were concentrated in upper fractions (density 1.15 g/ml) (Fig. 2B, lanes d and e). The upper band corresponds to *wt* and DI-A genomes, and the lower band to DIs B and C (data not shown). These data indicate that DI RNAs are independently packaged into defective virions that differ in density from the *wt.*

Determination of DI-B and DI-C genetic structures

Based on gel electrophoretic mobility, the sizes of DI-B and DI-C were estimated to be about 10.6 and 9.7 kb, respectively. Due to these large sizes, D-RNAs could not be amplified by RT-PCR using primers specific for the ends of the genome. To circumvent this limitation, four pairs of primers were designed to amplify the D-RNAs in four separate reactions, giving overlapping fragments that were referred to as, from 5′ to 3′, *a, b, c,* and *d.* RNA from passage 41 purified virions was used as template. As a control, parallel amplification of *wt* genomic RNA was performed.

RT-PCR amplification with primers 1 and 2 (Table 1) from either passage 1 or passage 41 RNAs gave rise to a unique PCR product of 1.9 kb, indicating that fragment *a* was common to all DI-RNAs, and corresponded to the 5′ 1.9-kb region of the TGEV genome (Fig. 3). Amplifica-

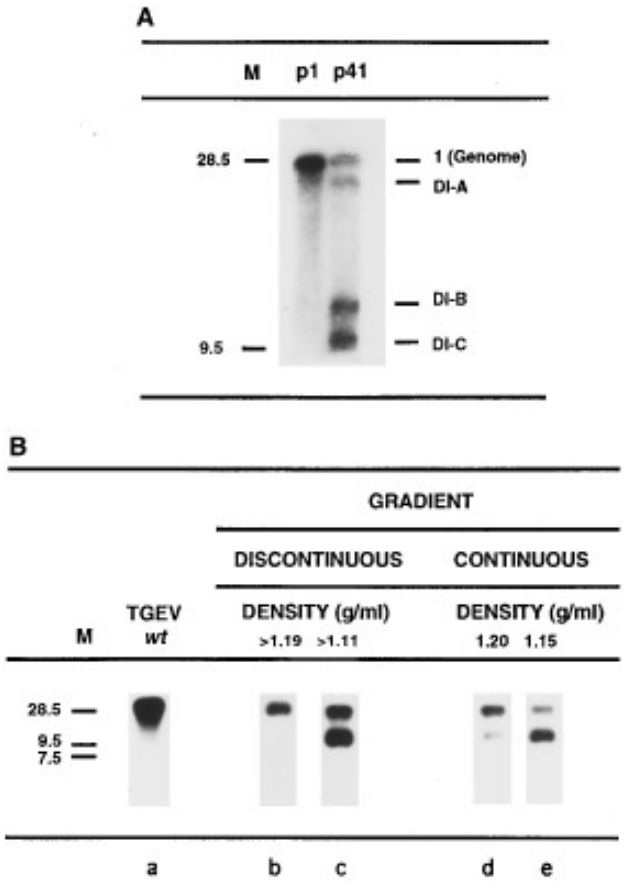

FIG. 2. Packaging of DI RNAs. (A) Virions from passages 1 and 41 were purified by centrifugation through 15% w/v sucrose. RNA was extracted and studied by Northern blot analysis. At passage 41 defective RNAs A, B, and C were detected in addition to *wt* RNA. Bars on the left indicate molecular weight markers in kb. (B) Virus from passage 41 was purified by centrifugation through either sucrose cushions or gradients. RNA from pelleted virus was analyzed by Northern blot with a leader-specific probe. RNA ladder (GIBCO) (M) and RNA from virions at passage 1 (lane a) were used as markers. Lanes b and c, RNA extracted from virus pelleted through 31 and 15% sucrose cushions, respectively. Lanes d and e, RNA extracted from virus separated in a continuous 15–42% sucrose gradient, corresponding to fractions with densities of 1.20 and 1.15 g/cm^3, respectively.

tion with primers 3 and 4 gave rise to a unique PCR product of 2.8 kb specific of DI-RNAs. As expected, no amplification of genomic RNA was observed with these primers, due to the large predicted size of the PCR product (12 kb). With primers 5 and 6 two different PCR products of 3.5 and 4.6-kb were obtained from p41 RNA. The 4.6-kb product was also obtained from passage 1 RNA, indicating that it could be derived from the helper RNA present in the DI-RNA stock (without excluding that it could be also present in one DI RNA). PCR amplification with primers 7 and 8 of genomic RNA from passage 1 did not result in a band since they map 9.5 kb apart in the full-length TGEV genome. In contrast, two dominant bands of 1.9 and 2.1 kb were generated by PCR from passage 41 RNA. In addition, minor less abundant bands were observed in agarose gels around the 1.9-kb band, indicating the presence of high heterogeneity in this fragment (results not shown). To assign fragments to the defective genomes, RNA from passage 41-purified virus was resolved in agarose gels until a clear separation between genomic, DI-A, DI-B, and DI-C RNAs was achieved. Bands containing each of these four RNAs were independently cut and used as templates for RT-PCR amplification with primers 8 and 9 (Table 1). No amplification occurred from genomic RNA. A predominant PCR product of 1.1 kb was obtained from the amplification of either DI-A or DI-B RNAs, although other less abundant DNAs of varying size were also detected with DI-B RNA, indicating again the presence of heterogeneity at the 3′ end of DI-B RNA. DI-C RNA amplification resulted in a 1.4-kb product. These results allowed the assignment of cloned fragments of 1.9- and 2.1-kb fragments to DI-B and DI-C, respectively. The fragments of 3.5 and 4.6 kb obtained using primers 5 and 6 were assigned to D-RNAs C and B, respectively, since the sum of fragments *a* to *d* resulting from this assignment accounts for the total size observed for DI-B and DI-C RNAs (Fig. 1). The assignment of these fragments was subsequently confirmed by Northern hybridization using oligonucleotides that mapped in the regions of DI-B not present in DI-C, and vice versa (data not shown).

The four DI-C PCR-derived overlapping fragments, *a* (1.9 kb), *b* (2.8 kb), *c* (3.5 kb), and *d* (2.1 kb) were cloned, and at least two clones from independent RT-PCR reactions were sequenced for each fragment. The sequences were compared with those of strains PUR46-PAR (Eleouet *et al.,* 1995) and PUR46-MAD of TGEV (Gebauer *et al.,* 1991; Sánchez *et al.,* 1995, 1992; A. Méndez, M. L. Ballesteros, and L. Enjuanes, unpublished results). Sequencing data indicated that DI-C RNA was composed of four noncontiguous regions of the full-length genome (Fig. 4A): (I) the 5′-terminal 2144 nt of the genome; (II) an internal region of 4540 nt corresponding to positions 12195 to 16734 in the TGEV genome which comprises the ORF1a–ORF1b overlapping region; (III) a region of 2531 nt corresponding to positions 17843 to 20372 of the genome, comprising the 3′ half of ORF1b and the first 8 nt from the S gene; and (IV) the 3′-terminal 493 nt. The sequence of DI-C RNA showed 14 nucleotide differences with the TGEV PUR46-PAR strain (Table 2). The areas containing these differences were then sequenced in the strain PUR46-MAD (i.e., the DI parental virus). The sequences of DI-C and PUR46-MAD were identical with the exception of three nucleotide differences (Table 2) and one nucleotide insertion at the 3′ end of the genome (nt 9189), where no ORF was identified in DI-B or DI-C genomes.

DI-B genome is closely related to DI-C (Fig. 4B). Both D-RNAs have identical region I. The second continuous region of DI-B includes a few nucleotides of ORF1a, the complete ORF1b, and a few nucleotides of the S gene.

TABLE 1

Characteristics of the Primers Used to Amplify DI RNA by RT-PCR

Oligonucleotide	Sequence[a]	Polarity	Coronavirus[b] ORF	Binding site in DI-C	Restriction site
1	5′ GTGAGTGTAGCGTGGCTATATCTCTTC 3′	+	TGEV leader	15–41	—
2	5′ CCGTTGTGGTGTCACATTAAC 3′	−	FIPV ORF1a	1874–1887	—
3	5′ GCCTCTAGAGGAGCTTTGTGGTTCACTCACAC 3′	+	TGEV ORF1a	1524–1550	*Xba*I
4	5′ GCTCTAGAGCGTTTGAATCAACCCCCAAAAGC 3′	−	TGEV ORF1b	4365–4389	*Xba*I
5	5′ GGAATTCCGGGACTATCCTAAGTGTG 3′	+	HCV229E ORF1b	4097–4114	*Eco*RI
6	5′ GGAATTCCAGCAATACTATTATCAA 3′	−	TGEV ORF1b	7633–7652	*Eco*RI
7	5′ TTGATAATAGTATTGCTGGC 3′	+	TGEV ORF1b	7633–7650	—
8	5′ GGACTAGTATCACTATCAAAAGG 3′	−	TGEV 3′ UTR	9691–9707	*Spe*I
9	5′ GATGGATGTTGTGGTGTGAG 3′	+	TGEV ORF1b	8251–8270	—

[a] Underlined sequences indicate restriction endonuclease cleavage sites that were included to facilitate cloning of the amplified fragments.
[b] The position of the oligonucleotide within the corresponding ORF is described under Materials and Methods.

This second region is equivalent to the sum of DI-C region II, the sequences located in the *wt* genome between DI-C regions II and III (δ, 1109 nt), and DI-C region III. The deletion between regions III and IV has a variable size starting at the 5′ end of S gene and ending at ORF 7 (Fig. 5A). The most abundant DI-B clone has 6 nt of the S gene.

According to the sequences assigned to either DI-C or DI-B RNAs these genomes have predicted large ORFs of 6370 and 10003 nt, respectively, at nt 315 of wt virus. The large ORF of DI-C RNA ends at the stop codon generated in the rearrangement site at position 6685, where the ORF1b internal deletion took place, while the large ORF of DI-B ends at the stop codon of ORF1b, close to the beginning of the S gene. In addition, in both defective RNAs there are two other small ORFs, one encoding a 3 amino acids peptide, which is identical in both defective RNAs, and another encoding a 16 amino acids peptide in DI-C, and a peptide of variable length in the DI-B (Figs. 4 and 5). Although there are three potential ORFs only two

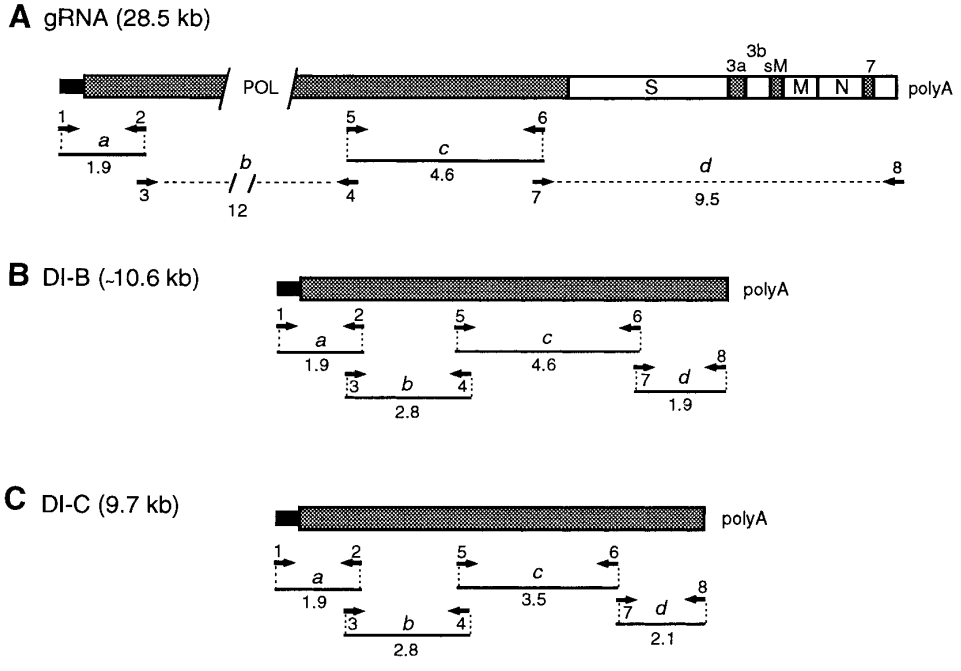

FIG. 3. Strategy to clone DI-B and DI-C defective RNAs. Schematic representation of the cDNA fragments obtained by RT-PCR amplification using as templates full-length genomic RNA (A), subgenomic DI-B (B), and subgenomic DI-C (C). Discontinuous lines indicate absence of the predicted fragment due to its large size. The subgenomic RNAs were cloned into four partially overlapping fragments (*a, b, c,* and *d*) represented by thin lines; numbers under these lines indicate the size of the fragment in kb, as determined by agarose gel electrophoresis. The primers used and their polarity are indicated by arrows and numbers close to them. Primer sequences are shown in Table 1. Stippled and open boxes in (A) indicate the relative position of viral genes: pol, polymerase; S, M, sM, and N, structural genes; 3a, 3b, and 7, small ORFs. Thin full rectangles indicate leader sequences.

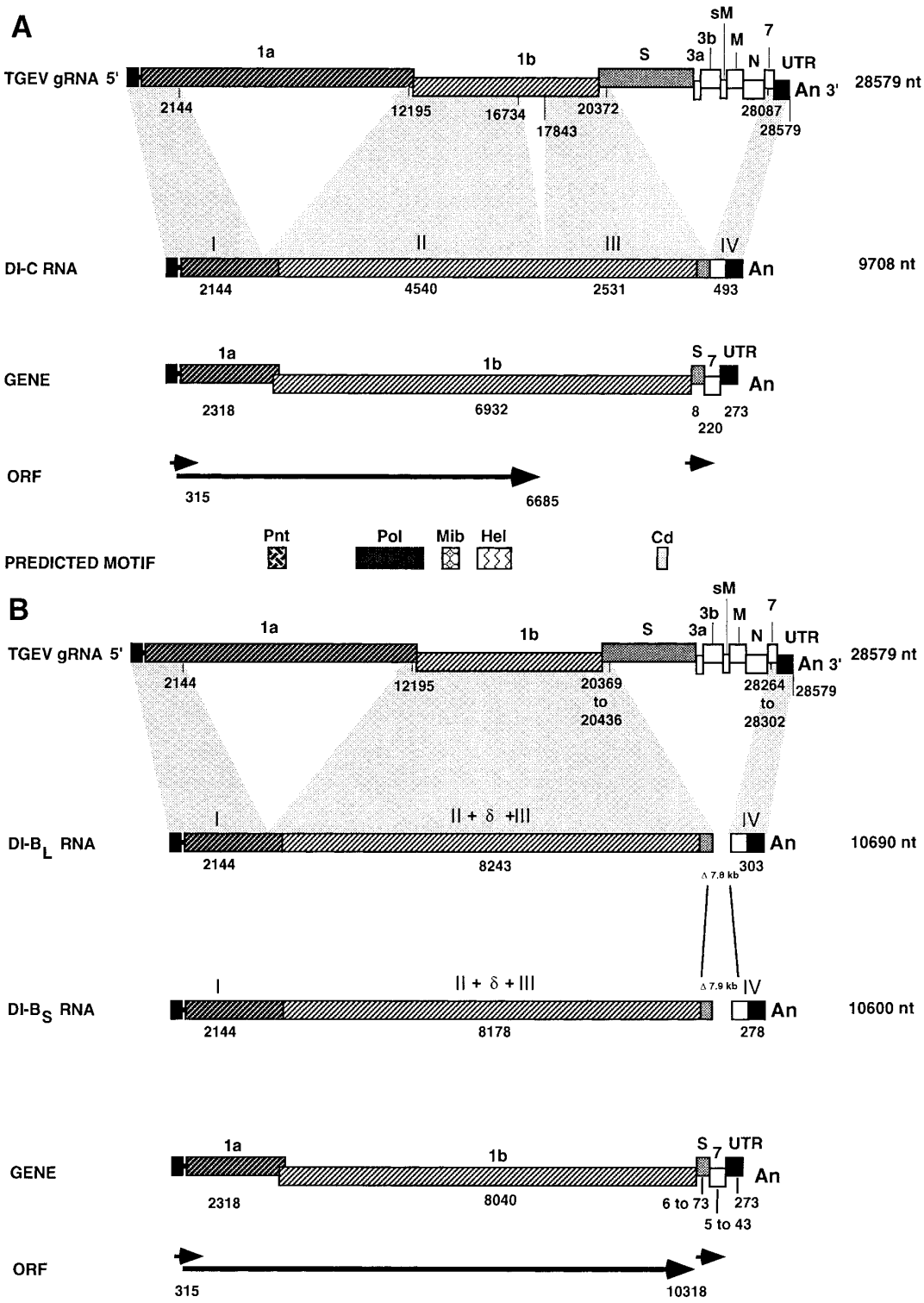

FIG. 4. Diagram summarizing the structure of TGEV DI RNAs. The structure of defective RNAs DI-C (A) or DI-B (B) is shown. The overall genome length is shown to the right of the boxes. (A) Defective DI-C RNA contains four discontinuous regions (I, II, III, and IV) of the TGEV genome. They comprise 2.1 kb from the 5′ end, almost complete ORF 1b including the 1a–1b overlapping region, the beginning of S gene, incomplete ORF 7, and the 3′ UTR. Letters or numbers above the first box indicate viral genes. Numbers below this box indicate positions of the flanking nt in the TGEV genome, taking as a reference the sequence of TGEV PUR46-PAR strain (Eleouet *et al.,* 1995). On the second box, the length of the four discontinuous regions is indicated in nt. On the third box, the number of nt derived from each viral gene is indicated, taking into account the 43 nt overlap between ORF1a and ORF1b. Predicted ORFs are indicated by arrows or arrowheads as determined by computer analysis. Pnt, pseudoknot; Pol, polymerase; Mib, metal ion binding; Hel, helicase; Cd, conserved domain. (B) DI-B contains three discontinuous regions of the genome. These regions have been designated I, (II + δ + III), and IV. Region I has 2144 nt. The second region (8178 nt) includes regions II and III of DI-C, plus the sequences (δ) located between regions II and III in the *wt* genome. Region IV has a variable size ranging from 278 to 303 nt. Description of numbers and arrows is as in (A). DI-B shows heterogeneity in the deletion between regions (II + δ + III) and IV. The structures of the two DI-B genomes with the largest (DI-B_L) and smallest (DI-B_S) DI-B RNAs are shown.

TABLE 2

Nucleotide Differences between DI-C and PUR46 RNA

Nucleotide position in DI-C	DI-C	PUR46[a] MAD	PUR46[b] PAR
637	T	G	G
2029	T	T	C
3499	G	G	A
4762	G	G	A
5146–47	TA	TA	—
5375–76	—	—	T
5607–08	—	—	A
6089	C	C	G
6397	A	G	G
6485	A	G	G
7315	G	G	A
8417–18	AT	AT	TA
8433–34	TG	TG	GT
9189–90	CC[c]	C	C

[a] Nucleotide positions differing between DI-C and PUR46-PAR (Eleouet *et al.,* 1995) were also determined on PUR46-MAD by RT-PCR and the fmol DNA Sequencing System.

[b] As reported by Eleouet *et al.* (1995).

[c] Sequence from DI-C nt 9190 was decided after many sequencing attempts by different procedures. Nevertheless, it is not considered definitive.

consensus promoter sequences CUAAAC (Rasschaert *et al.,* 1987) have been found along the full length of DI-B and DI-C. One consensus sequence precedes the first two ORFs, while the second one is close to the start of S gene.

The sequence of the 5′ end was determined by direct RNA sequencing and primer extension (C. Smerdou and L. Enjuanes, manuscript in preparation). This sequence was 5′-NCUUUUAAAG-3′. The nature of the first nt remains to be determined. The 3′ ends of *wt* TGEV and the three DI RNAs (A, B, and C) were polyadenylated, because all these RNAs could bind to Poly(A)Ttract paramagnetic particles (results not shown). The sequence of the 3′ end preceding the poly(A) was determined in four different clones. Both DI-B and DI-C contained the same 3′ end sequence (TAG-TGATACA$_n$) as the parental virus PUR46-MAD.

High heterogeneity at the most 3′-end deletion of defective genomes

RNA fragments including the junction between regions III and IV from DI-B and DI-C were RT-PCR amplified using primers 8 and 9 (Table 1) and cloned into pGEM-T plasmids. Twenty two and 5 clones derived from DI-B and DI-C, respectively, were sequenced. All DI-C-derived fragments had identical sequence (Fig. 5B). In contrast, the 22 clones derived from DI-B RNA showed up to 9 different sequences with variable lengths (Fig. 5A). A predominant sequence, however, was present in 9 of the 22 studied clones. The sequence AAACT was found flanking the deletions.

Structure of DI-A

DI-A contains a defective genome of 22 kb. To determine the regions of the *wt* genome involved in its formation Northern blot analysis with a collection of 14 oligonucleotides (Fig. 6) complementary to the leader, the 3′ UTR, and 5 ORFs was performed. DI-A RNA hybridized with all the oligonucleotides complementary to the leader and 3′ UTR, and to ORFs 1a and 1b, but not with the oligonucleotides specific for the structural genes, indicating that DI-A probably contains the complete ORFs 1a and 1b, but lacks the genes coding for the structural proteins and possibly for the other nonstructural proteins (see Fig. 6).

DISCUSSION

The isolation and characterization of three DI RNAs derived from TGEV is described. Two of these RNAs, of 9.7 and around 10.6 kb, are defective interfering RNAs that are composed of four and three discontinuous genomic regions, respectively. In addition, DI-B RNA shows high heterogeneity at the 3′ end.

The isolation of defective genomes was attempted using a variety of conditions: infecting at high m.o.i. with cloned and uncloned stocks of TGEV and PRCV, and using two types of cell lines (ST and IPEC-1). In most cases virus was passaged more than 35 times. Nevertheless, defective RNAs were detected only during the passage of one TGEV strain (PUR46-MAD) in ST cells. This low frequency of DI isolation using TGEV is in contrast with the high frequency of DI genome generation during infection with the murine coronavirus MHV (Makino *et al.,* 1984; Van der Most *et al.,* 1991). Both generation of DI and recombination seem to happen at a lower frequency in TGEV (Ballesteros *et al.,* 1995) than in MHV (Lai, 1990), possibly due to a higher accuracy in the replication of TGEV RNA.

The amount of both full-length and subgenomic viral mRNAs decreased in the presence of the identified defective genomes, indicating that one or more of these defective genomes interferes with TGEV replication. Accordingly, titers of TGEV in the absence of the DI genomes ranged between 1 and 10×10^8 PFU/ml while in their presence titers lowered 40-fold.

DI-B and DI-C genomes, of 10.6 and 9.7 kb, respectively, consisted of three and four noncontiguous genomic regions, including the 5′ end, all or most of the ORF1b (in DI-B or DI-C, respectively), the beginning of the S gene, and the 3′ end of the genomic RNA. This structure closely resembles that observed for an IBV-derived DI genome (CD-91) of 9.1 kb (Penzes *et al.,* 1994). Similarly to IBV CD-91 RNA, TGEV DI-B and DI-C RNAs contain ORF1 sequences, including the predicted motifs for frame shifting (with the slippery sequence UUUAAAC and a pseudoknot structure located 3′ to this sequence), and encoding for the polymerase, metal ion binding, and

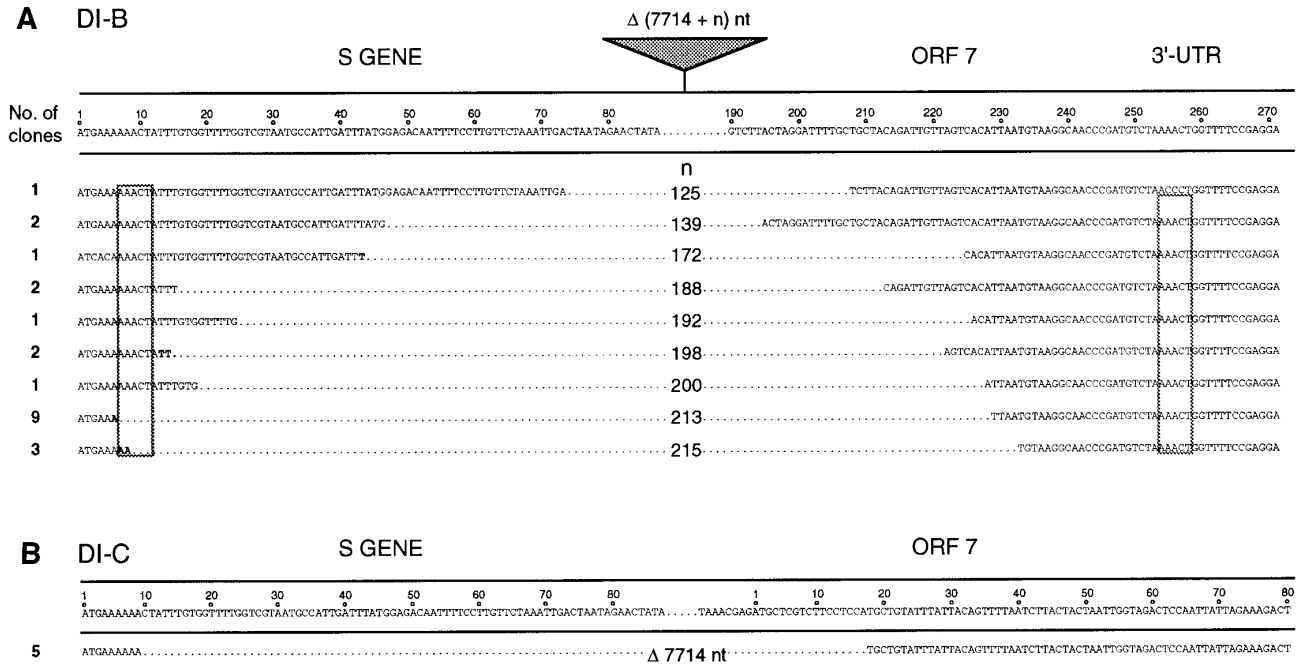

FIG. 5. Sequences flanking the 3′ end deletion. An alignment of the sequences flanking the deletions (Δ) of the different DI-B (A) and DI-C (B) sequenced genomes is shown. The number of clones giving the same sequence are indicated in the column to the left. All DI-C sequenced clones showed the same deletion of 7714 nt between regions III and IV. In contrast, DI-B genomes showed a larger deletion with variable size. Numbers above the nucleotide sequences indicate residue position in the indicated ORF. n, number of nucleotides to be added to 7714 to make the size of the deletion between regions III and IV. Small direct repeats flanking the deletions are indicated inside the squares. Letters in bold refer to nucleotides that could be assigned either to the beginning or to the end of the deletion.

helicase (Bredenbeek *et al.,* 1990; Eleouet *et al.,* 1995; Herold *et al.,* 1993; Lee *et al.,* 1991; Penzes *et al.,* 1994). TGEV DI RNAs also must contain packaging signals since they were efficiently passaged in cell culture and were present in purified virion preparations. In contrast to the structure of TGEV DI RNAs, small DI RNAs derived either from MHV A59 (DI-a, 5.5 kb) (Van der Most *et al.,* 1991) or MHV JHM (DIssF, 3.4 kb) (Kim and Makino, 1995; Makino *et al.,* 1990) contain the packaging signal but not the polymerase, metal ion, and helicase motifs. Albeit the presence in the three defective RNAs of an ORF containing the polymerase, metal ion binding, and heli-

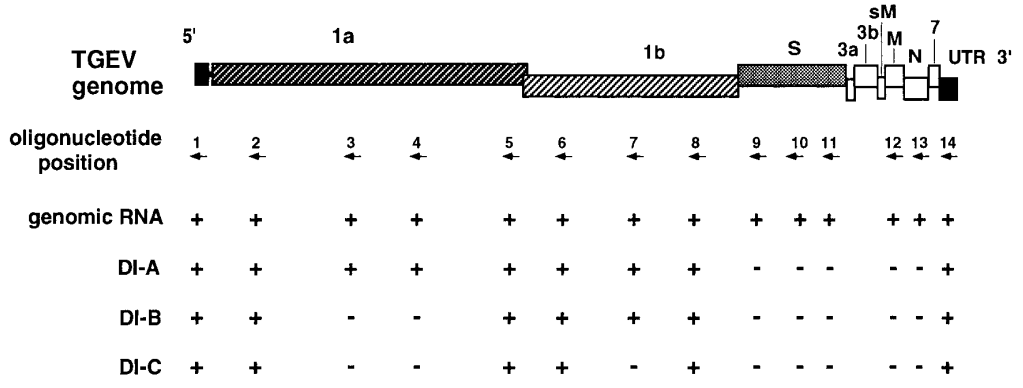

FIG. 6. Mapping of DI-A by Northern hybridization. RNA from PUR46 passage 41 was electrophoresed in agarose gel until complete separation of genomic RNA and defective RNAs A, B, and C was achieved. RNA was probed with various ^{32}P-labeled oligonucleotides that hybridized to the *wt* genome or to the DI RNAs (+) or that did not hybridized to these RNAs (−). The approximate localization of the complementary sequence recognized by the oligonucleotides in the *wt* genome is indicated by arrows. Letters and numbers above boxes indicate viral genes in the *wt* genome. The oligonucleotide (1) complementary to the leader sequence maps at 66 to 91 nucleotides from the TGEV genome 5′-end. The seven oligonucleotides (2 to 8) complementary to gene 1 map at positions: 2151 to 2170; 6121 to 6140; 8684 to 8703; 12261 to 12280; 14148 to 14167; 17363 to 17381; and 18792 to 18811, from the 5′-end, respectively. The oligonucleotides used to determine the presence of the structural genes (9 to 13) map at positions: 1055 to 1074; 1980 to 1999; and 3600 to 3619, from the 5′-end of S gene; 97 to 116, from the 5′-end of M gene; and 5 to 24, from the 5′-end of N gene, respectively. The oligonucleotide mapping in the UTR (number 14) corresponds to nucleotides 28524 to 28543 from the 5′-end of genomic RNA. All probes hybridized to *wt* genomic RNA and gave the expected results with DI-B and DI-C RNAs.

case sequence motifs, these subgenomic RNAs do not seem to have self-replicating activity, because they were lost after infecting at low m.o.i. However, DI-A RNA in contrast to DI-B and DI-C RNAs was not lost after passage into two other cell lines (IPEC and PM) under the same conditions. Initial virus passage on these cell lines involved a reduction in the m.o.i., due to the lower virus titers obtained in the first passages. It is tempting to speculate that DI-A RNA could self-replicate in those cells that did not receive helper virus and DI-A particles could be recovered after subsequent infection from a helper virus shed by neighboring cells. In contrast, DI-B and DI-C genomes probably were degraded by the time the helper virus infected the cells carrying them. The inability of DI-B and DI-C to self-replicate is likely due to the absence in these genomes of ORF1a sequences, which encode several protease activities that are required to cleave the ORF1-coded polyprotein in a functional replicase (Dong and Baker, 1994; Snijder *et al.,* 1994). DI-A RNA, however, does contain ORF1a sequences (determined by Northern blot analysis) and might be functionally similar to the self-replicating DIssA of MHV (Makino and Lai, 1989; Makino *et al.,* 1988b). Further work is required to confirm this hypothesis.

The junction between TGEV DI regions III and IV is highly variable because 10 different sequences were identified in the 27 clones that were sequenced. In all of these fragments a deletion is present at the beginning of S gene, starting from nt 7 to nt 74. These deletions end at ORF 7, from nt 18 to nt 233, almost 8 kb apart. It is striking that in the generation of both European and American PRCVs in field conditions four different deletions have been identified, and all of them also occurred at the beginning of S gene, in positions ranging from nt 45 to nt 745 (Sánchez *et al.,* 1992; Vaughn *et al.,* 1994; Wesley *et al.,* 1991). In PRCV the size of the deletion ranges from 622 to 682 nt, which is quite different from the approximately 8 kb deleted in this region of TGEV to generate DIs. In addition, it has been reported (Ballesteros *et al.,* 1995) that during the selection of TGEV recombinants by coinfecting cells with enteric and respiratory strains of TGEV about 25% of the recombination events occurred between nt 655 and 1755 of the S gene, when the selective pressure favored crossing-over between nt 1755 of the S gene and ORF1 sequences located more than 7 kb upstream. Taken together these data suggest that the 5′ half of the S gene is an area of high recombination frequency. However, these observations may simply reflect a selection advantage. Targeted recombination in the S gene seems unlikely because of the lack of specific features close to the recombination spot, such as repeated sequences or RNA secondary structure. Nevertheless, a role in recombination of direct repeats of limited length, as those present in areas flanking the DI-B 3′ deletion (Fig. 5) cannot be completely excluded.

DI-C was more abundant than DI-B, probably reflecting the higher replication efficiency and/or stability of DI-C. It could be speculated that sequences from ORF7, which are present in DI-C but missing from DI-B, are responsible for this effect. The observation that DI-B RNA is larger, more heterogeneous, and less abundant than DI-C RNA could suggest that DI-B is an intermediate in the generation of DI-C. However, this is unlikely because DI-C contains sequences, located in the junction between regions III and IV, that are missing in DI-B (Fig. 5).

Coronavirus N protein has been implicated in the replication and encapsidation of genomic RNA (Spaan *et al.,* 1988). TGEV DI includes the 3′ end of *wt* genome, but not N gene sequences, since in this virus N gene is not the most 3′ end gene. In contrast, MHV- and IBV-derived DIs have incorporated N gene sequences, present at the most 3′ end of these viruses (Van der Most *et al.,* 1992; Penzes *et al.,* 1994). These data indicate that N gene sequences *in cis* are neither essential for the replication of coronavirus defective genomes nor for their packaging.

The three DI genomes described in this report must contain a packaging signal since they are efficiently maintained by passage in tissue culture and are present in purified virion preparations. It has been shown that TGEV *wt* and DI genomes do not coencapsidate in the same virion. *A priori,* it is not possible to exclude that more than one DI-B or DI-C genome could have been encapsidated into the same particle. In fact, in other viral systems (i.e., adenovirus and SV-40) genomes of up to 105% of the *wt* size can be encapsidated (Bett *et al.,* 1993; Tai *et al.,* 1972), and three copies of DI-C RNA make 29.1 kb, a size representing 102% of the full-length TGEV genome. Nevertheless, encapsidation of several subgenomic RNAs into one virion seems very unlikely because virions with DI-B and DI-C genomes had a significantly lower density than the *wt* genome in sucrose gradients.

The isolated DI RNAs can be very useful to manipulate coronavirus genomes in order to study gene function, the molecular basis of porcine coronavirus tropism, to interfere with virus replication, or to design new expression vectors based on defective coronavirus genomes.

ACKNOWLEDGMENTS

This work has been supported by grants from the Consejo Superior de Investigaciones Científicas, the Comisión Interministerial de Ciencia y Tecnología, La Consejería de Educación y Cultura de la Comunidad de Madrid from Spain, Laboratorios Sobrino (Cyanamid), and the European Union (Projects Science and Biotech). A.M. received a fellowship from the Department of Education and Culture from the Community of Madrid, and F.G. received a fellowship provided by the Spanish Ministry of Education and Science. A.I. received a fellowship from the Department of Education, University and Research of the Gobierno Vasco (Spain).

REFERENCES

Ballesteros, M. L., Sánchez, C. M., Méndez, A., and Enjuanes, L. (1995). Recombination between transmissible gastroenteritis coronavirus related isolates which differ in tropism and virulence. *Adv. Exp. Med. Biol.* 380, 557–562.

Baric, R. S., Stohlman, S. A., and Lai, M. M. C. (1983). Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chains. *J. Virol.* 48, 633–640.

Barrett, A. D. T., and Dimmock, N. J. (1986). Defective interfering viruses and infections of animals. *Curr. Topics Microbiol. Immunol.* 128, 55–84.

Berschneider, H. M., and Powell, D. W. (1992). Fibroblasts modulate intestinal secretory responses to inflammatory mediators. *J. Clin. Invest.* 89, 484–489.

Bett, A. J., Prevec, L., and Graham, F. L. (1993). Packaging capacity and stability of human Adenovirus type 5 vectors. *J. Virol.* 67, 5911–5921.

Boursnell, M. E. G., Brown, T. D. K., Foulds, I. J., Green, P. F., Tomley, F. M., and Binns, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* 68, 57–77.

Bredenbeek, P. J., Pachuk, C. J., Noten, A. F. H., Charite, J., Luytjes, W., Weiss, S. R., and Spaan, W. J. M. (1990). The primary structure and expression of the 2nd open reading frame of the polymerase gene of the coronavirus MHV-A59- A highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. *Nucleic Acids Res.* 18, 1825–1832.

Brian, D. A., Chang, R. Y., Hofmann, M. A., and Sethna, P. B. (1994). Role of subgenomic minus strand RNA in coronavirus replication. *Arch. Virol.* 9(suppl.), 173–180.

Cavanagh, D., Brian, D. A., Enjuanes, L., Holmes, K. V., Lai, M. M. C., Laude, H., Siddell, S. G., Spaan, W., Taguchi, F., and Talbot, P. (1993). Revision of the taxonomy of the *Coronavirus, Torovirus,* and *Arterivirus* genera. *Arch. Virol.* 135, 227–237.

Cavanagh, D., Brian, D. A., Enjuanes, L., Holmes, K. V., Lai, M. M. C., Laude, H., Siddell, S. G., Spaan, W. J. M., Taguchi, F., and Talbot, P. J. (1990). Recomendations of the Coronavirus Study Group for the nomenclature of the structural proteins, mRNAs and genes of coronaviruses. *Virology* 176, 306–307.

Chang, R. Y., Hofmann, M. A., Sethna, P. B., and Brian, D. A. (1994). A cis-acting function for the coronavirus leader in defective interfering RNA replication. *J. Virol.* 68, 8223–8231.

Dong, S., and Baker, S. (1994). Determinants of the p28 cleavage site recognized by the first papain-like cysteine proteinase of murine coronavirus. *Virology* 204, 541–549.

Eleouet, J. F., Rasschaert, D., Lambert, P., Levy, L., Vende, P., and Laude, H. (1995). Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. *Virology* 206, 817–822.

Enjuanes, L., and Van der Zeijst, B. A. M. (1995). Molecular basis of transmissible gastroenteritis coronavirus (TGEV) epidemiology. *In* "The Coronaviridae" (S. G. Siddell, Ed.), pp. 337–376. Plenum, New York.

Fichot, O., and Girard, M. (1990). An improved method for sequencing of RNA templates. *Nucleic Acids Res.* 18, 6162.

Fosmire, J. A., Hwang, K., and Makino, S. (1992). Identification and characterization of a coronavirus packaging signal. *J. Virol.* 66, 3522–3530.

Gebauer, F., Posthumus, W. A. P., Correa, I., Suñé, C., Sánchez, C. M., Smerdou, C., Lenstra, J. A., Meloen, R., and Enjuanes, L. (1991). Residues involved in the formation of the antigenic sites of the S protein of transmissible gastroenteritis coronavirus. *Virology* 183, 225–238.

Herold, J., Raabe, T., Schelle-Prinz, B., and Siddell, S. G. (1993). Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. *Virology* 195, 680–691.

Hofmann, M. A., Sethna, P. B., and Brian, D. A. (1990). Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. *J. Virol.* 64, 4108–4114.

Jeong, Y. S., and Makino, S. (1992). Mechanism of coronavirus transcription: Duration of primary transcription initiation activity and effects of subgenomic RNA transcription on RNA replication. *J. Virol.* 66, 3339–3346.

Jeong, Y. S., and Makino, S. (1994). Evidence for coronavirus discontinuous transcription. *J. Virol.* 68, 2615–2623.

Joo, M., and Makino, S. (1995). The effect of two closely inserted transcription consensus sequences on coronavirus transcription. *J. Virol.* 69, 272–280.

Kim, K. H., and Makino, S. (1995). Two murine coronavirus genes suffice for viral RNA synthesis. *J. Virol.* 69, 2313–2321.

Lai, M. M. C. (1990). Coronavirus — Organization, replication and expression of genome. *Ann. Rev. Microbiol.* 44, 303–333.

Lai, M. M. C., Patton, C. D., Baric, R. S., and Stohlman, S. A. (1983). Presence of leader sequences in the mRNA of mouse hepatitis virus. *J. Virol.* 46, 1027–1033.

Lee, H. J., Shieh, C. K., Gorbalenya, A. E., Koonin, E. V., Lamonica, N., Tuler, J., Bagdzhadzhyan, A., and Lai, M. M. C. (1991). The complete sequence (22 Kilobases) of murine coronavirus gene-1 encoding the putative proteases and RNA polymerase. *Virology* 180, 567–582.

Liao, C. L., Zhang, X., and Lai, M. M. C. (1995). Coronavirus defective-interfering RNA as an expression vector: The generation of a pseudorecombinant mouse hepatitis virus expressing hemagglutinin-esterase. *Virology* 208, 319–327.

Lin, Y. J., and Lai, M. M. C. (1993). Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontiguous sequence for replication. *J. Virol.* 67, 6110–6118.

Makino, S., Joo, M., and Makino, J. K. (1991). A system for study of coronavirus messenger RNA synthesis: A regulated, expressed subgenomic defective interfering RNA results from intergenic site insertion. *J. Virol.* 65, 6031–6041.

Makino, S., and Lai, M. M. C. (1989). High-frequency leader sequence switching during coronavirus defective interfering RNA replication. *J. Virol.* 63, 5285–5292.

Makino, S., Shieh, C.-K., Soe, L. H., Baker, S. C., and Lai, M. C. (1988a). Primary structure and translation of a defective interfering RNA of murine coronavirus. *Virology* 166, 550–560.

Makino, S., Shieh, C. K., Keck, J. G., and Lai, M. M. C. (1988b). Defective-interfering particles of murine coronaviruses: Mechanism of synthesis of defective viral RNAs. *Virology* 163, 104–111.

Makino, S., Taguchi, F., and Fujiwara, K. (1984). Defective interfering particles of mouse hepatitis virus. *Virology* 133, 9–17.

Makino, S., Yokomori, K., and Lai, M. M. C. (1990). Analysis of efficiently packaged defective interfering RNAs of murine coronavirus — Localization of a possible RNA-packaging signal. *J. Virol.* 64, 6045–6053.

Masters, P. S., Koetzner, C. A., Kerr, C. A., and Heo, Y. (1994). Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in the coronavirus mouse hepatitis virus. *J. Virol.* 68, 328–337.

McClurkin, A. W., and Norman, J. O. (1966). Studies on transmissible gastroenteritis of swine II. Selected characteristics of a cytopathogenic virus common to five isolates from transmissible gastroenteritis. *Can. J. Comp. Vet. Sci.* 30, 190–198.

Nunn, M., and Johnson, R. H. (1979). A simple technique for establishing cell lines from porcine blood. *Australian Vet. J.* 55, 446.

Pachuk, C. J., Bredenbeek, P. J., Zoltick, P. W., Spaan, W. J. M., and Weiss, S. (1989). Molecular cloning of the gene encoding the putative polymerase of mouse hepatitis coronavirus, strain A59. *Virology* 171, 141–148.

Page, K. W., Britton, P., and Boursnell, M. E. G. (1990). Sequence analysis of the leader RNA of two porcine coronaviruses: transmissible gastroenteritis coronavirus and porcine respiratory coronavirus. *Virus Genes* 4, 289–301.

Peng, D., Koetzner, C. A., McMahon, T., Zhu, Y., and Masters, P. (1995).

Construction of murine coronavirus mutants containing interspecies chimeric nucleocapsid proteins. *J. Virol.* 69, 5475–5484.

Penzes, Z., Tibbles, K., Shaw, K., Britton, P., Brown, T. D. K., and Cavanagh, D. (1994). Characterization of a replicating and packaged defective RNA of avian coronavirus infectious bronchitis virus. *Virology* 203, 286–293.

Rasschaert, D., Gelfi, J., and Laude, H. (1987). Enteric coronavirus TGEV: partial sequence of the genomic RNA, its organization and expression. *Biochemie* 69, 591–600.

Roux, L., Simon, A. E., and Holland, J. J. (1991). Effects of defective interfering viruses on virus replication and pathogenesis *in vitro* and *in vivo*. *Adv. Virus Res.* 40, 181–211.

Saif, L. J., and Wesley, R. D. (1992). Transmissible gastroenteritis. *In* "Diseases of Swine" (A. D. Leman, B. Straw, W. L. Mengeling, S. D' Allaire, and D. J. Taylor, Eds.), pp. 362–386. Iowa State Univ. Press, Ames, IA.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular cloning: A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sánchez, C. M., Ballesteros, M. L., and Enjuanes, L. (1995). Tropism, virulence and primary genome structure in a cluster of transmissible gastroenteritis coronavirus derived from the Purdue isolate. Submitted for publication.

Sánchez, C. M., Gebauer, F., Suñe, C., Méndez, A., Dopazo, J., and Enjuanes, L. (1992). Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. *Virology* 190, 92–105.

Sánchez, C. M., Jiménez, G., Laviada, M. D., Correa, I., Suñé, C., Bullido, M. J., Gebauer, F., Smerdou, C., Callebaut, P., Escribano, J. M., and Enjuanes, L. (1990). Antigenic homology among coronaviruses related to transmissible gastroenteritis virus. *Virology* 174, 410–417.

Sawicki, S. G., and Sawicki, D. L. (1990). Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J. Virol.* 64, 1050–1056.

Schaad, M., and Baric, R. S. (1994). Genetics of mouse hepatitis virus transcription: evidence that subgenomic negative strands are functional. *J. Virol.* 68, 8169–8179.

Sethna, P. B., Hung, S. L., and Brian, D. A. (1989). Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proc. Natl. Acad. Sci. USA* 86, 5626–5630.

Sethna, P. B., Hofmann, M. A., and Brian, D. A. (1991). Minus-strand copies of replicating coronavirus mRNAs contain antileaders. *J. Virol.* 65, 320–325.

Siddell, S., Wege H., and ter Meu, V. (1983). The biology of coronaviruses. *J. Gen. Virol.* 64, 761–776.

Snijder, E. J., Wassenaar, A. L. M., and Spaan, W. J. M. (1994). Proteolytic processing of the replicase ORF1A protein of equine arteritis virus. *J. Virol.* 68, 5755–5764.

Spaan, W., Cavanagh, D., and Horzinek, M. C. (1988). Coronaviruses: Structure and genome expression. *J. Gen. Virol.* 69, 2939–2952.

Spaan, W. H., Delius, H., Skinner, M., Armstrong, J., Rottier, P., Smeekens, S., van der Zeijst, B. A. M., and Siddell, S. G. (1983). Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J.* 2, 1839–1844.

Tai, H. T., Smith, C. A., Sharp, P. A., and Vinograd, J. (1972). Sequence heterogeneity in closed simian virus 40 deoxyribonucleic acid. *J. Virol.* 9, 317–325.

Van der Most, R. G., Bredenbeek, P. J., and Spaan, W. J. M. (1991). A domain at the 3′ end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. *J. Virol.* 65, 3219–3226.

Van der Most, R. G., De Groot, R. J., and Spaan, W. J. M. (1994). Subgenomic RNA synthesis directed by a synthetic defective interfering RNA of mouse hepatitis virus: A study of coronavirus transcription initiation. *J. Virol.* 68, 3656–3666.

Van der Most, R. G., Heijnen, L., Spaan, W. J. M., and Degroot, R. J. (1992). Homologous RNA recombination allows efficient introduction of site-specific mutations into the genome of coronavirus MHV-A59 via synthetic co-replicating RNAs. *Nucleic Acids Res.* 20, 3375–3381.

Van der Most, R. G., and Spaan, W. J. M. (1995). Coronavirus replication, transcription, and RNA recombination. *In* "The Coronaviridae" (S. G. Siddell, Ed.), pp. 11–31. Plenum, New York.

Vaughn, E. M., Halbur, P. G., and Paul, P. S. (1994). Three new isolates of porcine respiratory coronavirus with various pathogenicities and spike (S) gene deletions. *J. Clin. Microbiol.* 32, 1809–1812.

Wesley, R. D., Woods, R. D., and Cheung, A. K. (1991). Genetic analysis of porcine respiratory coronavirus, an attenuated variant of transmissible gastroenteritis virus. *J. Virol.* 65, 3369–3373.

Zhang, X., and Lai, M. M. C. (1994). Unusual heterogeneity of leader-mRNA fusion in a murine coronavirus: Implications for the mechanism of RNA transcription and recombination. *J. Virol.* 68, 6626–6633.