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 186, 426–434 (1992)

Cloning and Sequencing of the Matrix Protein (M) Gene of Turkey Rhinotracheitis Virus Reveal a Gene Order Different from That of Respiratory Syncytial Virus

Q. YU, P. J. DAVIS, J. LI, AND D. CAVANAGH[1]

Institute for Animal Health, Houghton Laboratory, Houghton, Huntingdon, Cambridgeshire PE17 2DA, United Kingdom

Received December 13, 1990; accepted October 10, 1991

Several biochemical properties and the sequence of the fusion glycoprotein (F) have indicated that turkey rhinotracheitis virus (TRTV) is a pneumovirus, subfamily Pneumovirinae of the Paramyxoviridae family. As TRTV was known to generate polycistronic mRNAs, cDNA was generated from TRTV strain UK/3BV/85-infected Vero cell mRNAs using an oligonucleotide primer corresponding to a region of the F gene. Sequencing of four cDNAs revealed that the gene adjacent to the beginning (3′ end) of the F gene was that for the matrix (M) protein, i.e., that TRTV had the partial gene order 3′-M-F-5′. This was unexpected as human respiratory syncytial (RS) virus, the type species of the genus *Pneumovirus,* has the partial gene order 3′-M-SH-G-F-5′, where SH and G are the small hydrophobic protein and attachment glycoprotein, respectively. Instead TRTV resembled the *Morbillivirus* and *Paramyxovirus* genera of the Paramyxoviridae (subfamily Paramyxovirinae) which have the partial gene order 3′-M-F-5′. Two further oligonucleotides, one corresponding to a sequence near the end of the M gene and the other (oligo B) to a sequence near the beginning of the F gene, with their 5′ ends spaced 300 nucleotides apart on the basis of the cDNA sequence, were used in a polymerase chain reaction (PCR) using genomic RNA as template. Only a PCR product of 0.3 kb was obtained. The same sized product was also obtained using these oligonucleotides and genomic RNA from three other TRTV strains (SA/91/78, UK/8544/85, and SA/2381/88) which had been grown in chicken tracheal organ cultures. In addition PCR was performed using genomic RNA from TRTV-3BV and SA/2381/88 with oligo B and another oligonucleotide near the 5′ end of the gene upstream from M, spaced 1141 nucleotides apart on the basis of the sequence data. Only a 1.14-kb PCR product was obtained. Larger products would have been expected if another gene had been situated between M and F. The absence of such larger products, plus the demonstration that infected cells contained M-F dicistronic mRNAs, supported the conclusion that in the TRTV genome the M gene is adjacent to the F gene in the order 3′-M-F-5′. The 5′ termini of the M and F mRNAs were confirmed by mRNA mapping. The TRTV M gene encoded a protein of 254 amino acids, very similar to that of RS virus (256 residues; 37% amino acid identity) but very different from that of the morbilliviruses and paramyxoviruses (approximately 350 residues). Thus, on the basis of the sequences of both the M and the F genes, TRTV is more closely related to the pneumoviruses than to the paramyxoviruses and morbilliviruses.

INTRODUCTION

Turkey rhinotracheitis virus (TRTV), as the name indicates, was first isolated from turkeys with respiratory disease, although it has subsequently been shown that chickens and other domestic birds can also be infected (McDougall and Cook, 1986; Jones *et al.*, 1987; Gough *et al.*, 1988; and references in Yu *et al.*, 1991). Biochemical analysis of the TRTV-induced proteins (Collins and Gough, 1988; Ling and Pringle, 1988; Cavanagh and Barrett, 1988) and RNAs (Cavanagh and Barrett, 1988) suggested that this virus was a member of the Paramyxoviridae family and more closely related to the genus *Pneumovirus* (subfamily Pneumovirinae; Pringle, 1991) than to the other two genera in this family, *Morbillivirus* and *Paramyxovirus* (subfamily Paramyxovirinae). Our recent cloning and sequencing of the TRTV fusion (F) protein gene has supported this view. Thus the TRTV F protein had 39% amino acid identity with that of human respiratory syncytial (RS) virus, a pneumovirus, but only about half that with members of the other two genera (Yu *et al.*, 1991). In order to identify the gene upstream (3′ proximal) from the F gene, we have used an oligonucleotide, based on the F gene sequence, to prime cDNA synthesis in the hope that some clones would be generated from polycistronic mRNA which comprised F and the 3′ proximal gene. This was indeed the case, although the gene adjacent to the 3′ terminus of the F gene was not that which had been expected from knowledge of the gene order of RS virus and pneumonia virus of mice (PVM) (Collins *et al.*, 1984; Chambers *et al.*, 1990).

MATERIALS AND METHODS

Cells and virus

The TRTV strain used for cloning was 3BV, plaque-purified clone R1.1, this being the UK/3B/85 field iso-

[1] To whom reprints requests should be addressed.

0042-6822/92 $3.00

late (McDougall and Cook, 1986) after passage in cell culture (Yu *et al.*, 1991). This virus was propagated in Vero cells (ATCC No. CCL 81; Flow Laboratories, Irvine, Scotland). The South African TRTV strains SA/91/78 and SA/2381/88 and strain UK/8544/85 were grown in chick embryo tracheal organ cultures (Cook *et al.*, 1976). The coronavirus infectious bronchitis virus (IBV), strain Beaudette, was grown in Vero cells.

Preparation of TRTV-induced RNAs

Flasks (75 cm^2) of Vero cells were inoculated with 1 ml of undiluted TRTV strain 3BV working stock which resulted in the infection of less than 1% of the cells. After 72 hr at 37° actinomycin D (1 μg/ml; Sigma) was added and 24 hr later the cells were lysed by the guanidinium isothiocyanate method as described by Chomczynski and Sacchi (1987), except that isopropanol was substituted by ethanol. Poly(A)$^+$ RNA was selected using Hybond mAP affinity paper according to the manufacturer's instructions (Amersham International). RNA of strains SA/91/78, UK/8544/85, and SA/2381/88 was extracted from 60 tracheal organ culture rings, using guanidinium isothiocyanate, which had been infected 3 days previously. Virions of TRTV-3BV were prepared by centrifugation of clarified medium from infected Vero cells at 35,000 g_{max} for 90 min in a Sorvall TFA-20.250 angle rotor. Nucleocapsids were sedimented from Nonidet P-40-produced extracts of infected Vero cells (Ward *et al.*, 1983) at 100,000 g_{max} for 40 min in a Sorvall TH641 swing-out rotor. The genomic RNA of both virions and nucleocapsids was extracted using guanidinium isothiocyanate. IBV-Beaudette mRNAs were extracted from Vero cells by the same method.

Preparation of cDNA

cDNA synthesis was performed with a cDNA synthesis kit (Amersham) using the poly(A)$^+$ RNA from approximately two 75-cm^2 flasks of infected cells. Reverse transcription was primed using a synthetic oligonucleotide corresponding to nucleotides 531–547 of the TRTV F gene (Yu *et al.*, 1991).

Cloning and sequencing

cDNA was cloned into *Eco*RV-cut phagemid pBluescript KS$^+$ (Northumbria Biologicals, Cramlington, Northumbria, UK). Colonies containing TRTV-specific inserts were identified by probing with a ^{32}P-labeled, random hexanucleotide (Pharmacia)-primed probe corresponding to a region (398 nucleotides long) near the 3′ part of the F gene, generated by *Pst*I restriction of clone F17 (Yu *et al.*, 1991). The clones, in pBluescript, were sequenced by dideoxy chain termination following the generation of a nested set of deletions, as described by Yu *et al.* (1991). Briefly, the recombinant phagemid was digested within the polylinker either with *Bst*XI and *Xba*I or with *Apa*I and *Hin*dIII (Boehringer-Mannheim) to generate deletions from opposite ends and deletions generated with exonuclease III (Northumbria Biologicals). Sequences were analyzed on a MicroVAX 3600 computer system using the Wisconsin program GAP.

Polymerase chain reaction (PCR)

The PCR was performed essentially as described by Sambrook *et al.* (1989) using genomic RNA from virions and nucleocapsids in the case of strain UK/3BV/85 and total infected tracheal organ culture RNA for the other three TRTV strains. Reverse transcription was done with 2 units of AMV reverse transcriptase (Anglia Biotechnology, Colchester, UK); cDNA was amplified using 2.5 units of Taq polymerase (Boehringer-Mannheim) in 25 cycles, the polymerization step lasting for 5 or 15 min, as described under Results. Two microliters from the 80-μl reaction volume was analyzed in 1% agarose gels, run at 100 V for 2 hr, and stained with ethidium bromide (5 μg/ml).

Oligonucleotides

The positions of the 25-mer oligonucleotides (Severn Biotech, Kidderminster, UK) used for priming DNA synthesis are shown in Fig. 2. Oligonucleotide A (positive sense) corresponded to nucleotides 604–629 of the M sequence (Fig. 3). Oligonucleotide B (negative sense) corresponded to nucleotides 53–78 in the F gene (Yu *et al.*, 1991). Oligonucleotide C (positive sense), which corresponded to part of the gene upstream (3′) from the M gene, had its 5′ terminus 237 nucleotides from the beginning of the M gene. Oligonucleotide D (negative sense) corresponded to nucleotides 54–79 of the M sequence (Fig. 3).

Primer extension

The 5′ terminus of the M mRNA was determined by primer extension using a 25-mer oligonucleotide (oligonucleotide D) complementary to nucleotides 54–79 of the M mRNA sequence shown in Fig. 3. The primer was 5′ end-labeled with polynucleotide kinase and [γ-^{32}P]ATP. The primer (100 ng, 5×10^4 cpm) was added to 15 μg of total infected cell extract RNA or 4 μg of poly(A)$^+$ RNA from infected cells. Reverse transcription was performed as described previously (Cavanagh and Davis, 1988). The same oligonucleotide was used to prime a sequencing reaction (Yu *et al.*, 1991) using pBluescript containing insert MF36 as substrate. All reactions were analyzed in the same sequencing gel

so that the 5′ terminus of the M mRNA could be deduced directly by comparison with the sequence determined from clone MF36. The 5′ terminus of the F mRNA was mapped in the same way using oligonucleotide B, whose 5′ terminus was located 78 nucleotides from the presumed start of the F mRNA, and using pBluescript containing insert MF37.

Detection of polycistronic mRNAs

Polycistronic mRNAs of TRTV were detected by Northern blot analysis (Sambrook *et al.*, 1989) using a restriction fragment in the case of M and a PCR product of F to make random hexanucleotide-primed ^{32}P-labeled probes. IBV-Beaudette mRNAs, all of which have been sequenced (Boursnell *et al.*, 1987), were used as RNA size markers. The IBV-specific probe comprised a complete cDNA copy of the nucleocapsid (N) gene of IBV, produced by the PCR. All IBV mRNAs were detected by this probe because all the mRNAs contain the N gene at their 3′ end (Boursnell *et al.*, 1987). TRTV mRNAs were electrophoresed before and after selection of poly(A)$^+$ RNA; IBV RNAs were used without poly(A)$^+$ selection. Before transfer of the RNA to nitrocellulose, the gel was exposed to uv light (302 nm) (Ultra-Violet Products, Inc., U.S.A.) for 2 min to nick the RNA sufficiently to improve the transfer of genome and anti-genome TRTV RNA and the large IBV mRNAs.

Sequence databases

The sequence of the M mRNA (nucleotides 1–824 in Fig. 3) will appear in the DDBJ, EMBL, and Genbank Nucleotide Sequence Databases under Accession No. X58639.

RESULTS

Cloning and sequencing of the gene 3′ proximal to the F gene

In order to identify the gene upstream (3′ proximal) from the F gene, an F gene-specific oligonucleotide was used to generate cDNA from poly(A)$^+$-infected cell RNA, and clones were selected using a radiolabeled F gene restriction fragment. This strategy was based on the knowledge, gained from probing Northern blots of infected cell RNA with F gene-specific probes, that polycistronic mRNAs were generated during transcription (Fig. 1, lanes 3 and 5) and the hope that some of these mRNAs contained F and the upstream gene. Six clones were obtained, two of which did not extend beyond the 5′ terminus of the F mRNA. The remaining four clones did extend into the 3′ proximal gene; indeed two of them continued into a further gene (Fig. 2). Unexpectedly one clone (MF35) extended 622 nucleotides

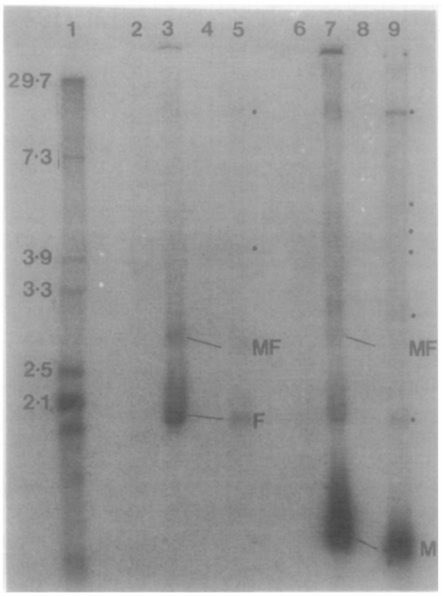

FIG. 1. Northern blots showing mono- and polycistronic TRTV mRNAs containing M and F gene sequences. RNA was extracted from TRTV-infected (lanes 3, 5, 7, and 9) and mock-infected (lanes 2, 4, 6, and 8) Vero cells and electrophoresed in a denaturing agarose gel before (lanes 2, 3, 6, and 7) and after (lanes 4, 5, 8, and 9) poly(A)$^+$ selection. The poly(A)$^+$-selected RNA migrated slightly faster than the nonselected RNA. After transfer to nitrocellulose the RNAs were probed with a TRTV F-specific (lanes 2–5) and M-specific (lanes 6–9) DNA probe. Lane 1 contains IBV mRNAs, the sizes (kilobases) of which are shown on the left. The monocistronic M and F mRNAs are identified by M and F, respectively, and dicistronic M-F mRNA by MF. The positions of other polycistronic mRNAs and genome-sized RNA are shown by filled circles.

beyond the primer in the direction of the 5′ end of the F gene, to nucleotide position 1170 in the F gene sequence (Yu *et al.*, 1991). Every nucleotide of the gene 3′ proximal to F was sequenced from at least two clones; 83% were sequenced from three clones. The entire sequence was determined from both strands of the cDNA and 96% of nucleotides were determined 5 to 16 times and the remaining 4% on 3 or 4 occasions.

The M gene is adjacent to the 3′ end of the F gene of TRTV

As described above the identity of the gene sequences contained within each clone was ascertained by sequencing. Clones MF35 and 36 encoded a large ORF of 254 amino acids. The next longest ORF would encode an ORF of only 51 residues, other ORFs being less than half that size. Comparison of the sequence of the longest ORF with a databank of Paramyxoviridae sequences revealed that it had 37% amino acid identity with the M protein of RS virus which comprises 256 residues (Fig. 3). The calculated molecular weights of

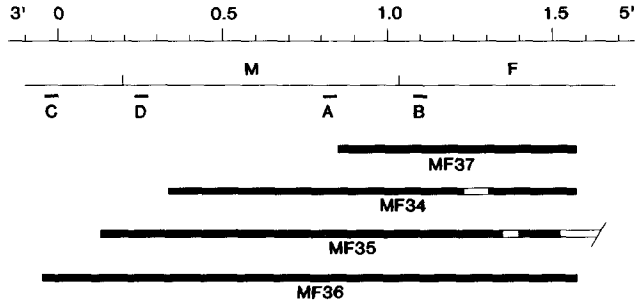

FIG. 2. Relationship of clones MF37, MF34, MF35, and MF36 to the 3'-M-F-5' region of the negative sense TRTV genome. The top line indicates the length of the clones in kilobases. The 3' ends (genome sense) of the clones were located thus: clones MF37 and MF34 are at nucleotide positions 672 and 140 in the M mRNA sequence (Fig. 3); clones MF36 and MF35 are at 237 and 88 upstream from nucleotide position 1 in Fig. 3. The 5' termini of clones MF37, MF34, and MF36 were located at nucleotide position 546 in the F gene sequence (Yu *et al.*, 1991). The 5' end of clone MF35 extended further into the F gene (to position 1170) than is shown in the figure. The solid areas show those parts of the clones that were sequenced. The short horizontal lines marked A, B, C, and D show the position of oligonucleotides A–D that were used for the PCR.

the M protein for TRTV and RS virus were 27,595 and 28,714, respectively.

Almost immediately upstream from the beginning of the ORF was the sequence GGGACAAGU (nucleotides 1–9, doubly underlined in Fig. 3) which was identical to the sequence just upstream from the F ORF (nucleotides 827–835, doubly underlined in Fig. 3) which we have suggested is the start sequence for the F mRNA (Yu *et al.*, 1991). Confirmation that these sequences did mark the beginnings of the M and F mRNAs was obtained by mapping the 5' end of the respective mRNAs. Gene-specific oligonucleotides were 5' end-labeled and then used for primer extension. Both the M and the F primer-extended products terminated to yield a double band, consistent with the ends of the respective mRNAs having been reached (Fig. 4). The same end-labeled oligonucleotides were used in sequencing reactions, with [^{35}S]ATP present, with pBluescript containing either the M or the F cDNA inserts. Electrophoresis of the reaction products in the same gels as those used for analysis of the cDNA generated by primer extension showed that the lower band of the doublet co-migrated with a G residue which was the beginning of the sequence GGGACAAGU. This was direct evidence that nucleotides 1–9 and 827–835 (Fig. 3) marked the beginning of the M and F mRNAs, respectively. The upper band of the doublet, which was sometimes of lower intensity, was considered most likely to be an artifact, as proposed previously (Jorgensen *et al.*, 1987). The presence of partial cross-banding in the pBluescript-generated sequence was a consequence of the presence of some $n - 1$ oligonucleotide in the primer preparations.

At the end of the TRTV F mRNA is the sequence AGUUA followed three nucleotides later by the poly(A) tail (Yu *et al.*, 1991). Immediately prior to the start of the F gene, i.e., near the end of the M sequence, was the similar sequence AGUCA (nucleotides 812 to 816, singly underlined, in Fig. 3) followed two nucleotides later by AAAAAAUU. This observation, coupled with the identification of the 5' terminus of the F mRNA (Fig. 4B), indicates that these sequences mark the end of the M mRNA, the UU residues being the M-F intergenic region (in the mRNA sense). Thus nucleotides 1 and 824 in Fig. 3 define the termini of the TRTV M mRNA.

Northern blot analysis

Northern blots of TRTV-infected cell RNA, before and after selection of poly(A)$^+$ RNA, were probed with F- and M-specific probes. In addition to the monocistronic F mRNA the F probe hybridized to an mRNA of approximately 2.63 kb (Fig. 1, lanes 3 and 5; size estimated using IBV mRNAs as markers), a mRNA of this length also being revealed by the M-specific probe (Fig. 1, lanes 7 and 9). The size estimate of 2.63 kb is close to that of 2.56 kb predicted for a M-F dicistronic mRNA based on the sequences for M and F of TRTV and assuming a poly(A) tail of approximately 100 nucleotides. The probes also detected small amounts of several other mRNA species and some genome-sized RNA; Afzal *et al.* (1990) detected both genome and anti-genome in poly(A)$^+$-selected mumps virus RNAs. TRTV clones MF35 and MF36, which extended into the gene upstream from M, could have been generated from one or more of these additional RNA species.

PCR evidence that the TRTV M gene is adjacent to the F gene

From the Northern blot analysis it was most likely that clones MF34–37 had been generated by reverse transcription of polycistronic mRNAs which faithfully reflected the gene order of the M and F genes in the genome. However, it was possible that in the TRTV genome there were one or more genes between M and F and that during the synthesis of our cDNA clones there might have been a fault leading to the deletion of intervening genes. Alternatively the virus RNA preparation may have contained deletion genomes which produced defective RNA from which the cDNAs were generated, although we have no evidence for the existence of such RNA. To investigate these possibilities we have used the PCR to generate cDNA, using UK/3BV/85 virion RNA as template, covering the proposed M-F junction. First oligonucleotides A and B (Fig. 2)

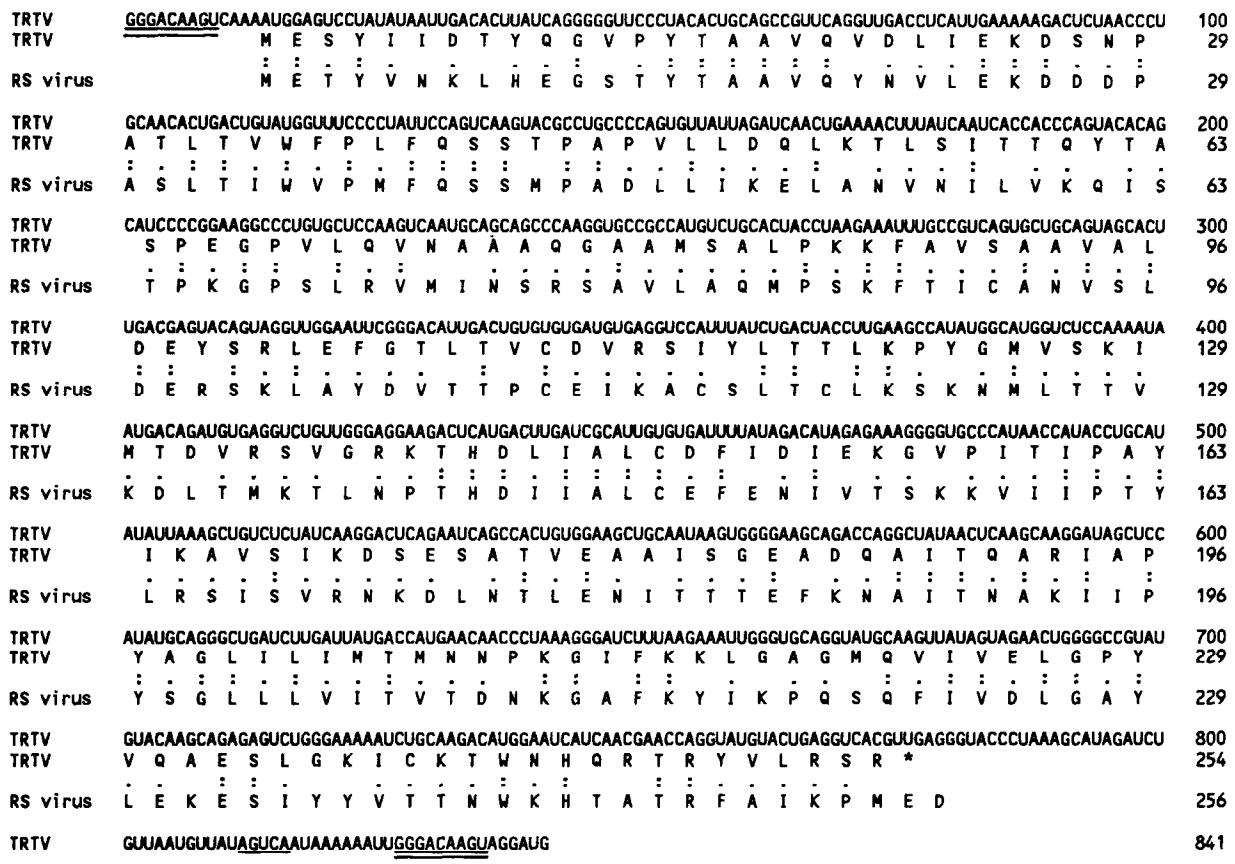

FIG. 3. Sequence of the TRTV M gene and comparison of the deduced amino acid sequence with that of RS virus strain A2 (Satake and Venkatesan, 1984). The mRNA start sequences for M (nucleotides 1 to 9) and F (nucleotides 827 to 835) have been underlined twice. The putative mRNA stop sequence for M (nucleotides 812 to 816) has been underlined once. The last three nucleotides in the figure (AUG) correspond to the translation start codon of the F gene (Yu *et al.*, 1991). A colon (:) indicates where TRTV and RS virus possessed exactly the same amino acid. A period (.) indicates where the amino acid differences were conservative. The translation stop codon for the M polypeptide is marked by an asterisk (*).

were used to prime cDNA synthesis. If the gene order M-F was correct, the only product generated by the PCR would be 300 nucleotides in length. This was indeed the case, irrespective of whether the elongation time in the PCR was 5 min (Fig. 5, lane b) or 15 min (Fig. 5, lane c). The long incubation time was included to permit the amplification of any genes which might have been located between M and F; no long PCR products corresponding to such a RNA population were observed. Sequencing of the 0.3-kb PCR product after cloning confirmed that the sequence was the same as that obtained from sequencing of clones MF34–37 (Fig. 3). The same PCR result was obtained with RNA that had been extracted from nucleocapsids and with RNA from tracheal organ cultures that had been infected with TRTV strains SA/91/78, UK/8544/85, and SA/2381/88 (data not shown).

A second PCR was done using oligonucleotides C and D whose 5′ ends were spaced 316 nucleotides apart (Fig. 2). In agreement with the analysis of the 3′ (genome sense) portion of clones MF35 and 36, the PCR produced only the predicted product of approximately 0.3 kb (Fig. 5, lanes d and e), confirming that the sequence upstream from the beginning of M was as had been deduced from sequencing clones MF35 and 36. Finally, the PCR was performed using oligonucleotides C and B with an extended polymerization time of 15 min. This produced only a single product of approximately 1.14 kb (Fig. 5, lane f) which was that predicted from the distance between oligonucleotides C and B based on the sequencing of clones MF35 and 36. The same result was obtained with RNA extracted from Vero cells which had been infected at very low multiplicity with TRTV strain 3BV which was only two passages beyond the plaque-purification stage (Fig. 5, lane g). In addition the PCR produced only one product, of 1.14 kb, when RNA of the tracheal organ culture-passaged strain SA/2381/88 was used (Fig. 5, lane h).

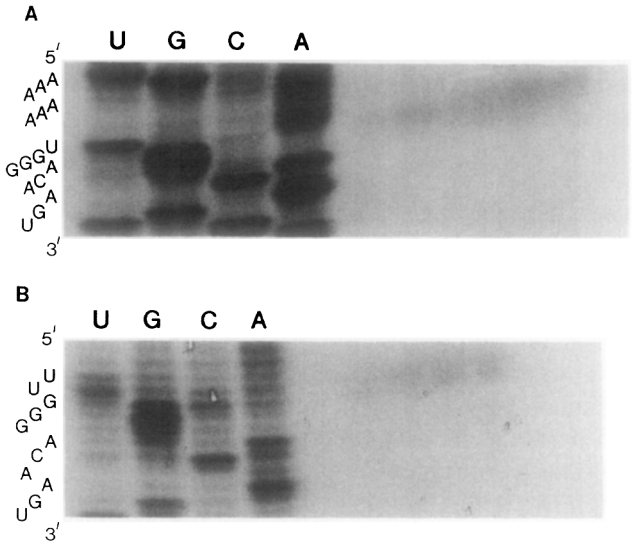

FIG. 4. Mapping of the 5′ end of the (A) M mRNA and (B) F mRNA by primer extension. (A) An end-labeled 25-mer oligonucleotide whose 5′ end corresponded to nucleotide position 79 in Fig. 3 was used to prime extension on M mRNA present in poly(A)$^+$-selected TRTV-infected Vero cell RNA. The same end-labeled oligonucleotide was used to prime a sequencing reaction using pBluescript containing insert MF36 as substrate. All reactions were analyzed in a 6% acrylamide sequencing gel. Primer extension of the M mRNA produced a product which migrated as a double band (right hand side four lanes), the lower of which comigrated with the first nucleotide G in the sequence GGGACAAGU derived from pBluescript containing clone MF36. (B) A 25-mer, end-labeled oligonucleotide whose 5′ end was located 78 nucleotides from the presumed start of the F mRNA was used to prime extension on F mRNA present in poly(A)$^+$-selected RNA. The oligonucleotide was also used to prime a sequencing reaction using pBluescript containing insert MF37 as substrate. Primer extension of the F mRNA produced a product which migrated as a double band (right hand side four lanes), the lower of which comigrated with the first base of the sequence GGGACAAGU. This confirmed that the sequence GGGACAAGU was at the start of both the M and the F mRNAs.

Taken together these results strongly support the view that the M and F genes of TRTV occur in the genome in the order 3′-M-F-5′.

DISCUSSION

The deduced sequence of the M protein of TRTV confirms the notion, based initially on superficial observations of TRTV-induced proteins and RNAs and more recently on the sequence of the F protein, that of the three genera within the Paramyxoviridae family, TRTV most closely resembles the *Pneumovirus* genus. First, TRTV and RS virus, the type species of the Pneumovirus genus, have M proteins of almost identical length, comprising 254 and 256 residues, respectively, in contrast to the much longer M polypeptides (335–377) possessed by paramyxoviruses (Chambers *et al.*, 1986; McGinnes and Morrison, 1987; Sato *et al.*, 1987; Blumberg *et al.*, 1984; Spriggs *et al.*, 1987; Luk *et al.*, 1987; Galinski *et al.*, 1987; Sakai *et al.*, 1987; Sheshberadaran and Lamb, 1990; Elliott *et al.*, 1989; Elango, 1989) and morbilliviruses (Bellini *et al.*, 1986; Limo and Yilma, 1990). Second, the M protein hydropathy plots for TRTV and RS virus are remarkably similar (Fig. 6).

The Wisconsin program GAP revealed that the M protein of TRTV had 37% amino acid identity with that of RS virus and 17–21% with members of the other two genera. However, this figure of 17–21% tends to exaggerate the extent to which the M proteins of the pneumoviruses resemble those of the other two genera since 7–11 gaps had to be inserted in order to produce even this low level of alignment.

The M protein of paramyxoviruses forms an ordered array at the inner surface of the virion envelope (Heggeness *et al.*, 1982; Büechi and Bächi, 1982) and plays an important role in the virus maturation process, the protein being believed to interact with the viral ribonucleoprotein (RNP) (Mountcastle *et al.*, 1974; Shimizu and Ishida, 1975; Yoshida *et al.*, 1976), the inner sur-

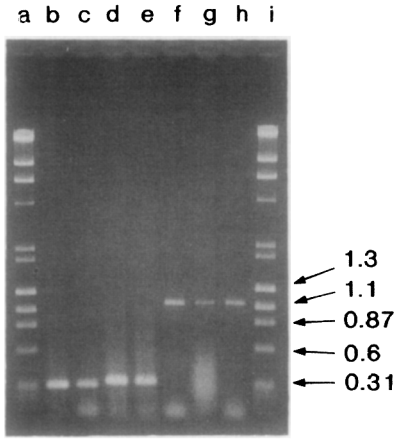

FIG. 5. Confirmation of the TRTV partial gene order 3′-M-F-5′. The PCR was used with various combinations of oligonucleotides A–D (see Fig. 2) to prime DNA synthesis using TRTV strain 3BV genomic RNA as template. The products were visualized in an ethidium bromide-stained agarose gel. (b,c) Oligonucleotides A (sequence near the 5′ end of the M gene) and B (near the 3′ end of the F gene); (d,e) oligonucleotides C (near the 5′ end of the gene upstream from the M gene) and D (near the 3′ end of the M gene); (f,g,h) oligonucleotides C and B; (g) the RNA had been extracted from virions derived after virus passage at very low m.o.i.; (h) the RNA was from tracheal organ culture-passaged TRTV strain SA/2381/88. The PCR polymerization time was either 5 min (b,d) or 15 min (c,e,f,g,h). Lanes (a) and (i) contained a *Hin*dIII digest of lambda DNA and a *Hae*III digest of ϕX174 DNA (BRL); the sizes in kilobases, of five of the DNA fragments are shown.

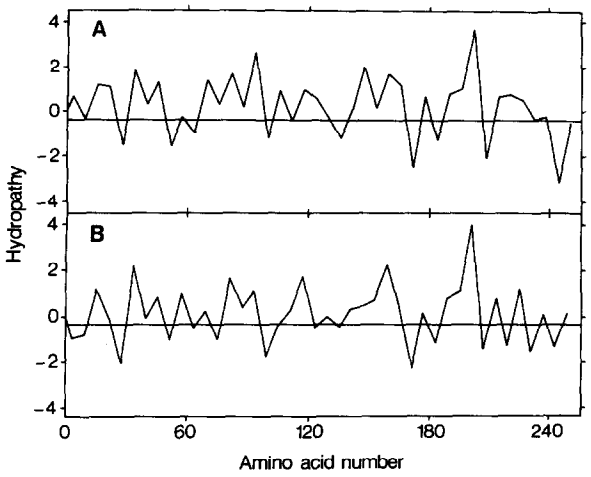

FIG. 6. Hydropathy profiles for the (A) TRTV and (B) RS virus M proteins. Positive values indicate hydrophobic regions and negative values indicate hydrophilic regions. The midpoint line represents a grand average of the hydropathy of the amino acid composition of a large number of sequenced proteins (Kyte and Doolittle, 1982). Each point on the profile represents the average hydropathy of a span of five residues.

face of the plasma membrane at areas associated with the viral glycoproteins (Caldwell and Lyles, 1986; Peeples and Bratt, 1984; Faaberg and Peeples, 1988) and actin (Giuffre *et al.*, 1982; Bohn *et al.*, 1986). In addition M has a role in regulating viral transcription (Marx *et al.*, 1974). The many interactions of M are believed to involve complex interactions of both hydrophobic and ionic natures (Caldwell and Lyles, 1986; Faaberg and Peeples, 1988). As for morbilliviruses and paramyxoviruses (Chambers *et al.*, 1986; Galinski *et al.*, 1987; Limo and Yilma, 1990) the hydropathy plots for TRTV and RS virus indicate that the M protein is moderately hydrophobic. While none of the sequences are obvious candidates for membrane-spanning structures, both TRTV and RS virus have a highly hydrophobic sequence of 14 residues centering around amino acid 200 (Figs. 3 and 6). This stretch in particular might be membrane associated. Bellini *et al.* (1986) have drawn attention to a sequence of 13 to 14 largely nonpolar residues situated about 90 residues from the carboxy terminus of the M protein of the morbilliviruses which might interact with the lipid environment of the virus envelope. With respect to ionic interactions the M proteins of paramyxoviruses and morbilliviruses are basic in nature, basic and acidic residues accounting for 12–14 and 8%, respectively, of M residues. Many (12–16) of the basic residues occur in pairs. In contrast, TRTV and RS virus M proteins have only a small preponderance of basic over acidic residues, these residues accounting for 9–11 and 8–9%, respectively, of total residues. TRTV and RS virus have only 6 and 2 basic residues, respectively, in pairs (Satake and Venkatesan, 1984). TRTV M protein does resemble that of the other two genera in that the amino terminal one-third of M has an acidic character, TRTV M having 7 and 2 acidic and basic residues, respectively, in this region. However, this is not a general observation for pneumoviruses as RS virus has almost equal numbers of acid and basic residues (8 and 7, respectively) in the first third of the molecule (Satake and Venkatesan, 1984).

All paramyxoviruses and morbilliviruses examined to date have the partial gene order 3′-M-F-(SH)-HN/H-5′; SH is in parentheses to indicate that this gene is not possessed by all viruses in these genera. In contrast two pneumoviruses, human RS virus and PVM, have the equivalent genes in the order 3′-M-SH-G-F-5′; i.e., they have, unlike the other two genera, the SH and cell attachment protein gene (G) on the 3′ side of the F gene (Collins *et al.*, 1984; Chambers *et al.*, 1990). It was a surprise, therefore, when we found that TRTV had the M gene immediately adjacent to the 3′ side of the F gene. There was no evidence from probing Northern blots that TRTV UK/3BV/85 had generated defective RNA which might have had genes deleted between M and F. Even if such defective RNA had been present, then in order to generate clones MF34–37 the deletion event would have had to be such as to delete genes in their entirety while at the same time preserving completely intact the M gene and the beginning of F. We have confirmed the beginning of both the M and the F mRNAs by 5′ mapping and have shown that M has mRNA start and end sequences very similar to those of F. The PCR data supported the gene order deduced from clones MF34–37. Although the PCR would tend to favor the amplification of small products, e.g., the 300-nucleotide product obtained using oligonucleotide pairs A/B, over larger ones, e.g., the product which would have been obtained if there had been a gene between M and F, there was no sign at all of larger PCR products to substantiate the view that there was a gene between M and F. Moreover, the gene order M-F was further supported by the PCR data obtained with three other TRTV strains, representing isolates from two continents, isolated over a 10-year period, which had different passage histories, and which had been propagated in chicken tracheal organ cultures.

Although the M protein of TRTV is clearly more closely related to that of RS virus than to that of the other two genera, the amino acid identity was only 37%. This is similar to the amino acid identity shared between the morbilliviruses and the human and bovine parainfluenza virus 3/Sendai virus group of paramyxoviruses, whereas within the morbillivirus genus the amino acid identity is about 75% (Rima, 1989). This suggests that the evolutionary pathways of TRTV and

RS virus diverged before those of members of the morbillivirus genus.

The subfamily Pneumovirinae is of particular interest because, based on the analysis of the genomes of human RS virus and PVM, the member viruses contain several genes that are not present in species of the Paramyxovirinae subfamily (Collins *et al.,* 1984; Chambers *et al.,* 1990). Given that TRTV has some pneumovirus characteristics but appears to have diverged a considerable time ago from the other known members of this genus, further analysis of the number, nature, and order of TRTV genes will be instructive with regard to whether those genes currently known to be unique to the pneumovirus genus are essential for replication. In addition, study of TRTV may help to clarify the evolutionary relationships between the genera of the Paramyxoviridae family.

ACKNOWLEDGMENTS

We thank Rachel Purvis for technical assistance, Dr. Jane Cook for providing isolates of TRTV, and Dr. Tom Barrett for advice during the course of this work. This work was supported by the Ministry of Agriculture, Fisheries and Food.

REFERENCES

AFZAL, M. A., ELLIOTT, G. D., RIMA, B. K., and ORVELL, C. (1990). Virus and host cell-dependent variation in transcription of the mumps virus genome. *J. Gen. Virol.* **71,** 615–619.

BELLINI, W. J., ENGLUND, G., RICHARDSON, C. D., ROZENBLATT, S. and LAZZARINI, R. A. (1986). Matrix genes of measles virus and canine distemper virus: Cloning, nucleotide sequences, and deduced amino acid sequences. *J. Virol.* **58,** 408–416.

BLUMBERG, B. M., ROSE, K., SIMONA, M. G., ROUX, L., COLOMBA, G., and KOLAKOFSKY, D. (1984). Analysis of the Sendai virus M gene and protein. *J. Virol.* **52,** 656–663.

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.

BÜECHI, M., and BÄCHI, TH. (1982). Microscopy of internal structures of Sendai virus associated with the cytoplasmic surface of host membranes. *Virology* **120,** 349–359.

CALDWELL, S. E., and LYLES, D. S. (1986). Dissociation of newly synthesized Sendai viral proteins from the cytoplasmic surface of isolated plasma membranes of infected cells. *J. Virol.* **57,** 678–683.

CAVANAGH, D., and BARRETT, T. (1988). Pneumovirus-like characteristics of the mRNA and proteins of turkey rhinotracheitis virus. *Virus Res.* **11,** 241–256.

CAVANAGH, D., and DAVIS, P. J. (1988). Evolution of avian coronavirus IBV: Sequence of the matrix glycoprotein gene and intergenic region of several serotypes. *J. Gen. Virol.* **69,** 621–629.

CHAMBERS, P., MILLAR, N. S., PLATT, S. G., and EMMERSON, P. T. (1986). Nucleotide sequence of the gene encoding the matrix protein of Newcastle disease virus. *Nucleic Acids Res.* **14,** 9051–9061.

CHAMBERS, P., BARR, J., PRINGLE, C. R., and EASTON, A. J. (1990). Molecular cloning of pneumonia virus of mice. *J. Virol.* **64,** 1869–1872.

CHOMCZYNSKI, P., and SACCHI, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* **162,** 156–159.

COLLINS, M. S., and GOUGH, R. E. (1988). Characterization of a virus associated with turkey rhinotracheitis. *J. Gen. Virol.* **69,** 909–916.

COLLINS, P. L., HUANG, Y. T., and WERTZ, G. W. (1984). Identification of a tenth mRNA of respiratory syncytial virus and assignment of polypeptides to the 10 viral genes. *J. Virol.* **49,** 572–578.

COOK, J. K. A., DARBYSHIRE, J. H., and PETERS, R. W. (1976). The use of chicken tracheal organ cultures for the isolation and assay of avian infectious bronchitis virus. *Arch. Virol.* **50,** 109–118.

ELANGO, N. (1989). Complete nucleotide sequence of the matrix protein mRNA of mumps virus. *Virology* **168,** 426–428.

ELLIOTT, G. D., AFZAL, M. A., MARTIN, S. J., and RIMA, B. K. (1989). Nucleotide sequence of the matrix, fusion and putative SH protein genes of mumps virus and their deduced amino acid sequences. *Virus Res.* **12,** 61–75.

FAABERG, K. S., and PEEPLES, M. E. (1988). Association of soluble matrix protein of Newcastle disease virus with liposomes is independent of ionic conditions. *Virology* **166,** 123–132.

GALINSKI, M. S., MINK, M. A., LAMBERT, D. M., WECHSLER, S. L., and PONS, M. W. (1987). Molecular cloning and sequence analysis of the human parainfluenza 3 virus gene encoding the matrix protein. *Virology* **157,** 24–30.

GIUFFRE, R. L. M., TOVELL, D. R., KAY, C. M., and TYRRELL, D. L. J. (1982). Evidence for an interaction between the membrane protein of a paramyxovirus and actin. *J. Virol.* **42,** 963–968.

GOUGH, R. E., COLLINS, M. S., COX, W. J., and CHETTLE, N. J. (1988). Experimental infection of turkeys, chickens, ducks, geese, guinea fowl, pheasants and pigeons with turkey rhinotracheitis virus. *Vet. Rec.* **123,** 58–59.

HEGGENESS, M. H., SMITH, P. R., and CHOPPIN, P. W. (1982). In vitro assembly of the nonglycosylated membrane protein (M) of Sendai virus. *Proc. Natl. Acad. Sci. USA* **79,** 6232–6236.

JONES, R. C., BAXTER-JONES, C., SAVAGE, C. E., KELLY, D. F., and WILDING, G. P. (1987). Experimental infection of chickens with a ciliostatic agent isolated from turkeys with rhinotracheitis. *Vet. Rec.* **120,** 301–302.

JORGENSEN, E. D., COLLINS, P. L., and LOMEDICO, P. T. (1987). Cloning and nucleotide sequence of Newcastle disease virus hemagglutinin–neuraminidase mRNA: Identification of a putative sialic acid binding site. *Virology* **156,** 12–24.

KYTE, J., and DOOLITTLE, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105–132.

LING, R., and PRINGLE, C. R. (1988). Turkey rhinotracheitis virus: In vivo and in vitro polypeptide synthesis. *J. Gen. Virol.* **69,** 917–923.

LIMO, M., and YILMA, T. (1990). Molecular cloning of the rinderpest virus matrix gene: Comparative sequence analysis with other paramyxoviruses. *Virology* **175,** 323–327.

LUK, D., MASTERS, P. S., SÁNCHEZ, A., and BANERJEE, A. K. (1987). Complete nucleotide sequence of the matrix protein mRNA and three intergenic junctions of human parainfluenza virus type 3. *Virology* **156,** 189–192.

MARX, P. A., PORTNER, A., and KINGSBURY, D. W. (1974). Sendai virion transcriptase complex: Polypeptide composition and inhibition by virion envelope proteins. *J. Virol.* **13,** 107–112.

MCDOUGALL, J. S., and COOK, J. K. A. (1986). Turkey rhinotracheitis: Preliminary investigations. *Vet. Rec.* **118,** 206–207.

MCGINNES, L. W., and MORRISON, T. G. (1987). The nucleotide sequence of the gene encoding the Newcastle disease virus membrane protein and comparisons of membrane protein sequences. *Virology* **156,** 221–228.

MOUNTCASTLE, W. E., COMPANS, R. W., LACKLAND, H., and CHOPPIN, P. W. (1974). Proteolytic cleavage of subunits of the nucleocapsid of the paramyxovirus simian virus 5. *J. Virol.* **14,** 1253–1261.

PEEPLES, M. E., and BRATT, M. A. (1984). Mutation in the matrix protein of Newcastle disease virus can result in decreased fusion

glycoprotein incorporation into particles and decreased infectivity. *J. Virol.* **51,** 81–90.

PRINGLE, C. R. (1991). Paramyxoviridae. Arch. Virol. Supplement 2, 242–246.

RIMA, B. K. (1989). Comparison of amino acid sequences of the major structural proteins of paramyxo- and morbilliviruses. *In* "Genetics and Pathogenicity of Negative Strand Viruses" (D. Kolakofsky and B. W. J. Mahy), pp. 254–263. Elsevier, Amsterdam.

SAKAI, Y., SUZU, S., SHIODA, T., and SHIBUTA, H. (1987). Nucleotide sequence of the bovine parainfluenza 3 virus genome: Its 3′ end and the genes of NP, P, C and M proteins. *Nucleic Acids Res.* **15,** 2927–2944.

SAMBROOK, J., FRITSCH, E. F., and MANIATIS, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

SATAKE, M., and VENKATESAN, S. (1984). Nucleotide sequence of the gene encoding respiratory syncytial virus matrix protein. *J. Virol.* **50,** 92–99.

SATO, H., OH-HIRA, M., ISHIDA, N., IMAMURA, Y., HATTORI, S., and KAWAKITA, M. (1987). Molecular cloning and nucleotide sequence of P, M and F genes of Newcastle disease virus avirulent strain D26. *Virus Res.* **7,** 241–255.

SHESHBERADARAN, H., and LAMB, R. A. (1990). Sequence characterization of the membrane protein gene of paramyxovirus simian virus 5. *Virology* **176,** 234–243.

SHIMIZU, K., and ISHIDA, N. (1975). The smallest protein of Sendai virus: Its candidate function of binding nucleocapsid to envelope. *Virology* **67,** 427–437.

SPRIGGS, M. K., JOHNSON, P. R., and COLLINS, P. L. (1987). Sequence analysis of the matrix protein gene of human parainfluenza virus type 3: Extensive sequence homology among paramyxoviruses. *J. Gen. Virol.* **68,** 1491–1497.

WARD, K. A., LAMDEN, P. R., OGILVIE, M. M., and WATT, P. J. (1983). Antibodies to respiratory syncytial virus polypeptides and their significance in human infection. *J. Gen. Virol.* **64,** 1867–1876.

YOSHIDA, T., NAGAI, Y., YOSHII, S., MAENO, K., and MATSUMOTO, T. (1976). Membrane (M) protein of HLV (Sendai virus): Its role in virus assembly. *Virology* **71,** 143–161.

YU, Q., DAVIS, P. J., BARRETT, T., BINNS, M. M., BOURSNELL, M. E. G., and CAVANAGH, D. (1991). Deduced amino acid sequence of the fusion (F) glycoprotein of turkey rhinotracheitis virus has greater identity with that of human respiratory syncytial virus, a pneumovirus, than that of paramyxoviruses and morbilliviruses. *J. Gen. Virol.* **72,** 75–81.