



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.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Genomic Organization and Expression of the 3' End of the Canine and Feline Enteric Coronaviruses

H. VENNEMA,¹ J. W. A. ROSSEN, J. WESSELING, M. C. HORZINEK, AND P. J. M. ROTTIER

Department of Virology, Faculty of Veterinary Medicine, State University of Utrecht, Yalelaan 1, P.O. Box 80.165, 3508 TD Utrecht, The Netherlands

Received May 5, 1992; accepted July 1, 1992

The genomic organization at the 3' end of canine coronavirus (CCV) and feline enteric coronavirus (FECV) was determined by sequence analysis and compared to that of feline infectious peritonitis virus (FIPV) and transmissible gastroenteritis virus (TGEV) of swine. Comparison of the latter two has previously revealed an extra open reading frame (ORF) at the 3' end of the FIPV genome, lacking in TGEV, which is currently designated ORF 6b. Both CCV and FECV possess 6b-related ORFs at the 3' ends of their genomes. The presence of ORF 6b in three of four viruses in this antigenic cluster strongly suggests that TGEV has lost this ORF by deletion. The CCV ORF 6b is collinear with that of FIPV, but the predicted amino acid sequences are only 58% identical. The FECV ORF 6b contains a large deletion compared to that of FIPV, reducing the collinear part to 60%. The sequence homologies were highest between CCV and TGEV on the one hand and between FECV and FIPV on the other. Previously, we showed that the expression product of the FIPV ORF 6b can be detected in infected cells by immunoprecipitation (Vennema *et al.*, 1992). In the present study we have performed similar experiments with CCV and FECV. In infected cells both viruses produced proteins related to but different from the FIPV 6b protein. © 1992 Academic Press, Inc.

INTRODUCTION

Canine coronavirus (CCV), feline enteric coronavirus (FECV), feline infectious peritonitis virus (FIPV), and transmissible gastroenteritis virus (TGEV) of swine belong to one antigenic cluster within the family *Coronaviridae* (Siddell *et al.*, 1983). Sequence analysis revealed a close genetic relatedness between FIPV and TGEV (Jacobs *et al.*, 1987; de Groot *et al.*, 1988; Vennema *et al.*, 1991). FIPV contains an extra open reading frame (ORF) in the 3'-terminal region of its genome (de Groot *et al.*, 1988). It is the second ORF of mRNA 6 which is currently designated ORF 6b, according to the new nomenclature (Cavanagh *et al.*, 1990). The first ORF of mRNA 6, designated ORF 6a, is the counterpart of TGEV ORF 7, previously called ORF-X3 (Kapke and Brian, 1986; Rasschaert *et al.*, 1987) or ORF-4 (Britton *et al.*, 1988). Comparison of ORFs 6a and 7 revealed that the FIPV genome contained an in frame insertion of 69 nucleotides (de Groot *et al.*, 1988). The protein product of TGEV ORF 7 was identified in TGEV-infected cells (Garwes *et al.*, 1989). Recently, the 6b gene was shown to be expressed in FIPV-infected cells (Vennema *et al.*, 1992). It is a glycoprotein which is released into the extracellular medium and is not stably associated with virus particles. In cats it induces antibodies during natural and experimental FIPV infections. Therefore, the 6b protein provides an antigenic distinction between FIPV and TGEV. The aim of the

present study was to examine whether this distinction could be extended to CCV and FECV. We determined the genomic organization of the 3' end of the viral genomes and investigated whether these viruses produced proteins related to the 6b protein of FIPV.

MATERIALS AND METHODS

Cells and viruses

FIPV strain 79-1146, FECV strain 79-1683 (McKeirnan *et al.*, 1981; obtained from Dr. J. Evermann), CCV strain K378 (Dutch field isolate, obtained from Dr. H. Flore), and CCV strain I-71 (Binn *et al.*, 1975) were grown in *Felis catus* whole fetus cells (fcwf-D; obtained from Dr. N. C. Pedersen) and Crandell feline kidney cells (CrFK). Recombinant vaccinia virus vTF7-3 (Fuerst *et al.*, 1986; obtained from Dr. B. Moss) infections were carried out in HeLa cells. Cells were maintained in Dulbecco's modified Eagle's medium (GIBCO Laboratories) containing 5% fetal bovine serum.

Cloning and sequence analysis of the 3' end of CCV

cDNA libraries were prepared of intracellular poly(A)-containing RNA from CCV-infected fcwf-D cells as described elsewhere (Wesseling *et al.*, manuscript in preparation). Clones containing sequences derived from the 3' end of the genome were selected by colony hybridization with restriction fragments of FIPV cDNA clone B12 (de Groot *et al.*, 1988) as probes. Nucleotide sequencing was performed by the dideoxy chain termi-

¹ To whom reprint requests should be addressed.

nation procedure (Sanger *et al.*, 1977) using double-stranded DNA and a bacteriophage T7 DNA polymerase based kit (Pharmacia, LKB). Sequence data were analyzed using the computer programs of Devereux *et al.* (1984).

cDNA synthesis and PCR amplification of the 3' end of FECV

Total intracellular RNA was isolated from FECV-infected fcwf-D cells as described (Chomczynski and Sacchi, 1987). Synthesis of cDNA on total RNA was performed as described (Kawasaki and Wang, 1989) by priming specifically with synthetic oligonucleotide 179 (5'-CCAGTTTTAGACATCGGG-3', reverse complement of nucleotides 1006–1023; de Groot *et al.*, 1988) which binds to a sequence in the 3' noncoding region of FIPV, downstream of ORF 6b. Oligonucleotide 185 (5'-GATCCAGACGTTAGCTC-3', reverse complement of nucleotides 1202–1218; de Groot *et al.*, 1988), was used to prime cDNA synthesis from a position closer to the 3' end. Amplification of cDNA was performed by the polymerase chain reaction (PCR) as described (Kawasaki and Wang, 1989), after the addition of synthetic oligonucleotide 178, 5'-GATGACACACAGGTTGAG-3', which is identical to the carboxyl-terminus of the nucleocapsid (N) protein gene of FIPV (nucleotides 1945–1962; Vennema *et al.*, 1991). The positions of the primers are indicated in Fig. 5. PCR-amplified FECV cDNA fragments were isolated from agarose gel, oligo-dC tailed with terminal transferase and cloned after annealing with oligo-dG-tailed pUC9 (Pharmacia LKB). Sequence analysis was performed as described above for CCV cDNA clones, with pUC/M13 primers and with the primers used for PCR amplification.

Recloning of the CCV and FECV 6b genes

The CCV 6b gene was isolated from cDNA clone C16 as a *Hpa*II–*Pst*I fragment and recloned in pBlue-script SK⁻ (Stratagene) digested with *Acc*I and *Pst*I, yielding pBSC6b. The 6b gene fragment was recloned from this construct as a *Xho*I–*Pst*I fragment into the vector part of pTF6b, (Vennema *et al.*, 1992) digested with the same enzymes. The final construct was designated pTC6b. The FECV 6b gene was recloned from cDNA-PCR clone FE1 as a *Spe*I–*Pst*I fragment in the vector part of pTF6b, prepared by digestion with *Spe*I and *Pst*I. This construct was designated pTE6b. All constructs were used in the transient T7 expression system with recombinant vaccinia virus vTF7-3 producing the T7 RNA polymerase (Fuerst *et al.*, 1986; obtained from Dr. B. Moss).

Radio immunoprecipitation assays (RIPA)

Lysates from coronavirus- or vaccinia virus-infected cells were prepared after metabolic labeling with L-

[³⁵S]cysteine (ICN Biomedicals, Inc.). Lysis, RIPA with ascites fluid from a field case of FIP, and endo-β-N-acetylglucosaminidase H (endo H; Boehringer-Mannheim Biochemicals) treatment were carried out as described (Vennema *et al.*, 1990). The ascites fluid contains antibodies to all viral proteins of FIPV identified so far, which are cross-reactive with proteins from CCV, FECV, and TGEV. Endo-β-N-acetylglucosaminidase F (endo F; Boehringer-Mannheim Biochemicals) digestions were carried out for 16 hr at 30° in 50 mM potassium phosphate buffer, pH 6.8, 20 mM EDTA, 1% Triton X-100, 0.2% sodium dodecyl sulphate (SDS), 1% 2-mercaptoethanol. Analysis by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described (Laemmli, 1970).

RESULTS

Sequence analysis of the 3' end of the CCV genome

CCV cDNA clones were prepared and selected as described under Materials and Methods. Most clones were derived from CCV strain K378, a field isolate from The Netherlands (Fig. 1). A clone derived from the laboratory strain I-71 (Binn *et al.*, 1975) was partially analyzed. The nucleotide sequences from these strains were more than 99% identical. We obtained a contiguous sequence of strain K378, extending 2.5 kb from the poly(A)-tail in the 5' direction (Fig. 2). Translation of the nucleotide sequence revealed three ORFs corresponding to the N protein and ORFs 6a and 6b of FIPV. The organization of the 3' end of the CCV genome was compared to that of FIPV and TGEV (Fig. 5). CCV contains an ORF 6b and an ORF 6a with 69 additional nucleotides which are lacking in ORF 7 of TGEV. The amino acid sequences were aligned with the corresponding sequences of FIPV and TGEV to determine the percentages of identical amino acid residues (Table 1). The N and 6a amino acid sequences have a higher level of identity when compared between CCV and TGEV than between CCV and FIPV. The same was found when the nucleotide sequences were compared. The 6b amino acid sequences of CCV and FIPV were only 58% identical and several small insertions in the CCV sequence were found. Nevertheless, the hydrophobicity plots of the putative 6b proteins were remarkably similar (data not shown), including a short hydrophobic amino-terminus, which may function as a signal sequence. In contrast to the FIPV ORF 6b, the CCV ORF 6b contains no potential N-glycosylation site.

PCR amplification of cDNA derived from the 3' end of the FECV genome

The nucleotide sequences flanking the ORFs 6a and 6b of FIPV and CCV and ORF 7 of TGEV were aligned

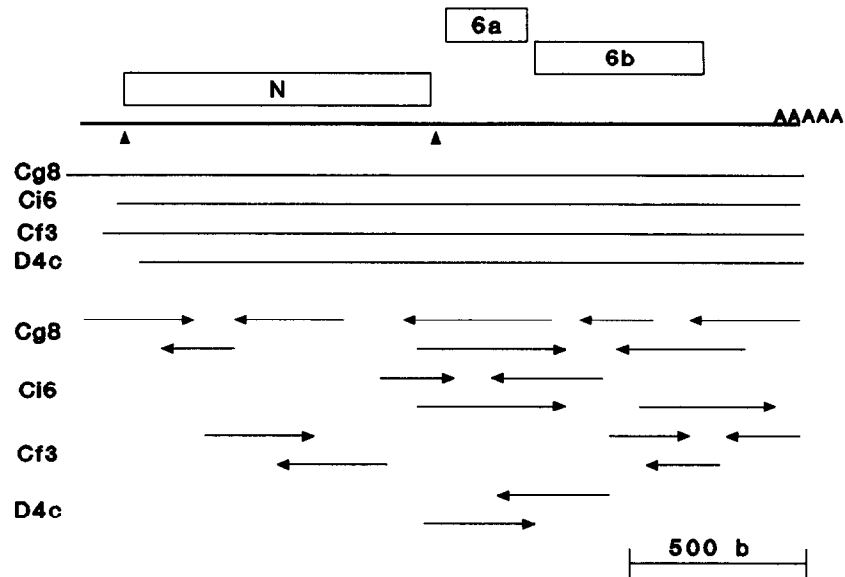


Fig. 1. Sequence strategy for CCV cDNA clones. Horizontal lines represent the relevant parts of cDNA clones Cg8, Cf3, and Ci6 which were all derived from CCV strain K 378. Clone D4c, derived from laboratory strain I-71, was partially analyzed. The direction and extent of the sequences obtained for each clone is represented by arrows in the lower part of the figure. Boxes indicate the open reading frames which were named in analogy to those of FIPV. The solid bar represents the genome. The poly(A) tail is indicated on the right. The arrowheads indicate the positions of conserved intergenic sequences (5'-CTAAAC-3').

to design primers for cDNA synthesis and polymerase chain reaction (PCR) amplification. The positions of the primers are indicated in Fig. 3. Synthetic oligonucleotide 179 was used to prime cDNA synthesis on total RNA isolated from FECV and CCV-infected cells. The cDNA synthesis was followed by PCR amplification after addition of primer 178. The CCV product had the expected size of approximately 1 kbp (data not shown). The FECV product was considerably smaller, being approximately 750 bp. The controls with RNA from mock-infected cells and without RNA were both negative. The sequence analysis presented below revealed that the FECV ORF 6b extended into the sequences used to design the PCR primers. Therefore, cDNA synthesis and PCR amplification were repeated with synthetic oligonucleotides 185 and 178, resulting in a fragment of approximately 950 bp (data not shown).

Analysis of genomic sequences of FECV near the 3' end

The PCR products of FECV were cloned and sequenced. Several independent clones were analyzed (Fig. 3) leading to a contiguous sequence of 957 nucleotides which ends approximately 36 nucleotides upstream of the poly(A)-tail (Fig. 4). Comparison with the corresponding sequence of the FIPV genome showed a single deletion of 238 nucleotides and an overall sequence identity of 93.6%. Translation of the nucleotide sequence revealed the presence of two

ORFs similar to the FIPV ORFs 6a and 6b. The 6a polypeptides are virtually identical, with only one amino acid difference. The deletion is located in ORF 6b. The FECV and FIPV 6b sequences are collinear for the amino-terminal 123 amino acid residues. In this part the identity is 89%. The deletion results in a shift to the -1 reading frame which extends 53 codons. The ORF specifies a polypeptide with a total length of 176 amino acid residues and a predicted mol wt of 20,300. The protein sequence predicts a short hydrophobic amino-terminus, probably acting as a signal sequence and one N-glycosylation site. The genomic organization of the FECV 3' end is similar to FIPV and CCV, which all contain an ORF 6b and the extra 69 nucleotides in ORF 6a as compared to TGEV ORF 7 (Fig. 5). Paired alignments of the collinear parts of the amino acid sequences with the corresponding sequences of FIPV, CCV, and TGEV revealed that FECV is more closely related to FIPV than to CCV and TGEV (Table 1). Similar homologies were found by comparing the nucleotide sequences (data not shown).

Identification of the CCV and FECV 6b proteins

The FIPV 6b protein was readily detected in lysates of FIPV-infected cells (Vennema *et al.*, 1992). It comigrated in SDS-PAGE with an expression product of the cloned 6b gene which was prepared by using the recombinant vaccinia virus T7 RNA polymerase expres-

10 30 50 70 90 110
AAGTCCGGTCCATATGTAATACTAAAGCTGGTACTCAACAGATGCACGAAGTATAATTTGAGTGAGCAAGAAAAATTATACATATGGTATAACTAACTTCTAAATGGCCAAAC
130 150 170 190 210 230
CAAGGCAACCGCTTCAGTGGGGAGTAGATATCACCACAAAAACGTGGTCTTTCCAATTCCTGGCGGGGAAGAATAACTACTACCTCTTTTATTCTTCAACCCCATTCACCTCCAAACA
Q G Q R V Q W G V D I T K K R G L S N S R G R K N N T I P L S F F N P I T L Q Q
250 270 290 310 330 350
GGTTCAAAATTTGGAAGTATGTCGAGAGACTTTGTACCCAAAGGAATAGGTAACAAGGATCAACAGATGGTTATTGGAATAGACAAAGCCGCTATCGCATGGTGAAGGTCAGCGT
G S K F W N L C P R D F V P K G I G N K D Q Q I G Y W N R Q S R Y R M V K G Q R
370 390 410 430 450 470
AAAGAGCTTCTGAAAGTGGTCTTCTACTACTAGTACTGGTCTCATGCTGATGCTAAATTTAAAGATAGAATAGATGGAGTGTCTGGGTTGCCAAGGATGGTCCATGAATAA
K E L P E R W F F Y Y L G T G P H A D A K F K D R I D G V V W V A K D G A M N K
490 510 530 550 570 590
CCAACACTACTGGTAATCGTGGTCTAATAATGAATCCAAAGCTTTGAAATTCGATGGTAAAGTACCAGGAGAATTTCAACTTGAAGTGAACCAATCAAGGCAAACTCAAGATCACCC
P T T L G N R G A N N E S K A L K F D G K V P G E F Q L E V N Q S R D N S R S P
610 630 650 670 690 710
TCTCAATAGATCTCAGTCTAGAAATAGATCTCAATCTAGAGGAAGGCAACAATCCAATAACAAGAAGGATGACAGTGTAGAACAAGCTGTCTGTGCACTCAAAAAGTTAGGTGT
S Q S R S Q S R N R S Q S R G R Q Q S N N K K D D S V E Q A V L A L K L G V
730 750 770 790 810 830
GACACAGAAAACACCAACACGCTCTGTTCTAAATCTAAAGACGTAGCAACTCTAAGCAAGAGACACTACACCTAAGAATGAAAACACACCTGGAAGAGACTGCAGGTAAA
D T E K Q Q Q R S R S K S K E R S N S K T R D T T P K N E N K H T W K R T A G K
850 870 890 910 930 950
GGTGTGATGACAAAATTTTGGAGCTAGAAGTACTTCCAGCAATTTGGTGACAGCGATCTCGTTGCCAATGGGAACGGTCCCAAGCAATACCCACAACCTGGCTGAATGTCTTCCACT
Q D V T K F Y G A R S S S A N F G D S D L V A N G A K H Y P Q L A E C V P S
970 990 1010 1030 1050 1070
GTATCTAGACTCTGTTTGGAAAGTATTGGACTGCAAAGGATGGCAGCAGATTGAAGTACACATACACATAAATACCCTGCAAGGATGATCCTAAGACTGACAAATTCCTT
V S S I L F G S Y W T A K E D G D Q I E V T F T H K Y H L P K D D P K T G Q F L
1090 1110 1130 1150 1170 1190
CAGCAGATTAATGCCTATGCTCCATCAGAGGTGGCAAAAGCAAGAGCAACGTAAGCTCGTTCAAAATCTGACAGAAAGGGTAGAGCAAGAGGTTGACCTGATGCAATTAACAGAA
Q Q I N A Y A R P S E V A K E Q R Q R K A R S K S A E R V E Q E V V P D A L T E
1210 1230 1250 1270 1290 1310
AATATACAGATGTGTTGATGACACAGGTTGAGATATTGATGAGGTAAACGAACTAAACGAACTGCTGTTTCTCCATGCTGATTTATTACAGTTTTACTACTACTAATTTGG
N Y T D V F D D T Q V E I I D E V T N * M L L V F L H A V F I T V L L L L I G
1330 1350 1370 1390 1410 1430
TAGACTCCAATATTAGAAAGATTACTTAATCATTCTCAATCTTAAACTGTCAATATGTTTGGTGTGACTGACACCGGTTGAAAGTAAATGCTTACAGCTTTTGAACCC
R L Q L L E R L L L N H S L N L K T V N N V L G V T D T G L K V N C L Q L L K P
1450 1470 1490 1510 1530 1550
AGACTGTCTGATTTAACATCTTACATAGGAGTTGGCAGAAACAGATTACTAAAAGTAGTACTTCGAGTAATCTTCTAGTCTTACTAGGGTTTGTGCTACAGATTTGTAGTCAC
D C L D F N I L H R S L A E T R L L K V V L R V I F L V L L G F C C Y R L L V T
1570 1590 1610 1630 1650 1670
ATTAATTTAACATCAGAAAGTTTTGATTTTGTACTGTCTTCTTGTGACCGGATATGGCATTAGAAGAAGCATACAAGAATATGACCCAAAAGAGTCCCATGAACATCCAAC
L I * M K F L I F V L C L S L V N G Y G I R R S I Q E Y D P K E S H E H P T M
1690 1710 1730 1750 1770 1790
TGACATGGGAATATTGGAAAAATTTGGAAGTACTTTGATACATCAACAACAGATCCTACTCTACCAATAGGAGCAGAAATCTATTGTGATGAGATTGAAGGATTCATGTT
T W E L L E K F V G S T L Y I T T N Q I L S L P L G A E I Y C D E I E G F P C S
1810 1830 1850 1870 1890 1910
CTTGGCTGGTTATAAAGCTTACGCTTAAACCATATTGATTATCATTGATCTTCCAGTCCGTTCTATTCTTTGATAGCAAAATCTACATTTCTTTAGGTGATAGAGAAGAAAA
W P G Y K A Y A Y N H I D Y H F D L S S P F Y S F V D K F Y I S L G D R E E K I
1930 1950 1970 1990 2010 2030
TTCATCTAGATGGTGGTGGCAACCAAAAGATAAAGATGACTGTTGGTGTGCAACATCTTCTCAGTAACTTCAATTTGAACTCAGATTACCATGACAAGATATGCAAC
H L R V V G A T P K D K R L T V G C R T S F S V N L P I G T Q I Y H D K D M Q H
2050 2070 2090 2110 2130 2150
ATCTTGTGAGGAGACATCTTGAATGTGCTCACAGATTACTTGTGAAGTATTGCAAGCCATTCACATGGTATTGCTTTAAGCAAGCTAAAGGTTTATGATCTGAAGCGCT
L V E G R Y L E C A H R V F V K Y C P S H S H Y C F K D K L K V Y D L K R V
2170 2190 2210 2230 2250 2270
TCAAAGTAGGAAGTGTGTTGAGAAAATCAACCAACAGAAAAGTGAAGTGAAGCAACCCGATGATAAAAAGTGGTTTTCCGAGGAATTACTGGTCATCGCGTCTACTCTCTG
K S R K C F E K I N Q Q K S E L *
2290 2310 2330 2350 2370 2390
TACAGAATGTAAGCAGTGTAAATAGGAGGTACAAGCCACCTATTGCATATTAGGAAGTTAGATTTGATTTGGCAATGCTAGATTAGTAAATTTAGAGAAGTTTAAAGATCCGCTATG
2410 2430 2450 2470 2490
ACGAGCCAACAATGGAAGAGCTACAGCTGGATCTAGTGATTGTTAAAATGTAATAATGTTTGAATAATTTCTCTTTGATAGTATACACAAAAA

Fig. 2. Nucleotide sequence of a 2.5-kb segment of the CCV K378 3' terminal region. The predicted amino acid sequences of the polypeptides encoded by the major open reading frames are presented below the nucleotide sequence in single-letter code. The sequence data have been submitted to the EMBL Data Library and are available under the Accession Number X66717.

sion system (Fuerst *et al.*, 1986). To identify the CCV and FECV 6b proteins in a similar way, their 6b ORFs were recloned in a T7 expression vector. The resulting constructs pTC6b and pTE6b, respectively, and pTF6b containing the FIPV 6b gene (Vennema *et al.*, 1992) were used to transfect HeLa cells infected with recombinant vaccinia virus vTF7-3, which produces T7 RNA polymerase (Fuerst *et al.*, 1986). The expression products were analyzed by metabolic labeling with [³⁵S]cysteine, RIPA, and endo H treatment followed by SDS-PAGE (Fig. 6). The CCV and FECV 6b proteins appeared to be slightly smaller than the FIPV 6b protein, the FECV 6b protein being the smallest. Digestion with endo H which cleaves high-mannose N-linked oligosaccharides, resulted in an approximately 2000 mol wt reduction of the FECV and FIPV 6b protein. The CCV 6b protein, however, was not affected and was also

insensitive to digestion by endo F, which cleaves complex N-linked sugars (data not shown). These observations indicate that FECV and FIPV are glycoproteins while CCV 6b is not. The shift in molecular weight of the FECV and FIPV 6b proteins is consistent with the removal of one sugar side chain (Neuberger *et al.*, 1972). This is in agreement with the predicted numbers of glycosylation sites in the amino acid sequences. The observed molecular weights of the CCV 6b protein and of the FECV 6b protein after deglycosylation are also in agreement with those predicted from the amino acid sequences. The recombinant expression products were compared to the proteins produced in CCV-, FECV-, and FIPV-infected cells, which were analyzed similarly (Fig. 6). All three matched with a protein in the sample from the corresponding coronavirus-infected cell lysate. In the latter samples the

TABLE 1

PAIRED COMPARISONS OF THE COLLINEAR PARTS OF THE AMINO ACID SEQUENCES, IN PERCENTAGES IDENTICAL RESIDUES

	FIPV		CCV			TGEV	
	6a	6b	N	6a	6b	N	7
FECV							
6a	99	—	—	80	—	—	77
6b	—	89	—	—	51	—	—
FIPV							
N	—	—	78	—	—	76	—
6a	100	—	—	79	—	—	77
6b	—	100	—	—	58	—	—
CCV							
N	—	—	100	—	—	93	—
6a	—	—	—	100	—	—	96

three known structural proteins also appeared: the membrane (M), nucleocapsid (N), and spike (S) proteins. In all cases the M protein was partially resistant to endo H digestion. The same was observed for the 6b protein in FIPV-infected cells. The lanes of FECV were overexposed to reveal the 6b protein band. This indicates that the expression level of the FECV 6b protein was much lower than that of CCV and FIPV.

DISCUSSION

The genomic organization of the 3' end of FIPV differs from that of TGEV in that it contains an additional ORF (de Groot *et al.*, 1988). Recently, we identified the expression product of this extra gene, designated 6b (Vennema *et al.*, 1992). These observations prompted us to study CCV and FECV of the same antigenic cluster. Sequence analysis showed that their genomic organization in the 3' terminal region is similar to that of FIPV. Therefore, the presence rather than the absence of ORF 6b appears to be the common theme, suggesting that TGEV has lost the corresponding ORF by deletion. The same inference probably holds true for the 69 nucleotides that are present in all 6a ORFs but not in the corresponding ORF 7 of TGEV. Deletions appear to happen frequently during coronavirus evolution. They occur, for example, in the S and HE protein genes of murine hepatitis virus (MHV; Parker *et al.*, 1989; La Monica *et al.*, 1991), in the S protein gene of porcine respiratory coronavirus (PRCV; Rasschaert *et al.*, 1990), and in the region between the S and M protein genes of MHV strain S (Yokomori and Lai, 1991) and of PRCV (Rasschaert *et al.*, 1990). Among the feline coronaviruses another deletion was found recently in FECV 79-1683 as compared to FIPV in the region between the S and M protein genes (E. Lewis and H.

Vennema, unpublished data). Alignment of nucleotide and amino acid sequences of TGEV, CCV, FIPV, and FECV revealed a close relationship among the strains of this cluster (Table 1). The four strains could be divided into two pairs on the basis of their homologies, TGEV and CCV on the one hand and FIPV and FECV on the other.

Both in CCV- and in FECV-infected cells 6b proteins are produced. Their characterization showed that the FECV 6b protein is glycosylated, like the FIPV 6b protein, while the CCV 6b protein is not. The observation that FIPV, CCV, and FECV induce the synthesis of 6b proteins appears to compromise their antigenic distinction. However, the differences between the 6b proteins may allow discrimination using, e.g., monoclonal antibodies. The deletion of 238 nucleotides in FECV 79-1683 was revealed by cDNA-PCR, allowing discrimination from CCV (Fig. 4) and FIPV 79-1146 (data not shown). It remains to be determined whether this is a universal distinguishing property of FIPV and FECV.

The feline coronaviruses used in this study are almost the same with respect to growth in tissue culture, protein composition, and antigenicity (Boyle *et al.*, 1984; Fiscus and Teramoto, 1987). Small differences were also reported; the N protein of FECV is slightly smaller than that of FIPV (Tupper *et al.*, 1987). Recently, a monoclonal antibody specific for the S protein

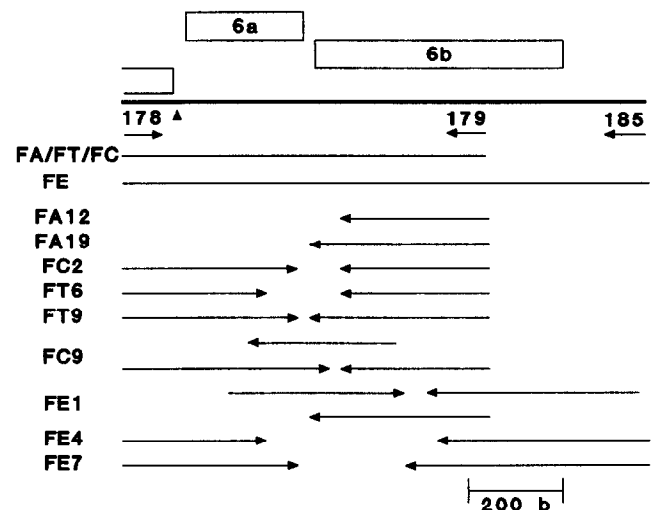


FIG. 3. Sequence strategy for FECV cDNA clones. The positions of primers 178, 179, and 185, used for cDNA synthesis and PCR amplification are indicated. Horizontal lines represent the cloned PCR fragments. FA, FT, and FC series of cloned PCR fragments are derived from primers 178 and 179 and series FE from primers 178 and 185. The direction and extent of the sequences obtained for each clone is represented by arrows in the lower part of the figure. Boxes indicate the open reading frames which were named in analogy to those of FIPV. The solid bar represents the genome. The position of a conserved intergenic sequence (5'-CTAAAC-3') is indicated by an arrowhead.

```

10      30      50      70      90
GATGACACACAGGTTGAGATGATCGATGAGGTAACGAACAAACGCATGCTCGTTTTCTCCATGCTGTACTTGTAAACAGCTTAACTTACTACTAATT
110     130     150     170     190
GGTAGAATCCAATTAAGAAAGTTGTTACTCAGTCTGCTTAATCTCAGCAGCAGTGTAGTAAATGTTTGGGTGTCGACAGTAGCCTGCCAGTAA
G R I Q L L E R L L S H L L N L T T V S N V L G V P D S S L R V N
210     230     250     270     290
ACTGCTTACAGCTTTTGAACACAGACTGCCTTGATTTAAACATCTTACATAAAGTTTACAGAAACAGATTACTAGTAGTAGTACTGCCAGTGATCTT
C L Q L L K P D C L D F N I L H K V L A E T R L L V V V L R V I F
310     330     350     370     390
TCTAGTTCCTTAGGGTTTTCTGCTATACATTGTTGGGTGCATTATTTAACATTATGATTGTTGAGTCTTGTGTATCTTTTGGCTAATGGAAT
L V L L G F S C Y T L L G A L F * M I V V V L V C I F L A N G I
410     430     450     470     490
TAAAGCTACTGCTGTGCAACCTGACCTTCATGAACATCCAGTCCCTTACTGGGACTTGTGCAACATTTTGTGGAAACACCCCTTACATCACCACACAC
K A T A V Q P D L H E H P V L T W D L L Q H F V G N T L Y I T T H
510     530     550     570     590
CAGGCTTAGCACTACCTCTGGGTACAGTGTGGAGTGTGAGAGTGTGATGGTTTCAATTGCACACGGCCTGGCTTCAAATTCGCACATGATCATA
Q V L A L P L G S R V E C E S V D G F N C T R P G F Q N S A H D H I
610     630     650     670     690
TTGATTTCTATTTGATCTTTCTAATCTTTCTATCTTTGTTGGATAATTTTATATTGATGTTGAAGGAAATCAAAAATCAATCTCAGATTAGTTGG
D F Y F D L S N P F Y S F V D N F Y I V E G N Q K I N L R L V G
710     730     750     770     790
TGCTGTGCCAAAACAAAGAGATTAACACTGAGTTATAAGCAACCCGATGTTTAAACTGGTCTTCCGAGGAATTACTGGTCATCGCGTGTCTACTC
A V P K Q K R L K L S Y K A T R C L K L V F P R N Y W S S R C L L
810     830     850     870     890
TTGTACAGAATGGTAAGCACGTGTAATGGGAGGTACAAGCAACCCATTCATATTAGGAAGTTTAGATTTGATTGGCAATGCTAGATTAGTAATTTA
L Y R M V S T C N G R Y K Q P H C I L G S L D L I W Q C *
910     930     950
GAGAAGTTTAAAGATCCGCTATGACGAGCCAAACAATGGAAGAGCTAACGCTCTGGACT
    
```

Fig. 4. Sequence of 957 nucleotides bases of the FECV 79-1683 3' terminal region. Sequences used to design the primers for cDNA synthesis and PCR amplification are underlined. The 5' and 3' terminal 17 bases represent the primer sequences. The predicted amino acid sequences of the polypeptides encoded by the major open reading frames are presented below the nucleotide sequence in single-letter code. The region where the deletion occurred has been indicated with arrowheads. The sequence data have been submitted to the EMBL Data Library and are available under the Accession Number X66718.

of strain 79-1146 was characterized (Hohdatsu *et al.*, 1991). The close resemblance and the low incidence of FIP despite the high proportion of feline coronavirus (FCV) seropositive cats have led to the hypothesis that FECV carriers are the source of FIPV which is generated *de novo* from FECV by minor mutations (Pedersen *et al.*, 1984; Pedersen 1989). Our sequence comparisons show that FIPV 79-1146 and FECV 79-1683 are related more closely to each other than they are to CCV. Therefore, FIPV 79-1146 did not originate from FECV 79-1683 by insertion of genetic information; it would be too much of a coincidence that an insertion with 77% identity is present in exactly the same position in CCV. An alternative mechanism for *de novo* generation of FIPV could be recombination, which is an established phenomenon for coronaviruses (Lai, 1992). Avirulent FIPV-strains and/or FECV-strains are

apparently widespread. These strains could be derived from FIP-inducing strains which attenuated by losing a virulence factor. Recombination of two attenuated viruses lacking different virulence factors during the rare event of a mixed infection could result in the reappearance of a virulent FIPV strain.

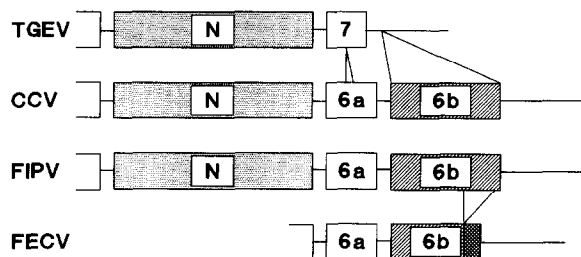


Fig. 5. Schematic comparison of the genomic organization of the 3' terminal regions of TGEV, CCV, FIPV, and FECV. Boxes represent the relevant open reading frames. The cross-hatched box at the C-terminal end of FECV ORF 6b indicates the divergent part.

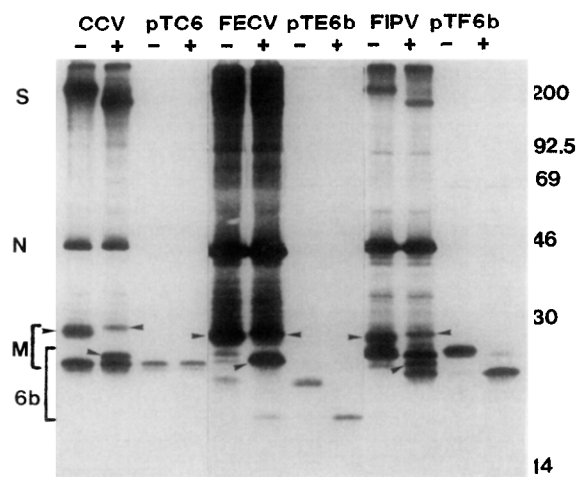


Fig. 6. Radio immunoprecipitation and SDS-PAGE analysis of lysates from CCV-, FECV-, FIPV-, and vTF7-3-infected cells. Recombinant vaccinia virus vTF7-3-infected cells were transfected with the plasmid DNAs pTC6b, pTE6b, and pTF6b, as indicated above the lanes. One-half of each sample was treated with endo H, the other half was mock treated (indicated with + and -, respectively). Structural proteins (S, N, and M) are indicated. In addition, the M protein bands are indicated with arrowheads. The region of the gel in which the 6b protein bands appear overlaps with that of M protein bands.

ACKNOWLEDGMENTS

The authors thank Raoul de Groot and Willy Spaan for critical reading of the manuscript. J.W. was supported by a grant from Solvay-Duphar B.V., Weesp, The Netherlands.

REFERENCES

- BINN, L. N., LAZER, E. C., KEENAN, K. P., HUXSOLL, D. L., MARCHWICKI, R. H., and STRANO, A. J. (1975). Recovery and characterization of a coronavirus from military dogs with diarrhoea. *Proc. 78th Annu. Meeting U.S. Anim. Health Assoc.* 359–366.
- BOYLE, J. F., PEDERSEN, N. C., EVERMANN, J. F., MCKEIRNAN, A. J., OTT, R. L., and BLACK, J. W. (1984). Plaque assay, polypeptide composition and immunochemistry of feline infectious peritonitis virus and feline enteric coronavirus. *Adv. Exp. Med. Biol.* 173, 133–147.
- BRITTON, P., CARMENES, R. S., PAGE, K. W., GARWES, D. J., and PARRA, F. (1988). Sequence of the nucleoprotein gene from a virulent British field isolate of transmissible gastroenteritis virus and its expression in *Saccharomyces cerevisiae*. *Mol. Microbiol.* 2, 89–99.
- 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. J. (1990). Recommendations of the Coronavirus Study Group for the nomenclature of the structural proteins mRNAs, and genes of coronaviruses. *Virology* 176, 306–307.
- CHOMCZYNSKI, P., and SACCHI, N. (1987). Single step method of RNA isolation by acid guanidium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* 162, 156–159.
- DE GROOT, R. J., ANDEWEG, A. C., HORZINEK, M. C., and SPAAN, W. J. M. (1988). Sequence analysis of the 3' end of the feline coronavirus FIPV 79-1146 genome: Comparison with the genome of porcine coronavirus TGEV reveals large insertions. *Virology* 167, 370–376.
- DEVEREUX, J., HAEBERLI, P., and SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12, 387–395.
- FISCUS, S. A., and TERAMOTO, Y. A. (1987). Antigenic comparison of feline coronavirus isolates: Evidence for markedly different peplomer glycoproteins. *J. Virol.* 61, 2607–2613.
- FUERST, T. R., NILES, E. G., STUDIER, F. W., and MOSS, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA* 83, 8122–8126.
- GARWES, D. J., STEWART, F., and BRITTON, P. (1989). The polypeptide of *M_r* 1400 of porcine transmissible gastroenteritis virus: Gene assignment and intracellular location. *J. Gen. Virol.* 70, 2495–2499.
- HOHDATSU, T., SASAMOTO, T., OKADA, S., and KOYAMA, H. (1991). Antigenic analysis of feline coronaviruses with monoclonal antibodies (Mabs): Preparation of Mabs which discriminate between FIPV strain 79-1146 and FECV strain 79-1683. *Vet. Microbiol.* 28, 13–24.
- JACOBS, L., DE GROOT, R. J., HORZINEK, M. C., VAN DER ZEIJST, B. A. M., and SPAAN, W. J. M. (1987). The nucleotide sequence of the peplomer gene of porcine transmissible gastroenteritis virus (TGEV): Comparison with the sequence of the peplomer protein of feline infectious peritonitis virus (FIPV). *Virus Res.* 8, 363–371.
- KAPKE, P. A., and BRIAN, D. A. (1986). Sequence analysis of the porcine transmissible gastroenteritis corona virus nucleocapsid protein gene. *Virology* 151, 41–49.
- KAWASAKI, E. S., and WANG, A. M. (1989). Detection of gene expression. In "PCR Technology: Principles and Applications for DNA Amplification, (H. A. Erlich, Ed.), pp. 89–97. Stockton Press, New York.
- LAEMMLI, U. K. (1970). Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
- LAI, M. M. C. (1992). RNA recombination in animal and plant viruses. *Microbiol. Rev.* 56, 61–79.
- LA MONICA, N., BANNER, L. R., MORRIS, V. L., and LAI, M. M. C. (1991). Localization of extensive deletions in the structural genes of two neurotropic variants of murine coronavirus JHM. *Virology* 182, 883–888.
- MANIATIS, T., FRITSCH, E. F., and SAMBROOK, J. (1982). "Molecular Cloning. A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- MCKEIRNAN, A. J., EVERMANN, J. F., HARGIS, A., MILLER, L. M., and OTT, R. L. (1981). Isolation of feline coronavirus from two cats with diverse disease manifestations. *Feline Pract.* 11, 16–20.
- NEUBERGER, A., GOTTSCHALK, A., MARSHALL, R. O., and SPIRO, R. G. (1972). Carbohydrate–peptide linkages in glycoproteins and methods for their elucidation. p. 450–490. In "The Glycoproteins: Their Composition, Structure and Function" (A. Gottschalk, Ed.). Elsevier, Amsterdam.
- PARKER, S. E., GALLAGHER, T. M., and BUCHMEIER, M. J. (1989). Sequence analysis reveals extensive polymorphism and evidence of deletions within the E2 glycoprotein gene of several strains of murine hepatitis virus. *Virology* 173, 664–673.
- PEDERSEN, N. C. (1989). Animal infections that defy vaccination: Equine Infectious Anemia, Caprine Encephalitis, Maedi–Visna, and Feline Infectious Peritonitis Virus. *Adv. Vet. Sci. Comp. Med.* 33, 413–428.
- PEDERSEN, N. C., EVERMANN, J. F., MCKEIRNAN, A. J., and OTT, R. L. (1984). Pathogenicity studies of feline coronavirus isolates 79-1146 and 79-1683. *Am. J. Vet. Res.* 45, 2580–2585.
- RASSCHAERT, D., DUARTE, M., and LAUDE, H. (1990). Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions. *J. Gen. Virol.* 71, 2599–2607.
- RASSCHAERT, D., GELFI, J., and LAUDE, H. (1987). Enteric coronavirus TGEV: Partial sequence of the genomic RNA, its organization and expression. *Biochimie* 69, 591–600.
- SANGER, F., NICKLEN, S., and COULSON, A. R. (1977). DNA sequencing with chain termination inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- SIDDEL, S., WEGE, H., and TER MEULEN, V. (1983). The biology of coronaviruses. *J. Gen. Virol.* 64, 761–776.
- TUPPER, G. T., EVERMANN, J. F., RUSSEL, R. G., and THOULESS, M. E. (1987). Antigenic and biological diversity of feline coronaviruses: Feline infectious peritonitis and feline enteritis virus. *Arch. Virol.* 96, 29–38.
- VENNEMA, H., HEIJNEN, L., ROTTIER, P. J. M., HORZINEK, M. C., and SPAAN, W. J. M. (1992). A novel glycoprotein of feline infectious peritonitis coronavirus contains a KDEL-like endoplasmic reticulum retention signal. *J. Virol.* 66, 4951–4956.
- VENNEMA, H., HEIJNEN, L., ZUIDERVELD, A., HORZINEK, M. C., and SPAAN, W. J. M. (1990). Intracellular transport of recombinant coronavirus spike proteins: Implications for virus assembly. *J. Virol.* 64, 339–346.
- VENNEMA, H., DE GROOT, R. J., HARBOUR, D. A., HORZINEK, M. C., and SPAAN, W. J. M. (1991). Primary structure of the membrane and nucleocapsid protein genes of feline infectious peritonitis virus and immunogenicity of recombinant vaccinia viruses in kittens. *Virology* 181, 327–335.
- WESLEY, R. D., WOODS, R. D., and CHEUNG, A. K. (1990). Genetic basis for the pathogenesis of transmissible gastroenteritis virus. *J. Virol.* 64, 4761–4766.
- YOKOMORI, K., and LAI, M. M. C. (1991). Mouse hepatitis virus S RNA sequence reveals that nonstructural proteins ns4 and ns5a are not essential for murine coronavirus replication. *J. Virol.* 65, 5605–5608.