JOURNAL OF VIROLOGY, Oct. 1982, p. 19–31
0022-538X/82/100019-13$02.00/0

Vol. 44, No. 1

Analysis of the *env* Gene of a Molecularly Cloned and Biologically Active Moloney Mink Cell Focus-Forming Proviral DNA

ROBERT A. BOSSELMAN, FLIP VAN STRAATEN. CHARLES VAN BEVEREN, INDER M. VERMA,* AND MARGUERITE VOGT

Tumor Virology Laboratory, The Salk Institute, San Diego, California 92138

Received 22 March 1982/Accepted 7 June 1982

A biologically active molecular clone of BALB/Moloney mink cell focus-forming (Mo-MCF) proviral DNA has been reconstructed in vitro. It contains the 5′ half of BALB/Moloney murine leukemia virus (Mo-MuLV) DNA and the 3′ half of BALB/Mo-MCF DNA. The complete nucleotide sequence of the *env* gene and the 3′ long terminal repeat (LTR) of the cloned Mo-MCF DNA has been determined and compared with the sequence of the corresponding region of parental Mo-MuLV DNA. The substitution in the Mo-MCF DNA encompasses 1,159 base pairs, beginning in the carboxyl terminus of the *pol* gene and extending to the middle of the *env* gene. The Mo-MCF *env* gene product is predicted to be 29 amino acids shorter than the parental Mo-MuLV *env* gene product. The portion of the *env* gene encoding the p15E peptide is identical in both viral DNAs. There is an additional *A* residue in the Mo-MCF viral DNA in a region just preceding the 3′ LTR. The nucleotide sequence of the 3′ LTR of Mo-MCF DNA is similar to that of the 5′ LTR of BALB/Mo-MuLV DNA with the exception of two single base substitutions. We conclude that the sequence substitution in the *env* gene is responsible for the dual-tropic properties of Mo-MCF viruses.

Murine leukemia viruses (MuLVs) induce neoplasia in experimental animals (20). Moloney MuLV (Mo-MuLV) is a replication-competent, nondefective type C retrovirus which, upon inoculation into newborn mice, gives rise to leukemias of mostly T-cell origin. Leukemogenesis is accompanied by somatic amplification and reintegration of Mo-MuLV DNA sequences in new chromosomal sites of tumor tissues (25, 49, 50). The preleukemic tissue appears to contain only authentic Mo-MuLV genomes, reintegrated randomly in new chromosomal sites, whereas the leukemic tissue displays both the parental Mo-MuLV and the recombinant mink cell focus-forming (MCF) viral DNA sequences (25, 49). The precise involvement of MCF viruses in the conversion from the preleukemic to the leukemic stage remains largely unknown.

The dual-tropic MCF viruses first described by Hartley et al. (22) have also been implicated in the genesis of spontaneous lymphomas and leukemias. The MCF MuLVs appear to have arisen by recombination between ecotropic and xenotrope-like envelope genes. Biochemical evidence based on tryptic peptide analysis of the major envelope glycoprotein gp70 (12), RNase T_1 oligonucleotide mapping (36, 39), and heteroduplex (4, 7, 9) and restriction endonuclease analyses (5, 6) support the notion that MCF viruses are ecotropic MuLVs which have acquired some sequences of their envelope gene from xenotrope-like *env* gene sequences . We have been studying the properties of MCF viruses isolated from thymomas occurring in BALB/Mo mice (54). This mouse strain is derived from a BALB/c129J preimplantation embryo infected with Mo-MuLV (24). In this BALB/Mo strain, Mo-MuLV is transmitted as a single Mendelian gene. Like the other MCF-MuLVs, BALB/Mo-MCF virus also exhibits a dual host range and contains a recombinant *env* gene. We have molecularly cloned a portion of the integrated form of BALB/Mo-MCF viral DNA in an attempt to understand the precise nature of the recombinant gene. This manuscript describes the biological and biochemical properties of the molecularly cloned proviral DNA. Furthermore, the complete nucleotide sequence of the recombinant *env* gene has been determined and compared with that of the parental Mo-MuLV *env* gene.

MATERIALS AND METHODS

Cloning. Mink lung fibroblasts (CCL64) were infected with a cloned isolate of Mo-MCF virus (Mo-MCF_{81} [54]). After three transfers at 5- to 7-day intervals, the

infected cells were cloned in Microtest II culture plates. One of the cell clones, MCF 147D11, was used as a source of high-molecular-weight DNA. Confluent cultures of MCF 1497D11 cells show a strong vacuolization and yield titers of approximately 10^5 cytopathic focus-forming units per ml of culture supernatant. About 1.0 mg of MCF 147D11 DNA was digested overnight with 500 U of restriction endonuclease *Hin*dIII (Bethesda Research Laboratories) and fractionated on a 0.7% agarose gel as described previously (2). A small portion of the agarose gel was cut across the entire length, and DNA was transferred onto a nitrocellulose filter by the Southern blotting technique (41). A number of bands ranging from 4 to 10 kilobase pairs (kbp) in size could be identified by hybridization to pMLV-1 DNA. A 5- to 6-kbp fraction was eluted from the agarose gel by the glass powder procedure (53). About 1 μg of purified 5- to 6-kbp fragment was ligated to 4 μg of *Hin*dIII-cleaved, bacterial alkaline phosphatase-treated Charon 27 phage DNA and packaged in vitro as described previously (52). The phage plaques were screened by using a labeled 3′-Mo-MuLV-specific probe (pMLV-41) and amplified as described by Blattner et al. (3). The subcloning in plasmid pBR322 was performed as described previously (52).

Construction of 3′-specific probe (pMLV-41). The recombinant plasmid pMLV-1, containing the unintegrated form of Mo-MuLV viral DNA (2), was cleaved with restriction endonuclease *Cla*I and, after ligation with phage T4 DNA ligase, was transferred to *Escherichia coli* C600 by standard transformation procedures. Since Mo-MuLV DNA in pMLV-1 is circularly permuted at the *Hin*dIII site (at 5.4 kbp from the 5′ end of the genome), digestion with *Cla*I (pBR322 has a *Cla*I site just adjacent to the *Hin*dIII site) followed by ligation generates circular molecules which are deleted of sequences lying 5′ to the Mo-MuLV *Hin*dIII site and 3′ to the Mo-MuLV *Cla*I site (8.05 kbp). Bacterial colonies sensitive to tetracycline and resistant to ampicillin were screened by filter hybridization with labeled *Hin*dIII (5.4 kbp) to *Cla*I (8.05 kbp)-cleaved pMLV-1 DNA fragment. This fragment contains a portion of the *pol* gene and a nearly complete *env* gene. A recombinant clone which hybridized to this probe but not to a probe containing the remainder of the viral DNA (i.e., *gag* gene, portion of the *pol* gene, and long terminal repeat [LTR]) was isolated, characterized, and used as the 3′-specific probe, pMLV-41.

Heteroduplex mapping. Cloned DNA fragments (1 μg/ml each) were incubated at 45°C for 1 to 3 h in a solution containing 0.01 M PIPES buffer [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)], pH 6.3, 0.4 M NaCl, 0.001 M EDTA, and 50% formamide (deionized with Amberlite MB-3). Heteroduplexes were spread with cytochrome *c*, adsorbed to Parlodion-coated grids, stained with uranyl acetate, shadowed with platinum-palladium, and examined in a Hitachi HU-11B transmission electron microscope. Selected molecules were photographed at a magnification of ×10,000, and contour lengths were determined with a Hewlett-Packard digitizer (4). Molecular lengths were determined by comparison with double-stranded polyoma DNA and single-stranded φX-174 DNA spread under identical conditions. The length ratio of double-stranded to single-stranded DNA was 1.00:1.11.

Transfection. Transfection of DNA on mink lung fibroblasts (CCL64) was carried out by Stow and Wilkie's modification (43) of the calcium phosphate coprecipitation technique of Graham and van der Eb (19). NIH/3T3 DNA was used as carrier at a concentration of 30 μg/ml. The transfections were done in 35-mm dishes seeded the previous day with 3×10^5 mink lung fibroblasts. Four hours after the transfection, the cells were treated for 10 min with 20% dimethyl sulfoxide. The following day, the cells from each dish were transferred to three 50-mm dishes. Two or all three cultures derived from each transfected dish were transferred twice at 5- to 7-day intervals, at which time the monolayers had reached confluence. Cultures from the third transfer were screened for the presence of cytopathic foci 7 to 9 days after seeding. Reverse transcriptase activity was determined for 24-h-old harvests as previously described (51).

Nucleotide sequence analysis. All DNA sequence analyses were performed by the partial chemical degradation procedure (29). An outline of the techniques used has been described (48).

RESULTS

Molecular cloning. To clone the BALB/Mo-MCF proviral DNA, we chose a mink lung fibroblast cell line (MCF 147D11) infected with a cloned viral isolate (Mo-MCF_{81}). This cell line carries several copies of BALB/Mo-MCF DNA. A preliminary physical map of the proviral DNA, however, indicated the absence of a restriction endonuclease site which could be conveniently manipulated for molecular cloning. Hence, we decided to clone the 3′ half of BALB/Mo-MCF proviral DNA and subsequently link it to the 5′ half of BALB/Mo-MuLV DNA to generate an infectious full-length viral DNA. The Mo-MCF proviral DNA contains a single *Hin*dIII site which maps within the *pol* gene at about 5.4 kbp from the 5′ end of the genome (49). This *Hin*dIII site lies outside the region of *env* gene substitution in the Mo-MCF proviral DNA. We used this cleavage site to generate 5′ and 3′ subgenomic fragments of Mo-MCF viral DNA integrated in the mink chromosomal DNA. Full-length proviral DNA was generated by ligating the molecularly cloned 3′ subgenomic fragment to the previously cloned 5′ half of BALB/Mo-MuLV DNA ($pMLV_I$-101), containing viral sequences to the left of the *Hin*dIII site (2).

A *Hin*dIII-cleaved cell DNA fraction of about 6 kbp in length, which hybridizes to a Mo-MuLV probe, was used for molecular cloning in the phage vector Charon 27. Three clones of recombinant phage which hybridized to a 3′-specific probe (pMLV-41) were purified by successive rounds of dilution. The DNA from one recombinant clone, λMo-MCF_I-16, was cleaved with restriction endonucleases *Hin*dIII, *Sac*I, *Kpn*I, *Eco*RI, *Sma*I, and *Bam*HI (Fig. 1B) and compared with the restriction endonuclease map of BALB/Mo-MCF proviral DNA (49). All of these restriction enzyme sites mapped at positions

similar to those described for BALB/Mo-MCF DNA. However, restriction endonuclease *Kpn*I, which was reported to cleave the recombinant *env* region of some isolates of BALB/Mo-MCF DNA (49), did not cleave in the substituted sequences of λMo-MCF$_I$-16 DNA (see Discussion). A composite restriction endonuclease map of the insert in λMo-MCF$_I$-16 is shown in Fig. 1C.

Heteroduplex analysis. The nature of the recombinant λMo-MCF$_I$-16 DNA was further characterized by electron microscopic analyses of duplexes formed with molecularly cloned unintegrated forms of Mo-MuLV DNA (pMLV-1 [2]). The insert DNA in λMo-MCF$_I$-16 was subcloned in plasmid pBR322 (pMo-MCF$_I$-16) and hybridized to pMLV-1 DNA after digestion with *Hin*dIII (Fig. 1D). An example of this duplex is shown in Fig. 1E and F. The double-stranded end of the molecule corresponds to the *Hin*dIII site at 5.4 kbp, whereas the two single-stranded tails located at the opposite end of the molecule correspond to the flanking mink cellular sequences present in pMo-MCF$_I$-16 and the remainder of the *pol* and *gag* gene sequences present in pMLV-1. A histogram displaying the distribution of the substitution loops with respect to the *Hin*dIII site at 5.4 kbp from the 5′ end of the genome is also shown in Fig. 1G. This analysis reveals two major regions and one minor region of nonhomology lying within the area of 6.3 to 8.1 kbp from the 5′ end of the genome. Each of the nonhomologies (β_1, β_2, and β_3) is bounded by regions of homology with the parental Mo-MuLV sequences. Furthermore, within the substituted region overall homology between Mo-MCF and Mo-MuLV sequences increases with distance from the 5′ end of the substitution. Differences in the degree of homologies within the substituted region of AKR MCF have also been reported (6).

In vitro reconstruction of an infectious full-length Mo-MCF viral DNA. Restriction endonuclease and heteroduplex analyses confirmed that the insert of λMo-MCF$_I$-16 DNA subcloned in pBR322 (pMo-MCF$_I$-16) contained the recombinant *env* gene sequences. The 5′ portion of the BALB/Mo-MuLV DNA obtained from the MOV-1 locus of the BALB/Mo mouse has previously been cloned (2). Plasmid pMLV$_I$-101 contained sequences to the left of the *Hin*dIII site at 5.4 kbp, including the 5′ LTR and adjacent mouse cellular sequences. The 6.3-kbp insert from λMo-MCF$_I$-16 and the 9.3-kbp insert from pMLV$_I$-101 were ligated and subsequently subcloned in the *Hin*dIII site of plasmid pBR322. The DNA of one plasmid, pMo-MCF$_I$-1, which released the diagnostic 2.1-kbp *Bam*HI fragment characteristic of 5′ and 3′ subgenomic fragments ligated in proper orientation, was further analyzed by digestion with restriction endonucleases *Sac*I, *Sma*I, *Hin*dIII, *Eco*RI, and *Bam*HI (Fig. 2). Each enzyme yielded the characteristic pattern of cleavage products predicted from previous restriction analysis of 5′ and 3′ subgenomic fragments and pBR322 vector DNA. The DNA fragments generated were present in molar ratios and added to a total length of about 19.9 kbp, the predicted total length of the 5′ (9.3 kbp) and 3′ (6.3 kbp) inserts plus the pBR322 DNA (4.36 kbp). The organization and partial restriction map of full-length pMo-MCF$_I$-1 DNA are shown in Fig. 2B.

Biological infectivity. Three types of recombinant DNA were found to induce cytopathic foci in mink lung cell fibroblasts (Table 1): (i) a ligation mixture containing *Hin*dIII-cleaved inserts from pMLV$_I$-101 (5′ half of BALB/Mo-MuLV) and λMo-MCF$_I$-16 DNA (experiment A); (ii) a ligation mixture containing *Hin*dIII-cleaved inserts from pMLV$_I$-101 and alkaline phosphatase-treated pMo-MCF$_I$-16 DNA (experiment B); and (iii) uncut pMo-MCF$_I$-1 DNA (recombinant plasmid containing full-length reconstructed Mo-MCF DNA; experiment C). The λMo-MCF$_I$-16 and pMo-MCF$_I$-16 by themselves are not infectious (experiment D). It has been shown previously that pMLV$_I$-101 DNA is also noninfectious (2). The rationale for doing experiments A and B separately was to ensure that the insert from λMo-MCF$_I$-16 subcloned in pBR322 (pMo-MCF$_I$-16) was infectious. This was particularly important since all of the nucleotide-sequencing studies were performed on pMo-MCF$_I$-16 DNA.

Most cytopathic foci were first seen at the second transfer of the transfected cultures. At the third transfer, all positive cultures were overcrowded with foci and the cells were strongly vacuolated. Twenty-four-hour harvests of the confluent cultures showed reverse transcriptase activity. The harvests contained virus infectious for mink lung, SC-1, and NIH/3T3 cells. Furthermore, the virus released from mink lung fibroblasts was able to induce a thymic lymphoma 15 weeks after injection of 1,000 infectious units into a newborn NIH Swiss mouse.

DNA sequence analysis of the recombinant *env* gene. The MCF-MuLVs have undergone a substitution in the *env* gene of parental ecotropic viral genomes. We wanted to study the precise nature of the recombinant *env* gene of Mo-MCF viral DNA. We were particularly interested in studying the degree and extent of *env* gene substitution. We have determined the complete nucleotide sequence of the *env* gene and the 3′ LTR of molecularly cloned pMo-MCF$_I$-16 DNA (Fig. 3). The sequence of the Mo-MCF DNA was compared with the nucleotide sequence of parental ecotropic Mo-MuLV. The complete

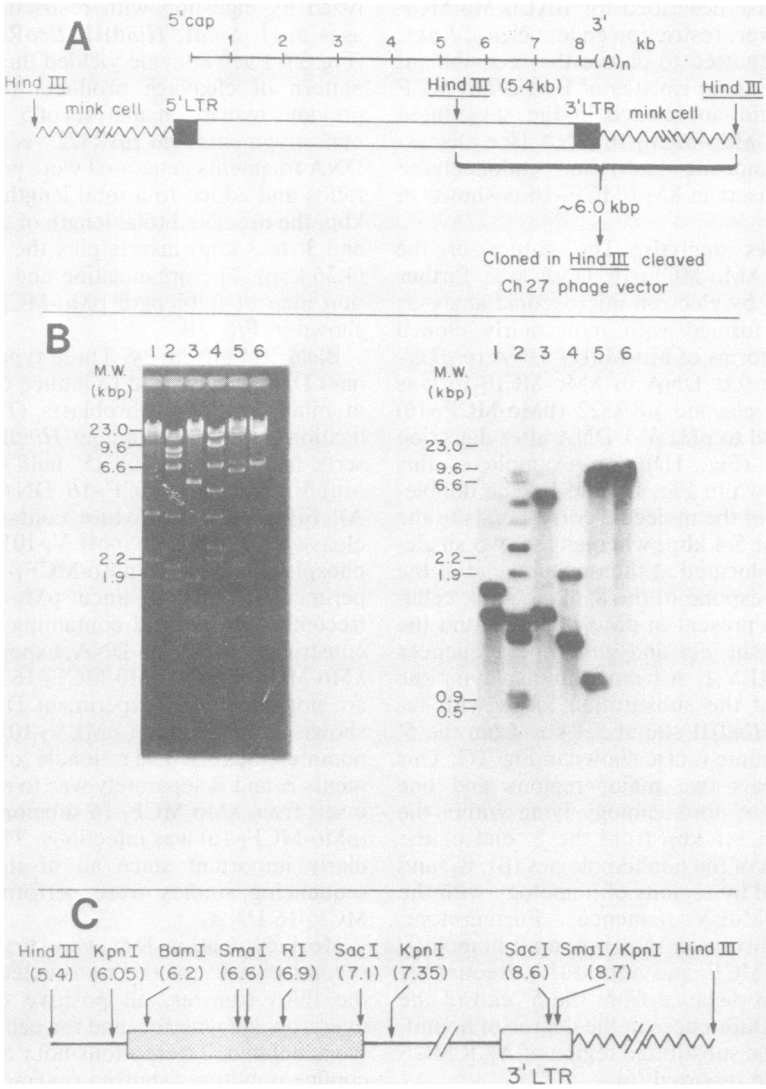
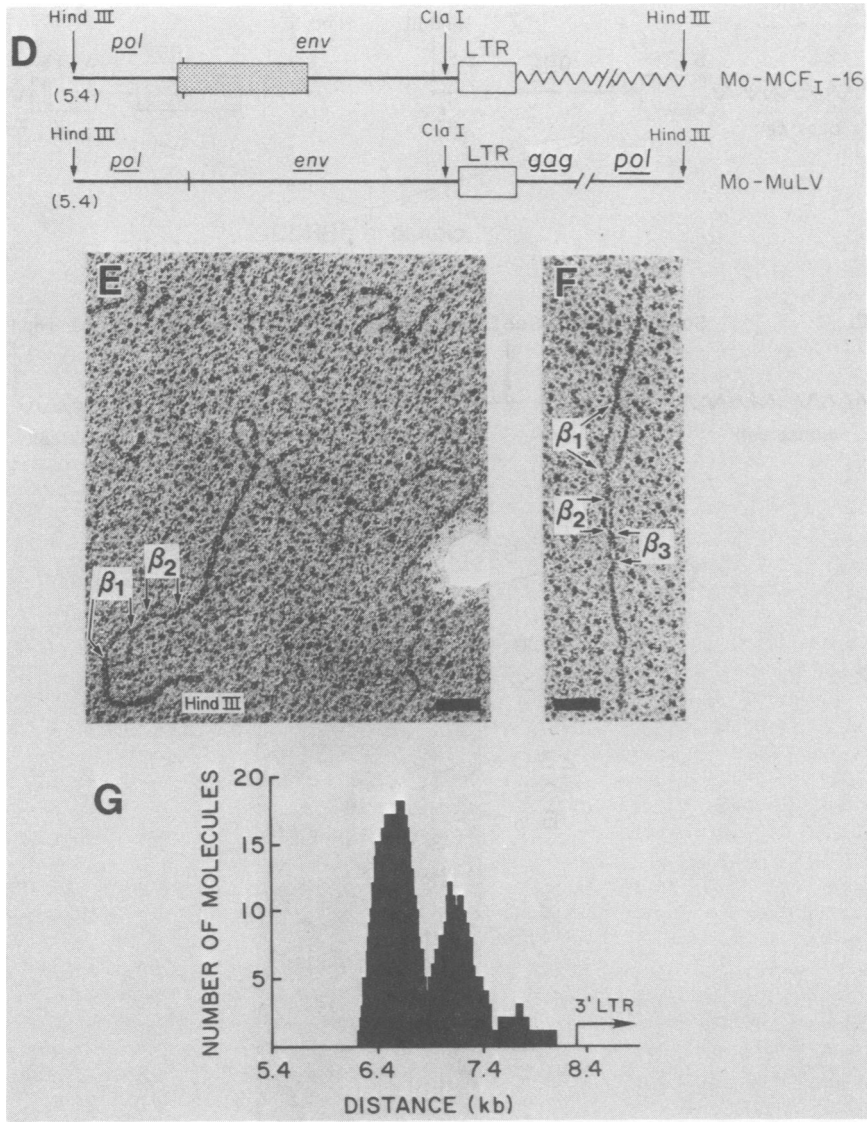


FIG. 1. Characterization of recombinant clone λMo-MCF$_1$-16. Both restriction endonuclease and heteroduplex analyses are shown. (A) Diagrammatic sketch of Mo-MCF proviral DNA based on our own analysis and that reported by van der Putten et al. (49). The region of the proviral DNA cloned by us is indicated. (B) Restriction endonuclease analysis of λMo-MCF$_1$-16. The recombinant phage DNA was cleaved with restriction endonucleases *Sac*I (lane 1), *Kpn*I (lane 2), *Eco*RI (lane 3), *Sma*I (lane 4), *Bam*HI (lane 5), and *Hin*dIII (lane 6): (left) ethidium bromide-stained pattern; (right) result of hybridization by 3′-specific probe to DNA from the same gel after Southern blot transfer (41). Lambda DNA fragments cleaved with *Hin*dIII were used as size markers. The *Kpn*I digest in lane 2 is only partially cleaved. (C) Physical map of λMo-MCF$_1$-16 based on restriction endonuclease analysis. The shaded area represents the putative MCF substitution region based on the restriction endonuclease analysis of Mo-MCF proviral DNA (49). (D) Diagrammatic sketch of inserts from pMo-MCF$_1$-16 and pMLV-1 to indicate the kind of expected heteroduplex molecules. The shaded area in pMo-MCF$_1$-16 insert DNA indicates the MCF substitution region. (C) and (D) are not drawn to scale. (E) A heteroduplex formed between inserts excised from λMo-MCF$_1$-16 and pMLV-1 DNAs is shown. The nonhomology loops β_1 and β_2 are indicated with arrows. Bar, 0.1 μm. (F) Another heteroduplex molecule where all three nonhomology loops (β_1, β_2, and β_3) are indicated. The heteroduplex molecule was formed between *Hin*dIII-*Cla*I fragments of pMo-MCF$_1$-16 and pMLV-1 DNA. (G) Histogram showing the frequency and position of substitution loops in the heteroduplex molecules analyzed. The 5′ end of the heteroduplex molecule is defined by the *Hin*dIII site, at 5.4 kbp of Mo-MCF proviral DNA (18). All nucelotide numbers in this figure include 5′ LTR; the 5′-cap nucleotide in this system is nucleotide 373.



nucleotide sequence of pMLV-1 (molecularly cloned, unintegrated form of Mo-MuLV [2]) has been derived by Shinnick et al. (40). Portions of the *env* gene and the complete sequence of the 3′ LTR of pMLV-1 and 5′ LTR of pMLV$_I$-101 were independently derived by us (46–48). Nucleotide 1 of pMo-MCF$_I$-16 in Fig. 3 corresponds to nucleotide 5,368 in pMLV-1 (40 [nucleotide numbers in this reference correspond to the viral genomic RNA]). In Fig. 3 the nucleotide number of the corresponding sequence of pMLV-1 is shown in parentheses. The results can be summarized as follows.

(i) The substitution in Mo-MCF DNA encompasses 1,159 bp and is located from nucleotide 323 (5,690) to nucleotide 1,482 (6,932). It should be noted that the substitution is not abrupt since sequences at the ends of the substitution show a high degree of homology with Mo-MuLV sequences (Fig. 4).

(ii) The substitution begins in the carboxyl terminus of the *pol* gene, about 90 nucleotides upstream from the initiator AUG of the Mo-MuLV *env* gene (5,777), and ends about 250 nucleotides before the first amino acid of p15E.

(iii) The computer-generated alignments of nucleotide sequences of Mo-MCF and Mo-MuLV *env* genes are shown in Fig. 4. The alignment score, an approximate measurement of the homology of two sequences (see legend, Fig. 4), was determined for each 100-nucleotide segment of the aligned sequences and plotted

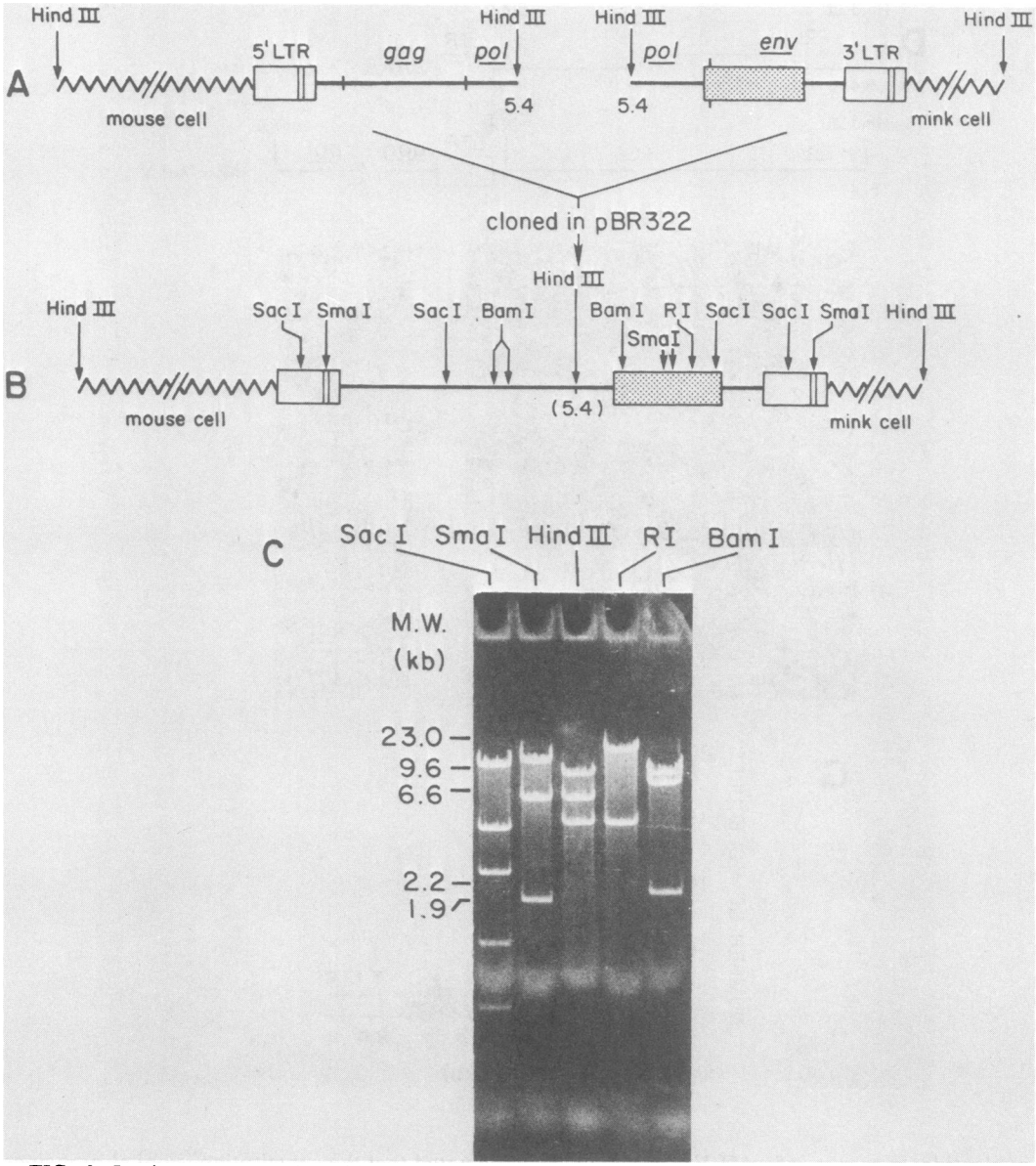


FIG. 2. In vitro reconstruction of full-length Mo-MCF proviral DNA. (A) Diagrammatic sketch of inserts of pMLV$_I$-101, 5′ half of BALB/Mo-MuLV containing adjacent mouse cellular sequences, and λMo-MCF$_I$-16 containing the 3′ half of Mo-MCF proviral DNA and adjacent mink cellular sequences. The shaded area in the λMo-MCF$_I$-16 insert DNA represents the MCF recombinant region. (B) Physical map of pMo-MCF$_I$-1 DNA. Although a more extensive restriction map is available upon request, only those enzymes shown in (C) are indicated. (C) Restriction endonuclease analysis of full-length pMo-MCF$_I$-1 DNA. The profile of the ethidium bromide-stained gel is displayed. The restriction endonucleases used were: *Sac*I (lane 1), *Sma*I (lane 2), *Hin*dIII (lane 3), *Eco*RI (lane 4), and *Bam*HI (lane 5).

versus the position of the segment in the alignment. The results indicate a gradual change in sequence homology at the 5′ and 3′ limits of the substitution, which suggests that the recombinant *env* gene of Mo-MCF may have been formed by homologous recombination (Fig. 4A). The two major regions of nonhomology are flanked by considerably more homologous regions. This observation is consistent with the heteroduplex analysis (Fig. 1G), which shows two major and one minor substitution loops. In contrast, the major substitution region of Mo-

TABLE 1. Biological infectivity of in vitro constructed Mo-MCF DNA molecules in transfected mink lung cells

Expt	Nature of DNA	Transfected culture	Proportion of cultures with cytopathic foci at the 3rd transfer
A	$pMLV_I$-101, + λMo-MCF_I-16	a	3/3
		b	3/3
		c	3/3
		d	3/3
		e	3/3
		f	3/3
B	$pMLV_I$-101 + pMo-MCF_I-16	a	0/2
		b	1/2
		c	1/2
		d	2/2
		e	0/1
		f	0/2
C[a]	pMo-MCF_I-1	a	0/3
		b	3/3
		c	3/3
		d	1/3
		e	1/3
		f	2/3
D	λMo-MCF_I-16 + pMo-MCF_I-16	a	0/3
		b	0/3
		c	0/3

[a] In experiment C, cultures d to f were transfected with a 1:10 dilution of the DNA used for transfection of cultures a to c.

loney murine sarcoma virus, which exhibits only one loop versus Mo-MuLV in heteroduplex analysis (10, 23), has a consistently low alignment score over its entire length (Fig. 4B).

(iv) The substituted sequence in Mo-MCF is 87 nucleotides (29 codons) shorter than the corresponding Mo-MuLV sequences. The two *env* genes share the same initiation and termination codons and are translated in the same reading frame.

(v) The nucleotide sequence predicts the Mo-MCF *env* gene product to be 636 amino acids long, which is 29 amino acids shorter than that predicted for the Mo-MuLV *env* gene product (40). The amino acid sequence differences between Mo-MCF and Mo-MuLV lie within the gp70 portion of the *env* gene (Fig. 5).

(vi) The putative signal peptide of the primary gene product of the Mo-MCF *env* gene is quite different from that of the parental Mo-MuLV DNA. Whereas the signal peptide cleavage site within the Pr85 *env* (primary gene product of Mo-MuLV) is known (S. Oroszlan, L. Henderson, and T. Copeland, personal communication; 48), sufficient change has occurred in the same region of the Mo-MCF *env* gene product to preclude any prediction of the signal peptide cleavage site or the N-terminal amino acid of the mature gp70 (Fig. 5).

(vii) There are seven canonical sequences, Asn-X-Thr and Asp-X-Ser (32), which can serve as glycosylation sites in the gp70 domains of both the Mo-MCF and Mo-MuLV *env* gene products. Six of the seven sites are common in both viral proteins. Previous work with G_{IX^+} and G_{IX^-} murine leukemia virus has indicated that the carbohydrate moiety is attached to the gp70 molecules at six or seven different sites, respectively (37).

(viii) The nucleotide sequences of Mo-MCF and Mo-MuLV *env* genes are identical after the substitution until nucleotide 2,358 (7,809), where an extra *A* residue is inserted in Mo-MCF. This region, just before the beginning of the 3′ LTR, has been implicated in the initiation of synthesis of a second strand of DNA during reverse transcription (18, 44).

(ix) The nucleotide sequence of the 3′ LTR of Mo-MCF may be compared with the previously published sequences of both the 5′ LTR of BALB/Mo-MuLV (46) and the LTR of unintegrated pMLV-1 (44, 47). Several murine LTRs contain duplication of sequences ranging from 72 to 110 bp (8, 47). The BALB/Mo-MuLV LTR does not contain these duplicated sequences. Hence, the BALB/Mo-MuLV LTR is about 70 to 100 bp shorter than other murine LTRs (47, 49) and, specifically, 75 bp shorter than that of its presumptive progenitor Mo-MuLV. If Mo-MCF arose from parental BALB/Mo-MuLV, one should expect its LTR sequence to be identical to the BALB/Mo-MuLV 5′ LTR sequences. The 3′ LTR of Mo-MCF DNA is, in fact, the same as that of the BALB/Mo-MuLV 5′ LTR (i.e., lacks 75-bp duplication), except for two changes at positions 2,730 (*T* for *C*) and 2,739 (*C* for *G*).

DISCUSSION

MuLVs can be classified into four related groups on the basis of their host range and interference properties in tissue culture: (i) ecotropic viruses, which replicate only in rodent cells; (ii) xenotropic viruses, which are unable to propagate in mouse cells but are capable of replication in nonmurine cells (1, 27); (iii) amphotropic viruses, which are unrestricted for growth in cells of either murine or nonmurine origin (21, 35); and (iv) MCF viruses, which have a dual host range similar to that of the amphotropic viruses. However, the growth of MCF viruses is interfered with by eco- and xenotropic viruses and not by amphotropic vi-

ruses (15, 16, 22, 45). Biochemical studies have clearly indicated that the MCF viruses are recombinants in which portions of the *env* gene are substituted by xenotrope-like *env* gene sequences. We have molecularly cloned the 3′ half of the BALB/Mo-MCF proviral DNA integrated in mink cell DNA. The molecularly cloned DNA (λMo-MCF$_I$-16) contains a portion of the *pol* gene, the recombinant *env* gene, the 3′ LTR, and adjacent mink cellular sequences. Restriction endonuclease analysis of the *env* gene portion of λMo-MCF$_I$-16 was in agreement with that of the previously published physical maps of Mo-MCF proviral DNA (49) with the exception that restriction endonuclease *Kpn*I, which cleaves the substituted sequences in Mo-MCF DNA, did not cleave the substituted sequences of the *env* gene in λMo-MCF$_I$-16. Thus, it appears that the extent of substitution may be different in various isolates of Mo-MCF viruses. Recently, it has been shown by Chattopadhyay et al. (5, 6) that different AKR MCF viruses have different extents of substitution, as judged by restriction endonuclease digestion patterns.

A biologically active full-length Mo-MCF DNA was constructed by ligation of the 3′ half (obtained from λMo-MCF$_I$-16 DNA) to the previously cloned 5′ half obtained from BALB/Mo-MuLV DNA. The full-length Mo-MCF DNA contained at its 5′ end cellular sequences derived from BALB/c mouse DNA, whereas the 3′-flanking sequences contained mink cellular sequences. Transfection of purified supercoiled plasmid DNA onto mink lung fibroblasts induced cytopathic foci, and harvests from confluent cultures were positive for reverse transcriptase activity. Furthermore, the virus released was infectious on mink lung fibroblasts, NIH/3T3 cells, and SC-1 cells, and the virus released from mink lung fibroblasts was able to induce a thymic lymphoma in a newborn NIH Swiss mouse. Controls involving transfection by the 5′- and 3′-cloned fragments alone did not produce any foci and showed no reverse transcriptase activity. In a previously reported experiment, the 5′ half of BALB/Mo-MuLV DNA was ligated to the 3′ half of the molecularly cloned unintegrated form of Mo-MuLV DNA and transfected onto NIH/3T3 cells to produce infectious Mo-MuLV virus (2). Thus, it appears that the information for the expanded host range and cytopathic effects of Mo-MCF virus lies in the 3′ half of Mo-MCF DNA.

The mature *env* gene products of murine retroviruses are comprised of two polypeptides, gp70 and p15E (13, 30). The two peptides are synthesized as a common precursor polyprotein, with the leader peptide and gp70 constituting the amino end and p15E constituting the carboxyl terminus. Nucleotide sequence analysis of the *env* gene of Mo-MCF DNA shows that recombination occurred in the NH_2-terminal portion of the envelope protein. In fact, the substitution begins in the COOH-terminal region of the *pol* gene of parental ecotropic viral DNA. Whether substitution occurred by homologous recombination or template switch during reverse transcription remains unclear, particularly in the absence of a known xenotrope-like *env* gene sequence that formed the substitution. The recombinant *env* gene product is 29 amino acids shorter than the parental ecotropic Mo-MuLV *env* gene product. The carboxyl termini of gP70 and the p15E of the recombinant and Mo-MuLV *env* genes are identical. Thus, it appears that the Mo-MCF-specific function lies in the NH_2 terminus of the gp70 moiety. Differences have been observed in the kinetics of processing of the *env* gene polyprotein of ecotropic, xenotropic, and dual-tropic AKR MCF MuLVs (14).

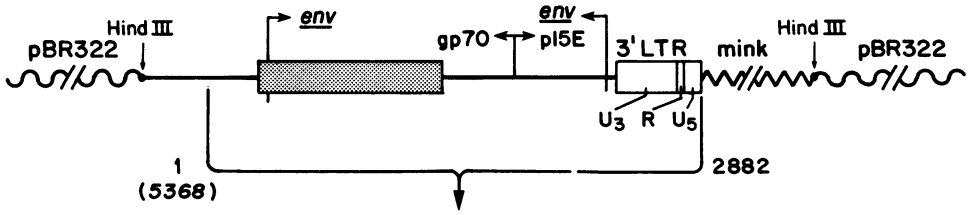
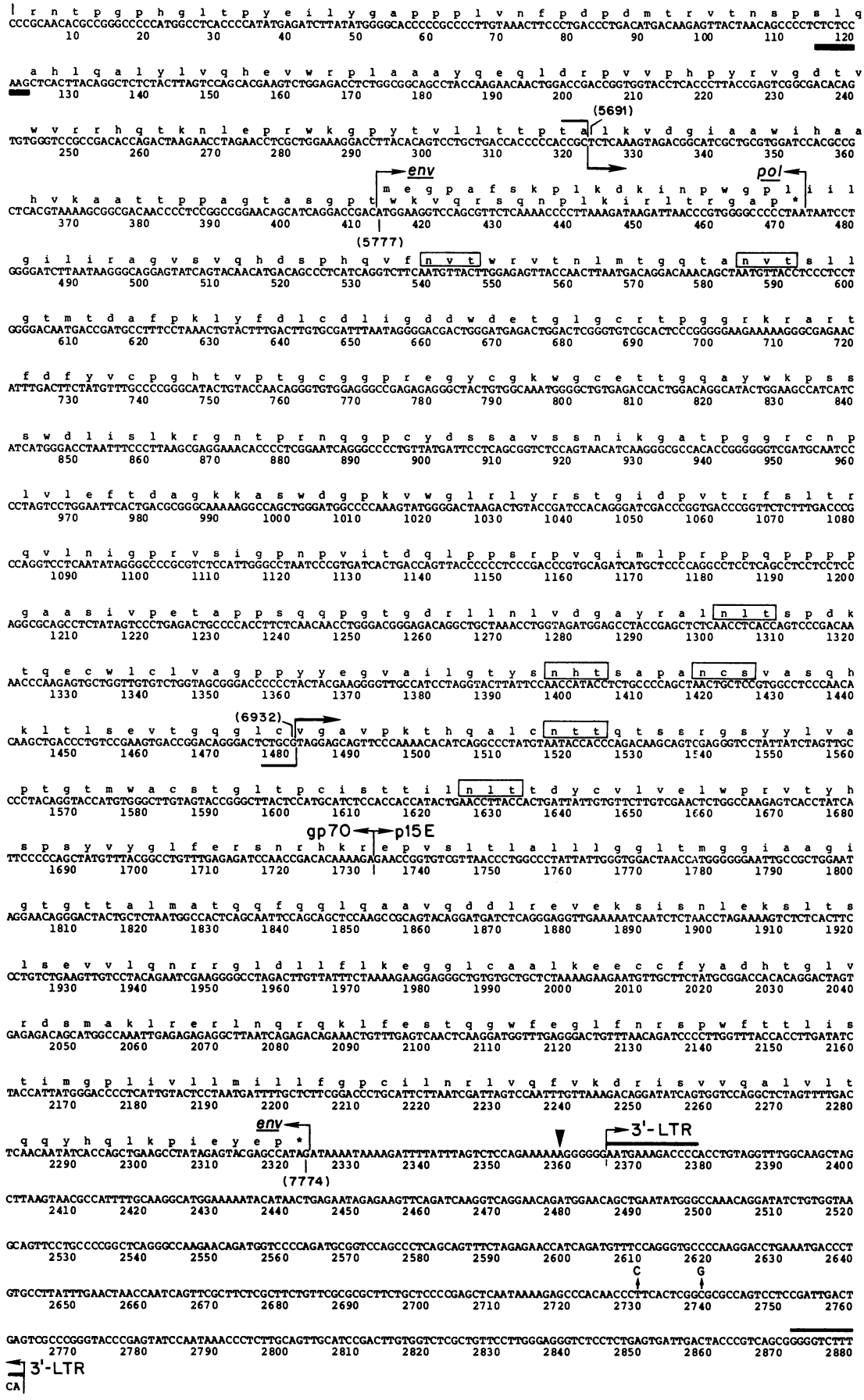


FIG. 3. Nucleotide sequence of Mo-MCF *env* gene and 3′ LTR. The nucleotide sequence of a portion of the insert in pMo-MCF$_I$-16 is presented, with the translations of the *pol* and *env* gene products given in single-letter code above. Nucleotide 1 corresponds to nucleotide 5,368 in the pMLV-1 sequence of Shinnick et al. (40; in the sequence of pMLV-1, the 5′-cap nucleotide of the genomic RNA is nucleotide 1). Nucleotide numbers in parentheses refer to positions in the sequence of pMLV-1. The features indicated include: closed box, splice acceptor site in Mo-MuLV *env* mRNA (W. N. Burnette, C. Van Beveren, and H. Fan, unpublished data); zig-zag arrows, beginning and end of changes in Mo-MCF (versus pMLV-1) in the *pol-env* region; open boxes, possible glycosylation sites in gp70; closed triangle, extra *A* residue in Mo-MCF (versus pMLV-1); overbar, inverted repeat at ends of LTR; short arrows, base changes in Mo-MCF LTR (versus pMLV$_I$-101 5′ LTR). Nucleotide sequence data were displayed by using the computer programs of Staden (42).


..pol
(5368)
(5691)
env
pol
(5777)
(6932)
gp70
p15E
env
3'-LTR
(7774)
3'-LTR

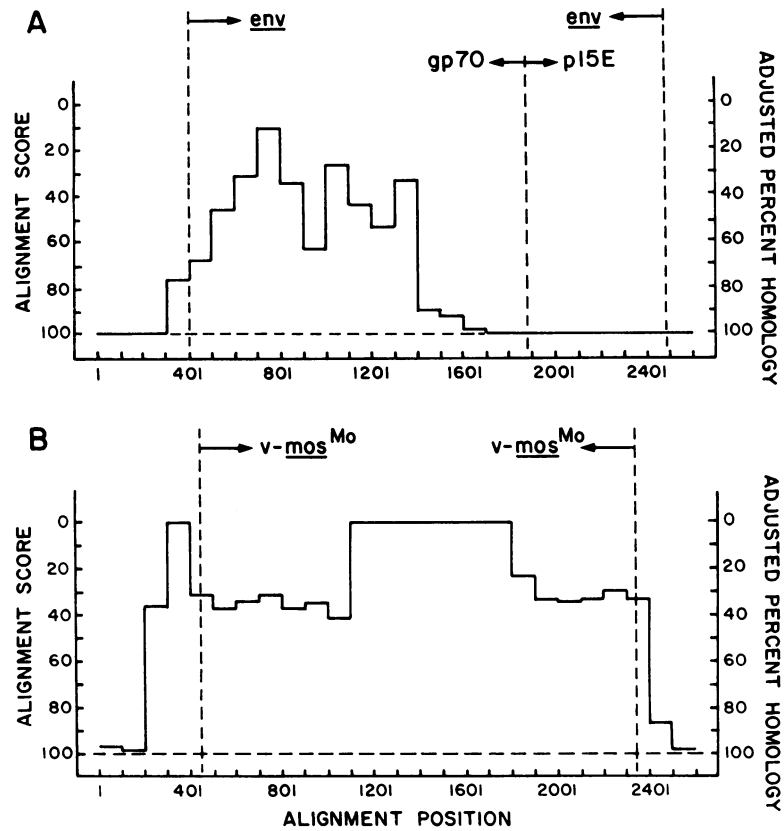


FIG. 4. Comparison of Mo-MCF and Mo-MuLV nucleotide sequences. Nucleotide sequences were compared by using the ALIGN program (National Biomedical Research Foundation). A unitary matrix and a gap penalty of 3 were used. The alignment score (total number of identities minus three times the number of gaps) was determined for each 100-nucleotide segment of the aligned sequences. (The aligned sequences include gaps; thus, the aligned nucleotide numbers are not necessarily the same as the numbers in Fig. 3 [see Fig. 5].) The alignment score for each 100-nucleotide segment was plotted versus the computer-assigned alignment position. The alignment score can be considered an approximate measure of the degree of homology of two sequences. (A) Comparison of nucleotide sequences of pMO-MCF$_r$-16 (Fig. 3, nucleotides 1 to 2,354) and pMLV-1 (40; nucleotides 5,368 to 7,805). (B) Comparison of nucleotide sequences of pMSV-1L (48; nucleotides 3,621 to 5,236) and pMLV-1 (40; nucleotides 5,330 to 7,807). In this case, to optimize the alignment, the 193-bp deletion of Mo-MuLV (deletion V in pMSV-1L; 48) was removed from the pMLV-1 sequence. Also, 712 bp of pMLV-1 sequence were removed arbitrarily from the middle of the substitution loop, permitting proper alignment of the beginning and end of the substitution. The deleted regions were reinserted after alignment, and the aligned sequences were renumbered appropriately.

The MCF *env* gene polyprotein is processed more slowly and is accessible to surface labeling. The differences in the amino acid sequences of the leader peptide of MCF and parental ecotropic virus may contribute towards the slow processing of the *env* gene polyprotein. It is, however, formally possible that the amino acid changes in the carboxyl terminus of the *pol* gene (10 of 49 amino acids) may have some role to play in the leukemogenesis by Mo-MCF viruses.

One of the salient features of the MCF viruses appears to be the different extent of substitution in different MCF viruses (28). However, all MCF viruses generally appear to have substitutions in the amino-terminal half of the *env* gene. On the basis of oligonucleotide fingerprinting data, it has been suggested that, in AKR MCF viruses, differences in the carboxyl terminus of p15E and the 3′ LTR may be responsible for the enhanced leukemogenic potential (28). The analysis of the Mo-MCF *env* gene reported here shows that the deduced amino acid sequence of the Mo-MCF p15E is identical to the parental Mo-MuLV p15E polypeptide. The only difference in the sequence of Mo-MCF and Mo-MuLV DNA prior to the 3′ LTR is the presence of an additional *A* residue in the region immediately preceding the LTR. This is the region

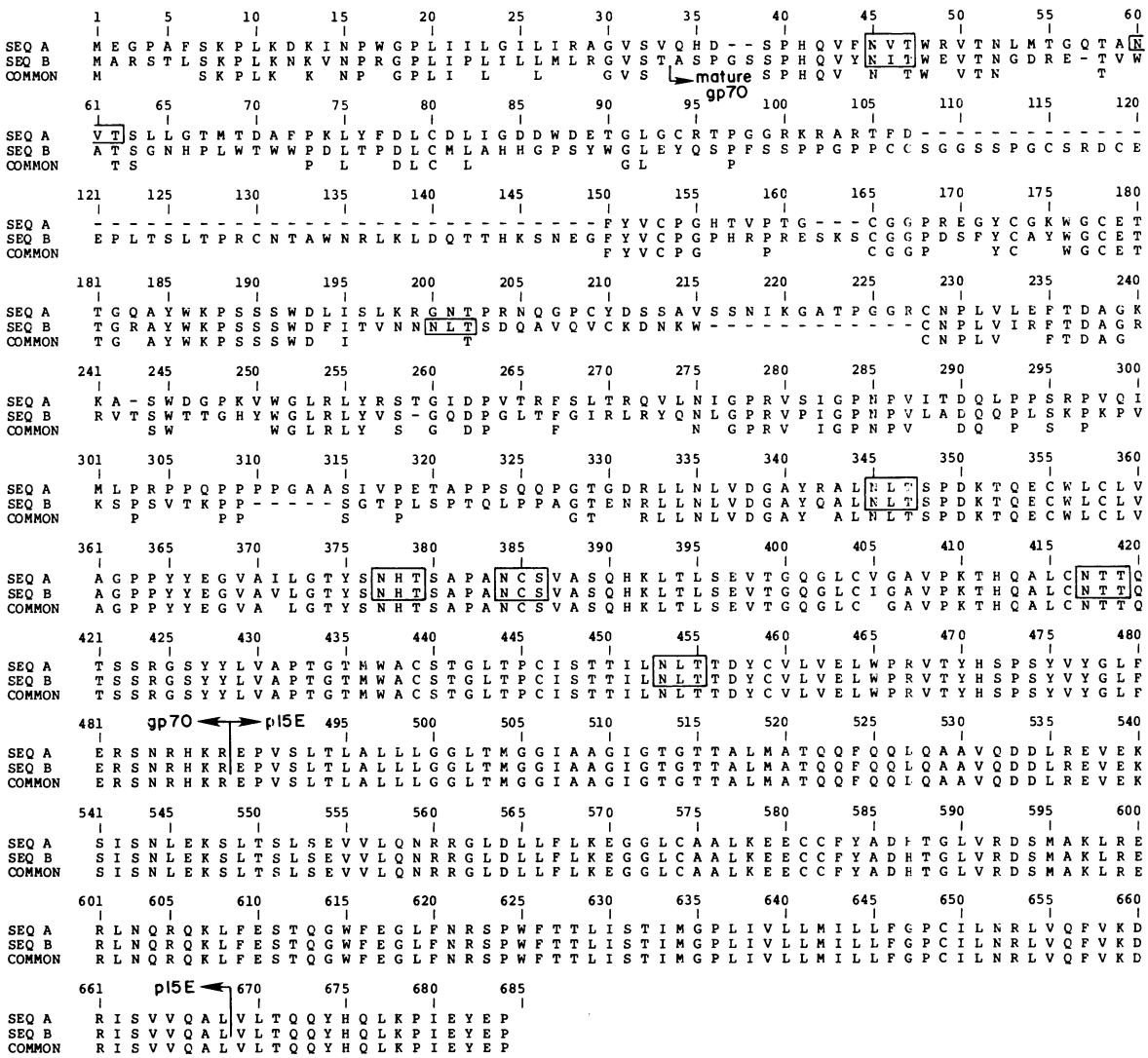


FIG. 5. Comparison of Mo-MCF and Mo-MuLV *env* gene products. The predicted amino acid sequences of the *env* gene products from pMo-MCF$_1$-16 and pMLV-1 were compared, using the ALIGN program. With the exception of cysteines (scored as 2), all identities were scored as 1 (i.e., unitary matrix), and the penalty for introduction of a gap was set at 3 (11). The N-terminus of the mature gp70 of Mo-MuLV (Oroszlan et al., personal communication) is shown (arrow), as are possible glycosylation sites (open boxes) of the two gp70s. The numbering, set by computer after the introduction of gaps to optimize the fit, refers to the total aligned sequence and does not necessarily indicate the amino acid number in either sequence. SEQ A, Mo-MCF *env* gene product; SEQ B, Mo-MuLV *env* gene product.

where the synthesis of the plus strand of viral DNA may initiate and conceivably could make a difference in the synthesis of double-stranded viral DNA. It will be interesting to determine whether the additional *A* residue is also found in other MCF isolates. The two single base changes in the 3′ LTR of Mo-MCF compared with the 5′ LTR of BALB/Mo-MuLV DNA do not appear to affect the control elements, such as the "TATA-like" box, "CAT" box, or poly-adenylate addition signals (17, 46). It should be pointed out, however, that there seems to be a fundamental difference between spontaneous tumors of AKR mice and spontaneous tumors of BALB/Mo mice. Unlike the tumors in BALB/Mo mice (25, 49), the tumors in AKR mice show no amplification of the ecotropic parental DNA (5, 34, 55).

The precise origin and role of MCF viruses in tumorigenesis remain elusive. The endogenous

nonecotropic virus-like sequences that presumably participate in the formation of the recombinant MCF *env* gene have not yet been identified. The consistent association of MCF sequences in MuLV-induced tumors suggests that this specific substitution of endogenous sequences is an important, and perhaps an obligatory, step in the onset of the disease. The MCF viruses are not always oncogenic (38), which indicates that mere acquisition of recombinant sequences is not sufficient for tumorigenesis. Other features, such as the growth of the recombinant virus in the target tissue, may also play an important role in the process of leukemogenesis. In light of the promoter insertion model of avian leukosis virus-induced leukemogenesis (31, 33), it can be argued that the MCF viruses induce the disease in a similar fashion by integrating in the vicinity of a cellular *onc* gene. Sustained efforts to demonstrate MCF-induced promotion of cellular genes have not been successful. It is our current working hypothesis that the formation of MCF recombinants is the first step in the onset of the disease. The actual neoplasia may be caused by other events involving chromosomal aberrations (26).

ACKNOWLEDGMENTS

We thank Susan Swift for technical assistance.

R.A.B. was supported by a Senior Dernham Fellowship from the California Division of the American Cancer Society. This work was supported by Public Health Service grants from the National Cancer Institute (CA-21408 and CA-16561 to I.M.V. and CA-13608 to M.V.) and a grant from the American Cancer Society (MCV-100B to I.M.V.).

LITERATURE CITED

1. **Benveniste, R. E., M. M. Lieber, and G. J. Todaro.** 1974. A distinct class of inducible murine type-C viruses that replicates in the rabbit S1RC cell line. Proc. Natl. Acad. Sci. U.S.A. **71**:602–606.
2. **Berns, A. J. M., M. H.-T. Lai, R. A. Bosselman, M. A. McKennett, L. T. Bacheler, H. Fan, E. C. Robanus-Maandag, H. Van der Putten, and I. M. Verma.** 1980. Molecular cloning of unintegrated and a portion of integrated Moloney murine leukemia viral DNA in bacteriophage lambda. J. Virol. **36**:254–263.
3. **Blattner, F. R., A. E. Blechl, K. Denniston-Thompson, H. E. Faber, J. E. Richards, J. L. Slightom, P. W. Tucker, and O. Smithies.** 1978. Cloning human fetal γ globin and mouse α-type globin DNA: preparation and screening of shotgun collections. Science **202**:1279–1284.
4. **Bosselman, R. A., L. J. L. D. van Griensven, M. Vogt, and I. M. Verma.** 1979. Genome organization of retroviruses. VI. Heteroduplex analysis of ecotropic and xenotropic sequences of Moloney mink cell focus-inducing viral RNA obtained from either a cloned isolate or a thymoma cell line. J. Virol. **32**:968–978.
5. **Chattopadhyay, S. K., M. W. Cloyd, D. L. Linemeyer, M. R. Lander, E. Rands, and D. R. Lowy.** 1982. Cellular origin and role of mink cell focus-forming viruses in murine thymic lymphomas. Nature (London) **295**:25–31.
6. **Chattopadhyay, S. K., M. R. Lander, S. Gupta, E. Rands, and D. R. Lowy.** 1981. Origin of mink cytopathic focus-forming (MCF) viruses: comparison with ecotropic and xenotropic murine leukemia virus genomes. Virology **113**:465–483.
7. **Chien, Y.-H., I. M. Verma, T. Y. Shih, E. M. Scolnick, and N. Davidson.** 1978. Heteroduplex analysis of the sequence relations between the RNAs of mink cell focus-inducing and murine leukemia viruses. J. Virol. **28**:352–360.
8. **Dhar, R., W. L. McClements, L. W. Enquist, and G. F. Vande Woude.** 1980. Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. Proc. Natl. Acad. Sci. U.S.A. **77**:3937–3941.
9. **Donoghue, D. J., E. Rothenberg, N. Hopkins, D. Baltimore, and P. A. Sharp.** 1978. Heteroduplex analysis of nonhomology region between Moloney MuLV and the dual host range derivative HIX virus. Cell **14**:959–970.
10. **Donoghue, D. J., P. J. Sharp, and R. A. Weinberg.** 1979. Comparative study of different isolates of murine sarcoma virus. J. Virol. **32**:1015–1027.
11. **Doolittle, R. F.** 1981. Similar amino acid sequences: chance or common ancestry? Science **214**:149–159.
12. **Elder, J. H., J. W. Gautsch, F. C. Jensen, R. A. Lerner, J. W. Hartley, and W. P. Rowe.** 1977. Biochemical evidence that MCF murine leukemia viruses are envelope (*env*) gene recombinants. Proc. Natl. Acad. Sci. U.S.A. **74**:4676–4680.
13. **Famulari, N. G., D. L. Buchhagen, H. D. Klenk, and E. Fleissner.** 1977. Presence of murine leukemia virus envelope proteins gp70 and p15(E) in a common polyprotein of infected cells. J. Virol. **20**:501–508.
14. **Famulari, N. G., J.-S. Tung, P. V. O'Donnell, and E. Fleissner.** 1980. Murine leukemia virus *env* gene expression in preleukemic thymocytes and leukemia cells of AKR strain mice. Cold Spring Harbor Symp. Quant. Biol. **44**:1281–1287.
15. **Fischinger, P. J., J. N. Ihle, F. de Noronha, and D. P. Bolognesi.** 1977. Oncogenic and immunogenic potential of cloned HIX virus in mice and cats. Med. Microbiol. Immunol. **164**:119–129.
16. **Fischinger, P. J., S. Nomura, and D. P. Bolognesi.** 1975. A novel murine oncornavirus with dual eco- and xenotropic properties. Proc. Natl. Acad. Sci. U.S.A. **72**:5150–5155.
17. **Fuhrman, S., C. Van Beveren, and I. M. Verma.** 1981. Identification of an RNA polymerase II initiation site in the long terminal repeat of Moloney murine leukemia viral DNA. Proc. Natl. Acad. Sci. U.S.A. **78**:5411–5415.
18. **Gilboa, E., S. W. Mitra, S. Goff, and D. Baltimore.** 1979. A detailed model of reverse transcription and tests of crucial aspects. Cell **18**:93–100.
19. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
20. **Gross, L.** 1970. Oncogenic viruses, 2nd ed. Pergamon Press, New York.
21. **Hartley, J. W., and W. P. Rowe.** 1976. Naturally occurring murine leukemia viruses in wild mice: characterization of a new "amphotropic" class. J. Virol. **19**:19–25.
22. **Hartley, J. W., N. K. Wolford, L. J. Old, and W. P. Rowe.** 1977. A new class of murine leukemia virus associated with development of spontaneous lymphomas. Proc. Natl. Acad. Sci. U.S.A. **74**:789–792.
23. **Hu, S., N. Davidson, and I. M. Verma.** 1977. A heteroduplex study of the sequence relationships between the RNAs of M-MSV and M-MLV. Cell **10**:469–477.
24. **Jaenisch, R.** 1977. Germ line integration of Moloney leukemia virus: effect of homozygosity at the M-MuLV locus. Cell **12**:691–696.
25. **Jähner, D., H. Stuhlmann, and R. Jaenisch.** 1980. Conformation of free and of integrated Moloney leukemia virus proviral DNA in preleukemic and leukemic Balb/Mo mice. Virology **101**:111–123.
26. **Klein, G.** 1981. The role of gene dosage and genetic transpositions in carcinogenesis. Nature (London) **294**:313–318.
27. **Levy, J. A.** 1973. Xenotropic viruses: murine leukemia viruses associated with NIH Swiss, NZB and other mouse strains. Science **182**:1151–1153.

28. **Lung, M. L., C. Hering, J. W. Hartley, W. P. Rowe, and N. Hopkins.** 1980. Analysis of the genomes of mink cell focus-inducing murine type-C viruses: a progress report. Cold Spring Harbor Symp. Quant. Biol. **44**:1269–1274.
29. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
30. **Naso, R. B., L. J. Arcement, W. L. Karsnin, G. A. Jamjoom, and R. B. Arlinghaus.** 1976. A fucose-deficient glycoprotein precursor to Rauscher leukemia virus gp69/71. Proc. Natl. Acad. Sci. U.S.A. **73**:2326–2330.
31. **Neel, B. G., W. S. Hayward, H. L. Robinson, J. Fang, and S. M. Astrin.** 1981. Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. Cell **23**:323–334.
32. **Neuberger, A., A. Gottschalk, R. D. Marshall, and R. G. Spiro.** 1972. Carbohydrate-peptide linkages in glycoproteins and methods for their elucidation, p. 450–490. *In* A. Gottschalk (ed.), Glycoproteins: their composition, structure and function, 2nd ed. Elsevier, Amsterdam.
33. **Payne, G. S., S. A. Courtneidge, L. B. Crittenden, A. M. Fadly, J. M. Bishop, and H. E. Varmus.** 1981. Analysis of avian leukosis virus DNA and RNA in bursal tumors: viral gene expression is not required for maintenance of the tumor state. Cell **23**:311–322.
34. **Quint, W., W. Quax, H. van de Putten, and A. Berns.** 1981. Characterization of AKR murine leukemia virus sequences in AKR mouse substrains and structure of integrated recombinant genomes in tumor tissue. J. Virol. **39**:1–10.
35. **Rasheed, S., M. B. Gardner, and E. Chan.** 1976. Amphotropic host range of naturally occurring wild mouse leukemia viruses. J. Virol. **19**:13–18.
36. **Rommelaere, J., D. V. Faller, and N. Hopkins.** 1978. Characterization and mapping of RNase T1-resistant oligonucleotides derived from the genomes of Akv and MCF murine leukemia viruses. Proc. Natl. Acad. Sci. U.S.A. **75**:495–499.
37. **Rosner, M. R., J.-S. Tung, N. Hopkins, and P. W. Robbins.** 1980. Relationship of G_{IX} antigen expression to the glycosylation of murine leukemia virus glycoprotein. Proc. Natl. Acad. Sci. U.S.A. **77**:6420–6424.
38. **Rowe, W. P., M. W. Cloyd, and J. W. Hartley.** 1980. Status of the association of mink cell focus-forming viruses with leukemogenesis. Cold Spring Harbor Symp. Quant. Biol. **44**:1265–1268.
39. **Shih, T. Y., M. O. Weeks, D. H. Troxler, J. M. Coffin, and E. M. Scolnick.** 1978. Mapping host range-specific oligonucleotides within genomes of the ecotropic and mink cell focus-inducing strains of Moloney murine leukemia virus. J. Virol. **26**:71–83.
40. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliff.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) **293**:543–548.
41. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
42. **Staden, R.** 1979. A strategy of DNA sequencing employing computer programs. Nucleic Acids Res. **6**:2601–2610.
43. **Stow, N. D., and N. M. Wilkie.** 1976. An improved technique for obtaining enhanced infectivity with herpes simplex virus type I DNA. J. Gen. Virol. **33**:447–458.
44. **Sutcliffe, J. G., T. M. Schinnick, I. M. Verma, and R. A. Lerner.** 1980. Nucleotide sequence of Moloney leukemia virus: 3′ end reveals details of replication, analogy to bacterial transposons, and an unexpected gene. Proc. Natl. Acad. Sci. U.S.A. **77**:3302–3306.
45. **Troxler, D. H., E. Yuen, D. Linemeyer, S. Ruscetti, and E. M. Scolnick.** 1978. Helper-independent mink cell focus-inducing strains of Friend murine type-C virus: potential relationship to the origin of replication-defective spleen focus-forming virus. J. Exp. Med. **148**:639–653.
46. **Van Beveren, C., J. G. Goddard, A. Berns, and I. M. Verma.** 1980. Structure of Moloney murine leukemia viral DNA: nucleotide sequence of the 5′ long terminal repeat and adjacent cellular sequences. Proc. Natl. Acad. Sci. U.S.A. **77**:3307–3311.
47. **Van Beveren, C., E. Rands, S. K. Chattopadhyay, D. R. Lowy, and I. M. Verma.** 1982. Long terminal repeat of murine retroviral DNAs: sequence analysis, host-proviral junctions and preintegration site. J. Virol. **41**:542–556.
48. **Van Beveren, C., F. van Straaten, J. A. Galleshaw, and I. M. Verma.** 1981. Nucleotide sequence of the genome of a murine sarcoma virus. Cell **27**:97–108.
49. **van der Putten, H., W. Quint, J. van Raaij, E. Robanus-Maandag. I. M. Verma, and A. Berns.** 1981. M-MuLV-induced leukemogenesis: integration and structure of recombinant proviruses in tumors. Cell **24**:729–739.
50. **van der Putten, H., E. Terwindt, A. Berns, and R. Jaenisch.** 1979. The integration sites of endogenous and exogenous Moloney murine leukemia virus. Cell **18**:109–116.
51. **Van Griensven, L. J. L. D., and M. Vogt.** 1980. Rauscher "mink cell focus-inducing" (MCF) virus causes erythroleukemia in mice: its isolation and properties. Virology **101**:376–388.
52. **Verma, I. M., M. H.-T. Lai, R. A. Bosselman, M. A. McKennett, H. Fan, and A. Berns.** 1980. Molecular cloning of unintegrated Moloney mouse sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **77**:1773–1777.
53. **Vogelstein, B., and D. Gillespie.** 1979. Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. U.S.A. **76**:615–619.
54. **Vogt, M.** 1979. Properties of "mink cell focus-inducing" (MCF) virus isolated from spontaneous lymphoma lines of BALB/c mice carrying Moloney leukemia virus as an endogenous virus. Virology **93**:226–236.
55. **Yoshimura, F. K., and M. Breda.** 1981. Lack of AKR ecotropic provirus amplification in AKR leukemia thymuses. J. Virol. **39**:808–815.