

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.

## Sequence Analysis of the 3'-End of Feline Calicivirus Genome

YUKINOBU TOHYA, YASUKO TANIGUCHI, EIJI TAKAHASHI, ETSUKO UTAGAWA, \* NAOKAZU TAKEDA, \* KIKUKO MIYAMURA, \* SHUDO YAMAZAKI, \* AND TAKESHI MIKAMI'

Department of Veterinary Microbiology, Faculty of Agriculture, The University of Tokyo, Bunkyo-ku, Tokyo 113, and 'Central Virus Diagnostic Laboratory, National Institute of Health, Musashimurayama-shi, Tokyo 190-12, lapan

Received February 6, 199 1; accepted April 18, 199 1

The nucleotide sequence of the 3'-end of the Japanese F4 strain of feline calicivirus (FCV) RNA was determined from a cloned cDNA of 3.5 kbp. We found three open reading frames (ORFs). The largest ORF encoded a 668-amino acid protein of 73,588 Da, which was presumably the capsid precursor protein of FCV and had significant amino acid sequence homology with the VP3 of picornaviruses. A small ORF at the extreme 3'-end was compared with that of the F9 strain of FCV, a vaccine strain originally from the U.S. Highly conserved amino acid sequences were shown, suggesting that this ORF might be functional and encode a putative 106-amino acid protein of 12,153 Da. The other ORF in the 5'-flanking region of the cDNA had consensus amino acid sequences conserved among the RNA-dependent RNA polymerases. @ 1991 Academic Press, Inc.

stranded, positive-sense RNA viruses whose spherical geling (10). Recently, Neill (14) has determined the virions are 35-40 nm in diameter and consist of a sin- nucleotide sequence of the central region of the FCV gle major capsid protein of 60-70 kDa (1). Members genome and found a large open reading frame (ORF) belonging to this group are feline calicivirus (FCV), ve-<br>which terminated 2.4 kb from the 3'-end of the gesicular exanthema of swine virus (VESV), and San Mi-<br>nome. The deduced amino acid sequence of the reguel sea lion virus (1). Although the caliciviruses had gion was shown to encode three picornavirus-like nonbeen provisionally classified as a genus of the family structural proteins, RNA-dependent RNA polymerase, Picornaviridae (2) , they were reclassified as a new fam- cysteine protease, and 2C polypeptides. Although the ily Caliciviridae because of, in addition to their distinct data were not published, Neill ( $14$ ) also mentioned the morphology of the virion (3, 4) and possession of a existence of another large ORF in the 3'-terminal 2400 single capsid protein  $(5-7)$ , the presence of a subge-<br>bases of the genomic RNA, which was entirely ennomic RNA found in calicivirus-infected cells (8, 9). coded by 2.4-kb subgenomic RNA.

Relatively few studies with regard to the molecular biology of the caliciviruses have been carried out. Neill and Mengeling ( 10) first cloned cDNAs from the CFI strain of FCV, which was isolated in the U.S. (11), and analyzed the FCV-specific RNAs formed in the virus-infected cells using the cDNAs as probes. They indicated that an 8.2-kb genomic RNA and three subgenomic RNAs of 4.8, 4.2, and 2.4 kb synthesized in the cells were nested, coterminal transcripts with common 3'-ends. Carter (12) reported the sequence of the 3'-terminal 505 nucleotides (nt) of FCV RNA using the F9 vaccine strain whose origin was the U.S. ( 13). The  $\overline{D}$  vacuus used as a probe in Northern block and  $\overline{D}$ 

Sequence data from this article have been deposited with the DDBJ and EMBL/GenBank Data Libraries under Accession No.<br>D90357.  $\overline{S}$  to whom correspondence and representation  $\overline{S}$  should be ad-

dressed.

Caliciviridae is a family composed of single- confirm and extend the observations of Neill and Men-

Previously, we have identified seven neutralizing epitopes of the F4 strain of FCV, the prototype strain of FCV in Japan, with the aid of an extensive panel of neutralizing monoclonal antibodies and neutralizationresistant variants ( 15, 16). Four of the seven epitopes were shown to exist on the capsid protein of the virus by immunoblot analysis ( 76). To further understand by minimizive analysis (19). To further and stown lyze the nucleotide sequence of the gene for the capsio protein. protein.<br>In this paper we report the molecular cloning and the

sequence analysis of a cloned cDNA representing the 30 years of FCV F4 RNA. Within the second of the sequence of t  $3$ -chu bi i ovi i + ntv $\sim$ . Vitaminale Sequence of and 3'-terminal 3516 nt we found three ORFs. The deduced amino acid sequences of the two ORFs at the extreme 3'-end and in the 5'-flanking region of the cDNA were compared with those of other FCV strains reported previously by Carter (12) and Neill (14). The other ORF which located between the two ORFs was suggested to encode the capsid precursor protein having significant amino acid sequence homology with the VP3 of picornaviruses.

Crandell feline kidney (CRFK) cells were grown in Dulbecco's modified Eagle's medium supplemented with 8% fetal calf serum and antibiotics. The FCV F4 strain ( 17) was propagated in CRFK cells and plaquepurified to prepare working stocks of the virus without defective interfering particles. Following adsorption for 90 min at 37", CRFK monolayers infected with the FCV F4 strain were further incubated for 24 hr, and the cell culture fluid was harvested. The cell debris was removed by centrifugation at 8000 rpm for 30 min at 4° (Sorvall GS-3 rotor, U.S.). The resulting supernatant was layered over 25% (w/w) sucrose in phosphatebuffered saline containing 0.1% Triton X-100 (PBS-TX) and centrifuged at 24,000 rpm for 3 hr at 4° (Hitachi RPS25-2 rotor, Japan), The pellet was resuspended in PBS-TX, layered on a 5 to 35% (w/w) sucrose density gradient, and centrifuged at 24,000 rpm for 2 hr at 4°C (RPS25-2 rotor). Fractions were collected from the bottom of tubes. Infectivity and optical density at 260 and 280 nm of the fractions were measured. The peak fractions were pooled and then SDS was added to a final concentration of 0.5%.

Genomic RNA was extracted with phenol-chloroform from the pooled fractions and precipitated with ethanol. The RNA was placed on 15-30% (w/w) sucrose gradient in TNE buffer ( 10 mM Tris-HCI (pH 7.4), 100 mM NaCI, 1 mM EDTA) containing 0.5% SDS and centrifuged at 19,000 rpm for 18 hr at 22" (Hitachi RPS40T rotor). The peak fractions in optical density at 260 nm were pooled and precipitated with ethanol. The purified viral genomic RNA was initially used for cDNA cloning using oligo (dT) as a primer for reverse transcription. cDNA cloning was performed using a modified method of Gubler and Hoffman ( 18). The dCtailed double-stranded DNA was annealed with Pstl-digested and dG-tailed pBR322 plasmid (BRL, Gaithersburg, MD). The annealed molecules were used to transform Escherichia coli MC1061 and screened by plating on Lucrician and Tool and Colochol by  $m$  on terms brown agaily places containing i.e.,  $\mu$ g ml tetracycline. The recombinant plasmids from tetracycline-resistant colonies were screened to determine the length of the inserted molecules. See all recombined molecules. nant plasmids with inserted increduce. Octobal 100011101 nant plasmids with inserts larger than 1 kbp were selected and subjected to restriction enzyme map analysis. Five clones had common EcoRI and PstI sites, which were located approximately in 350 and 1000 bp from the end of the inserts, respectively. The 650-bp WI-fcoRl fragment of pFCV 1 19, one

 $\frac{1}{2}$  the 600-00  $rsn$ - $ccon$  in agment 01  $\mu$  by  $\sigma$  in  $s$ , or of the five clones, was labeled with <sup>32</sup>P by nick-translation and hybridized with total RNAs from FCV F4-in-<br>fected CRFK cells, which were isolated at 5 hr postinfection by the acid guanidinium thiocyanate-phenolchloroform method ( 19), separated on formaldehyde denaturing agarose gels, and transferred to nitrocellulose filters essentially according to the procedures described previously (20). The result of Northern blot analysis revealed that the cDNA probe hybridized to 8.4, 4.3, and 2.5 kb FCV-specific RNAs (data not shown). These bands seemed to represent the fulllength genomic and the subgenomic mRNAs of FCV, which were reported previously (10, 12). Because subgenomic mRNAs of FCV were found to form a 3' coterminal nested set (10, 12), we postulated that pFCV 1 19 contained 3.5 kbp of cDNA insert which represented the sequence of the 3'-end of FCV F4 RNA.

To determine the nucleotide sequence of the 3'-end of FCV F4 genome, various restriction enzyme-digested or deletion fragments of pFCV 119 were subcloned into M13mp10 and mp11. The sequencing of both strands was performed by the dideoxy sequencing method (21) using the 7-Deaza sequence kit (U.S. Biochemical). Sequence analyses were carried out using SDC-GENETYX program, a software package for gene analysis developed by Software Development Co. (Japan). The complete nucleotide sequence of the cDNA representing the 3'-termina13.5 kb of the FCV F4 RNA genome is shown in Fig. 1. Although the insert of pFCV 119 contained only four adenylic acids at the end, the sequencing analysis of other clones indicated that the adenylic acids were a part of a poly(A) stretch at the 3'-end of FCV F4 genome. Figure 2 shows a schematic diagram of the possible ORFs obtained from the predicted amino acid sequences. A large ORF (2004 nt) in frame 3 and a small ORF (318 nt) in frame 2 were found. There was another ORF (1143 nt) in the 5'-flanking region of frame 1, which remained open at the extreme 5'-end.

The short ORF in frame 2 and noncoding region at the 3'-end (nucleotide positions 3 152 to 35 16 in Fig. 1) were compared with those reported by Carter (12) (Fig. 3). Both sequences were identical in nucleotide length and had ATG and TGA codons for the small ORF (nucleotide positions 1 to 318 in Fig. 3) at the same positions. In the coding region, there were 47 nt (14.8%) changes in 318 nt, but only 8 nt changes were accompanied by amino acid changes. The amino acid secompanies by annie abia origingle. The annie abia beductions in this original manner original conserved (32.07% between the strains of different origins, suggesting that this ORF might be functional and encode a protein of 12,153 Da. For this ORF, a corresponding intracellular RNA of 550 bases was detected as a candidate band by Northern blot analyses (12). It will be interesting to determine whether such a protein is synthesized<br>in FCV-infected cells.





Fig. 1. The nucleotide sequence of cloned FCV cDNA representing the 3'-end of FCV F4 RNA and the amino acid sequences of the three open reading frames. The conserved amino acid sequences among the RNA-dependent RNA polymerases are boxed. Underlined region was compared with the amino acid sequences of VP3 of picornaviruses in Fig. 4.

o

When the deduced amino acid sequence of the ORF in the 5'-flanking region of frame 1 was compared with the RNA polymerase-like sequence of FCV reported by Neill (14), strong homology (89.5%) was found between the sequences of the Japanese F4 and the American CFI strains of FCV. It was also found that the sequence in this region of FCV had consensus amino acid sequences conserved among the RNA-depen-



FIG. 2. Visualization of open reading frames within the first 3516 nucleotides of FCV F4 RNA from the 3'-end. All of the three possible frames are shown. The vertical lines indicate the posrtion of terminal codons. Open triangles indicate proximal ATGs in the corresponding frame.

dent RNA polymerases  $(22)$ , including the KDELR, GM(L)PSG, YGDD, and FLKR sequences (boxed sequences in Fig. 1), as indicated by Neill  $(14)$ . Therefore, the ORF in frame 1 was suggested to contain a part of the coding sequence for the RNA polymerase of FCV.

The largest ORF which could encode a protein of 668 amino acids was found in frame 3 (Fig. 2). The position of this ORF in the FCV genome was thought to locate in the region for the most abundant subgenomic mRNA of FCV, one of the mRNAs to form a 3' coterminal nested set, whose size was reported to be 2.4  $(10)$  or 2.7 kb  $(12)$ . We also detected a 2.5-kb RNA by Northern blot analysis (data not shown) using the cDNA probe prepared from this ORF. This mRNA has been considered to produce the capsid protein of FCV  $(8-10, 23)$ . If translation is performed from the first



sig. 3. Companison of the nucleotide and the deduced arritrito actor sequences for the 3'-end of FCV F4 and F9 RNAs. Dashes in the FCV F9 sequences represent identity with the FCV F4 sequences.



FIG. 4. Alignment of amino acid sequences of the VP3-like region of FCV (an underlined amino acid sequence in Fig. 1) and VP3 of picornaviruses. The consensus sequence was calculated to represent areas where a majority of the sequences of many picornaviruses conserve specific residues or a series of very similar residues (26). The asterisks indicate the amino acid residues of FCV conserved in the consensus sequence. The plus signs indicate amino acid residues that are present in all the sequences of the three viruses but not in the consensus sequence. The colons indicate amino acid residues of FCV conserved in one of the two picornaviruses.

ATG in this ORF (nucleotide positions 1 149 to 3152 in Fig. 1) to the termination codon, a protein with a molecular weight of 73,588 is expected to be produced. This molecular weight is similar to that of the 76-kDa capsid precursor protein of FCV estimated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (24). From these observations, it is strongly suggested that this ORF encodes the capsid precursor protein of FCV F4. Preliminary experiments indicated that the capsid protein was blocked to Edoman degradation. The presence of a blocked amino terminus was also indicated in attempts to dansylate the amino terminus of the capsid protein of VESV (25). Further studies will be necessary to identify the cleavage site of the capsid precursor protein.<br>Protein. A computer-aided search of known amino acid se-

quences as contained in the NBRF database revealed significant homology between the deduced amino acid significant nonrology bottoon the acadeca anime acid sequences of the capsid precursor protein of FCV and<br>the capsid protein VP3 of picornaviruses. The amino acid sequences of the VP3 of poliovirus Sabin 1 and add ocquisou of the virus of pollovirus Cubility and region of FCV (and underlined sequence in Fig. 1) were in Fig. 1) region of FCV (an underlined sequence in Fig. 1) were aligned essentially according to the alignments for picornaviral capsid proteins described by Palmenberg  $(26)$  (Fig. 4). Homology observed between FCV and

two representative picornaviruses suggested that these proteins have similar function(s) and have been diverged from a common ancestral protein. Similarities observed not only in nonstructural proteins between FCV and picornaviruses ( 14), but also in a structural protein between these viruses support the grouping of caliciviruses into the superfamily II, picornavirus-like families (27).

It has been suggested that the scheme of transcription of FCV is similar to that employed by the coronaviruses and the alphaviruses ( 10, 14, 28). Our nucleotide and amino acid sequence data appear to confirm the similarity in the genomic structure of FCV with that of the alphaviruses, since both viruses have the gene(s) for structural protein(s) at the 3' region of their genome and the polymerase gene at the center of each genome. However, the junction sequences between the three ORFs of FCV were short (5 nt, nucleotide positions 1 144 to 1 148 in Fig. 1) or absent (1 nt overlapped, nucleotide position 3152 in Fig. 1). The junction sequences in the genomes of the coronaviruses and the alphaviruses are regarded to be of importance in transcriptions of their mRNAs (27,29). The strategy of transcription in FCV replication may differ from that of the alphaviruses or the coronaviruses.

## ACKNOWLEDGMENTS

We are indebted to Dr. lunzo Norimine, Department of Veterinary Microbiology, The University of Tokyo, for Northern blot analysis and to Dr. Joselito A. Limcumpao for help in the preparation of this manuscript. This work was supported by grants from the Ministry of Education, Science and Culture of Japan.

## **REFERENCES**

- 1. SCHAFFER, F. L., BACHRACH, H. L., BROWN, F., GILLESPIE, J. H.. BURROUGHS, 1. N., MADIN. S. H., MADELEY, C. R., POVEY, R. C., SCOTT, F., SMITH, A. W., and STUDDERT, M. J., Intervirology 14,  $1 - 6$  (1980).
- 2. COOPER, P. D., AGOL, V. I., BACHRACH, H. L., BROWN, F., GHEN-DON, Y., GIBBS, A. J., GILLESPIE, J. H., LONBERG-HOLM, K., MAN-DEL, B., MELNICK, J. L., MOHANTY, S. B., POVEY, R. C., RUECK-ERT, R. R., SCHAFFER, F. L., and TYRRELL, D. A. J., Intervirology 10, 165-180 (1978).
- 3. ALMEIDA, J. D., WATERSON, A. P., PRYDIE, J., and FLETCHER, E. W. L., Arch. Gesamte Virusforsch. 25, 105-l 14 (1968).
- 4. PETERSON, J. E., and STUDDERT. M. J., Arch. Gesamte Virusforsch. 32, 249-260 (1970).
- 5. BACHRACH, H. L., and HESS, W. R.. Biochem. Biophys. Res. Commun. 55,141-149 (1973).
- 6. BURROUGHS, J. N., and BROWN, F., J. Gen. Virol. 22, 281–26 (1974).
- 7. SCHAFFER, F. L., and SOERGEL, M. F., J. Virol. 19, 925-9 (1976).
- 8. EHRESMANN, D. W., and SCHAFFER, E. L., *Utrol*. 22, 572-5 (1977).
- 9. BLACK, D. N., BURROUGHS, J. N., HARRIS, T. J. R., and BROWN, F. Nature (London) 274, 614-6 15 ( 1978).
- 10.  $N$ FILLER. D. and MENGELING,  $W \perp N$ rus Res. 11,59-72 (1988).
- 11. CRANDELL, R. A., and MADIN, S. H., Am. J. Vet. Res. 21,551-556 (1960).
- 12. CARTER, M. J., Arch. Virol. 114, 143-152 (1990)
- 13. BITTLE, J.J. VORK, C. J. NEWBERNE, J.W., and MARTIN, M., Am. I. Vet. Res. 21, 547-550 (1960).
- 14. MEILL, D., *Virus Ros*. 17, 145-160 (1990)
- 15. TOHYA, Y., TANIGUCHI, Y., TSUEIAKIMOTO, M., TAKAHASHI, E., and MIKAMI, T., Jpn. J. Vet. Sci. 52, 251-256 (1990).
- 16. TOHYA, Y., MASUOKA, K.. TAKAHASHI, E., and MIKAMI, T., Arch. *Virol.*, in press (1991).
- $17.$  Takahashi, E., Konishi, S., and Ogata. M.,  $\ln l$ ,  $1/6t$ , Soi. 33 81-87 (1971).
- $18.$  Gubien, H., and Horfman, B. J., Gene 25, 263, 269 (1983).
- 10. CHOMOZANSKI, P., and SACCHI, N... Anal. Biochem. 162, 156, 156 (1987).
- 20. SAMBROOK, J., FRITSCH, E. E., and MANIATIS, T., "Molecular Clo ing: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.
- 21. SANGER, F.. NICKLEN, S., and COULSON, A. R., Proc. Nat/. Acad. Sci. USA 74, 5463-5467 (1977).
- 22. Aggres. P.. Kamer, G., Nicklass, M. J. H., and Williams E., Nucle Acids Res. 12,7251-7267 (1984).
- 23. FRETZ, M., and SCHAFFER, F. L., Virology 89, 318-321 (1978).
- 24. CARTER, M. J., Arch. Viral. 108, 69-79 (1989).
- 25. BURROUGHS, J. N., DOEL, T. R., SMALE, C. J., and BROWN, F., J. Gen. Virol. 40, 161-174 (1978).
- $\overline{R}$   $\overline{$ tion and Detection" (B. L. Semler and E. Ehrenfeld, Eds.), pp. 211-241. Am. Soc. Microbiol., Washington, DC, 1989.
- $27.87.$  Strauss, J. H., and Strauss, E. G., Annu. Rev. Microbial. 42, 657-683 (1988).
- $\overline{R}$  . Such that  $\overline{R}$  is the virology (H)  $\overline{R}$  is the computational set of  $\overline{R}$  is the computation of  $\overline{R}$  $\frac{1}{2}$ . The case of the Horzine Company,  $\frac{1}{2}$  and  $\frac{1}{2}$ Conrat, and R. R. Wagner, Eds.), Vol. 14. pp. 249-284. Ple-Conrat, and R. R. Wagner, Eds.), Vol. 14, pp. 249-284. Ple-<br>num, New York, 1979.
- 2939-2952 (1988).