JOURNAL OF VIROLOGY, Jan. 1990, p. 331–338
0022-538X/90/010331-08$02.00/0

Vol. 64, No. 1

A 3′-Coterminal Nested Set of Independently Transcribed mRNAs Is Generated during Berne Virus Replication

ERIC J. SNIJDER,* MARIAN C. HORZINEK, AND WILLY J. M. SPAAN

Institute of Virology, Veterinary Faculty, State University of Utrecht, Yalelaan 1, 3584 CL Utrecht, The Netherlands

Received 23 May 1989/Accepted 25 September 1989

By using poly(A)-selected RNA from Berne virus (BEV)-infected embryonic mule skin cells as a template, cDNA was prepared and cloned in plasmid pUC9. Recombinants covering a contiguous sequence of about 10 kilobases were identified. Northern (RNA) blot hybridizations with various restriction fragments from these clones showed that the five BEV mRNAs formed a 3′-coterminal nested set. Sequence analysis revealed the presence of four complete open reading frames of 4743, 699, 426, and 480 nucleotides, with initiation codons coinciding with the 5′ ends of BEV RNAs 2 through 5, respectively. By using primer extension analysis and oligonucleotide hybridizations, RNA 5 was found to be contiguous on the consensus sequence. The transcription of BEV mRNAs was studied by means of UV mapping. BEV RNAs 1, 2, and 3 were shown to be transcribed independently, which is also likely—although not rigorously proven—for RNAs 4 and 5. Upstream of the AUG codon of each open reading frame a conserved sequence pattern was observed which is postulated to function as a core promoter sequence in subgenomic RNA transcription. In the area surrounding the core promoter region of the two most abundant subgenomic BEV RNAs, a number of homologous sequence motifs were identified.

Berne virus (BEV) has been characterized as a representative of a new group of enveloped, positive-stranded RNA viruses (29), for which a family status has been proposed. The *Toroviridae* (12) contain this equine virus as the family prototype. Antigenically related bovine (Breda virus [30]) and human (1) viruses can be considered additional members to the proposed family.

Analysis of purified BEV has revealed the presence of four structural proteins: a 19-kilodalton nucleocapsid (N) protein, two nonglycosylated membrane-associated proteins of 21 (E) and 37 (M) kilodaltons, and an N-glycosylated peplomer (P) protein in the range of 80 to 100 kilodaltons, forming club-shaped projections on the virion surface (for reviews, see references 11 and 28).

The BEV genome consists of a single infectious and polyadenylated RNA molecule of at least 20 kilobases (kb) (21). In BEV-infected cells the presence of five virus-specific, polyadenylated RNAs of 0.8, 1.4, 2.1, 7.5, and ≥20 kb has been demonstrated (21). In vitro translation studies have shown that RNA 5 (0.8 kb) and RNA 3 (2.1 kb) encode the N and E proteins, respectively. Translation of RNA 2 (7.5 kb) resulted in the synthesis of a 151-kilodalton product, possibly the intracellular precursor of the peplomer protein (21).

The sizes of the translation products of the different RNA species are in good agreement with the theoretical nonoverlapping coding capacities of a 3′-coterminal nested set. To study the genome organization and replication strategy, we cloned the BEV mRNAs and used UV transcription mapping to investigate the mechanism of their synthesis. In this report we describe the organization of 10 kb (extending from the 3′ end) of the BEV genome and the presence of a 3′-coterminal nested set. In addition, we show that the BEV RNAs 1, 2, and 3 are transcribed independently.

* Corresponding author.

MATERIALS AND METHODS

Cells and virus. BEV strain P138/72 was grown in embryonic mule skin (EMS) cells as described previously (21).

Purification of genomic and virus-specific mRNAs. The purification of BEV from tissue culture medium and the isolation of genomic RNA have been described previously (21). Intracellular RNA was isolated from infected EMS cells, and mRNAs were purified by using oligo(dT)-cellulose chromatography (21). About 300 μg of poly(A)-containing RNA could be isolated at 13 h postinfection from 2.7×10^8 EMS cells infected with a multiplicity of infection of 7 (21).

cDNA synthesis and cloning. Poly(A)-selected RNA from BEV-infected EMS cells was denatured with methylmercury hydroxide (14). cDNA synthesis was primed with oligo(dT) or calf thymus pentanucleotides. Procedures for first- and second-strand syntheses were similar to those described by Gubler and Hoffman (8). The cDNA was tailed with dC residues and annealed to *Pst*I-digested, dG-tailed pUC9 DNA (Pharmacia Fine Chemicals). Annealed DNA was used for transformation of *Escherichia coli* XL1-blue (Stratagene) as described by Hanahan (10). Bacteria were grown in the presence of 50 μg of ampicillin per ml. White colonies were selected for further analysis.

Preparation of probes and hybridizations. Fragments of BEV genomic RNA and of RNA from uninfected EMS cells were prepared by incubation in 50 mM Na_2CO_3 (at 50°C for 80 min.). RNA fragments, $oligo(dT)_{12-18}$, and synthetic oligonucleotides were labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase (15). Gel-purified restriction fragments were labeled by using pentanucleotides as primers and *E. coli* DNA polymerase I as described by Feinberg and Vogelstein (4). Filters for colony hybridizations were prepared as described by Maniatis et al. (15). Colony hybridizations with RNA fragments were carried out at 42°C in 5× SSPE containing 50% formamide, 5× Denhardt reagent, 0.1% sodium dodecyl sulfate, and 100 μg of yeast tRNA per ml; SSPE and Denhardt reagent were prepared as described by Maniatis et al. (15). Hybridizations with labeled oligo(dT) (at

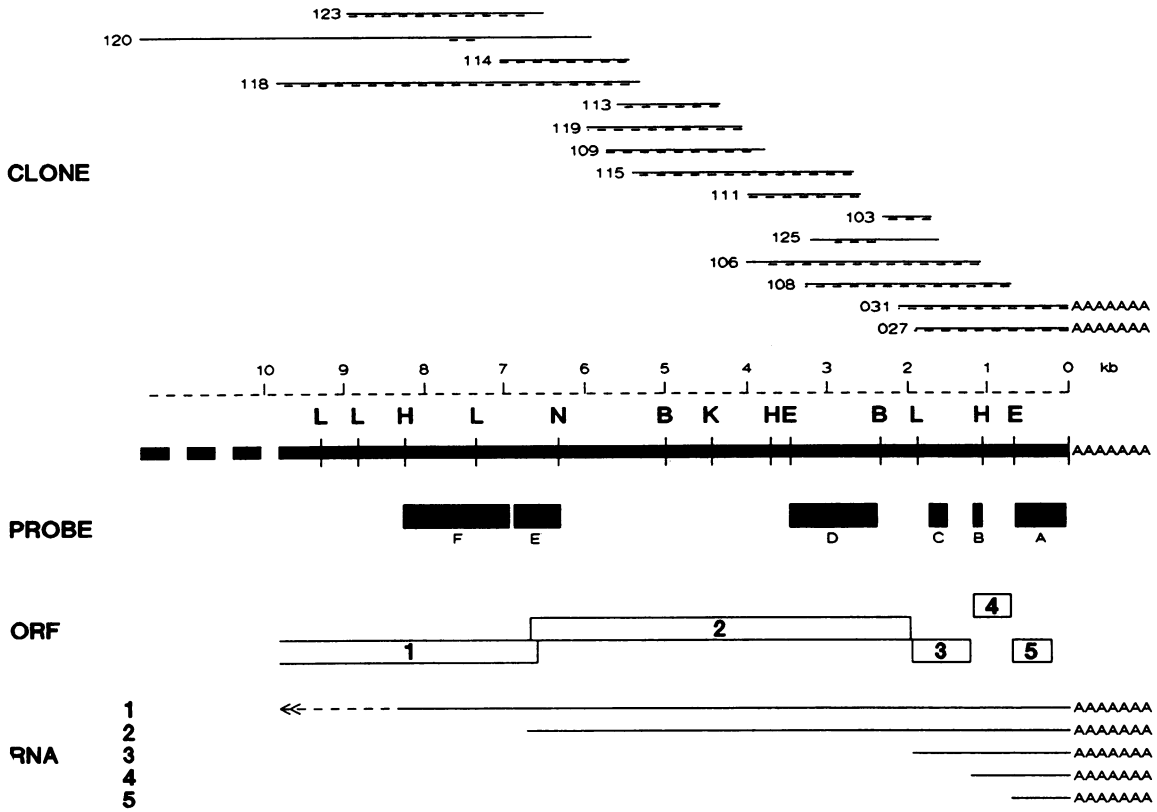

FIG. 1. Cloning and sequencing strategy of BEV mRNA sequences. A consensus restriction map of the 10 kb upstream of the poly(A) tail is presented. Restriction sites: L, *Bal*I; H, *Hin*dIII; N, *Nde*I; B, *Bam*HI; K, *Kpn*I; E, *Eco*RI. The positions of 15 cDNA clones in the consensus sequence are shown in the upper part of the figure. Dashed lines indicate the parts of clones used for sequence analysis. The lower part of the figure shows the ORFs (□), as determined by sequence analysis. The approximate mapping of the five BEV mRNAs in the consensus sequence is indicated. Restriction fragments (■) were used as probes in Northern blot hybridizations (Fig. 2). The following probes were used [numbers indicate the distance of the first nucleotide of the restriction site to the first nucleotide of the poly(A) tail]: A, *Eco*RI (678 nt) to *Hae*III (30 nt); B, *Bgl*II (1,171 nt) to *Hin*dIII (1,074 nt); C, *Rsa*I (1,756 nt) to *Taq*I (1,512 nt); D, *Eco*RI (3,478 nt) to *Pst*I (2,401 nt); E, *Pvu*II (6,889 nt) to *Nde*I (6,341 nt); F, *Hin*dIII (8,254 nt) to *Pvu*II (6,907 nt).

room temperature) and DNA probes (at 42°C) were carried out as described by Meinkoth and Wahl (17) for synthetic oligonucleotides and nick-translated probes, respectively. Northern (RNA) blot hybridization with kinase-labeled oligonucleotides (at 5°C below the melting temperature) and DNA probes were performed as described by Meinkoth and Wahl (17) and de Groot et al. (3), respectively.

Hybridizations in dried agarose gels were carried out under conditions identical to those for Northern blot hybridization with 10 ng of kinase-labeled oligonucleotide per ml of hybridization mixture.

DNA sequencing. Restriction fragments were isolated from agarose gels by binding to NA-45 paper (Schleicher & Schuell Co.) and subsequently subcloned in bacteriophage M13 vectors. Sequencing was carried out by using the dideoxynucleotide-chain termination procedure with Klenow DNA polymerase and [α-^{32}P]dATP as the label (19) or by using a bacteriophage T7 polymerase sequencing kit (Pharmacia) and [α-^{35}S]dATP as recommended by the manufacturer. Sequence data were analyzed by using the computer programs of Staden (23).

RNA electrophoresis, hybridization in gel, and Northern blot hybridization. Formaldehyde-formamide RNA electrophoresis in 1% agarose gels was performed as described by Meinkoth and Wahl (17). For direct hybridization in gel, gels were washed twice in a 10 mM phosphate buffer and dried at 60°C. Northern blotting was carried out as described by de Groot et al. (9).

UV transcription mapping. EMS cells were infected at a multiplicity of infection of 20 as described previously (21). At 8 h postinfection, dactinomycin (1 μg/ml) was added to the medium. At this stage dactinomycin does not inhibit viral replication (21). After an incubation of 30 min, the medium was removed, and cells were UV irradiated for various intervals with a dose rate of 50 erg s^{-1} mm^{-2}, as described previously for mouse hepatitis virus by Jacobs et al. (13). After irradiation, cultures were incubated for 45 min in medium containing 100 μCi of [^{3}H]uridine and 1 μg of dactinomycin per ml. A quantitative analysis of total RNA synthesis after irradiation was performed by using 2-cm^2 wells. After cell lysis, the amount of incorporated [^{3}H]uridine was determined by trichloroacetic acid precipitation and scintillation counting. To study the effect of UV irradiation on the synthesis of individual BEV RNA species, RNA was isolated and electrophoresed as described previously (21). Incorporation of label into individual RNA species was quantitated by excision of bands from dried gels and scintillation counting.

Primer extension. A kinase-labeled synthetic oligonucleotide with the sequence 5′ ATGCAACAACCTGAG 3′ (oligonucleotide A [oligo A], complementary to nucleotides 634 to 620; see Fig. 3A) (10 ng) was annealed to 0.5 μg of poly(A)-

selected RNA from BEV- and mock-infected EMS cells in the presence of 10 mM methylmercury hydroxide. After incubation for 10 min at room temperature, first-strand cDNA synthesis was performed under the same conditions as those used for the generation of the cDNA libraries. Analysis of the extension products was performed on a 6% acrylamide–urea sequencing gel, using as a marker oligo A-primed sequencing samples made from single-stranded M13 DNA containing the BEV sequence from the poly(A) tail to position 1074.

RESULTS

cDNA cloning of BEV intracellular RNAs and mapping of recombinant plasmids. Poly(A)-selected RNA from BEV-infected EMS cells was used to prepare oligo(dT)- and random-primed cDNA libraries as described in Materials and Methods. Recombinant transformants containing viral sequences were selected by colony hybridization with kinase-labeled fragments of alkali-digested genomic RNA isolated from purified BEV. Approximately 30% of the recombinant transformants from the oligo(dT)-primed cDNA library contained a virus-specific insert. None of these positive transformants hybridized to RNA fragments from mock-infected EMS cells. To map the inserts of the virus-specific cDNA clones, a hybridization with kinased oligo(dT) was employed to identify cDNA representing the 3′ end of BEV RNAs. The insert size of oligo(dT)-positive cDNA clones was determined by restriction enzyme analysis. Restriction fragments from recombinants p027 and p031, containing the largest inserts, were purified and used to probe the random-primed cDNA library (Fig. 1). The inserts of positive transformants were analyzed by restriction enzyme analysis and used to select transformants, which would extend the cloned DNA toward the 5′ end of the viral RNA. This "mRNA walking" resulted in a continuous map of over 10 kb (Fig. 1).

BEV RNAs form a 3′-coterminal nested set. BEV intracellular RNAs 3 and 5 are the most abundant viral RNAs found in BEV-infected EMS cells (21). Restriction enzyme analysis of the inserts of oligo(dT)-positive cDNA clones revealed identical restriction maps. This strongly suggested sequence homologies between the BEV RNAs 3 and 5 and, possibly, all BEV subgenomic (sg) RNAs. To verify this hypothesis, several restriction fragments were used as probes in Northern blot analyses. The position of probes A through F on the restriction map are indicated in Fig. 1. The corresponding Northern blot hybridizations to poly(A)-selected RNA from BEV-infected EMS cells are shown in Fig. 2. Probe A, containing the 3′ end of the cDNA insert of clone p031, hybridized to all sg intracellular BEV RNAs, whereas probe B, located at positions 1,074 to 1,171 nucleotides (nt) upstream of the poly(A) tail, failed to recognize RNA 5 but did hybridize to all the other RNA species. Hybridizations with probes from upstream regions were employed to map the positions of the other BEV RNAs (Fig. 2, lanes A through D). The results obtained with probes E and F, which come from a separate experiment, indicate that the region unique to RNA 1 starts at approximately 7 kb upstream of the poly(A) tail (Fig. 2, lanes E and F). The hybridization data confirmed the presence of a 3′-coterminal nested set (Fig. 1).

Due to competition for binding of the probe, the hybridization signals in Fig. 2 reflect the relative molarities of the BEV RNAs (21). The poor RNA 1 signal in lanes A through C can therefore be explained by the low relative molarity of this RNA and the fact that probes A through C did also react

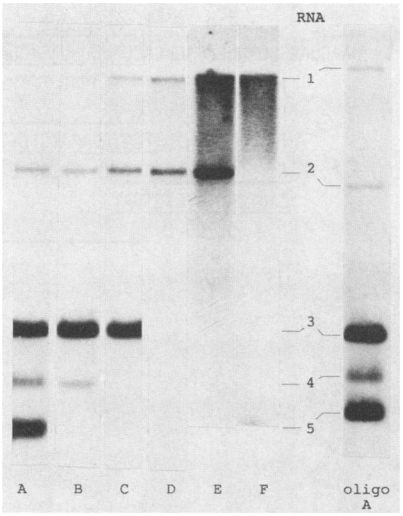

FIG. 2. Northern blot hybridizations of various restriction fragments to poly(A)-selected mRNAs from BEV-infected cells. The restriction fragments A through F (explained in Fig. 1) were used to map the BEV mRNAs, separated in a 1% agarose gel. RNAs are indicated by their numbers (21). From lanes E and F, which come from a different experiment than lanes A through D, it can be concluded that probe F is the first probe located in the unique region of RNA 1. The lane on the right shows the result of a hybridization in a dried agarose gel by using oligo A (Fig. 3) as a probe. Because of the poor blotting efficiency of RNA 1, a hybridization in gel provides a more accurate image of relative RNA molarities.

with the abundant small BEV RNA species. In addition, the blotting efficiency of large RNA molecules is notoriously poor, preventing an accurate determination of relative RNA molarities from Northern blot hybridization. To circumvent this technical problem, a direct hybridization in a dried agarose gel was performed, using as a probe an oligonucleotide from the RNA 5 region [oligo A; see below]. The result [Fig. 2; oligo(A) lane] allowed a comparison of relative RNA molarities with the values obtained previously on the basis of [^{3}H]uridine incorporation and estimated RNA lengths (21). Only the relative amounts of RNAs 3 and 5 were found to be slightly different from the values calculated before. Relative molarities of 3, 3, 39, 13, and 42% were now determined for RNAs 1 through 5, respectively.

Nucleotide sequence analysis and identification of ORFs. The 15 cDNA clones shown in the upper part of Fig. 1 were used for sequence analysis. Over a distance of about 9 kb, extending from the beginning of the poly(A) tail to the 5′ end, the nucleotide sequence was determined on at least two independent cDNA clones (Fig. 1). The consensus sequence (9,829 nt in total) revealed the presence of four open reading frames (ORFs) and the 3′ end of a potential fifth ORF (Fig. 1). Relative to the poly(A) tail, ORFs were located at positions 680 through 200 (480 nt), 1150 through 724 (426 nt), 1910 through 1211 (699 nt), and 6684 through 1941 (4,743 nt). The fifth ORF (3,026 nt so far) terminated at position 6603 and was the only one that partially overlapped one of the other reading frames.

The distances from the initiation codons of the three (small) ORFs at the 3′ end to the poly(A) tail were in good agreement with the estimated sizes of the three smallest BEV RNAs. Subtracting an average of 150 nt for the poly(A) tail, these sizes were 650, 1,250, and 1,950 nt for BEV RNAs 5, 4, and 3, respectively (21). The numbers of nucleotides computed from the consensus sequence were 680, 1,150, and

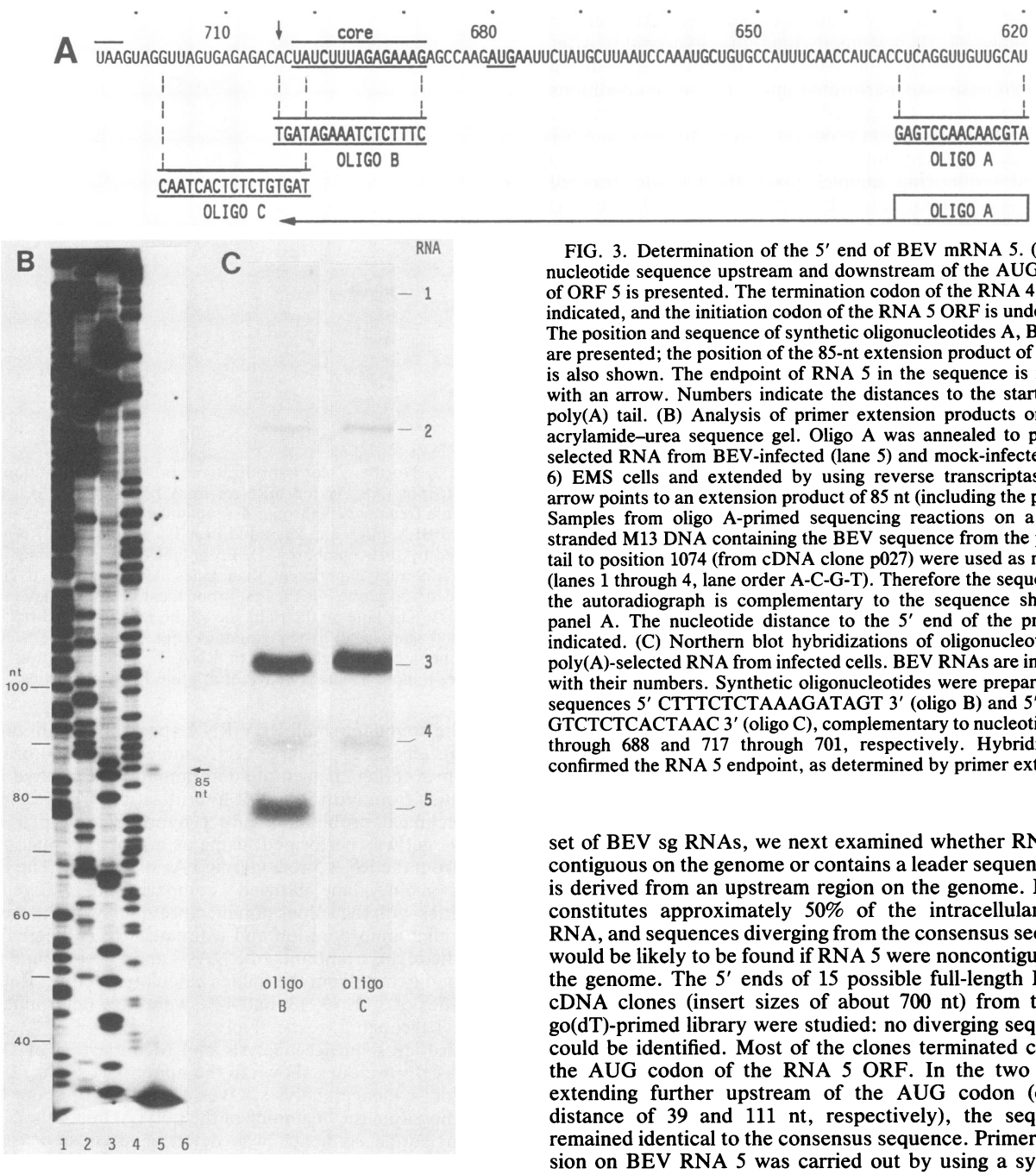

FIG. 3. Determination of the 5′ end of BEV mRNA 5. (A) The nucleotide sequence upstream and downstream of the AUG codon of ORF 5 is presented. The termination codon of the RNA 4 gene is indicated, and the initiation codon of the RNA 5 ORF is underlined. The position and sequence of synthetic oligonucleotides A, B, and C are presented; the position of the 85-nt extension product of oligo A is also shown. The endpoint of RNA 5 in the sequence is marked with an arrow. Numbers indicate the distances to the start of the poly(A) tail. (B) Analysis of primer extension products on a 6% acrylamide–urea sequence gel. Oligo A was annealed to poly(A)-selected RNA from BEV-infected (lane 5) and mock-infected (lane 6) EMS cells and extended by using reverse transcriptase. The arrow points to an extension product of 85 nt (including the primer). Samples from oligo A-primed sequencing reactions on a single-stranded M13 DNA containing the BEV sequence from the poly(A) tail to position 1074 (from cDNA clone p027) were used as markers (lanes 1 through 4, lane order A-C-G-T). Therefore the sequence on the autoradiograph is complementary to the sequence shown in panel A. The nucleotide distance to the 5′ end of the primer is indicated. (C) Northern blot hybridizations of oligonucleotides to poly(A)-selected RNA from infected cells. BEV RNAs are indicated with their numbers. Synthetic oligonucleotides were prepared with sequences 5′ CTTTCTCTAAAGATAGT 3′ (oligo B) and 5′ TAGT GTCTCTCACTAAC 3′ (oligo C), complementary to nucleotides 704 through 688 and 717 through 701, respectively. Hybridizations confirmed the RNA 5 endpoint, as determined by primer extension.

1,910, respectively. For RNA 2 we predicted a size of approximately 7.35 kb, excluding the poly(A) tail; from the sequence data a size of 6,684 nt was calculated. Full details on the nucleotide sequence of these ORFs and the characterization of their translation products will be presented elsewhere (E. J. Snijder, J. A. den Boon, W. J. M. Spaan, G. M. G. M. Verjans, M. C. Horzinek, J. Gen. Virol., in press; E. J. Snijder, J. A. den Boon, W. J. M. Spaan, M. Weiss, M. C. Horzinek, submitted for publication; E. J. Snijder, J. A. den Boon, P. J. Bredenbeck, M. C. Horzinek, R. A. Rijnbrand, W. J. M. Spaan, submitted for publication; E. J. Snijder et al., manuscript in preparation; J. A. den Boon, manuscript in preparation).

RNA 5 does not contain a diverging leader sequence. Having established the presence of a 3′-coterminal nested set of BEV sg RNAs, we next examined whether RNA 5 is contiguous on the genome or contains a leader sequence that is derived from an upstream region on the genome. RNA 5 constitutes approximately 50% of the intracellular BEV RNA, and sequences diverging from the consensus sequence would be likely to be found if RNA 5 were noncontiguous on the genome. The 5′ ends of 15 possible full-length RNA 5 cDNA clones (insert sizes of about 700 nt) from the oligo(dT)-primed library were studied: no diverging sequences could be identified. Most of the clones terminated close to the AUG codon of the RNA 5 ORF. In the two clones extending further upstream of the AUG codon (over a distance of 39 and 111 nt, respectively), the sequences remained identical to the consensus sequence. Primer extension on BEV RNA 5 was carried out by using a synthetic oligonucleotide. Kinase-labeled oligonucleotide A (Fig. 3A) was annealed to poly(A)-selected RNA from BEV-infected cells and extended by using reverse transcriptase as described in Materials and Methods. Electrophoresis on a denaturing sequence gel revealed the presence of an extension product with a length of 85 nt, mapping the 5′ end of RNA 5 at position 704, 24 nt upstream of the AUG codon (Fig. 3B). The amount of radioactivity in the 85-nt primer extension product was too low to perform a Maxam and Gilbert sequence analysis. To obtain direct evidence that the sequence of the 5′ end of RNA 5 was indeed identical to the consensus sequence up to position 704, we hybridized the intracellular BEV RNAs to synthetic oligo B (complementary to positions 704 through 688). RNA 5 did hybridize to oligo B (Fig. 3C). Because of the nested set structure, all other BEV RNAs hybridized to oligo B as well. In contrast,

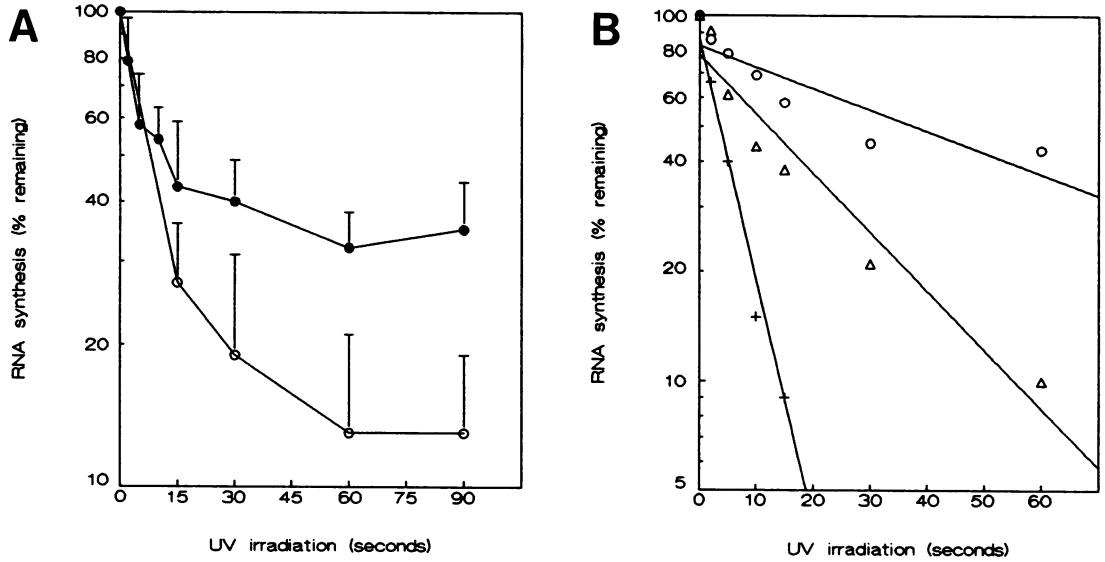

FIG. 4. Quantitative analysis of BEV mRNA UV transcription mapping. (A) Absence of repair was tested by [^{3}H]uridine labeling immediately after UV irradiation (●; 8.5 to 9.25 h postinfection) and from 45 to 90 min after irradiation (○; 9.25 to 10 h postinfection) in the presence of dactinomycin. After trichloroacetic acid precipitation, the total amount of incorporated label was determined by scintillation counting. Bars indicate the standard deviations; $n = 4$. (B) Quantitative analysis for individual BEV RNAs 1 (+), 2 (△), and 3 (○). Bands were excised from the gel shown in Fig. 5B, and incorporation was determined by scintillation counting. Graphs were fitted by linear regression analysis. Correlation coefficients were 0.99, 0.96, and 0.88 for RNAs 1, 2, and 3, respectively.

oligo C, which overlapped oligo B for 4 nt and extended in the 5′ direction, did not hybridize to RNA 5, whereas the other RNAs were found to be positive. Even under conditions of low-stringency hybridization, BEV RNA 5 remained negative with oligo C. These data illustrate that the 5′ end of RNA 5 does not diverge from the consensus sequence and that the presence of a noncontiguous leader sequence on RNA 5 is unlikely.

UV transcription mapping. The multiple virus-specific RNAs could arise either by processing of larger precursor molecules or from independent transcription units. We used UV transcription mapping to distinguish between these two possibilities. UV irradiation induces uracil dimers in RNA molecules. It is assumed that formation of one dimer will stop transcription of an RNA, that repair is slow or absent, and that the number of hits at a given dose is proportional to the time of irradiation and to the length of the template (for a review, see reference 20). The distance between the promoter and termination site of a transcription unit, i.e., the target size T, was determined by measuring the UV inactivation kinetics of the synthesis of its RNA product. Processing of one precursor into smaller RNAs would result in an equal T for all products, whereas independent initiation of RNA transcription would make UV sensitivity proportional to the physical size of the RNA transcript. EMS cells were infected at a multiplicity of infection of 20 and UV irradiated with increasing doses at 8.5 h postinfection, just before viral RNA synthesis reaches its maximum (21). The effect of UV irradiation on the total RNA synthesis during 45 min postirradiation was determined by labeling the cells with [^{3}H]uridine in the presence of dactinomycin. A nonlinear dose-response curve was obtained (Fig. 4A) which was similar to the curve obtained for the UV inactivation of the RNA synthesis of the coronavirus MHV-A59 (13). The nonlinearity of the curve already suggested the presence of a composite template. To demonstrate the absence of repair, the UV-irradiated infected cells were labeled between 45 and 90 min after irradiation. A similar curve, but with decreased levels of remaining RNA synthesis, was obtained, indicating the absence of repair (Fig. 4A).

Analysis of the UV inactivation of the synthesis of the individual RNAs was performed by agarose gel electrophoresis of labeled BEV RNAs. The kinetics of UV inactivation of RNA synthesis were shown to be related to the size of the RNA: the smaller the RNA, the more resistant to UV was its synthesis (Fig. 5A).

Because of the relatively low molarity of RNA 4 and the small size of RNA 5, the amount of [^{3}H]uridine incorporated into these RNA species was low. This prevented a solid quantitative analysis of the data from RNAs 4 and 5. However, target sizes could be calculated for the three larger RNA species. For this purpose, bands were excised from the gel shown in Fig. 5B, and radioactivity was quantitated by scintillation counting (Fig. 4B). It can be concluded from Fig. 4B that the logarithms of the rate of remaining synthesis of RNAs 1, 2, and 3 are linearly correlated with the UV doses, indicating that the Poisson distribution applies. The slope of the curve decreased with the size of the RNA. The value of the slope was used to calculate the target sizes for RNAs 1, 2 and 3 (Table 1). The target sizes of these RNAs were very close to their physical sizes.

Identification of possible RNA promoter sites. The independent initiation of transcription, as demonstrated by UV transcription mapping, suggests the presence of four RNA promoters on the complementary template. ORFs 2 through 5 (Fig. 1) were each separated by short intergenic regions. The 5′ end of RNA 5 was mapped at position 704 of the consensus sequence (see above), located in the middle of the intergenic region between ORFs 4 and 5. The nucleotide sequence surrounding the potential initiation site of RNA 5 showed a high degree of similarity to sequences in areas of the other intergenic regions (Fig. 6A). A potential core promoter sequence of 5′ UCUUUAGA 3′ could be identified; especially the areas surrounding the core promoter sequences of RNA 3 and 5 showed striking sequence homologies. An alignment of these regions is presented in Fig. 6B.

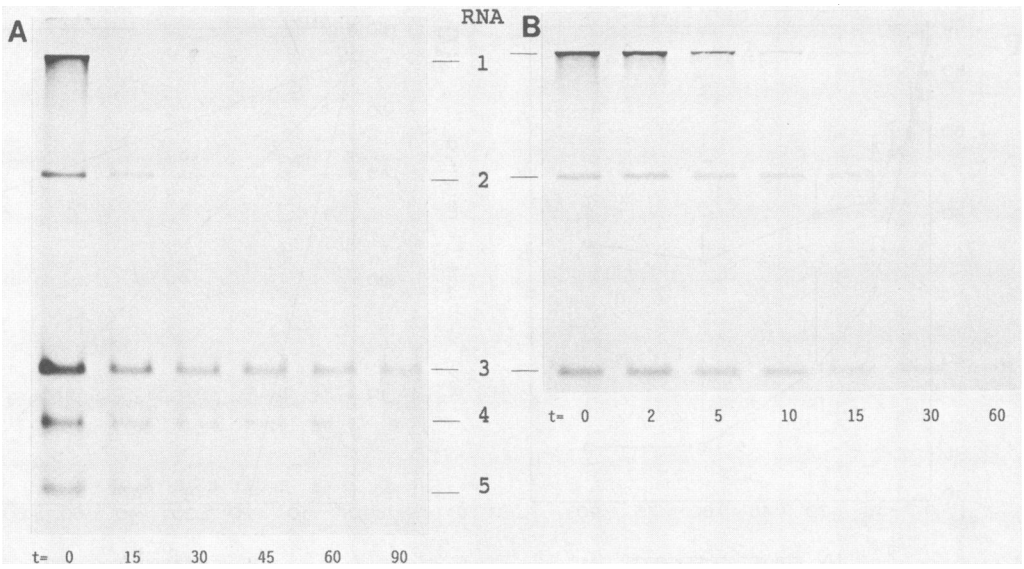

FIG. 5. Electrophoretic analysis of BEV mRNA UV transcription mapping. (A) EMS cells were infected at a multiplicity of infection of 20 and UV irradiated at 8.5 h postinfection for 0, 15, 30, 45, 60, and 90 s. Immediately after UV treatment, cells were labeled with [^{3}H]uridine in the presence of dactinomycin for 45 min. RNA was analyzed on a 1% agarose gel as described before (21). (B) UV transcription mapping was performed as described for panel A, but shorter irradiation times were used (0, 2, 5, 10, 15, 30, and 60 s) for a better analysis of UV inactivation kinetics of the larger RNA species. Bands were excised from gel and used for a quantitative analysis, as described in the text and in the legend to Fig. 4B.

In addition to the putative core promoter region, a second box of homology was detected. A 14-nt sequence motif, starting at 124 nt upstream of the ORF 3 AUG codon, was also present at 111 nt upstream of the ORF 5 initiation codon. In the area surrounding the two boxes of homology, additional sequence similarities were found, including an (A+U)-rich stretch starting approximately 50 nt upstream of the core promoter sequence. The 3′ end of ORF 1 did overlap with ORF 2, meaning that an intergenic region was formally absent. However, a sequence that was similar to the core promoter region of the other RNAs was identified at approximately 50 nt upstream of the ORF 2 AUG codon (Fig. 6A).

DISCUSSION

Studies on torovirus replication are hampered by the difficulty of obtaining sufficient viral material. BEV is the only torovirus so far that can be grown in cultured cells, but titers are low. The limited amounts of viral genomic RNA led us to the use of mRNA from infected cells, containing both viral genomic RNA and subgenomic RNAs, as a template for cDNA synthesis. In this report we have shown that the intracellular BEV RNAs form a 3′-coterminal nested set. Four ORFs were identified, and their initiation codons were shown to coincide with the approximate mapping of the 5′ ends of the four sg mRNAs. Because of the absence of a genomic library, direct evidence for colinearity of the nested set of mRNAs and the 3′ end of the BEV genome is still lacking. However, clones containing the region unique to the genomic RNA have now been identified with the aid of Northern blot hybridizations, and still no observations at variance with colinearity have been made.

TABLE 1. Comparison of UV target sizes and estimated physical sizes for BEV RNAs 1, 2, and 3

RNA	$K \times T$[a] (s^{-1})	Physical size[b] (kb)	Target size[c] of template (kb)	
			A	B
1	0.160	>20	28.9	25.7
2	0.038	6.8*	6.8*	6.1
3	0.013	2.1*	2.4	2.1*

[a] $K \times T$ was calculated from the relationship $\ln (N_t/N_o) = -K \times T \times t$. N_t is the amount of label incorporated into RNA after t seconds of UV irradiation, T is the target size, and K is a constant.

[b] The physical size of RNA 1 was estimated from agarose gels (21). For RNAs 2 and 3, values derived from sequence analysis were used, adding an average of 150 nt for the poly(A) tail.

[c] BEV RNA 2 or 3 was used as an internal standard to calculate the target sizes of the other RNAs. By substituting the values indicated by asterisks (derived from sequence analysis) for the target sizes of RNA 2 (A) or RNA 3 (B), the value of K was calculated as 1.60×10^{-8} or 1.81×10^{-8}, respectively. By using these values, the target sizes for the other RNAs were calculated.

In vitro transcription and translation experiments have demonstrated that ORFs 3 and 5 encode the BEV E and N proteins, respectively (Snijder et al., unpublished data). The size of the potential translation product of ORF 2 and the fact that it contains a potential signal peptide and transmembrane anchor suggest that ORF 2 encodes the peplomer protein.

The BEV genome expression, involving multiple, 3′ colinear sg mRNAs containing an ORF at their unique 5′ end, is not without precedent. A similar translation strategy has been described for coronaviruses (for a review, see reference 22) and arteriviruses (25, 26). Like coronaviral mRNAs, the mRNAs of BEV are transcribed independently, as shown herein. Poor [^{3}H]uridine incorporation during UV transcription mapping of BEV RNAs precluded reliable target size determination for the smallest RNA species, RNAs 4 and 5. The correlation coefficient of linear regression analysis on the RNA 3 data (Fig. 4B) was acceptable, allowing comparison of the target sizes of RNAs 1, 2, and 3. On the basis of sequencing data, RNA lengths were estimated for RNAs 2 and 3 and used as internal standards in the calculation of the target sizes for the other RNA species. With either of these standards, the obtained target size

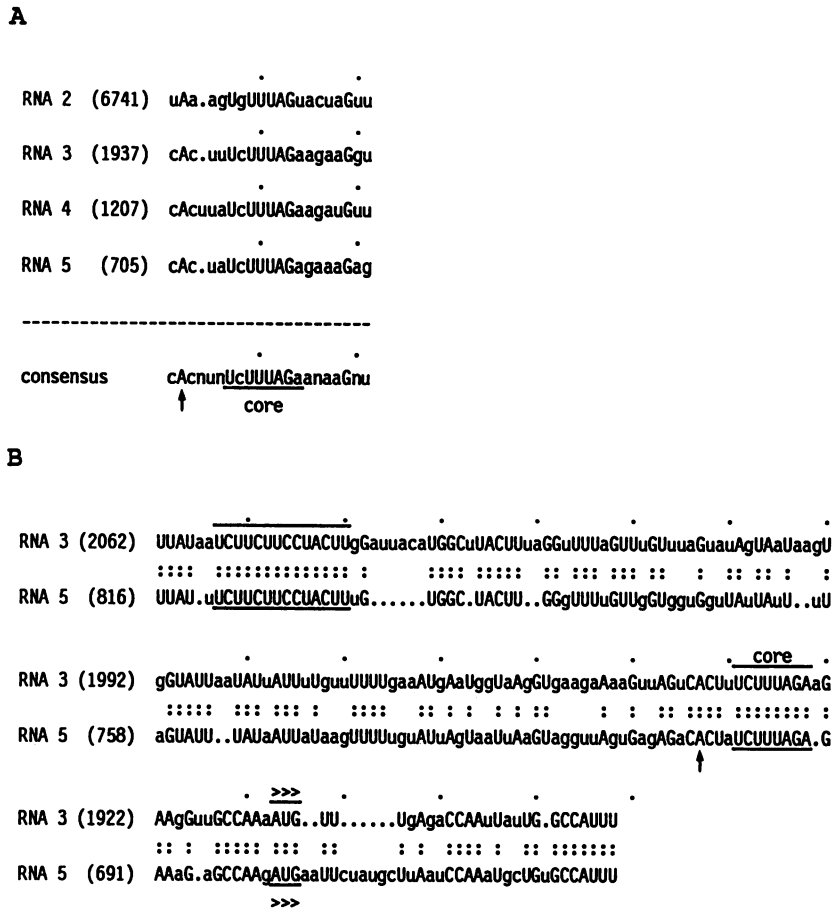

FIG. 6. Nucleotide sequences possibly involved in the transcription of BEV subgenomic RNAs. Numbers indicate the distance to the start of the poly(A) tail. (A) A conserved nucleotide sequence was observed upstream of each BEV ORF. Nucleotides conserved in all four of these postulated core promoter regions are represented in capital letters. Spaces introduced for alignment are indicated (. . .); the arrow indicates the first nucleotide of BEV RNA 5, as determined by primer extension (see the text). (B) Alignment of the nucleotide sequences surrounding the core promoter regions of BEV RNAs 3 and 5. Matches are represented by colons; spaces introduced for alignment are indicated (. . .). The upstream 14-nt match is underlined; the core promoter sequence, AUG codons of ORFs 3 and 5 (>>>), and the first nucleotide of RNA 5 (arrow) are indicated.

values justify the conclusion of independent initiation of transcription for RNAs 1, 2, and 3. Although the target sizes for the templates of RNAs 4 and 5 could not be calculated, it is evident that neither RNA is the result of processing of a large precursor. However unlikely it may be, we cannot formally exclude the possibility that both RNAs are generated by splicing of a small precursor molecule.

The coronaviral sg RNAs contain a common sequence at their respective 5′ ends due to a leader-primed transcription process. No evidence for the presence of such a leader sequence in BEV RNAs has been obtained. For RNA 5, the results of primer extension analysis and oligonucleotide hybridizations suggest a 5′ noncoding sequence of only 24 nucleotides, which does not diverge from the consensus sequence of the nested set. It is not likely that the extension of oligo A on RNA 5 was terminated prematurely due to secondary structure of the RNA molecule; the experiment was carried out in the presence of methylmercury hydroxide, and electrophoresis demonstrated that the primer extension product was of homogeneous length. In addition, computer analysis (Fold program; Computer Genetics Group, University of Wisconsin) does not predict a strong secondary structure near the 5′ end of RNA 5.

Sequence motifs from the area surrounding the 5′ end of RNA 5 are (partly) repeated at three other positions, just upstream of each ORF. Their position near the 5′ end of each RNA suggests a role in sg mRNA transcription. The observed putative core promoter sequence 5′ UCUUUAGA 3′ occurs only in the regions upstream of ORFs 3, 4, and 5 and has not been found elsewhere in the consensus sequence obtained so far. The sequence 5′ UGUUUAG 3′, located immediately upstream of ORF 2, is the best candidate for the core promoter sequence involved in RNA 2 transcription. This sequence may be less unique (it is found at four other positions in the consensus sequence), but the area surrounding it shows the highest similarity to the regions surrounding the putative core promoters of ORFs 3, 4, and 5.

For a number of animal and plant RNA viruses that are known to produce sg mRNAs during their replication (7, 24), RNA promoter sequences have been postulated. However, in most cases these promoters have not been thoroughly studied. In various alphaviruses a conserved 21-nt sequence block is present on the genomic RNA near the transcription initiation site used in sg RNA synthesis (18). The complement of this sequence has been postulated to be the recognition site used by the viral transcriptase. Sequence motifs

homologous to parts of the alphavirus core promoter sequence have been observed near the sg RNA transcription initiation sites of several plant RNA viruses (5, 16). The internal sg RNA promoter region of brome mosaic virus has been characterized in more detail (5, 16). In addition to the core promoter region, surrounding (A+U)-rich sequences and an internal poly(A) tract have been shown to play a role in the transcription of sg brome mosaic virus RNA 3. Besides the interviral homologies described above, nucleotide sequences are available from the sg RNA promoter regions of several RNA plant viruses that produce more than one sg RNA, like the tobamoviruses (6, 27) and the tobraviruses (2, 9). These data allow a comparison of multiple promoter regions from the same viral genome but do not change the overall picture: sg RNA promoters seem to consist of a more or less conserved core sequence that is surrounded by similar, probably transcription-modulating, sequences. No sequence homologies were observed between the promoter regions described above and the BEV sequences that may be involved in sg RNA transcription.

In the case of brome mosaic virus RNA 3 transcription, sequences upstream of the core promoter region seem to play an enhancerlike role and determine the level of transcription. For BEV a similar modulating effect is suggested by the presence of extensive homologies in the regions surrounding the core promoters of the two most abundant BEV RNAs (RNAs 3 and 5). A detailed study of the BEV promoter regions described above is required for a better understanding of the sequences and mechanisms involved in BEV sg RNA production.

Although toroviruses clearly differ in virion structure and mechanism of sg RNA transcription, their genome organization and expression are similar to those of coronaviruses. Recently, the classification of positive-stranded RNA virus families into alphaviruslike and picornaviruslike superfamilies was proposed (24). The coronavirus family did not fit into any of these clusters. An analysis of the genetic information of BEV, especially of the unique region of the genome which most likely will encode the viral polymerase, might provide a basis for the classification of toroviruses and coronaviruses into a third superfamily. Experiments to make such a comparison possible are in progress in our laboratory.

ACKNOWLEDGMENTS

We thank Peter Bredenbeek for many helpful suggestions and stimulating discussions. The assistance of Johan den Boon, Georges Verjans and Marianne Weiss in cDNA synthesis and sequence analysis is gratefully acknowledged.

This study was supported by the Division for Health Research TNO in cooperation with the Foundation for Medical Research MEDIGON (project 900-502-081).

LITERATURE CITED

1. **Beards, G. M., C. Hall, J. Green, T. H. Flewett, F. Lamouliatte, and P. du Pasquier.** 1984. An enveloped virus in stools of children and adults with gastroenteritis that resembles the Breda virus of calves. Lancet **ii**:1050–1052.
2. **Boccara, M., W. D. O. Hamilton, and D. C. Baulcombe.** 1986. The organisation and interviral homologies of genes at the 3′ end of tobacco rattle virus RNA1. EMBO J. **5**:223–229.
3. **De Groot, R. J., R. J. ter Haar, M. C. Horzinek, and B. A. M. van der Zeijst.** 1987. Intracellular RNAs of the feline infectious peritonitis coronavirus strain 79-1146. J. Gen. Virol. **68**:995–1002.
4. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
5. **French, R., and P. Ahlquist.** 1988. Characterization and engineering of sequences controlling in vivo synthesis of brome mosaic virus subgenomic RNA. J. Virol. **62**:2411–2420.
6. **Goelet, P., G. P. Lomonossoff, P. J. G. Butler, M. E. Akam, N. J. Gait, and J. Karn.** 1982. Nucleotide sequence of tobacco mosaic virus RNA. Proc. Natl. Acad. Sci. USA **79**:5818–5822.
7. **Goldbach, R., and J. Wellink.** 1988. Evolution of plus-strand RNA viruses. Intervirology **29**:260–267.
8. **Gubler, U., and B. J. Hoffman.** 1983. A simple and very efficient method for generating cDNA libraries. Gene **25**:263–269.
9. **Hamilton, W. D. O., M. Boccara, D. J. Robinson, and D. C. Baulcombe.** 1987. The complete nucleotide sequence of tobacco rattle virus RNA-1. J. Gen. Virol. **68**:2563–2575.
10. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.
11. **Horzinek, M. C., T. H. Flewett, L. J. Saif, W. J. M. Spaan, M. Weiss, and G. N. Woode.** 1987. A new family of vertebrate viruses: Toroviridae. Intervirology **27**:17–24.
12. **Horzinek, M. C., and M. Weiss.** 1984. Toroviridae: a taxonomic proposal. Zentralbl. Veterinaermed. Reihe B **31**:649–659.
13. **Jacobs, L., W. J. M. Spaan, M. C. Horzinek, and B. A. M. van der Zeijst.** 1981. Synthesis of subgenomic mRNA's of mouse hepatitis virus is initiated independently: evidence from UV transcription mapping. J. Virol. **39**:401–406.
14. **Lenstra, J. A., R. J. de Groot, L. Jacobs, J. G. Kusters, H. G. M. Niesters, and B. A. M. van der Zeijst.** 1988. Synthesis of long cDNA from viral RNA template. Gene Anal. Tech. **5**:57–61.
15. **Maniatis. T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Marsh, L. E., T. W. Dreher, and T. C. Hall.** 1988. Mutational analysis of the core and modulator sequence of the BMV RNA 3 subgenomic promoter. Nucleic Acids Res. **16**:981–995.
17. **Meinkoth, J., and G. Wahl.** 1984. Hybridization of nucleic acids immobilized on solid supports. Anal. Biochem. **138**:267–284.
18. **Ou, J.-H., C. M. Rice, L. Dalgarno, E. G. Strauss, and J. H. Strauss.** 1982. Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA. Proc. Natl. Acad. Sci. USA **79**:5235–5239.
19. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
20. **Sauerbier, W., and K. Hercules.** 1978. Gene and transcription unit mapping by radiation effects. Annu. Rev. Genet. **12**: 329–363.
21. **Snijder, E. J., J. Ederveen, W. J. M. Spaan, M. Weiss, and M. C. Horzinek.** 1988. Characterization of Berne virus genomic and messenger RNAs. J. Gen. Virol. **69**:2135–2144.
22. **Spaan, W., D. Cavanagh, and M. C. Horzinek.** 1988. Coronaviruses: structure and genome expression. J. Gen. Virol. **69**: 2939–2952.
23. **Staden, R.** 1986. The current status and portability of our sequence handling software. Nucleic Acids Res. **14**:217–233.
24. **Strauss, J. H., and E. G. Strauss.** 1988. Evolution of RNA viruses. Annu. Rev. Microbiol. **42**:657–683.
25. **Van Berlo, M. F., M. C. Horzinek, and B. A. M. van der Zeijst.** 1982. Equine arteritis virus-infected cells contain six polyadenylated virus-specific RNAs. Virology **118**:345–352.
26. **Van Berlo, M. F., P. J. M. Rottier, M. C. Horzinek, and B. A. M. van der Zeijst.** 1986. Intracellular equine arteritis virus (EAV) specific RNAs contain common sequences. Virology **152**:492–496.
27. **Watanabe, Y., T. Meshi, and Y. Okada.** 1984. The initiation site for transcription of the TMV 30-kDa protein messenger RNA. FEBS Lett. **173**:247–250.
28. **Weiss, M., and M. C. Horzinek.** 1987. The proposed family Toroviridae: agents of enteric infections. Arch. Virol. **92**:1–15.
29. **Weiss, M., F. Steck, and M. C. Horzinek.** 1983. Purification and partial characterization of a new enveloped RNA virus (Berne virus). J. Gen. Virol. **64**:1849–1858.
30. **Woode, G. N., L. J. Saif, M. Quesada, N. J. Winand, J. F. L. Pohlenz, and N. Gourley.** 1985. Comparative studies on three isolates of Breda virus of calves. Am. J. Vet. Res. **46**:1003–1010.