JOURNAL OF VIROLOGY, July 1983, p. 178–184
0022-538X/83/070178-07$02.00/0

Vol. 47, No. 1

Apparent Recombinants Between Virus-Like (VL30) and Murine Leukemia Virus-Related Sequences in Mouse DNA

AHUVA ITIN AND ELI KESHET*

Department of Virology, The Hebrew University-Hadassah Medical School, Jerusalem, Israel

Received 10 January 1983/Accepted 8 April 1983

VL30 elements are a dispersed multigene family that is ubiquitous in all murine cells. Despite not sharing nucleic acid sequence homology with natural retroviruses (exogenous or endogenous), VL30 elements are distinguished by several retrovirus-like features. By screening a mouse embryonic library, we have cloned DNA units that contain VL30 sequences linked to MuLV-related sequences. Using blot hybridization with the aid of specific subgenomic probes and heteroduplex analyses, we have established that the DNA element is composed of two VL30 long terminal repeat (LTR) units, a limited subset of VL30 information adjacent to both 5′ and 3′ LTRs, and an enclosure of MuLV-related information that shares homology primarily with MuLV *gag* and *pol* determinants (but lacks MuLV-related LTRs). This sequence arrangement is reciprocal in nature to the recombinations between MuLV and rat VL30 that generated the genomes of the Harvey and Kirsten strains of mouse sarcoma virus and most likely is the consequence of recombination between VL30 and MuLV-related elements and the subsequent deposition of the putative recombinant DNA in the mouse genome.

VL30 genes represent a dispersed multigene family that is ubiquitous in all murine cells. VL30 units are distinguished by several retrovirus-like features. Most notably, they exhibit (i) basic structure in common with retrovirus proviruses and in particular the presence of long terminal repeat (LTR) sequences at the boundaries of the 5.2-kilobase (kb) unit (11, 12); (ii) the efficient encapsidation of 30S RNA transcripts of VL30 genes in type C virions (1, 8, 16); and (iii) the subsequent capacity of VL30 information to participate in cell-to-cell transmission (14). The origin of VL30 information is unknown because no homology has been detected between VL30 and the nucleic acids of known retroviruses (exogenous or endogenous) (1, 8, 16), no proteins encoded by VL30 genes have been identified (15), and no infectious extracellular counterpart of VL30 elements has been encountered. Data exist supporting the thesis that VL30 elements are of a relatively early evolutionary origin and that they were established before speciation of the genus *Mus* (3).

Genetic elements similar to murine VL30 are also present in rat DNA (6). Rat VL30 sequences were found incorporated in the genomes of two independent isolates of recombinant transforming viruses, namely, the Harvey and the Kirsten strains of mouse sarcoma virus (Ha-MSV and Ki-MSV). The genomes of these viruses are composed from the genetic information of the infecting nontransforming murine leukemia virus (MuLV), VL30 sequences, and the oncogene (5, 20). The nature of the recombinations from which these apparent tripartite genomes are derived is not known. The selection applied during isolation of Ha-MSV and Ki-MSV was for transduction of the oncogene. Therefore, it is possible that the concomitant acquisition of VL30 sequences reflects a role for VL30 sequences in the capture of the oncogene.

We decided to examine the thesis that the aforementioned recombinations reflect a more general participation of VL30 sequences in recombinations with retrovirus-related genes as well as other cellular genes. Prompted by the fact that VL30 RNA is packagable in virions of type C retroviruses and by the proposal that packaging of heterodimeric RNA plays a key role in recombination between retroviruses (2, 9) we first searched for DNA units that contained VL30 sequences linked to sequences of murine proviruses. By screening mouse embryonic libraries, we have encountered DNA units with a sequence organization that appears to be the consequence of recombination between VL30 and MuLV-related proviruses. This paper describes the strategy for isolating these putative

recombinant proviruses and a detailed analysis of a representative DNA clone by blot hybridization and heteroduplex analyses.

MATERIALS AND METHODS

Recombinant DNA clones. Recombinant phage containing VL30-related sequences were cloned from a library of embryonic BALB/c DNA. The library was constructed through ligation of DNA fragments obtained by partial digestion of mouse DNA with *Eco*RI in a λ charon 4A vector, as previously described (12).

Mouse DNA fragments containing VL30-reactive sequences were transferred onto plasmids by codigestion with *Eco*RI (or other combinations of restriction enzymes) of equimolar amounts of pBR322 and the respective clone of DNA in λ phage followed by ligation, transformation of *Escherichia coli* HB101 cells, and identification of recombinant plasmids by in situ colony hybridization. Specific subgenomic VL30 clones were obtained as described in the figure legends.

pBR322 plasmid bearing *Sal*I-permuted MuLV 8.8-kb DNA was the gift of D. Linemeyer.

Mapping of restriction enzyme cleavage sites was carried out by the use of different combinations of restriction enzymes and subsequent size analysis of resultant fragments.

Blot hybridization. Blotting of DNA fragments onto nitrocellulose filter paper was carried out by the Southern procedure (18). Filter hybridization was performed as previously described (10). To serve as hybridization probes, DNA fragments were eluted from agarose gel by absorption onto glass beads (19) and were labeled by in vitro nick translation (13) to specific activities of 2×10^8 to 4×10^8 cpm per μg of DNA. Labeling of DNA fragments cloned in M13 mp8 was carried out by priming the synthesis of complementary DNA strands with specific oligomeric primer and by the use of the large fragment of DNA polymerase I. The specific activity thus obtained was 10^9 cpm per μg of DNA.

Heteroduplex analysis. Heteroduplexes were prepared and mounted for electron microscopy by the formamide procedure of Davis et al. (4). Samples were examined in a Philips 300 electron microscope. ϕX174 DNAs were included as internal size markers.

RESULTS

Strategy for isolating putative recombinants between VL30 and endogenous MuLV-related sequences. Because a functional selection for VL30-containing sequences was not available, we screened a mouse embryonic gene library for the presence of DNA fragments containing VL30 sequences physically linked to MuLV-related sequences. The rationale of this approach was that certain recombination events that occurred in the past between VL30 and MuLV-related endogenous proviruses could have led to the stable deposition in the mouse genome of recombinant genetic units in which VL30 and MuLV-specific sequences are linked.

Approximately a genome equivalent of recombinant phage from BALB/c embryo DNA in a λ charon 4A vector was plated, and plaque imprints were obtained on duplicate nitrocellulose filters. One filter was hybridized with VL30 LTR sequences (subcloned in pBR322), and the other filter was hybridized with a Moloney (Mo)-MuLV cDNA clone. As expected from the lack of hybridization between VL30 elements and MuLV, most positive hybridization signals were unique either to the VL30 probe or to the MuLV probe. Yet, a few colonies hybridized to both probes, suggesting a linkage between VL30 LTRs and at least some MuLV-related DNA sequences in these particular mouse DNA fragments. These clones were plaque purified and further analyzed for the sequence organization of both VL30 and MuLV-related information. The detailed analysis of a representative clone (designated VM-1) is described below.

Incomplete representation of both VL30 and MuLV-related sequences in VM-1 DNA. It was necessary to distinguish between the possibility that VM-1 DNA is composed of a VL30 unit that resides adjacent to an endogenous MuLV-related provirus and the possibility that this apparent linkage is the result of recombination between these two genetic units which might be accompanied by loss of some VL30 as well as MuLV-related sequences. The following experiment was designed to determine which segments of a VL30 genome and of a MuLV genome are retained by VM-1 DNA. DNA fragments generated by digestion with restriction endonucleases of a well-mapped VL30 unit and of a MuLV genome were electrophoretically resolved on agarose gels, blotted, and hybridized with VM-1 DNA that had been radiolabeled by in vitro nick translation (Fig. 1). Some VL30 sequences were missing from VM-1 DNA (Fig. 1A). VM-1 DNA hybridized with the VL30 LTR-containing fragments (fragments A and I), with the fragment adjacent to the 3′ LTR (fragment D), and, to a much lesser extent, with the fragment adjacent to the 5′ LTR (fragment H, detected only in longer exposures). On the other hand, the centrally located VL30 fragments (C, E, and J) failed to hybridize. Representation of MuLV-related sequences in VM-1 DNA was also incomplete (Fig. 1B). The strongest hybridization was obtained with *gag*-related sequences (and in particular with sequences encoding the amino terminal portion of the *gag* polyprotein). Hybridization was also detected with *pol*-related sequences, and little hybridization was detected with *env*-related sequences. On the other hand, MuLV fragments containing the LTRs failed to hybridize with VM-1 DNA. To be sure that MuLV LTR sequences were indeed absent from VM-1 DNA, we performed the reciprocal experiment. Specifically, VM-1 DNA fragments were blotted and hybridized with a MuLV LTR

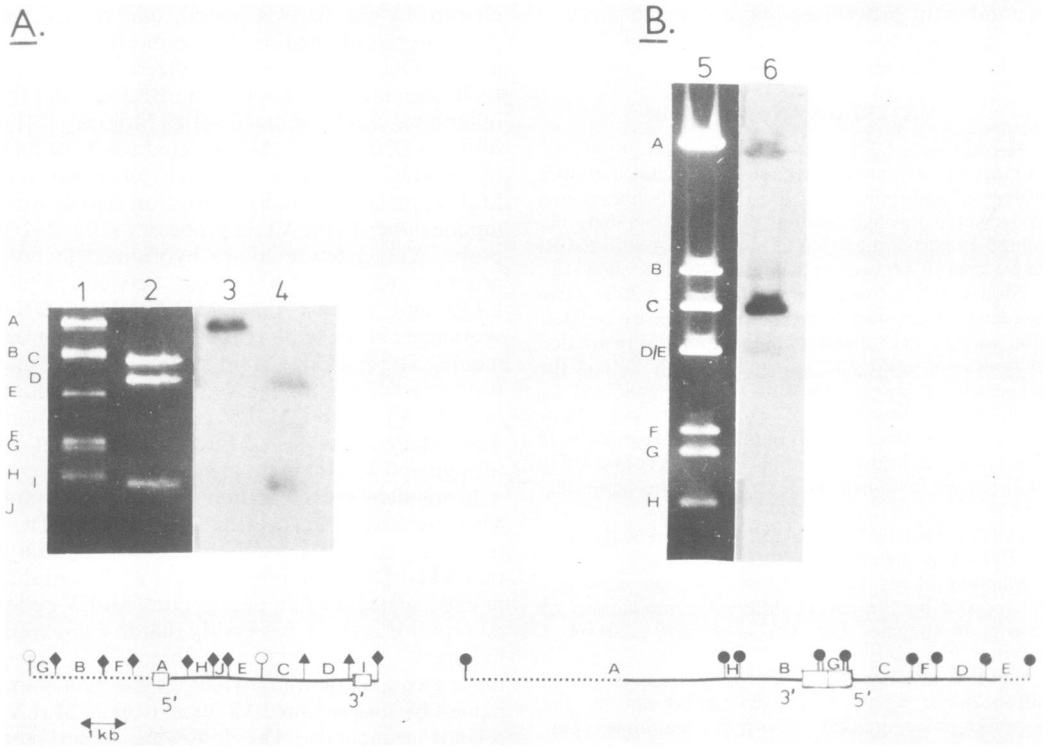

FIG. 1. Identification of VL30 and MuLV sequences present in clone VM-1. (A) A schematic representation of a clone containing a complete VL30 unit and at least some flanking DNA sequences at both sides is shown at the bottom. To distribute VL30 sequences as electrophoretically separable fragments, subclones were obtained that contained sequences residing to the left or to the right of the single centeral *Eco*RI site (○). Each subclone was digested with the restriction enzymes indicated, and resultant fragments were electrophoresed through 1.2% agarose gel. A photograph of the ethidium bromide-stained gel is shown in lanes 1 and 2. (For illustrative purposes, the larger vector DNA fragments were cut out.) DNA fragments were blot hybridized with clone VM-1 DNA that had been labeled by nick translation and autoradiographed. (B) A schematic representation of a Mo-MuLV cDNA cloned at the *Sal*I site of pBR322 is shown at the bottom. In this clone, MuLV sequences are circularly permuted. DNA was digested with *Ava*I, electrophoresed, and blot hybridized with VM-1 DNA probe as in (A). Symbols: ——, VL30 or MuLV unique DNA sequences; □, LTRs; ----, mouse flanking DNA sequences (A) or plasmid DNA sequences (B). Restriction sites were: (○) *Eco*RI; (◆) *Pvu*II; (▲) *Hin*dIII; (●) *Ava*I.

probe. For this purpose, a *Xba*I fragment of MuLV DNA that contained all of the LTR sequences (with a circular permutation) was radiolabeled to high specific activity. No hybridization was detected when microgram amounts of VM-1 DNA fragments were challenged with this probe (data not shown), indicating that VM-1 DNA does not contain MuLV-related LTRs.

Delineation of VL30 and MuLV-related sequences on VM-1 DNA. Preliminary experiments have indicated that all VL30 and MuLV-reactive sequences are confined within a 8-kb *Eco*RI fragment of VM-1 DNA. This fragment was subcloned in pBR322, and restriction enzyme cleavage sites within this DNA (designated pVM-1) were mapped.

To elucidate the sequence organization of both VL30 and MuLV-related information within VM-1, DNA blot hybridization experiments were performed with the aid of defined subgenomic VL30 and MuLV probes. A typical experiment is shown in Fig. 2. pVM-1 DNA was digested with *Eco*RI and *Pst*I, electrophoresed, blotted onto a nitrocellulose filter, and hybridized with each of the following probes: VL30 LTR, VL30 fragment D, and three MuLV fragments corresponding to different regions of *gag* and *pol* (as specified in the legend to Fig. 2). The following observations were made: (i) VL30 LTRs hybridized with two nonadjacent fragments, indicating that VM-1 DNA possesses two VL30 LTR units 4.5 to 7.0 kb apart; (ii) MuLV-related sequences are enclosed between the VL30 LTRs; (iii) MuLV *gag*-related sequences hybridized with fragment D of pVM-1, whereas MuLV *pol*-related sequences hybridized with

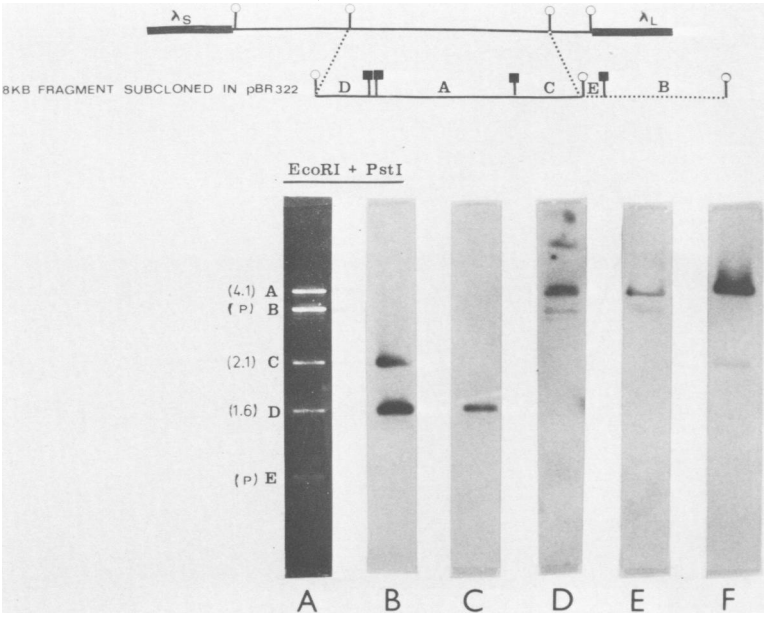

FIG. 2. Localization of VL30 and MuLV-related sequences within VM-1 DNA. To facilitate restriction enzyme mapping of clone VM-1, the middle 8-kb *Eco*RI fragment that accounts for all of the hybridization with both VL30 and MuLV DNA was transferred to pBR322 (broken line in top scheme). The plasmid clone was digested with *Eco*RI (○) and *Pst*I (■). Electrophoretically resolved fragments (lane A) were blotted and hybridized with the following probes. (B) VL30 LTR-specific probe. The DNA used was a VL30 subclone (in a phage vector) generated by deletion of all VL30 sequences residing between the LTRs during propagation in bacteria (2). (C) Fragment D of VL30 (Fig. 1A) subcloned in M13 mp8 vector. (D) A 1.6-kb *Pvu*II fragment of MuLV corresponding to MuLV map coordinates (numbers relate to the 5′ end of viral RNA according to Shinnick et al. [17]) 8,012 to 1,436 (part of the LTR, p15, p12, and part of p30). (E) A 0.75-kb *Hin*cII fragment of MuLV corresponding to map coordinates 2,949 to 3,705 (*pol*). (F) A 0.9-kb *Pvu*II fragment of MuLV corresponding to map coordinates 3,807 to 4,784 (*pol*). The indicated DNA fragments in lanes D to F were eluted from agarose gel by absorption onto glass beads and were labeled by nick translation. (The faint hybridization band B in lanes D and E is probably due to contaminating pBR322 sequences.)

fragment A of pVM-1, indicating organization of the MuLV-related sequences in the 5′ → 3′ orientation (see Fig. 4); (iv) a VL30-specific probe derived from VL30 sequences residing close to the 3′ LTR (fragment D) hybridized with both pVM-1 fragment C and fragment A, suggesting that VL30-reactive sequences in VM-1 DNA might be arranged in a noncontiguous fashion (see below).

Elucidation of VL30 sequence organization in VM-1 DNA by heteroduplexing. To determine more precisely the pattern in which VL30 sequences were organized in this apparent recombination element, we heteroduplexed VM-1 DNA with a standard VL30 unit. The heteroduplex structure (Fig. 3) confirmed the basic organization of two VL30-related segments (segments C and I) that are interrupted by large segments of non-VL30 information. Segments A and J are presumably mouse sequences that flank the recombinant element. Segments C and I, presumably the LTR-containing sequences, were of a size (estimated to be 1.2 and 0.55 kb, respectively) sufficient to accommodate a limited amount of VL30 information adjacent to both 5′ and 3′ LTRs.

A short segment of VL30-related sequences (segment F of about 300 base pairs) was embedded in non-VL30 information. This region hybridized exclusively with VL30 fragment D. This hybridization might reflect additional recombination sites between VL30 and a MuLV-related provirus. Alternatively, this cross-reactivity might be related to our previous findings that VL30 sequences contained in fragment D are dispersed in the mouse genome in additional molecular contexts besides VL30 elements (A. Itin et al., in press).

The data presented above are compatible with the putative recombinant DNA structure summarized schematically in Fig. 4. Noteworthy is the parallel 5′ → 3′ orientation of VL30 and MuLV-related sequences.

DISCUSSION

In the preceding sections, we have provided evidence that genetic units composed of VL30

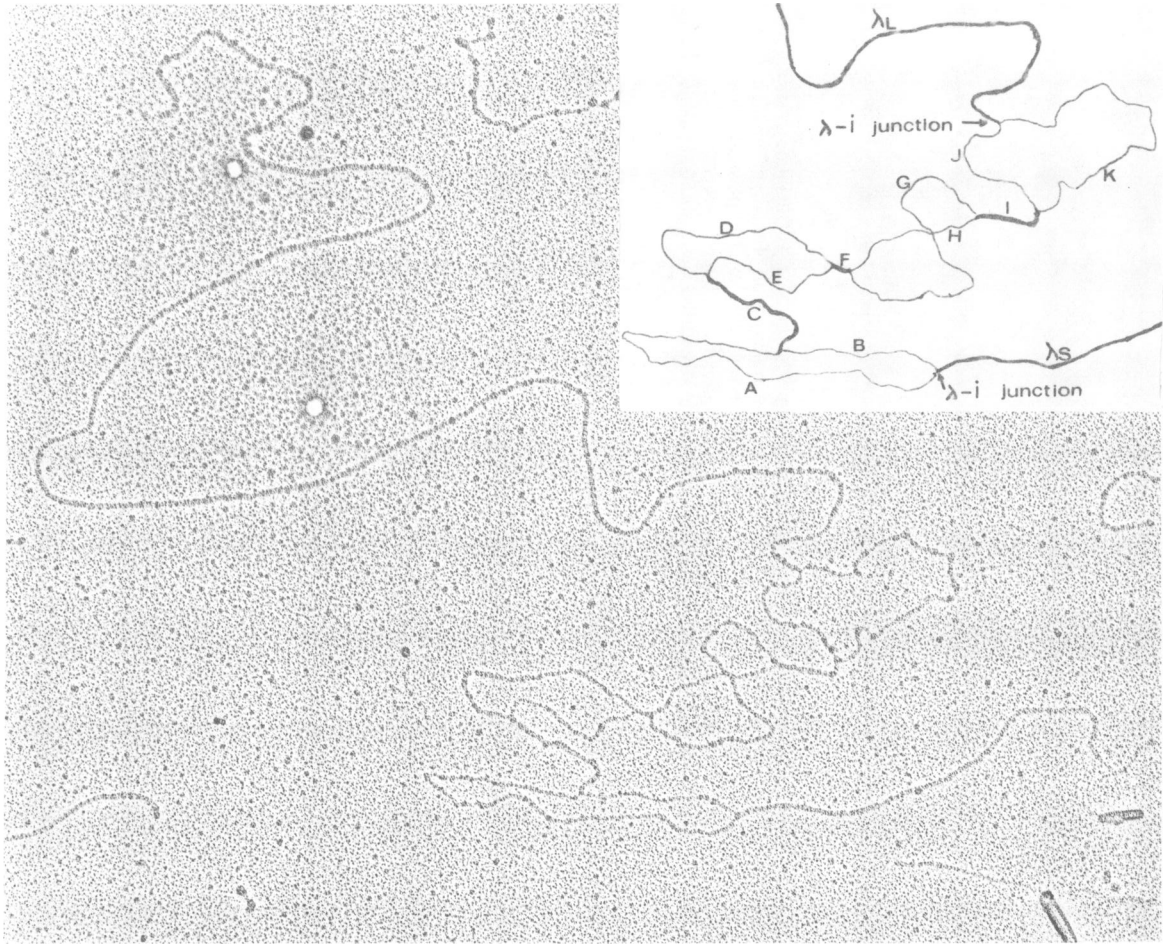

FIG. 3. Heteroduplex between clone VM-1 and a standard VL30 clone. VM-1 DNA (in a λ charon 4A vector) was heteroduplexed with a standard VL30 clone in the same vector under conditions specified in the text. The VL30 DNA used was composed of 5.3-kb VL30 sequences flanked by 1.0 kb of mouse DNA adjacent to the short λ arm and by 5.3-kb mouse DNA adjacent to the long λ arm (clone 2 in reference 10). λ_L and λ_S denote long and short lambda arms, respectively. λ-i junctions denote vector-insert junctions. Ten heteroduplex molecules were measured. ϕX174 DNAs and the lambda arms served as size standards. Lengths measured were (standard deviation in parentheses): A, 4.07 (±0.40); B, 1.10 (±0.23); C, 1.17 (±0.13); D, 2.39 (±0.18); E, 1.28 (±0.34); F, 0.30 (±0.12); G, 3.64 (±0.38); H, 1.63 (±0.35); I, 0.54 (±0.11); J, 1.49 (±0.23); K, 4.08 (±0.44).

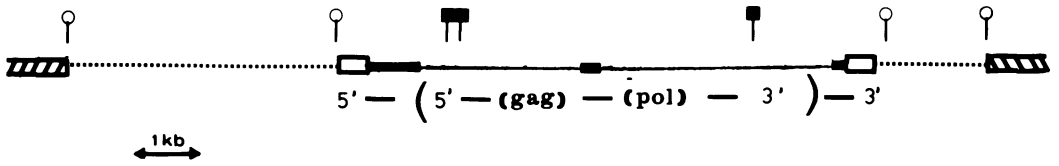

FIG. 4. Proposed sequence organization of VM-1 DNA: a schematic representation delineating VL30-specific and MuLV-related sequences on VM-1 DNA. Symbols: , lambda vector; ----, mouse flanking sequences; □, VL30 LTRs; heavy line, VL30-related sequences; thin line, non-VL30 sequences. *Eco*RI (○) and *Pst*I (■) sites were drawn for orientation. The nature and localization of MuLV-related sequences were not accurately determined. The scheme illustrates the maximal boundaries of MuLV-related information, points out that at least portions of *gag* and *pol* genes are represented in VM-1 DNA, and denotes the 5′ → 3′ arrangement of MuLV-related sequences.

sequences linked to MuLV-related sequences are present in normal mouse DNA. The DNA element described above contains two VL30 LTR units, a limited additional subset of VL30 information adjacent to both 5′ and 3′ LTRs, and an enclosure of MuLV-related information that shares homology primarily with Mo-MuLV *pol* and *gag* determinants (but lacks MuLV LTRs) (Fig. 4). The deposition of such a structure in mouse DNA could be explained as the consequence of (i) recombination between the genetic information of a VL30 unit and of a MuLV-related endogenous provirus and (ii) subsequent integration of the recombinant unit in mouse DNA in a manner that ensured its stable inheritance.

The apparent recombinant structure represented by VM-1 DNA is reciprocal in nature to the putative recombination event between MuLV and rat VL30 that occurred during the generation of the Harvey and Kirsten strains of mouse sarcoma virus. In both Ha-MSV and Ki-MSV, the recombinant virus retained MuLV LTRs together with a limited subset of MuLV sequences adjacent to both LTRs, whereas the bulk of MuLV sequences was lost concomitantly with the acquisition of rat VL30 information. (No data are available concerning the nature of VL30 sequences retained.) In addition, the restoration of highly transmissible Ha-MSV from 3′-deleted Ha-MSV (established through transfection) also occurred via recombination between MuLV and rat VL30 sequences of Ha-MSV in a fashion that restored the missing LTR units (7). The data presented above thus indicate a general mechanism of VL30-mediated recombination in which non-VL30 sequences are captured in a VL30 unit. In principle, capture of non-VL30 sequences may not be limited to capture of retrovirus-related genes; it may also include incorporation of other cellular sequences into VL30 units. In a similar fashion, the tripartite structure of Ha-MSV and Ki-MSV may have evolved via incorporation of rat cellular oncogene sequences (encoding p21) into a VL30 unit and subsequent recombination with MuLV. In this respect, the abundance of VL30 units (over 100 copies) in *M. musculus* may be a factor in the frequency at which cellular genes (e.g., oncogenes) are captured onto genomes of retroviruses.

Packaging signals of VL30 have not yet been characterized. It is possible that a genetic unit like VM-1 DNA retained VL30 packaging signals (particularly if, by analogy to retroviruses, packaging signals are located close to the 5′ LTR). In this case, incorporation of nonpackagable sequences onto VL30 might render them potentially transmissible. We are currently carrying out experiments aimed at determining whether VM-1-like DNA is expressed and whether these RNA transcripts are indeed transmissible.

The level at which recombination between VL30 and MuLV occurred is not known. The parallel orientation of both VL30 and MuLV-related information in VM-1 DNA is consistent with a mechanism of copackaging of a VL30 and a MuLV-related RNA transcript to form heterodimeric RNA within a virus particle and subsequent reverse transcription of portions of both VL30 and MuLV in a manner similar to models proposed to explain recombinations between retroviruses (2, 9). The type of LTR units retained by the recombinant molecule would depend upon the pattern of copy choice during reverse transcription. It should be emphasized that DNA structures such as VM-1 DNA are constituents of the normal mouse genome. Recombinations between VL30 and endogenous proviruses do not necessarily lead to the establishment of the recombinant DNA as an inheritable entity, for which integration in germ cells is essential. Therefore, the frequency of these putative recombination events could be underestimated if calculated from the frequency at which structures like VM-1 DNA are encountered in mouse DNA. The selection applied in this study (selection of VL30 LTR-containing DNA) has yielded additional DNA clones with VL30-MuLV sequence linkage that are distinct from VM-1 DNA (unpublished data). These DNAs

are currently under investigation to elucidate their sequence arrangement and to investigate the existence of preferred recombination sites.

ACKNOWLEDGMENTS

We thank N. Goldblum for his support, S. Kitov for technical assistance, G. Rotman and A. Asher for helpful suggestions, and D. Linemeyer for Mo-MuLV DNA.

This work was supported by a grant from the USA-Israel Binational Science Foundation.

LITERATURE CITED

1. **Besmer, P., U. Olshevsky, D. Baltimore, D. Dolberg, and H. Fan.** 1979. Virus-like 30S RNA in mouse cells. J. Virol. **29**:1168–1176.
2. **Coffin, J. M.** Structure, replication and recombination of retrovirus genomes: some unifying hypotheses. J. Gen. Virol. **42**:1–26.
3. **Courtney, M. G., P. K. Elder, D. Steffen, and M. J. Getz.** 1982. Evidence for an early evolutionary origin and locus polymorphism of mouse VL30 DNA sequences. J. Virol. **43**:511–518.
4. **Davis, R., M. Simon, and N. Davidson.** 1971. Electron microscopic heteroduplex methods in mapping regions of base sequence homology. Methods Enzymol. **21**:413–428.
5. **Ellis, R. W., D. DeFeo, J. M. Maryak, H. A. Young, T. Y. Shih, E. H. Chang, D. R. Lowy, and E. M. Scolnick.** 1980. Dual evolutionary origin for the rat genetic sequences of Harvey murine sarcoma virus. J. Virol. **36**:408–420.
6. **Goldberg, R. J., R. Levin, W. P. Parks, and E. M. Scolnick.** 1976. Quantitative analysis of the rescue of RNA sequences by mammalian type C viruses. J. Virol. **17**:43–50.
7. **Goldfarb, M. P., and R. A. Weinberg.** 1981. Generation of novel, biologically active Harvey sarcoma virus via apparent illegitimate recombination. J. Virol. **38**:136–150.
8. **Howk, R. S., D. H. Troxler, D. R. Lowy, P. H. Duesberg, and E. M. Scolnick.** 1978. Identification of a 30S RNA with properties of a defective type C virus in murine cells. J. Virol. **25**:115–123.
9. **Hunter, E.** 1978. The mechanism for genetic recombination in the avian retroviruses. Curr. Top. Microbiol. Immunol. **79**:295–309.
10. **Keshet, E., and A. Itin.** 1982. Patterns of genomic distribution and sequence heterogeneity of a murine "retrovirus-like" multigene family. J. Virol. **43**:50–58.
11. **Keshet, E., and Y. Shaul.** 1981. Terminal direct repeats in a retrovirus-like repeated mouse gene family. Nature (London) **289**:83–85.
12. **Keshet, E., Y. Shaul, J. Kaminchik, and H. Aviv.** 1980. Heterogeneity of virus-like genes encoding retrovirus-associated 30S RNA and their organization within the mouse genome. Cell **20**:431–439.
13. **Rigby, P. W., J. M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
14. **Scolnick, E. M., W. C. Vass, R. S. Howk, and P. H. Duesberg.** 1979. Defective retrovirus-like 30S RNA species of rat and mouse cells are infectious if packaged by type C helper virus. J. Virol. **29**:964–972.
15. **Scolnick, E. M., D. Williams, J. Maryak, W. Vass, R. J. Goldberg, and W. P. Parks.** 1976. Type C particle-positive and type C particle-negative rat cell lines: characterization of the coding capacity of endogenous sarcoma virus-specific RNA. J. Virol. **30**:570–582.
16. **Sherwin, S. A., U. R. Rapp, R. E. Benveniste, A. Sen, and G. J. Tadaro.** 1978. Rescue of endogenous 30S retroviral sequences from mouse cells by baboon type C virus. J. Virol. **26**:257–264.
17. **Shinnick, T. M., R. A. Lerner, and C. J. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) **293**:543–548.
18. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
19. **Vogelstein, B., and D. Gillespie.** 1979. Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. U.S.A. **76**:615–619.
20. **Wei, C. M., D. R. Lowy, and E. M. Scolnick.** 1980. Mapping of transforming region of the Harvey murine sarcoma virus genome by using insertion-deletion mutants constructed in vitro. Proc. Natl. Acad. Sci. U.S.A. **77**:4674–4678.