1

VRR 00454

Sequence analysis of nucleocapsid gene and leader RNA of human coronavirus OC43

Toshio Kamahora, Lisa H. Soe and Michael M.C. Lai

Departments of Microbiology and Neurology, University of Southern California,
School of Medicine, Los Angeles, California, U.S.A.

Summary

The nucleotide sequence of the 3′-end of the genomic RNA of human coronavirus OC43 (HCV-OC43) was determined from the cDNA clones of the intracellular virus-specific mRNAs. The nucleotide sequence and the predicted amino acid sequence of the main open reading frame (ORF), which represents the nucleocapsid (N) protein, were highly homologous to those of bovine coronavirus (BCV) Mebus strain. This ORF predicts a protein of 448 amino acids. Additional smaller ORFs are also present in a different reading frame. We have also determined the leader sequence present at the 5′-end of HCV-OC43 mRNAs by a primer extension study. This sequence is highly homologous to that of mouse hepatitis virus, particularly in the 3′-end of the leader sequence, which is postulated to be involved in the unique mechanism of leader-primed transcription. These data suggest that HCV-OC43 and BCV might have diverged from each other fairly recently and that the 3′-end of the leader sequence has significant functional roles.

Human coronavirus; Nucleocapsid sequence; Leader RNA

Human coronaviruses (HCV) are important human respiratory pathogens, responsible for almost 20% of common colds in winter seasons (McIntosh, 1974). They have also been associated with diarrhea of unknown etiology, such as infantile

Correspondence to: M.M.C. Lai, Departments of Microbiology and Neurology, University of Southern California, School of Medicine, Los Angeles, CA 90033, U.S.A.

necrotizing enterocolitis (Resta et al., 1985). Furthermore, coronaviruses have been isolated from autopsied brain of multiple sclerosis patients (Burks et al., 1980), although the identities of these viruses have been questioned (Weiss, 1983) and the work has not been repeated. HCV can be divided into two antigenically distinct groups: one is represented by HCV-OC43, which shares antigenicity with coronaviruses of some other species, such as bovine coronavirus (BCV) and mouse hepatitis virus (MHV) (Macnaughton et al., 1981). The other group is represented by HCV-229E, which shares antigenicity with porcine transmissible gastroenteritis virus (TGEV) and canine coronavirus (CCV) (Pedersen et al., 1978). The close homology between HCV-OC43 and BCV, in particular, has been demonstrated by oligonucleotide fingerprinting analysis of the genomic RNA of these two viruses, which showed remarkable similarity (Lapps and Brian, 1985), and also by im-munoprecipitation of individual viral proteins with specific antisera (Hogue et al., 1984). These studies suggested that HCV-OC43 and BCV may have diverged fairly recently.

Coronaviruses, as a group, contain a single-stranded, positive-sense RNA genome of 6×10^6–8×10^6 daltons (Lai and Stohlman, 1978; Lomniczi and Kennedy, 1977). In infected cells, the viral RNA genome is first transcribed into a full-length negative-strand RNA, which, in turn, is transcribed into a positive-sense genomic RNA and six subgenomic mRNAs (Lai et al., 1981). These mRNAs are arranged in a 3'-coterminal nested set structure, in which the sequence of every mRNA is contained within the sequence of the next larger mRNA (Lai et al., 1981). Further-more, each mRNA contains an approximately 70-nucleotide leader RNA which is derived from the 5'-end of the genomic RNA (Lai et al., 1982, 1983, 1984; Spaan et al., 1983). The derivation of the leader RNA takes place by a mechanism of "leader-primed transcription", in which the leader RNA is transcribed indepen-dently, dissociated from the template and then binds to the template at the downstream transcriptional start sites (Baric et al., 1983; Makino et al., 1986). This mechanism appears to involve recognition of a stretch of repeat sequences present at both the 3'-end of the leader sequence and the transcriptional start sites of the template RNA (Shieh et al., 1987; Makino et al., 1988).

Relatively little of the molecular biology of human coronaviruses has been studied. It has been shown that the size and species of virion RNA and intracellular viral RNAs of at least one human coronavirus, 229E, are comparable to those of other coronaviruses (MacNaughton, 1978; Weiss, 1983). Also, both HCV-OC43 and HCV-229E consist of at least three structural proteins: a nucleocapsid (N) protein of 52 kDa, an envelope peplomer (E2) protein of 190 kDa and a matrix (E1) protein of approximately 26 kDa (Schmidt and Kenny, 1982; Hogue et al., 1984). Several additional minor proteins, such as a glycoprotein of 140 kDa, have also been reported (Hogue et al., 1984; Hierholzer et al., 1972).

In this report, we studied the sequence of both the 3'- and 5'-ends of the genomic RNA of HCV-OC43. We determined the sequence of the nucleocapsid gene, which shows the presence of a very strong conservation between HCV-OC43 and BCV. We also found that the leader sequence of HCV-OC43, particularly at its 3'-end, is highly homologous to that of MHV.

The HCV-OC43 strain (McIntosh et al., 1967) was originally obtained from Dr. Marion Cooney of the University of Washington, Seattle, and was propagated on a human rectal tumor (HRT) cell line (Tompkins et al., 1974). The virus harvested from the medium of infected cell culture was purified according to published procedures (Makino et al., 1984) and viral RNA was extracted as described (Kamahora et al., 1979). The purified viral genomic RNA was initially used for cDNA cloning using oligo (dT) as a primer for reverse transcription. Subsequent cDNA cloning followed essentially the procedures of Gubler and Hoffman (1983), as modified by Shieh et al. (1987). Several cDNA clones specific for HCV-OC43 were obtained; one of these, T16, has a 0.6 kilobase (kb) insert and hybridized to all of the virus-specific subgenomic mRNA species in Northern blot analysis (data not shown). This result suggests that T16 represents a cDNA clone of the 3′-end of the genomic RNA since the subgenomic mRNAs of coronaviruses have a 3′-end co-terminal nested-set structure (Stern and Kennedy, 1980; Lai et al., 1981); thus, only the cDNA probe representing the 3′-end of the genomic RNA could hybridize to all of the subgenomic mRNA species.

Additional cDNA clones were obtained using poly(A)-containing intracellular RNA from HCV-OC43-infected HRT cells as template and oligo(dT) as primer. The virus-specific clones were identified with a nick-translated T16 as the probe. The positive clones were further tested by Northern blot analysis of the intracellular RNA. All of the cDNA clones selected hybridized specifically to HCV-OC43-specific mRNAs, similar to the mRNA patterns detected with T16 as the probe (data not shown). Thus, these cDNA clones represent sequences of the 3′-end of HCV-OC43 RNA, overlapping at least part of the 3′-most gene, which encodes the N protein.

Several representative cDNA clones were sequenced. The sequencing strategy is shown in Fig. 1. Sequencing was carried out by Sanger's dideoxyribonucleotide chain termination method (Sanger et al., 1977). In some cases, deoxyinosine triphosphate (Bankier and Barrell, 1983) was used in place of deoxyguanosine triphosphate to reduce GC compression. Clone A3 has a poly A tract of 19 nucleotides at its 3′-end and represents a nearly full-length cDNA clone of mRNA 7 (about 1700 bases). Clone M6 is 2.4 kb in length, the 3′-terminal 1.6 kb of which overlaps with the A3 clone. However, the 5′-end 0.8 kb is distinct from the corresponding region of A3. Thus, the clone M6 probably represents a cDNA clone of mRNA6. This interpretation is consistent with the presence of the consensus intergenic sequence in this clone, at the position upstream of the open reading frame (ORF) of mRNA 7 (see below).

The complete sequence of the 3′-terminal 1.7 kb of the HCV-OC43 RNA, covering the entire nucleocapsid gene, was obtained from both A3 and M6 clones. This sequence is shown in Fig. 2. Both A3 and M6 clones are identical starting from nucleotide 69 in reference to mRNA 7 sequence (see below), until the 3′-end. The entire sequence is translated in three possible reading frames (Fig. 3). The largest ORF can code for a protein of 448 amino acids. This predicted protein is highly homologous to the corresponding ORF of BCV Mebus strain (Lapps et al., 1987) (Fig. 2). These two proteins have identical predicted numbers of amino acids. There are 44 nucleotide differences and 11 amino acid substitutions between the two
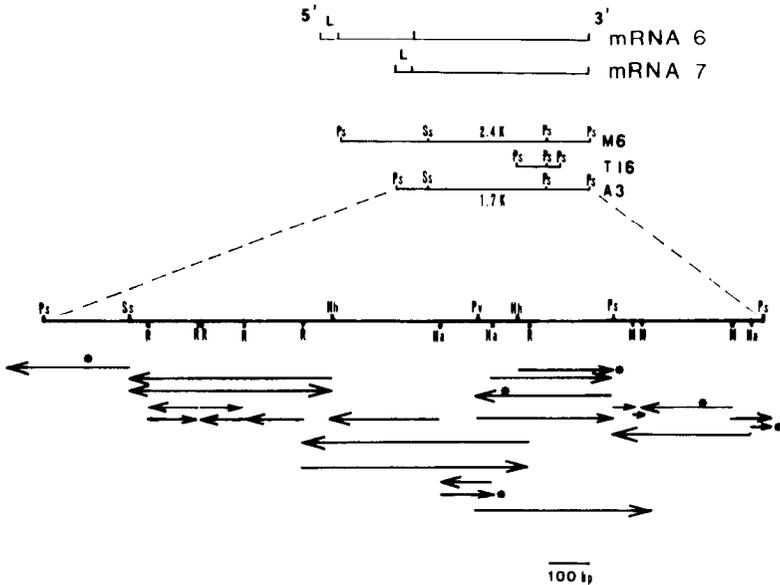
Fig. 1. Restriction endonuclease map and sequencing strategy for HCV-OC43 cDNA clones A3, T16 and M6. Restriction enzyme sites are: Ps, Pst I; Hh, Hha I; Pv, Pvu II; Ss, Sst I; R, Rsa I; Ha, Hae III; M, Mbo II. The arrows with the mark (*) indicate the sequences determined by using both dGTP and dITP. The Pst I sites at the ends of each fragment are the cloning sites on the vector. The corresponding mRNA structures are included for comparison. L represents the leader sequence.

strains. This protein is most likely the nucleocapsid protein of the virus. There are two additional smaller ORFs on the first reading frame. These two ORFs could potentially encode proteins of 60 and 108 amino acids, respectively. This structure is different from that of BCV RNA since the corresponding region of BCV RNA has a single larger continuous ORF capable of encoding a protein of 207 amino acids (Lapps et al., 1987). The functional significance of these ORFs is not clear.

The 5'-ends of the clones A3 and M6 are divergent as shown by the two branched sequences in Fig. 2. At the juncture of divergence, both clones contain a stretch of sequence, TCTAAAT (nucleotides 69-75 of A3), which is very similar to the postulated consensus leader RNA binding site of MHV (Shieh et al., 1987). In the case of MHV, this stretch of sequence is present at both the 3'-end of the leader RNA sequence and the intergenic regions of the genomic RNA and subgenomic RNAs (Shieh et al., 1987). Thus, the 5'-ends of these two clones may represent the sequences of leader RNA and part of mRNA 6-coding sequence, respectively. To determine the leader sequence of HCV-OC43 mRNA, we performed a primer-extension study using a ^{32}P-5'-end-labeled oligodeoxyribonucleotide (5'-CATCCT-TAAAATTTA-3') complementary to nucleotides 71 to 85 (the underlined region in Fig. 2) as the primer and mRNA 7 as the template. The cDNA product extended with reverse transcriptase was sequenced by the Maxam and Gilbert method (1980). The sequence determined by this method is identical to the 5'-end sequence of the clone A3. Similar primer extension studies have also been performed using the same

```
                                  intergenic region    ******ACACCGCATTGTTGAGAAATAAT
                                                                                   ↓
   1        5'TTGTGAGCGAAGTTGCGTGCGTGCATCCCGCTTCACCTGATCTCTTGTTAGATCTTTTTCTAATCTAATCTAAATTTTAAGG    82
```

```
  83  ATGTCTTTTACTCCTGGTAAGCAATCCAGTAGTAGAGCGTCCTCTGGAAATCGTTCTGGTAATCGCATCCTCAAGTGGGCCGATCAGTCC  172
   1   M  S  F  T  P  G  K  Q  S  S  S  R  A  S  S  G  N  R  S  G  N  G  I  L  K  W  A  D  Q  S   30
                F                                                                                BCV
```

```
 173  GACCAGGTTAGAAATGTTCAAACCAGGGGTAGAAGAGCTCAACCCAAGCAAACTGCTACCTCTCAGCAACCATCAGGAGGGAATGTTGTA  262
  31   D  Q  V  R  N  V  Q  T  R  G  R  R  A  Q  P  K  Q  T  A  T  S  Q  Q  P  S  G  G  N  V  V   60
             S                                                                L                  BCV
```

```
 263  CCCTACTATTCTTGGTTCTCTGGAATTACTCAGTTTCAAAAGGGAAAGGAGTTTGAGTTTGTAGAAGGACAAGGTGTGCCTATTGCACCA  352
  61   P  Y  Y  S  W  F  S  G  I  T  Q  F  Q  K  G  K  E  F  E  F  V  E  G  Q  G  P  P  I  A  P   90
                                                        A                                        BCV
```

```
 353  GGAGTCCCAGCTACTGAAGCTAAGGGGTACTGGTACAGACACAACAGAGGTTCTTTTAAAACAGCCGATGGCAACCAGCGTCAACTGCTG  442
  91   G  V  P  A  T  E  A  K  G  Y  W  Y  R  H  N  R  G  S  F  K  T  A  D  G  N  Q  R  Q  L  L  120
                                                                                                 BCV
```

```
 443  CCACGATGGTATTTTTACTATCTGGGAACAGGACCGCATGCTAAAGACCAGTACGGCACCGATATTGACGGAGTCTACTGGGTCGCTAGT  532
 121   P  R  W  Y  F  Y  Y  L  G  T  G  P  H  A  K  D  Q  Y  G  T  D  I  D  G  V  Y  W  V  A  S  150
                                                                                 F               BCV
```

```
 533  AACCAGGCTGATGTCAATACCCCGGCTGACATTGTCGATCGGGACCCAAGTAGCGATGAGGCTATTCCGACTAGGTTTCCGCCTGGCACG  622
 151   N  Q  A  D  V  N  T  P  A  D  I  V  D  R  D  P  S  S  D  E  A  I  P  T  R  F  P  P  G  T  180
                                  L                                                              BCV
```

```
 623  GTACTCCCTCAGGGTTACTATATTGAAGGCTCAGGAAGGTCTGCTCCTAATTCCAGATCTACTTCGCGCACATCCAGCAGAGCCTCTAGT  712
 181   V  L  P  Q  G  Y  Y  I  E  G  S  G  R  S  A  P  N  S  R  S  T  S  R  T  S  S  R  A  S  S  210
                                                                            A                    BCV
```

```
 713  GCAGGATCGCGTAGTAGAGCCAATTCTGGCAATAGAACCCCTACCTCTGGTGTAACACCTGACATGGCTGATCAAATTGCTAGTCTTGTT  902
 211   A  G  S  R  S  R  A  N  S  G  N  R  T  P  T  S  G  V  T  P  D  M  A  D  Q  I  A  S  L  V  240
                                                                                                 BCV
```

```
 903  CTGGCAAAACTTGGCAAGGATGCCACTAAACCTCAGCAAGTAACTAAGCATACTGCCAAAGAAGTCAGACAGAAAATTTTGAATAAGCCC  992
 241   L  A  K  L  G  K  D  A  T  K  P  Q  Q  V  T  K  H  T  A  K  E  V  R  Q  K  I  L  N  K  P  270
                                                Q                     I                          BCV
```

```
 993  CGCCAGAAGAGGAGCCCCAATAAACAATGCACTGTTCAGCAGTGTTTTGGTAAGAGAGGCCCTAATCAGAATTTTGGTGGTGGAGAAATG 1082
 271   R  Q  K  R  S  P  N  K  Q  C  T  V  Q  Q  C  F  G  K  R  G  P  N  Q  N  F  G  G  G  E  M  300
                                                                                                 BCV
```

```
1083  TTAAAAACTTGGAACTAGTGACCCACAGTTCCCCATCTTTGCAGAACTCGCACCCACAGCTGGTGCGTTTTTCTTTGGATCAAGATTAGAG 1172
 301   L  K  L  G  T  S  D  P  Q  F  P  I  L  A  E  L  R  T  A  G  A  F  F  G  S  R  L  E  330
                                                                                                 BCV
```

```
1173  TTGGCCAAAGTGCAGAATTTATCTGGGAATCCTGACGAGCCCCAGAAGGATGTTTATGAATTGCGCTATAACGGCGCAATTAGGTTTGAC 1262
 331   L  A  K  V  Q  N  L  S  G  N  P  D  E  P  Q  K  D  V  Y  E  L  R  Y  N  G  A  I  R  F  D  360
                                   L                                                             BCV
```

```
1263  AGTACACTTTCAGGTTTTGAGACCATAATGAAGGTGTTGAATGAGAATTTGAATGCATATCAACAACAAGATGGTATGATGAATATGAGT 1352
 361   S  T  L  S  G  F  E  T  I  M  K  V  L  N  E  N  L  N  A  Y  Q  Q  Q  D  G  M  M  N  M  S  390
                                                                                                 BCV
```

```
1353  CCAAAACCACAGCGTCAGCGTGGTCATAAGAATGGACAAGGAGAAAATGATAATATAAGTGTTGCAGTGCCCAAAGCCGCGTGCAGCAA 1442
 391   P  K  P  Q  R  Q  R  G  H  K  N  G  Q  G  E  N  D  N  I  S  V  A  V  P  K  S  R  V  Q  Q  420
                                                                           A                     BCV
```

```
1443  AATAAGAGTAGAGAGTTGACTGCAGAGGACATCAGCCTTCTTAAGAAGATGGATGAGCCCTATACTGAAGACACCTCAGAAATATAAGAG 1532
 421   N  K  S  R  E  L  T  A  E  D  I  S  L  L  K  K  M  D  E  P  Y  T  E  D  T  S  E  I  Z     448
                                                                                                 BCV
```

```
1533  AATGAACCTTATGTCGGCATCTGGTGGTAACCCCTCGCAGAAAAGTCGAGATAAGGCACTCTCTATCAGAATGGATGTCTTGCTGCTATA 1622
```

```
1623  ATAGATAGAGAAGGTTATAGCAGACTATAGATTAATTAGTTGAAAGTTTTGTGTTGTAATGTATAGTGTTGGAGAAAGTGAAAGACTTGC 1712
```

```
1713  GGAAGTAATTGCCGACAAGTGCCCAAGGGAAGAGCCCAGCATGTTAAGTTACCACCCAGTAATTAGTAAATGAATGAAGTTAATTATGGCC 1802
```

```
1803  AATTGGAAGAATCAC-POLY(A)-3'                                                                 1817
```

Fig. 2. The primary nucleotide sequence of clone A3 and the deduced amino acid sequences for the largest ORF. The region underlined (nucleotides 71 through 85) were used for preparing a complementary oligodeoxyribonucleotide to serve as a primer for primer-extension studies (Lai et al., 1984) and subsequent determination of the leader sequence. The sequence preceded by an asterisk and above the nucleotides 46 through 68 was obtained from clone M6 and represents the partial sequences of mRNA 6 and intergenic region between mRNAs 6 and 7. The individual amino acid codes beneath the translated continuous amino acids correspond to amino acid differences in the BCV Mebus strain (Lapps et al., 1987), as compared to the HCV-OC43 strain.
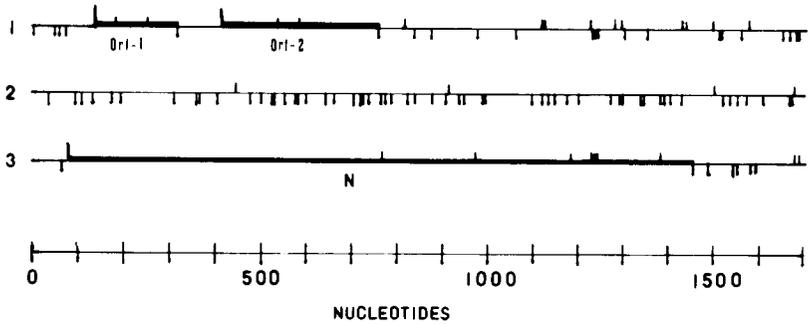
Fig. 3. Schematic diagram of possible open reading frames obtained when translating the A3 nucleotide sequence as virus-sense RNA. Vertical bars above the line represent the methionine codon that could serve as the initiation site for translation. Vertical bars below the line represent termination codons. N, sequence of the nucleocapsid protein gene. Orf-1, Orf-2, possible short open reading frames.

synthetic primer and mRNA 6 as the template. The sequence determined from the primer-extended cDNA product was identical to that determined from clone M6 (the sequence preceded by asterisks) (Fig. 2). Thus, the juncture where the sequences of M6 and A3 clones diverge represents the leader-mRNA fusion site.

The sequence analysis presented in this communication confirmed previous reports of close relationship between HCV-OC43 and BCV as revealed by serological analysis (Pedersen et al., 1978; Gerna et al., 1981), immunoprecipitation of virion structural proteins (Hogue et al., 1984) and oligonucleotide fingerprinting of genomic RNA (Lapps and Brian, 1985). Among the 448 amino acids predicted for the N protein, only 11 amino acids differ between the two viruses, which represents 97.5% homology. This result suggests that these two viruses have only recently diverged from each other. Although the remaining sequence of their genomic RNAs have not been compared, they are likely to be very homologous as well since their genomic RNA fingerprints are extremely similar (Lapps and Brian, 1985). Interestingly, these two viruses have different target cell specificity in vitro and infect different animals. Furthermore, HCV-OC43 causes mainly respiratory illness while BCV mainly affects the gastrointestinal system. Thus, these two viruses may provide a useful tool for understanding the molecular basis of these biological properties of coronaviruses.

An additional difference between the sequences of the 3′-end of the genomic RNA of the two viruses is that BCV has a second ORF on a different reading frame from that of the N gene. This ORF has the potential to code for a protein of 207 amino acids (Lapps et al., 1987). In contrast, the corresponding ORF in HCV-OC43 RNA is interrupted by a termination codon, leaving two smaller ORFs, potentially capable of coding for proteins of 60 and 108 amino acids, respectively. It is not clear whether these ORFs are functional; it is interesting to note that these ORFs have an optimum translation initiation signal, according to M. Kozak's consensus sequences (Kozak, 1984, 1986), in both BCV and HCV-OC43 RNA. It will be interesting to determine whether such proteins are synthesized in HCV-OC43 or BCV-infected cells.

A

```
          10        20        30        40        50        60        70        80
TTGTGAGCGAAGTT*GCGTGCGTGCAT*CCCGCTTCCACTGAT***CTCTTGTTAGATCTTTTTCTAATCTAATCTAAATTTTAAGGATG    HCV-OC43
  :  :  ::  :     ::  ::::  :::  :  :  :::  ::  :  :::       ::::::  :::  :  :    :::::::::::::::::::  :::::::::::
   TATAAGAGTGATTGGCGTCCGTACGTACCCTCT*CTACTCTAAAACTCTTG*TAG*TTTAAATCTAATCTAATCTAAACTTTAAGGATG   MHV-JHM
```

B

```
          10        20        30        40        50        60        70        80
TTGTGAGCGAAGTTGCGTGCGTGCATCCCGCTTCCACTGATCTCTTG**TTAGATCTTTTTCT*AATCTAATCTAAATTTTAAGGATG    HCV-OC43
  ::  ::  ::  ::          :  :::       ::  ::::  :  :  ::::       :::::  :::   ::  ::  :  ::  :      :   :::  :
ACTTAAGATAGATATT***AATATATATC**TATTACACT*AGC*CTTGCGCTAGATTTTTAACTTAA*CAAAGCAGGAGAC*AAGCAGA   IBV
```

Fig. 4. Comparison of leader sequences between HCV-OC43 and MHV-JHM (Shieh et al., 1987), and between HCV-OC43 and infectious bronchitis virus (IBV) (Brown et al., 1986). Asterisks represent nucleotides missing at the corresponding positions and were used to assist sequence alignment. The underlined region denotes UCUAA repeats.

Our data also show that the leader sequences of HCV-OC43 and MHV are very well conserved (73%) (Fig. 4). The sequence conservation is particularly notable in the 3'-half of the leader sequence, in which 26 of 27 nucleotides are perfectly matched (96% homology). It has been shown that this region contains the binding sites for the leader RNA to template RNA during the transcription of coronavirus mRNAs (Shieh et al., 1987; Makino et al., 1988). In contrast, the 5'-end of the leader RNA was not as well conserved (60% homology). This conservation pattern is consistent with the functional significance of the 3'-end of the leader RNA in coronavirus RNA transcription (Shieh et al., 1987). It is notable that only one UCUAA sequence (nucleotides 69–73) overlaps between mRNA 6 and mRNA 7, while there are three UCUAA repeats in the leader region of mRNA 7 (Fig. 2). The UCUAA sequence has been implicated in leader RNA binding (Makino et al., 1988). The presence of multiple UCUAA repeats in the leader RNA predicts that mRNA 7 of HCV-OC43 may be heterogeneous in number of UCUAA repeats, as has been shown for MHV mRNAs (Makino et al., 1988). When compared to IBV (Brown et al., 1986), which belongs to a separate antigenic group of coronaviruses, the sequence conservation is not as remarkable (52%) (Fig. 4). The significance of the leader sequence conservation will require further analysis of additional coronavirus strains.

We have previously shown that different strains of MHV could freely exchange the leader RNA on their subgenomic mRNAs during mixed infection (Makino et al., 1986). Furthermore, different strains of MHV could undergo RNA recombination at a very high frequency (Makino et al., 1986; Keck et al., 1988). The phenomena of leader RNA reassortment and RNA recombination have not been demonstrated between coronaviruses of different species. The finding presented in this report that the leader RNAs of HCV-OC43 and MHV are highly homologous, particularly in the 3'-terminal region where leader RNA binds to template RNA, suggests that the leader RNA of HCV-OC43 and MHV might be exchangeable during mixed infection. Furthermore, the sequence homology in the nucleocapsid

gene and possibly other genes as well suggests that these two viruses might be able to undergo interspecies RNA recombination between them. Such possibilities are currently being investigated in our laboratory.

Acknowledgments

References

Bankier, A. and Barrell, B.G. (1983) Shotgun DNA sequencing. In: R.A. Flavell (E.d.), Techniques in the Life Sciences (Biochemistry), vol. B5: Techniques in Nucleic Acid Biochemistry, pp. 1–34. Elsevier Biomedical Press, Amsterdam.

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

Brown, T.D.K., Boursnell, M.E.G., Binns, M.M. and Tomley, F.M. (1986) Cloning and sequencing of 5′-terminal sequences from avian infectious bronchitis virus genomic RNA. J. Gen. Virol. 67, 221–228.

Burks, J.S., DeVald, B.D., Jankovsky, L.C. and Gerdes, J. (1980) Two coronaviruses isolated from central nervous system tissue of two multiple sclerosis patients. Science 209, 933–936.

Gerna, G., Cereda, P.M., Grazia-Revello, M., Cattaneo, E., Battaglia, M. and Torsellini-Gerna, M. (1981) Antigenic and biological relationships between human coronavirus OC43 and neonatal calf diarrhea coronavirus. J. Gen. Virol. 54, 91–102.

Gubler, U. and Hoffman, B.J. (1983) A simple and very efficient method for generating cDNA libraries. Gene 25, 263–269.

Hierholzer, J.C., Palmer, E.L., Whitfield, S.G., Kaye, H.S. and Dowdle, W.R. (1972) Protein composition of coronavirus OC43. Virology 48, 516–527.

Hogue, B.G., King, B. and Brian, D.A. (1984) Antigenic relationships among proteins of bovine coronavirus, human respiratory coronavirus OC43, and mouse hepatitis coronavirus A59. J. Virol. 51, 384–388.

Kamahora, T., Sugiyama, H., Nomoto, A., Yoshida, M. and Toyoshima, K. (1979) RNA specific for the transforming component of avian erythroblastosis virus strain R. Virology 96, 291–294.

Keck, J.G., Soe, L.H., Makino, S., Stohlman, S.A. and Lai, M.M.C. (1988) RNA recombination of murine coronaviruses: recombination between fusion-positive MHV-A59 and fusion-negative MHV-2. J. Virol. 62, 1989–1998.

Kozak, M. (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. 12, 857–872.

Kozak, M. (1986) Bifunctional messenger RNAs in eukaryotes. Cell 47, 481–483.

Lai, M.M.C. and Stohlman, S.A. (1978) The RNA of mouse hepatitis virus. J. Virol. 26, 236–242.

Lai, M.M.C., Brayton, P.R., Armen, R.C., Patton, C.D., Pugh, C. and Stohlman, S.A. (1981) Mouse hepatitis virus A59: messenger RNA structure and genetic localization of the sequence divergence from the hepatotropic stain MHV-3. J. Virol. 39, 823–834.

Lai, M.M.C., Patton, C.D. and Stohlman, S.A. (1982) Further characterization of mouse hepatitis virus mRNAs: presence of common 5'-end nucleotides. J. Virol. 41, 557–565.

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

Lai, M.M.C., Baric, R.S., Brayton, P.R. and Stohlman, S.A. (1984) Characterization of leader RNA sequence on the virion and mRNAs of mouse hepatitis virus: a cytoplasmic virus. Proc. Natl. Acad. Sci. USA 81, 3626–3630.

Lapps, W. and Brian, D.A. (1985) Oligonucleotide fingerprints of antigenically related bovine coronavirus and human coronavirus OC43. Arch. Virol. 86, 101–108.

Lapps, W., Hogue, B.G. and Brain, D.A. (1987) Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein gene. Virology 157, 47–57.

Lomniczi, B. and Kennedy, I. (1977) Genome of infectious bronchitis virus. J. Virol. 24, 99–107.

Macnaughton, M.R. (1978) The genome of three coronaviruses. FEBS Lett. 94, 191–194.

Macnaughton, M.R., Madge, M.H. and Reed, S.E. (1981) Two antigenic groups of human coronaviruses detected by using enzyme-linked immunosorbent assay. Infect. Immun. 33, 734–737.

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

Makino, S., Stohlman, S.A. and Lai, M.M.C. (1986) Leader sequences of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription. Proc. Natl. Acad. Sci USA 83, 4204–4208.

Makino, S., Soe, L.H., Shieh, C.-K. and Lai, M.M.C. (1988) Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNAs. J. Virol. 62, 3870–3873.

Maxam, A.M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. In: (L. Grossman and K. Moldave, Eds.), Methods in Enzymology, Vol. 65, pp. 499–560. Academic Press, Orlando.

McIntosh, K. (1974) Coronaviruses: a comparative review. Curr. Top. Microbiol. Immunol. 63, 85–129.

McIntosh, K., Dees, J.H., Becker, W.B., Kapikian, A.Z. and Chanock, R.M. (1967) Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. Proc. Natl. Acad. Sci. USA 57, 933–940.

Pedersen, N.C., Ward, I. and Mengeling, W.L. (1978) Antigenic relationships of the feline infectious peritonitis virus to coronaviruses of other species. Arch. Virol. 58, 45–53.

Resta, S., Luby, J.P., Rosenfeld, C.R. and Siegel, J.D. (1985) Isolation and propagation of a human enteric coronavirus. Science 229, 978–981.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Schmidt, O.W. and Kenny, G.E. (1982) Polypeptides and functions of antigens from human coronaviruses 229E and OC43. Infect. Immun. 35, 515–522.

Shieh, C.-K., Soe, L.H., Makino, S., Chang, M.-F., Stohlman, S.A. and Lai, M.M.C. (1987) The 5'-end sequence of the murine coronavirus genome: Implications for multiple fusion sites in leader-primed transcription. Virology 156, 321–330.

Spaan, W., Delius, H., Skinner, M., Armstrong, J., Rottier, P., Smeekens, S., van der Zeijst, B.A.M. and Siddell, S.G. (1983) Coronavirus mRNA synthesis involves fusion of noncontiguous sequences. Eur. Mol. Biol. Organ. J. 2, 1839–1844.

Stern, D.F. and Kennedy, S.I.T. (1980) Coronavirus multiplication strategy. II. Mapping the avian infectious bronchitis virus intracellular RNA species to the genome. J. Virol. 36, 440–449.

Tompkins, W.A.F., Watrach, A.M., Schmale, J.D., Schulze, R.M. and Harris, J.A. (1974) Cultural and antigenic properties of newly established cell strains derived from adenocarcinomas of the human colon and rectum. J. Natl. Cancer Inst. 52, 1101–1106.

Weiss, S.R. (1983) Coronavirus SD and SK share extensive nucleotide homology with murine coronavirus MHV-A59, more than that shared between human and murine coronaviruses. Virology 126, 669–677.