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 194, 408–413 (1993)

Natural Evolution of Coronavirus Defective-Interfering RNA Involves RNA Recombination

TETSUYA FURUYA, THOMAS B. MACNAUGHTON,[1] NICOLA LA MONICA,[2] AND MICHAEL M. C. LAI[3]

Howard Hughes Medical Institute and Department of Microbiology, University of Southern California, School of Medicine, Los Angeles, California 90033-1054

Received November 20, 1992; accepted January 11, 1993

Defective-interfering (DI) RNAs of RNA viruses, in general, are generated and continue to evolve in size during serial undiluted passages of viruses. This evolution was thought to occur by independent generation of DI RNAs during virus passages and subsequent selection of new DI RNAs under new cellular conditions. Here we demonstrate that recombination between the old DI RNA and the helper viral RNA can be one of the mechanisms for natural DI RNA evolution. A mouse hepatitis virus (MHV) DI RNA, DIssE RNA, was transcribed *in vitro* and transfected into a mouse cell line infected with a different MHV strain (A59), which is distinguishable from the original natural helper MHV (JHM). During subsequent serial undiluted passages of the harvested virus, several novel DI RNA species were generated, while the original DIssE RNA disappeared by passage 11. cDNA cloning and sequence analysis of one of these novel DI RNAs, designated DI-2, revealed that it is composed of four discontinuous regions of the genomic sequence and is different from the structure of the original DIssE RNA. Sequence comparison among DI-2, DIssE, and helper MHV-A59 RNAs showed that DI-2 sequence is similar to DIssE in the first and second regions, but similar to the helper A59 virus in the third and last regions. Thus, this DI RNA was generated by RNA recombination between the original DIssE RNA and the helper viral RNA. These results indicate that recombination between DI RNA and helper virus RNA can be involved in the natural evolution of DI RNAs. © 1993 Academic Press, Inc.

Most RNA viruses generate defective-interfering (DI) RNAs upon serial passages at high multiplicity of infection. These DI RNAs usually consist of various noncontiguous regions of the wild-type viral RNA genome and are thought to be generated by RNA polymerase jumping during RNA transcription (*1, 2*). DI RNAs generally replicate more efficiently than the nondefective viral RNA and suppress the replication of the latter. However, upon continuous passage of virus and cells, the relative ratio of the DI RNA and nondefective viral RNA fluctuates, and new DI RNA species sometimes appear, replacing the original DI RNA. This type of evolution of DI RNA has been noted in vesicular stomatitis virus (*3, 4*) and mouse hepatitis virus (MHV) (*5*). The evolution of DI RNA species was thought to be the result of selection of a new DI RNA during serial passages, which altered the virus–cell equilibrium (*6*) and favored the new type of DI RNA. However, the mechanism of the generation of the new DI RNA was not clear. These new DI RNAs could be generated *de novo* from the helper viral RNA, independently of the existing DI RNA, or could be the result of recombination or rearrangement of sequences between the existing DI RNAs and helper virus RNA. These two possibilities could not be distinguished because the primary nucleotide sequences of the DI RNA and helper virus RNA are identical, making the origin of new DI RNAs impossible to determine. In this report, we addressed this issue by using an *in vitro* synthesized DI RNA of MHV to transfect into cells preinfected with a heterologous MHV, which is different from the original helper virus; thus, the sequences of transfected DI RNA and helper viral RNA were distinguishable. The evolution of DI RNA was followed during serial virus passages and was found to result from recombination between the transfected DI RNA and helper virus RNA. Thus, we conclude that RNA recombination plays a major role in the natural evolution of DI RNAs.

MHV, a member of Coronaviridae, is an enveloped virus containing a 31-kb, single-stranded, positive-sense genomic RNA (*7, 8*). MHV synthesizes seven to eight mRNAs that have a 3′-coterminal, nested-set structure (*9, 10*) and an identical 5′-end leader sequence of 72 to 77 nucleotides, which is derived from the 5′ end of the genome (*11–13*). The MHV mRNAs probably are synthesized by a discontinuous transcription mechanism involving polymerase jumping (*14*). Unlike other single-stranded RNA viruses, MHV under-

[1] Current address: Division of Medical Virology, Institute of Medical and Veterinary Science, Adelaide, South Australia.

[2] Current address: Instituto di Ricerche di Biologia Molecolare, Pomezia, Rome, Italy.

[3] To whom reprint requests should be addressed.

0042-6822/93 $5.00

goes RNA recombination at a very high frequency, which is almost as high as that of viruses with segmented RNA genomes (*15, 16*). This phenomenon suggests that discontinuous RNA synthesis and occasional copy-choice switching of RNA templates may also be involved in the replication of coronavirus RNA. Probably as a result of discontinuous RNA synthesis, DI RNA is readily generated from MHV during virus passages at high multiplicity of infection (*17*). Several different kinds of DI RNA have been identified, and the dominant DI RNA species varies with different passage levels (*5*). Some of these DI RNAs have been molecularly cloned and sequenced (*18–20*). All of them are composed of several noncontiguous genomic regions and contain the 5′ and 3′ ends of the parental MHV genome (*18–20*).

Two particular DI RNA species have been extensively studied; DIssE RNA was generated from JHM strain of MHV after 16 serial undiluted passages (*5*) in DBT cells, an astrocytoma cell line (*21*). This DI RNA is 2.2 kb long and is composed of three discontinuous regions of JHM genomic RNA (*18*). Another one, DIssF RNA, was generated also from JHM after five additional undiluted passages of DIssE-containing virus, when DIssE RNA became undetectable (*5*). This RNA is 3.6 kb long and contains five discontiguous regions, some of which, including the 5′- and 3′-ends and part of the second domain, are identical to DIssE (*19*). The sequence relationship between DIssE and DIssF RNAs, particularly the fact that they have the same junction site between the first and second domains in these two RNAs, suggests a possibility that DIssF may have been directly derived from DIssE rather than generated independently. However, this possibility could not be further examined since both RNA species were generated from the cells infected with the same helper virus. Nevertheless, these RNAs offer an experimental system to examine the mechanism of evolution of DI RNAs in MHV-infected cells.

DIssE RNA was transcribed from a pT7 vector, DE25 (*22*), by T7 RNA polymerase and transfected into DBT cells which had been infected with A59, an MHV strain different from the original helper virus JHM strain. Previously, it has been demonstrated that the *in vitro* transcribed DE25 RNA can replicate in the presence of a helper MHV and, despite the absence of a specific packaging signal, can be passaged with high efficiency (*22*). The viruses were harvested from the DE25-transfected and A59-infected cells and passaged serially in DBT cells without dilution of virus samples. A total of 11 passages was carried out. ^{32}P-labeled intracellular viral RNA patterns at selected passage levels were then examined after denaturation of RNA with glyoxal by agarose gel electrophoresis. Figure 1 shows that at passage 3, an RNA with a size equivalent to that of DIssE was the major RNA species, while the synthesis

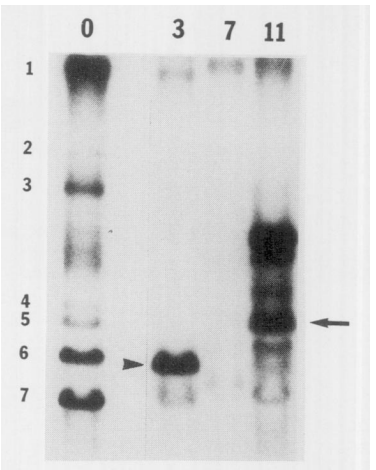

FIG. 1. Agarose gel electrophoresis of glyoxal-denatured intracellular RNAs of DIssE-transfected cells at different passage levels. Plasmid DE25 (*22*), which contains the T7 promoter sequence preceding the complete DIssE cDNA, was transcribed as described previously (*22*). The DIssE RNA was transfected into MHV A59-infected L2 cells by DEAE-dextran as previously described (*34*). The virus was harvested from the media and serial passages were carried out on DBT cells using undiluted virus. ^{32}P-labeled intracellular viral RNA was extracted as described previously (*35*). The arrowhead and arrow represent DIssE and DI-2 RNAs, respectively. The number over each lane denotes the passage level, and in the lane of passage level 0, only the helper virus, MHV A59, was propagated in the cells. The intracellular RNA species of the standard MHV A59 are denoted on the side of the gel.

of most of the MHV-A59-specific mRNAs was inhibited. At passage 7, very little DIssE RNA and helper virus mRNAs could be detected. Only a small amount of genomic RNA and a new RNA in between mRNAs 6 and 7 could be detected. Interestingly, at passage 11, large amounts of viral RNA synthesis resumed, and several new RNA species of different sizes were detected; three of them migrated between MHV-A59 mRNA 3 and 4, and another new RNA species (DI-2) was slightly smaller than MHV-A59 mRNA 5.

To study the origin of these new RNA species, we cloned several of them. We first focused on the most prominent RNA, DI-2 (arrow, Fig. 1), since it is the smallest and most prominent new RNA and has apparently different size from DIssE. The DI-2 RNA from the intracellular RNA at passage 11 was purified from the gel as previously described (*22*) and used for oligo(dT)-primed cDNA synthesis. The double-stranded DNAs were cloned into *Eco*RI-linker-ligated pBluescript according to the published procedures (*19*). cDNA clones were screened by colony hybridization using DIssE RNA as a probe (*19*). Several overlapping cDNA clones were isolated and their structure was analyzed by sequencing. Any sequence gaps and ambiguous regions were further examined by reverse transcription-polymerase chain reaction (RT-PCR) using specific primers to make cDNA clones covering the gaps. Finally, every sequence jumping site was further confirmed by RT-PCR

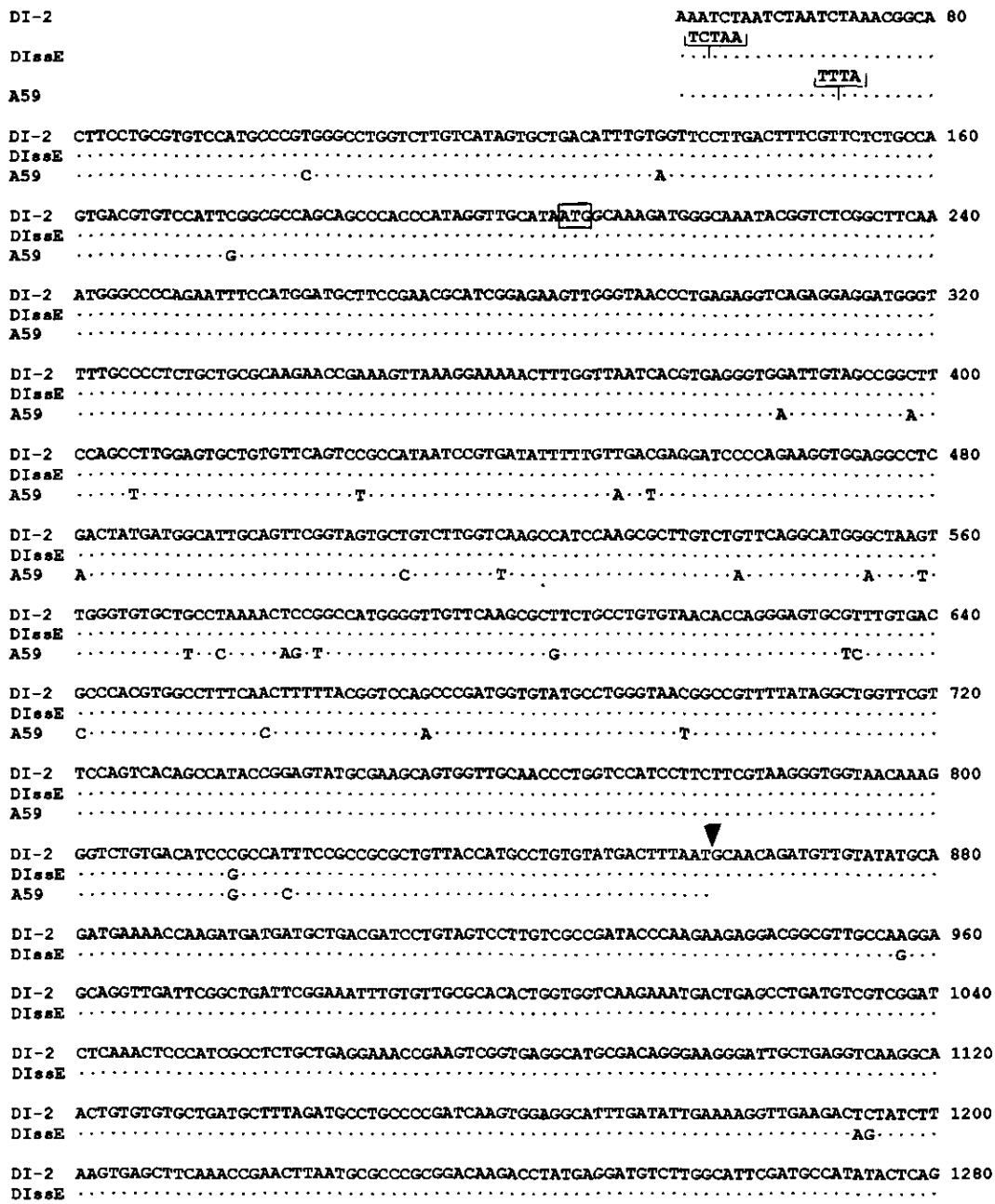

FIG. 2. Nucleotide sequence comparison of DI-2 with the corresponding regions of DIssE and MHV A59 genomic RNA. Sequences of DIssE, MHV A59 genes 1a, 1b, and gene 7 were derived from published data (*18, 24–26*). The open box indicates AUG initiation codon and the underline indicates termination codon of ORF of DI-2. Nucleotide insertions are identified with brackets underneath. Solid triangles indicate the sites where fusion of discontinuous sequences occurred. Numbers of nucleotides from the 5′-end are according to the DIssE sequence (*18*). cDNA cloning of DI-2 was carried out according to the procedure described previously (*19*) except for linkers and cloning vector. *Eco*RI linkers and *Eco*RI–digested plasmid pBluescript SK+ (Stratagene) were used for cDNA cloning. Recombinant DNAs were screened by colony hybridization (*36*) using a complete cDNA clone of DIssE as the probe (*22*). For RT-PCR, the entire RNA was made into two overlapping PCR products by using two pairs of primers separately for the 5′-half and 3′-half of the DI-2 RNA. cDNA cloning by PCR was carried out as described previously (*19*). Sequencing was carried out by Sanger's dideoxyribonucleotide chain termination method (*37*). Since one of the primers used for PCR cloning is complementary to the 5′-end of the leader of MHV-JHM RNA, the sequence of the very 5′ end of DI-2 could not be determined.

using primers across the junction sequences. Using these cDNA clones, the entire sequence of DI-2 except the very 5′ end was obtained (Fig. 2).

The sequence analysis of DI-2 RNA showed that this DI RNA is 2924 nucleotides long and composed of four regions derived from discontiguous regions of the MHV genome (Figs. 2 and 3). The entire third region and part of the fourth region of DI-2 are not present in DIssE. The first region (864 nucleotides) and the second region (445 nucleotides) represent a sequence corresponding to the same regions of DIssE RNA, and the junction between these two regions is identical to

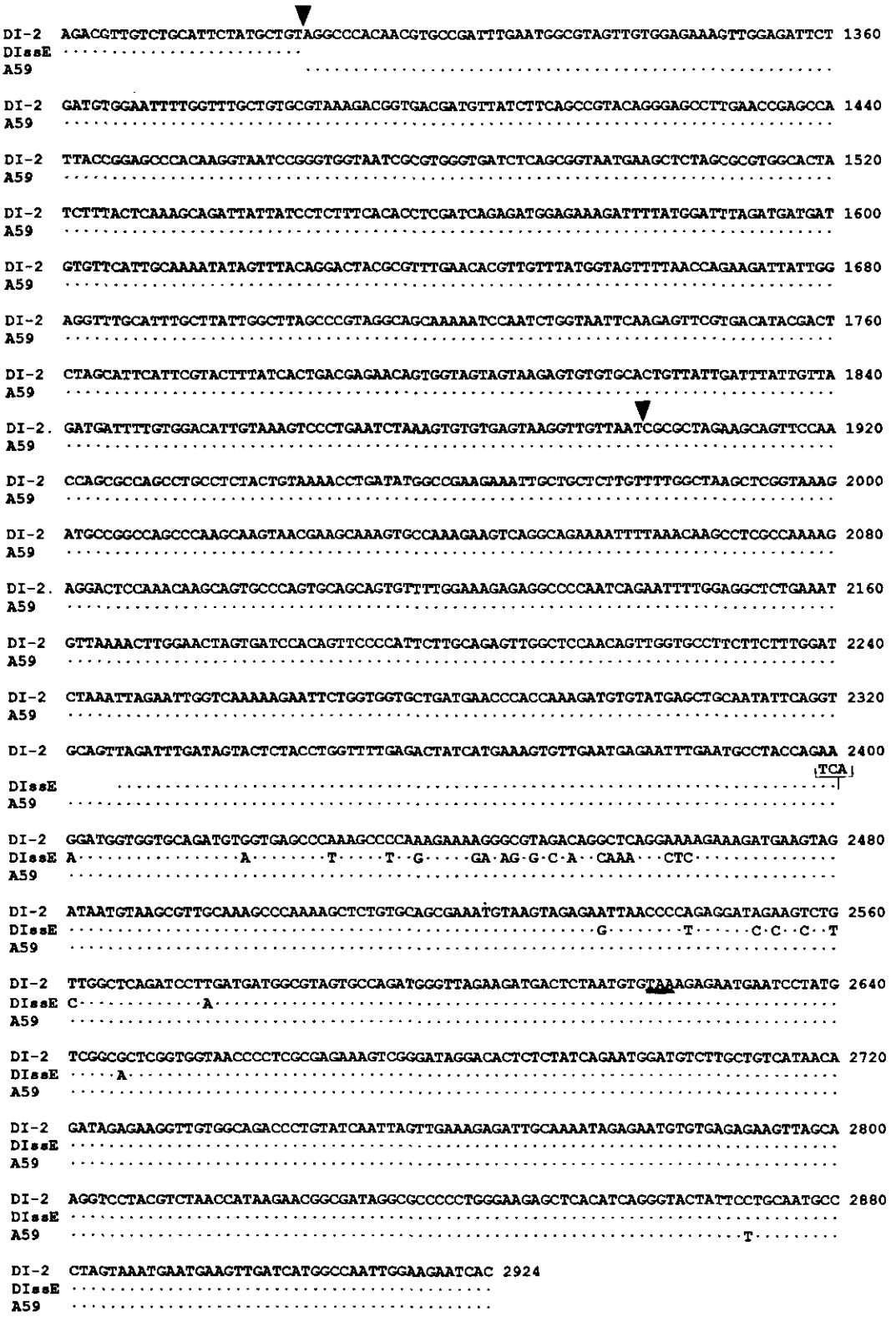

FIG. 2—*Continued*

that of DIssE. However, the second region of DI-2 is shorter than that of DIssE. The third region was derived from the sequence corresponding to 20.2 to 20.8 kb from the 5′ end of the genomic RNA. This region corresponds to the sequence containing the MHV RNA packaging signal as identified in DIssF RNA (*23*). The fourth region represents 1023 nucleotides from the 3′ end of the MHV genome.

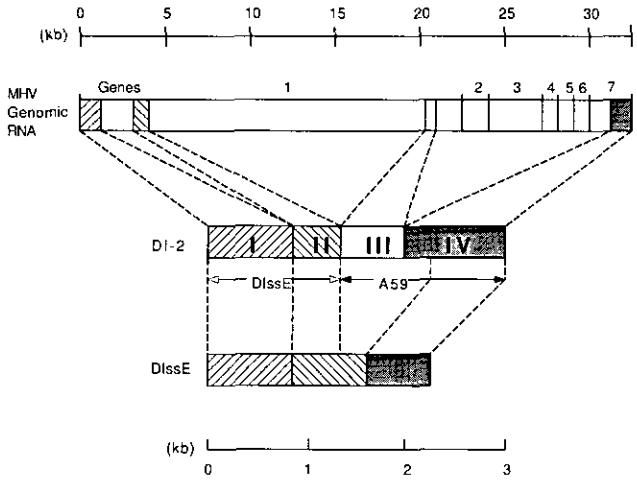

FIG. 3. Schematic diagram of the structure of DI-2 RNA. The structure of DI-2 is compared with the structure of the MHV A59 genomic RNA and DIssE RNA. Genes 1 through 7 represent the seven genes of MHV. Regions which represent DIssE sequence and MHV A59 genomic sequence are indicated by open arrows and solid arrows, respectively.

To examine the origin of DI-2 RNA, the DI-2 sequence was aligned with the previously published DIssE (*18*) and A59 (*24–26*) sequences (Fig. 2). Since the third region and the first 436 nucleotides of the last region of DI-2 are not present in DIssE, they were compared only to the A59 genomic RNA sequence. The results showed that the first region of DI-2 is almost identical to the corresponding sequence of DIssE, with an identical junction site. The corresponding A59 sequence has 29 nucleotide differences and four extra nucleotides. However, within the leader sequence, DI-2 RNA contains three UCUAA repeats, while the reported DIssE sequence contains four of them (*18*), and A59 genomic RNA contains two copies (*11, 24*). The heterogeneity of UCUAA copy numbers in the leader RNA has previously been reported for several DI RNAs (*27*), genomic RNAs (*28*), and mRNAs (*29*). This heterogeneity probably resulted from leader RNA jumping during DI RNA replication. Besides, one nucleotide in the first region at 815 nucleotides from the 5′ end is different from that of both DIssE and MHV-A59. This could be explained by mutations during DI RNA replication. The second region of DI-2 RNA is also identical to that of DIssE RNA and distinct from the corresponding A59 sequence, which differs by 32 nucleotides in this region (S. Weiss, personal communication). This region is smaller than the second region of DIssE RNA. These results suggest that the first and second regions of DI-2 RNA were directly derived from DIssE RNA. In contrast, the sequences of the third and fourth regions are identical to those of the MHV-A59 genome except for one nucleotide at 54 nucleotides from the 3′ end of the genome. The corresponding region of DIssE contains 28 nucleotide differences and 3 extra nucleotides. Since the third and fourth regions of DI-2 RNA contain an extra stretch of sequence which is not present in DIssE, and the sequence of the last region of DI-2 is identical to that of MHV-A59 genomic RNA and different from DIssE, it is likely that the third and fourth regions of DI-2 were derived from MHV-A59 genome RNA. Thus, DI-2 RNA was clearly derived by recombination between DIssE and helper A59 virus RNA with a cross-over site in the second domain of DIssE RNA, followed by additional polymerase jumping during copying of A59 RNA sequences. Since nucleotide 2870 of DI-2 is identical to that of DIssE, in contrast to A59, we cannot rule out the possibility that the 3′ end of DI-2 was derived from DIssE resulting from an additional recombination event.

This result indicates that RNA recombination is at least one of the mechanisms for the natural evolution of DI RNAs. The possible occurrence of recombination between DI RNA and helper viral RNA has also been demonstrated in a recent study, in which marker mutations in DI RNA can be incorporated into the viral genome by RNA recombination during virus replication (*30*). The DI-2 RNA is unique in that multiple recombination events between nonhomologous sequences of DI RNA and helper viral RNA (aberrant homologous recombination (*38*)) were involved. Thus, both nonhomologous and homologous recombination can occur between the DI and helper viral RNAs. At the present time, we do not know whether DI-2 RNA was directly derived from recombination between DIssE RNA and helper viral RNA by polymerase jumping during RNA synthesis, or indirectly from recombination between DIssE RNA and an unidentified A59-derived DI RNA species. If it is the former, several polymerase jumping events must have occurred during the synthesis of DI-2 RNA.

It is interesting to note that the third region of DI-2 RNA is not present in DIssE. However, DIssF RNA, which is packaged into virus particles efficiently, also contains 276 nucleotides corresponding to the 5′ end of the third region of DI-2 (*19*). This region contains the putative packaging signal for MHV RNA (*23*). Since DIssF and DI-2 were independently derived using two different helper viruses, the presence of similar stretches of sequence in both RNAs suggests that this region is a recombinational hot spot or that the presence of this sequence confers selective advantage. Similar sequences have also been found in another DI RNA derived from MHV-A59 (*20*).

Like DIssE, DI-2 has a single large open reading frame (ORF), which is 2415 bp long. The initiation and termination sites of the ORF are identical to those of ORF in DIssE, although the ORF of DI-2 contains different regions from those of DIssE. The same ORF is contained in both DI-2 and DIssE in the first and second regions. As reported previously, the first 215 amino

acids correspond to the N terminus of the MHV gene 1 product, and part of the protein is cleaved into a p28 protein (*31–33*). The significance of this ORF in DI RNA replication is not yet clear.

REFERENCES

1. LAZZARINI, R. A., KEENE, J. D., and SCHUBERT, M., *Cell* **26,** 145–154 (1981).
2. NAYAK, D. P., CHAMBERS, T. M., and AKKINA, R. K., *Curr. Top. Microbiol. Immunol.* **114,** 103–151 (1985).
3. DEPOLO, N. J., GIACHETTI, C., and HOLLAND, J. J., *J. Virol.* **61,** 454–464 (1987).
4. HORODYSKI, F. M., NICHOL, S. T., SPINDLER, K. R., and HOLLAND, J. J., *Cell* **33,** 801–810 (1983).
5. MAKINO, S., FUJIOKA, N., and FUJIWARA, K., *J. Virol.* **54,** 329–336 (1985).
6. AHMED, R., CANNING, W. M., KAUFFMAN, R. S., SHARPE, A. H., HALLUM, J. V., and FIELDS, B. N., *Cell* **25,** 325–332 (1981).
7. LAI, M. M. C., and STOHLMAN, S. A., *J. Virol.* **26,** 236–242 (1978).
8. LEE, H.-J., SHIEH, C.-K., GORBALENYA, A. E., KOONIN, E. V., LA MONICA, N., TULER, J., BAGDZHADZHYAN, A., and LAI, M. M. C., *Virology* **180,** 567–582 (1991).
9. LAI, M. M. C., BRAYTON, P. R., ARMEN, R. C., PATTON, C. D., PUGH, C., and STOHLMAN, S. A., *J. Virol.* **39,** 823–834 (1981).
10. LEIBOWITZ, J. L., WILHELMSEN, K. C., and BOND, C. W., *Virology* **114,** 39–51 (1981).
11. LAI, M. M. C., BARIC, R. S., BRAYTON, P. R., and STOHLMAN, S. A., *Proc. Natl. Acad. Sci. USA* **81,** 3626–3630 (1984).
12. LAI, M. M. C., PATTON, C. D., BARIC, R. S., and STOHLMAN, S. A., *J. Virol.* **46,** 1027–1033 (1983).
13. SPAAN, W., DELIUS, H., SKINNER, M., ARMSTRONG, J., ROTTIER, P., SMEEKENS, S., VAN DER ZEIJST, B. A. M., and SIDDELL, S. G., *EMBO J.* **2,** 1839–1844 (1983).
14. LAI, M. M. C., *Annu. Rev. Microbiol.* **44,** 303–333 (1990).
15. MAKINO, S., FLEMING, J. O., KECK, J. G., STOHLMAN, S. A., and LAI, M. M. C., *J. Virol.* **57,** 729–737 (1986).
16. BARIC, R. S., FU, K., SCHAAD, M. C., and STOHLMAN, S. A., *Virology* **177,** 646–656 (1990).
17. MAKINO, S., TAGUCHI, F., and FUJIWARA, K., *Virology* **133,** 9–17 (1984).
18. MAKINO, S., SHIEH, C. K., SOE, L. H., BAKER, S. C., and LAI, M. M. C., *Virology* **166,** 550–560 (1988).
19. MAKINO, S., YOKOMORI, K., and LAI, M. M. C., *J. Virol.* **64,** 6045–6053 (1990).
20. VANDER MOST, R. G., BREDENBEEK, P. J., and SPAAN, W. J. M., *J. Virol.* **65,** 3219–3226 (1991).
21. HIRANO, N., FUJIWARA, K., HINO, S., and MATSUMOTO, M., *Arch. Gesamte Virusforsch.* **44,** *298–302* (1974).
22. MAKINO, S., and LAI, M. M. C., *J. Virol.* **63,** 5285–5292 (1989).
23. FOSMIRE, J. A., HWANG, K., and MAKINO, S., *J. Virol.* **66,** 3522–3530 (1992).
24. PACHUK, C. J., BREDENBEEK, P. J., ZOLTICK, P. W., SPAAN, W. J. M., and WEISS, S. R., *Virology,* **171,** 141–148 (1989).
25. PARKER, M. M., and MASTERS, P. S., *Virology* **179,** 463–468 (1990).
26. BREDENBEEK, P. J., PACHUK, C. J., NOTEN, A. F. H., CHARITE, J., LUYTJES, W., WEISS, S. R., and SPAAN, W. J. M., *Nucleic Acids Res.* **18,** 1825–1832 (1990).
27. LAI, M. M. C., MAKINO, S., SOE, L. H., SHIEH, C.-K., KECK, J. G., and FLEMING, J. O., *Cold Spring Harbor Symp. Quant. Biol.* **52,** 359–365 (1987).
28. MAKINO, S., and LAI, M. M. C., *Virology* **169,** 227–232 (1989).
29. MAKINO, S., SOE, L. H., SHIEH, C. K., and LAI, M. M. C., *J. Virol.* **62,** 3870–3873 (1988).
30. VANDER MOST, R. G., HEIJNEN, L., SPAAN, W. J. M., and DE GROOT, R. J., *Nucleic Acids Res.* **20,** 3375–3381 (1992).
31. BAKER, S. C., SHIEH, C.-K., SOE, L. H., CHANG, M.-F., VANNIER, D. M., and LAI, M. M. C., *J. Virol.* **63,** 3693–3699 (1989).
32. DENISON, M. R., and PERLMAN, S., *J. Virol.* **60,** 12–18 (1986).
33. SOE, L. H., SHIEH, C.-K., BAKER, S. C., CHANG, M.-F., and LAI, M. M. C., *J. Virol.* **61,** 3968–3976 (1987).
34. MAKINO, S., SHIEH, C.-K., KECK, J. G., and LAI, M. M. C., *Virology* **163,** 104–111 (1988).
35. MAKINO, S., TAGUCHI, F., HIRANO, N., and FUJIWARA, K., *Virology* **139,** *138–151* (*1984*).
36. SHIEH, C.-K., SOE, L. H., MAKINO, S., CHANG, M.-F., STOHLMAN, S. A., and LAI, M. M. C., *Virology* **156,** 321–330 (1987).
37. SANGER, F., NICKLEN, S., and COULSON, A. R., *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467 (1977).
38. LAI, M. M. C., *Microbiol. Rev.* **56,** 61–79, (1992).