JOURNAL OF VIROLOGY, Sept. 1987, p. 2659–2669
0022-538X/87/092659-11$02.00/0

Vol. 61, No. 9

The Four Classes of Endogenous Murine Leukemia Virus: Structural Relationships and Potential for Recombination

JONATHAN P. STOYE AND JOHN M. COFFIN*

Cancer Research Center and Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts 02111

Received 26 March 1987/Accepted 12 May 1987

The process by which leukemogenic viruses are generated during the lifetime of certain strains of mice is poorly understood. We have therefore set out to define all the murine leukemia virus-related endogenous proviruses of HRS/J mice. We have cloned 34 different proviral fragments and their flanking cellular sequences. These have been characterized by restriction enzyme analysis, by fingerprinting in vitro-synthesized RNA, and by DNA sequencing. We conclude that all the proviruses can be assigned into one of four different classes: the previously characterized ecotropic, xenotropic, and polytropic viruses, as well as a new class we have termed modified polytropic viruses. The xenotropic, polytropic, and modified polytropic classes are closely related to one another, but as a group they differ considerably from the ecotropic class. Sequence analyses show that both polytropic and modified polytropic sequences can contribute *env* sequences to recombinant viruses.

All inbred strains of mice contain numerous endogenous type C retroviruses related to murine leukemia virus (MLV) whose expression is both strain and tissue specific (reviewed in references 10, 51, 56). These elements are likely to be of considerable importance to the biology and genetics of these mice; however, much confusion remains about the nature and number of these endogenous proviruses. They are conveniently classified by the host range conferred by their *env* genes. The best-characterized sequences are those encoding infectious ecotropic viruses which can replicate on murine cells only and xenotropic viruses which can replicate on nonmurine but not on murine cells (12, 13, 18, 20, 25, 43). Most mice contain only one or no ecotropic viruses (26), whereas the number of xenotropic viruses is usually somewhat higher (44). The remaining endogenous viruses, present at between 12 and 30 copies (44), have not yet been systematically studied, although a number of isolated proviruses have been examined in varying detail (28, 30, 52). These viruses differ from the inducible ecotropic and xenotropic viruses in a number of respects, including long terminal repeat (LTR) size (29, 45), tRNA primer-binding site (42), and the sequences of their *env* genes (28, 32). Studies of cloned proviruses have not revealed major differences among the nonecotropic proviruses; however, analyses of genomic DNA have provided suggestive evidence for the presence of two classes of nonecotropic, nonxenotropic viruses differing slightly in the distribution of restriction endonuclease cleavage sites (15, 55).

Endogenous proviruses have also been implicated in spontaneous leukemogenesis. High leukemia strains of mice such as AKR, C58, and HRS/J express high levels of ecotropic virus from birth (for a review, see reference 51). Xenotropic as well as polytropic viruses, which replicate on murine and nonmurine cells, can be isolated from lymphoid tissues of older animals (19, 21, 58). Injection of some polytropic viruses into newborn leukemia-prone animals accelerates the leukemogenic process, and thus the polytropic viruses appear to be the proximal viral inducers of disease. They are generated by recombination between ecotropic and one or more endogenous nonecotropic viruses, resulting in substitutions both within the gp70-coding region of *env* and within the p15E/U_3 region (7, 19, 24, 27, 30, 35). Structural analyses of a number of recombinant viruses imply a recombination with the same parental sequence but at different positions (35). The p15E/U_3 substitution occurs at a different time and thus seems to result from recombination with yet another unique endogenous sequence (47, 58).

Spontaneous leukemogenesis thus appears to be an extremely complex virological process involving the inheritance of the appropriate proviral parents, their expression, and a series of recombination events to generate the leukemogenic virus followed by infection of the target tissue and induction of the disease. To study this process, we wished to identify and characterize the participating endogenous sequences. In addition, we wanted to clarify the relationships between the endogenous proviruses themselves as well as their relationship with the host genome. With these aims in mind, we set out to obtain a panel of endogenous viral clones which is as complete and representative as possible and to develop methods for their rapid analysis. We report here the isolation and characterization of two different classes of endogenous virus which appear to contribute different *env* genes to the recombinant viruses derived from distinct parental viruses.

MATERIALS AND METHODS

Molecular cloning of endogenous proviruses. High-molecular-weight DNA extracted from the brains of newborn HRS/J mice, progeny of a cross between HRS/J *hr*/*hr* and HRS/J *hr*/+, was digested with *Hin*dIII, ligated to lambda bacteriophages Charon 27 (50) or Charon 35 (33), and plated, without amplification, on *Escherichia coli* LE392 and DH1, respectively. Plaque lifts (2) were screened by using nick-translated, gel-purified p247-1b DNA (27). p247-1b contains an 8.2-kilobase (kb) *Hin*dIII fragment corresponding to an intact, permuted copy of mink cell focus-forming virus 247

* Corresponding author.

TABLE 1. Characteristics of nonecotropic proviruses

Nonecotropic provirus	Restriction enzyme site[a]		27-bp deletion in gp70 at 7.1 kb	150- to 190-bp insertion in LTR	No. in HRS/J genome[b]	Titer[c]
	*Hin*dIII at 5.6 kb	*Bam*HI at 6.2 kb				
Xenotropic	–[d]	–	–	–	5	1
Modified polytropic	+	+	+	+	11	0.002
Polytropic	–	+	–	+	18	0.1

[a] Relative to the predicted 5′ end of the genome.
[b] Estimated by using class-specific oligonucleotide probes (Stoye and Coffin, in preparation).
[c] Ratio of the endpoint titers of the respective virus class in mink cells relative to the titers in *Mus dunii* cells (Stoye and Coffin, in preparation).
[d] –, Absent; +, present.

(MCF-247). Positive clones were purified by two further rounds of plaque purification. Plate lysate preparations of DNA (37) were analyzed by digestion with *Hin*dIII and *Bam*HI. Unique isolates were subcloned into pBR322, pGEM1, or pGEM2 (Promega Biotec) or one of the Bluescribe vectors (Stratagene, Inc.) for further restriction mapping and hybridization analysis using viral fragments of pMCF-2471b or p623 (34), that were subcloned into M13 and labeled by DNA polymerase-Klenow fragment extension of the hybridization probe primer (New England BioLabs, Inc., Beverly, Mass.). The subgenomic clones were (i) a 380-base-pair (bp) *Pst*I-*Sma*I LTR fragment of p247-1b (0.1 to 0.4 kb), (ii) a 1.9-kb *Bam*HI *gag-pol* fragment of p623 (1.8 to 3.7 kb), (iii) a 0.8-kb *Cla*I-*Bam*HI *pol* fragment of p247-1b (5.6 to 6.2 kb), and (iv) a 0.6 kb *Bam*HI-*Eco*RI *env* fragment of p247-1b (6.2 to 6.9 kb). All restriction enzymes came from New England BioLabs, and digestions were performed as recommended by the manufacturer.

DNA sequencing. The *env* genes and LTRs of pMX27 and pMX33 were sequenced by the dideoxy-chain terminator method (54), using buffer gradient gels (3) and [α-^{32}P]dATP (New England Nuclear Corp., Boston, Mass.). Appropriate restriction fragments were cloned into M13 vectors mp8, mp9, and mp19 (40). Nested sets of deletions were prepared as described by Dale et al. (14) with reagents purchased from International Biotechnologies, Inc. The sequences of MX11, -12, -13, and -17 were obtained by sequencing mp19 clones containing *Eco*RI-*Kpn*I fragments from the middle of the *env* gene.

In vitro RNA synthesis. Labeled RNA was prepared by slight modifications of a published procedure (39). pGEM1 or Bluescribe M13 plus DNA (1 μg) containing the subcloned proviral fragments was incubated for 45 min at 37°C in 25 μl of 40 mM Tris hydrochloride (pH 7.5)–6mM $MgCl_2$–2 mM spermidine–10 mM dithiothreitol–0.5 mM ATP–0.5 mM GTP–0.5 mM CTP–0.05 mM UTP containing 5 U of SP6 RNA polymerase or 5 U of T7 RNA polymerase, chosen so as to synthesize positive-strand RNA–25 U of RNasin–100 μCi of [α-^{32}P]UTP (800 Ci/mMol). Then, 1 μg of DNase I and $MgCl_2$ to 20 mM was added, and the incubation was continued for 30 min at room temperature before adding 50 μg of carrier yeast RNA. Newly synthesized RNA was separated from unincorporated nucleotides on a 1-ml Sephadex G100 column in 100 mM NaCl–10 mM Tris hydrochloride (pH 7.5)–1 mM EDTA, ethanol precipitated, and suspended in 100 μl of 0.5 M NaCl–10 mM Tris hydrochloride (pH 7.5). Single-stranded M13 DNA clones (10 μg) containing the negative strand of MCF-247 from the *Cla*I (5.6 kb) to the *Sma*I (6.5 kb) and the *Sma*I fragment (6.6 to 8.7 kb) were added, and the samples were boiled for 5 min, chilled on ice, and incubated at 65°C overnight. Samples were then treated with 10 U of RNase T_1, and hybridized, protected RNA was purified by chromatography on Sephadex G100 as previously described (11). Samples were then denatured and digested to completion with RNase T_1 and separated on two-dimensional polyacrylamide gels (11, 59). Selected oligonucleotides were further analyzed by digestion with RNase A and high-voltage ionophoresis on DEAE-chromatography paper, as described previously (5, 11).

RESULTS

Identification and classification of cloned proviruses. To isolate proviruses containing intact *env* genes for structural and functional characterization, we chose to clone proviral fragments from mouse DNA digested with *Hin*dIII, an endonuclease which does not cut within the substituted sequence in most polytropic viruses (8). We have cloned 34 unique *Hin*dIII fragments, designated MX1 to MX34, from HRS/J DNA containing retroviral and associated cellular sequences by using the cloned polytropic virus MCF-247 as a probe (24, 27). A full description of these clones will be published elsewhere (J. P. Stoye and J. M. Coffin, manuscript in preparation). In brief, the cloned fragments include the 5′ and 3′ halves of two ecotropic viruses endogenous to HRS/J, *Emv-1* and *Emv-3* (26); three solo LTRs; six proviral

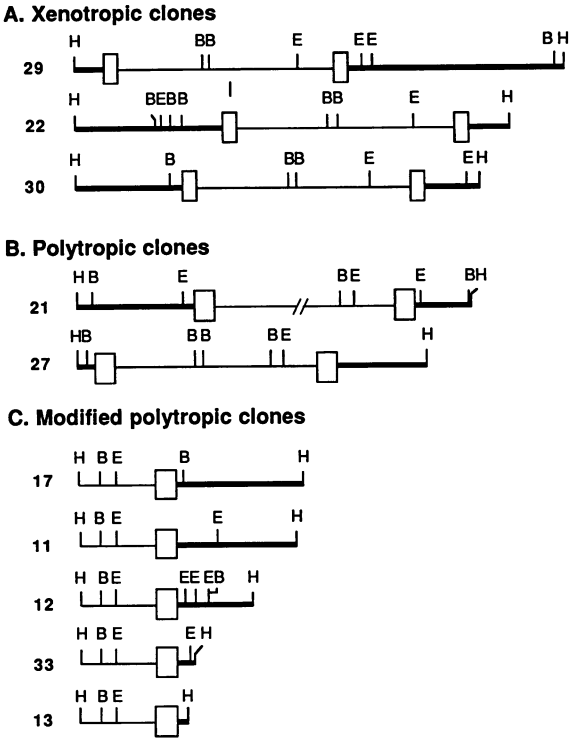

FIG. 1. Cloned proviruses and their flanking sequences. Clones were mapped by digestion with a number of restriction enzymes (including some not shown on the figure) and by hybridization with a variety of subgenomic viral probes (see Materials and Methods). The LTRs of the polytropic and modified polytropic sequences are 150 to 190 bp longer than the xenotropic proviral clones. Symbols: □, LTRs; ▬, flanking sequences; —, viral sequences; //, a deletion. Restriction enzyme sites shown: E, *Eco*RI; B, *Bam*HI; H, *Hin*dIII. Numbers at the left are the MX designations of each clone.

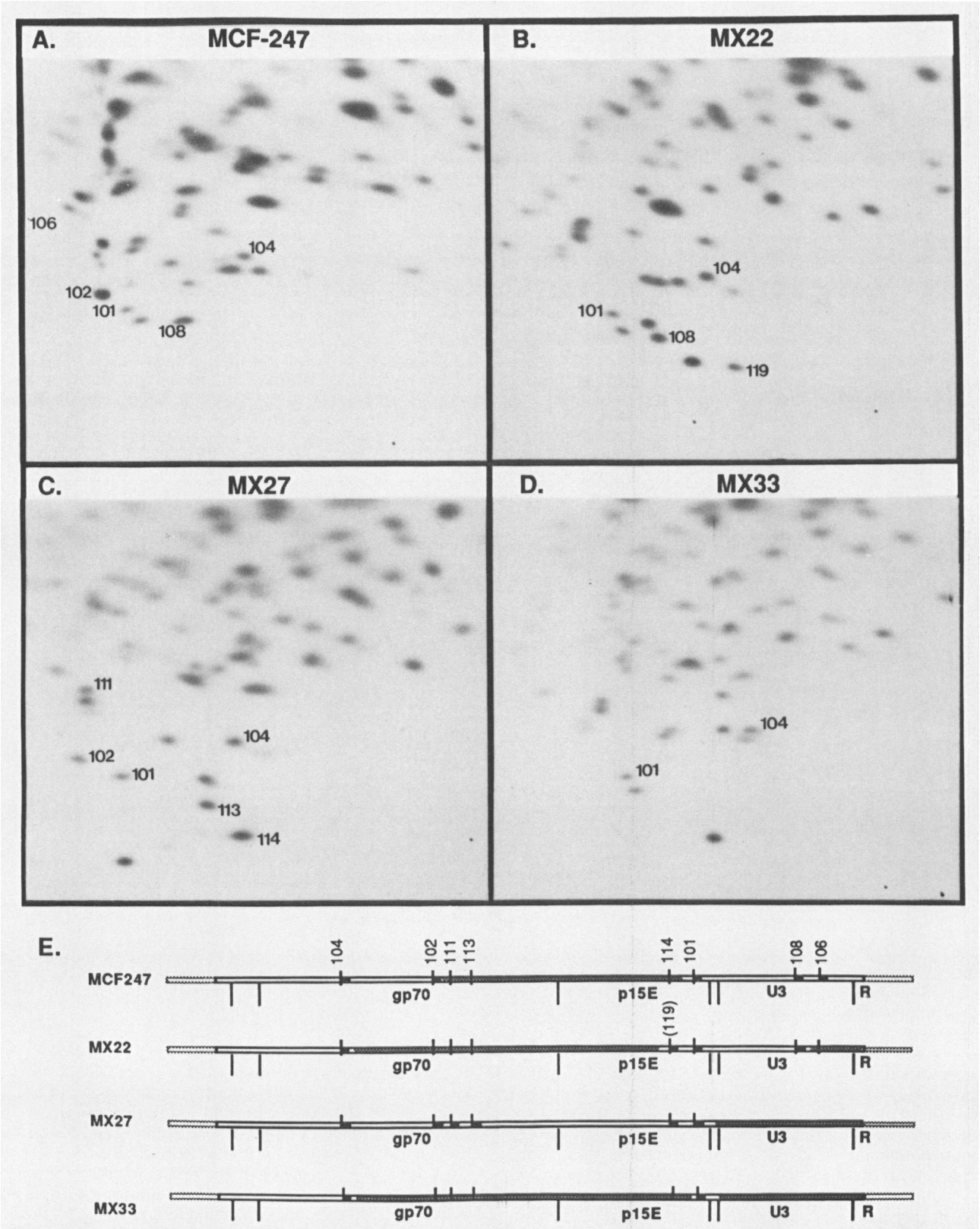

FIG. 2. RNA fingerprints of representative clones. Fingerprints of the *env* genes and LTRs of representative clones MCF-247 (A), MX22 (B), MX27 (C), and MX33 (D). Oligonucleotides were numbered according to previously published conventions (19, 35, 53, 59). Only the oligonucleotides previously characterized in polytropic viruses have been numbered. (E) A schematic representation of the *env* and LTR regions of the four viruses, showing the boundaries of gp70, p15E, U_3, and R and marking the positions of the different oligonucleotides. The presence of a given oligonucleotide within a virus is indicated by a small filled black box, the absence of a given oligonucleotide and the areas of divergent sequence are shown in the shaded regions. 114 and 119 are allelic forms of the same oligonucleotide.

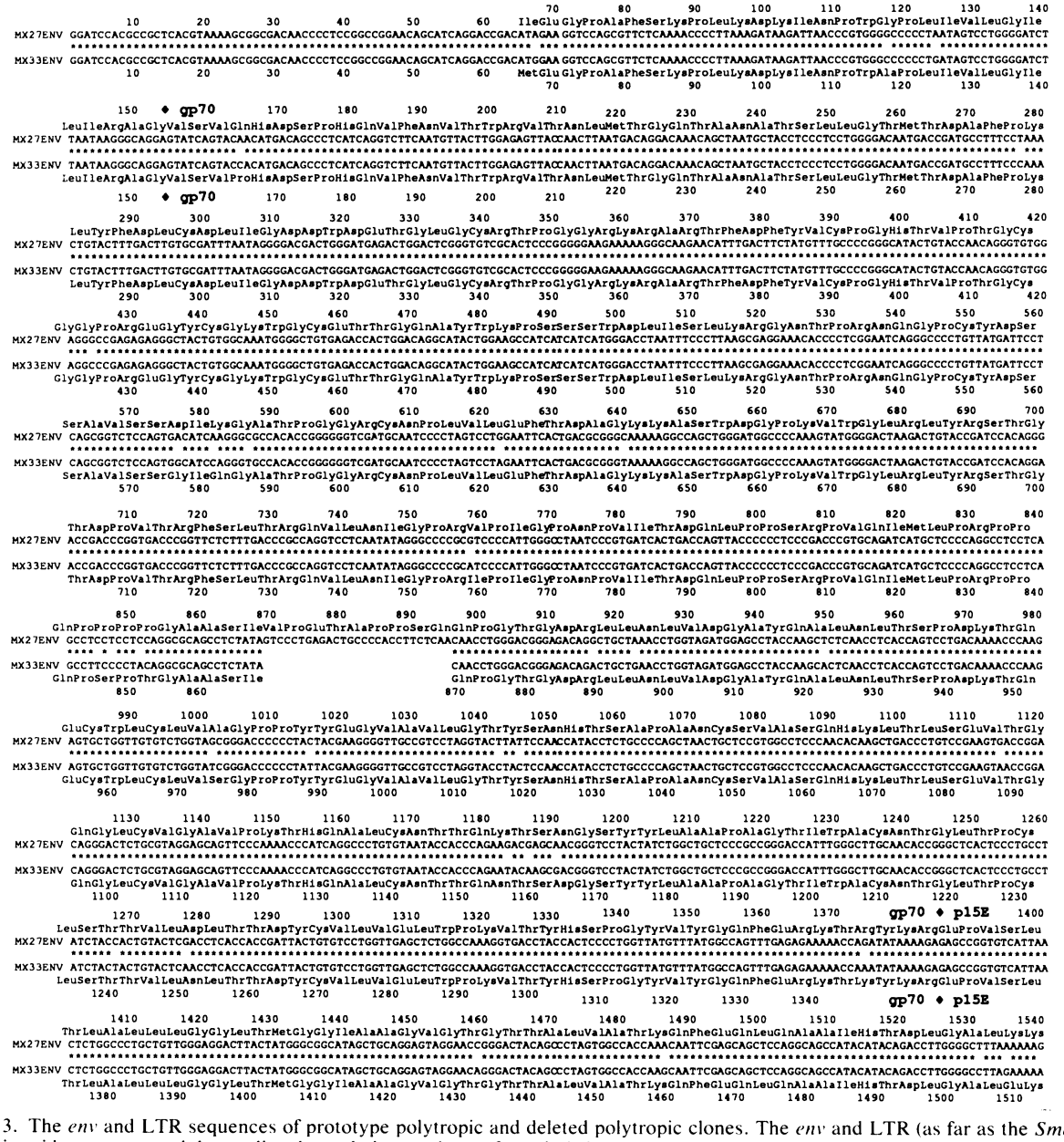

FIG. 3. The *env* and LTR sequences of prototype polytropic and deleted polytropic clones. The *env* and LTR (as far as the *Sma*I site in R) nucleic acid sequences and the predicted translation products of proviral clones MX27 and MX33 are shown. Cleavage sites of the *env* gene product (◆) and the boundaries of U_3 (⇒ and ⇐) are indicated. Bases 2252 through 2442 (MX27) and 2269 through 2459 (MX33) correspond to the 190-bp insert not present in ecotropic and xenotropic viruses (29, 45).

structures with two LTRs; as well as 11 5′ and 10 3′ halves of proviruses which contain *Hin*dIII sites near the end of the *pol* gene (data not shown). As expected, on the basis of experience with other endogenous proviruses, many of the clones contained deletions compared with replication-competent MLV; in particular, several appeared to have lost all or almost all of their *env* genes. Others contained deletions in *gag* or *pol* or both. In the course of the analyses described in this report, it became clear that the nonecotropic clones could be classified into three groups with the properties shown in Table 1.

The structures of the nonecotropic clones containing apparently intact *env* genes are shown in Fig. 1. They were initially classified into three groups based on several linked characteristics. First is the presence of a *Hin*dIII site near the 3′ end of *pol*. Second is the variation in size of the LTR as judged by digestion with two restriction enzymes (*Pst*I and *Kpn*I) that cut at conserved sites in U_3 and R (data not shown). Three of the complete proviruses have short LTRs (300- to 380-bp *Pst*I-*Kpn*I fragments) comparable in size to replication-competent ecotropic and xenotropic viruses (60), whereas the remaining clones contain LTRs which are some 150 to 190 bp longer and are similar in size to those reported for a number of other endogenous clones (29, 45). Third, the three proviruses with short LTRs also lack the *Bam*HI site at the start of the *env* gene which is present in the other clones but absent in a number of xenotropic viruses (7). We thus divide the nonecotropic clones into three classes, one of

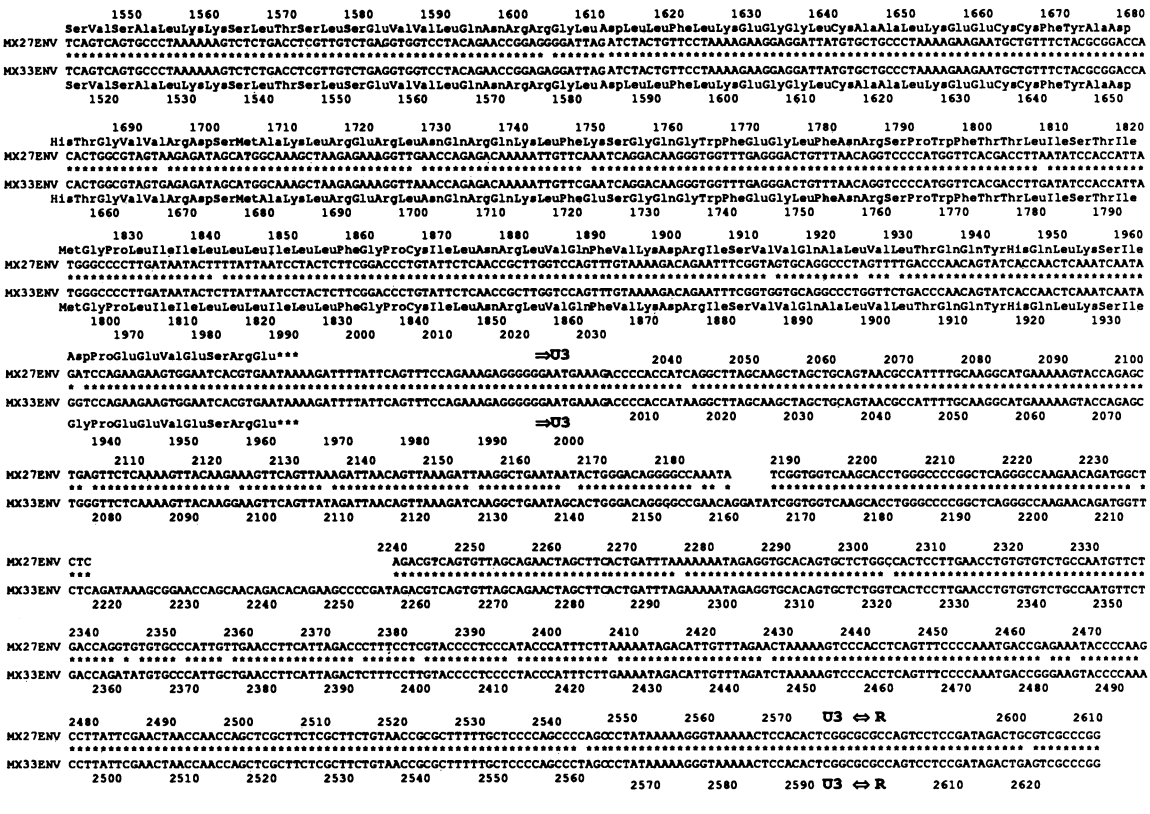

FIG. 3—*Continued.*

which resembles xenotropic viruses in *env* and in the LTR (MX22, -29, and -30) and two which differ only by the presence (MX11, -12, -13, -17, and -33) or absence (MX21 and -27) of a *Hin*dIII site in the *pol* gene.

We wished to examine more closely the relationship among these proviruses and to compare them with known viruses without having to resort to sequencing each clone. We therefore devised an in vitro scheme for obtaining RNA fingerprints from the proviral clones. This approach takes advantage of plasmids which contain promoters for the RNA polymerases of phages T7 and SP6. Viral DNA was subcloned into one of these plasmids, and RNA corresponding to the viral plus strand was synthesized by using the appropriate RNA polymerase. RNA homologous to *env* and LTR sequences from the prototype polytropic virus MCF-247 was selected and fingerprinted as described in Materials and Methods. A fingerprint of MCF-247 obtained in this way is illustrated in Fig. 2A. It showed all the oligonucleotides predicted from its sequence (24, 27).

The fingerprints obtained from representative members of the three different classes of nonecotropic virus are also shown in Fig. 2. Although some oligonucleotides are shared by all, e.g., 101 and 104, clear-cut differences among the three classes are visible. MX11, -12, -13, and -17 gave fingerprints virtually indistinguishable from that of MX33, while MX21 was very similar to MX27, but the lack of six oligonucleotides suggested a small deletion within the *env* gene. MX29 and -30 yielded fingerprints similar to that of MX22 (data not shown). These findings lend further weight to the assignment of the different clones to three different classes of nonecotropic sequences.

Some of the oligonucleotides shown in Fig. 2 are identified by number; these have previously been identified in xenotropic and polytropic viruses. The class of provirus represented by MX27 (Fig. 2C) most closely resembles polytropic virus, as is indicated by the presence of oligonucleotides 102, 104, 111, 113, and 114, all of which have been identified within the *env* genes of various AKR ecotropic virus (AKV)-derived recombinant viruses (35). Thus, one or more members of this class of endogenous sequence contributes gp70 coding sequences to the leukemogenic viruses isolated from AKR and HRS/J mice.

The class of viruses typified by MX33 (Fig. 2D), although clearly related to that of MX27, does not encode oligonucleotide 102, a marker for all AKV-derived recombinant viruses, and consequently these proviruses cannot be *env* donors for these viruses. However, at least some recombinant viruses lack oligonucleotide 102; thus, the MX33 class might represent the *env* parent for these viruses (see Discussion). MX22 shares a number of oligonucleotides uniquely found previously in xenotropic viruses or predicted from their sequences (19, 43, 59), confirming that this class of sequences encodes xenotropiclike viruses. A provirus of this class most likely contributed LTR sequences to the polytropic recombinant virus (22, 47, 58; see Discussion). Although MX22 contains three oligonucleotides (101, 108, and 119) characteristic of the LTR region of HRS-derived polytropic virus genomes (59), neither its sequence nor those of MX29 and -30 are consistent with a role as donor for these sequences (data not shown).

DNA sequences of MX27 and -33. The fingerprint analyses showed that MX27 and -33 resembled the *env* parents predicted for recombinant viruses. To test this further, we determined the nucleotide sequences of a region including all of *env* and most of the LTR (Fig. 3). Although MX27 and -33 represent distinct classes of endogenous proviruses, they are

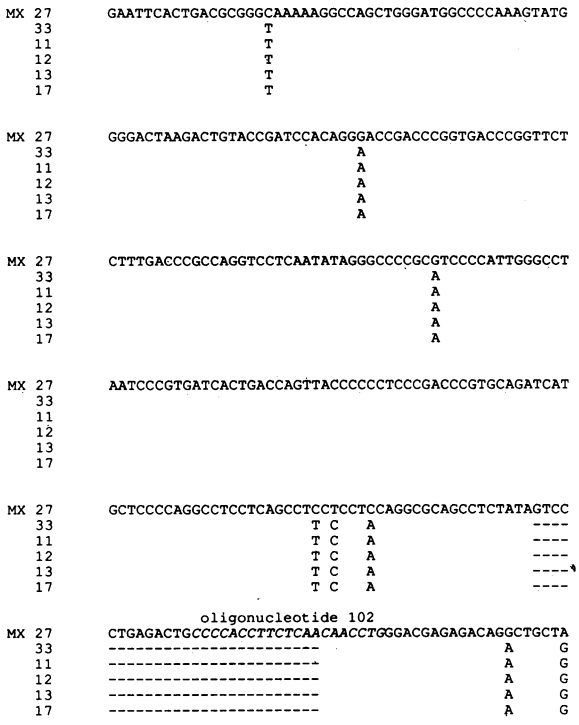

FIG. 4. *env* sequences of five mPT clones. The sequence differences between the PT clone, MX27, and five mPT clones (MX33, -11, -12, -13, and -17) are shown. The sequence shown corresponds to bases 623 through 900 of MX33 (Fig. 3).

closely related, differing by only 44 scattered base changes within *env* as well as by a 27-bp deletion in MX33 compared with MX27. These differences result in 18 amino acid differences out of 641 in the predicted *env* products as well as a nine-amino-acid deletion. These seemed no reason why MX33 should not encode a functional *env* gene; MX27, however, differs by a single base in the presumed initiator $tRNA^{Met}$ codon and presumably is translationally inactive. It was otherwise identical in sequence to the *env* gene parent predicted from the sequence of AKV-derived recombinant viruses.

That MX33 was representative of a class of proviruses was confirmed by sequencing the central region of *env* from four other similar clones (MX11, -12, -13, and -17). All clones containing the *Hin*dIII site near the 3′ end of *pol* had identical sequences in this region (bases 623 through 900 of MX33; Fig. 4). The 27-bp deletion overlaps the sequence of oligonucleotide 102 and thus accounts for its loss. Neither this deletion nor any of the 18 amino acid changes adversely affects the function of the *env* gene, since we have recently reconstructed infectious polytropic virus by substituting the gp70 gene of MCF-247 with that of MX11, -17, and -33 (unpublished observations), indicating that the *env* genes are functional.

The *env* genes of MX27 and -33 closely resemble those found in recombinant viruses, indicating that proviruses of the classes represented by these clones are the *env* gene donors (see Discussion). However, the LTR sequences of both MX27 and MX33 themselves differ somewhat from those found in recombinant viruses (27), indicating that these proviruses do not contribute this region to the leukemogenic recombinants. The LTRs are similar to one another and closely related to those reported for a variety of other endogenous sequences in that they contain an inserted sequence of 190 bp (29, 45).

DISCUSSION

Endogenous type C proviruses are associated with a number of important biological phenomena in mice. Despite this, they are poorly understood in terms of number, type, and distribution. To address these issues we have set out to obtain DNA clones representing all the proviruses present in a mouse genome. Although our collection of clones to date includes only about a third of the HRS/J proviruses as estimated by using class-specific oligonucleotide probes (see below), we have used it to identify all three classes of nonecotropic endogenous virus present in the mouse genome. These classes are most easily identified by the nature of their *env* genes, although they differ in other linked characteristics as well. These differences are summarized in Table 1. They include one class of provirus with an apparently xenotropic host range and two classes with a polytropic host range.

We believe that these viruses represent the whole spectrum of intact endogenous nonecotropic viruses of HRS/J mice for the following reasons. First, we have not isolated any viral clones which do not fall into one of these classes nor are we aware of any endogenous or recombinant type C MLVs whose *env* genes would represent a distinct class. Although we estimate that our collection of clones is missing most of the complete repertoire there is no reason to believe that we have systematically failed to clone any particular virus class. Second, we have recently prepared oligonucleotide probes which differentiate the classes (Stoye and Coffin, in preparation). With these probes we can show that HRS/J mice contain 5 xenotropic, 11 modified polytropic, and 18 polytropic endogenous proviruses, indicating that the modified polytropic sequences constitute a significant fraction of the endogenous MLV-related proviruses. Taken together, these probes react with all or almost all of the endogenous proviruses detectable with a broadly reactive *env* probe (data not shown). If there are any additional classes they must be present in very low numbers. We have however cloned four proviral fragments containing the internal *Hin*dIII site characteristic of the modified polytropic class which have lost their entire *env* genes (data not shown) and which presumably correspond to the deleted proviruses present in 129/J mice which were detected by cDNA cloning of in vivo-expressed viral RNA (32).

Distinguishing the two classes of endogenous polytropic sequences was greatly facilitated by the development of a technique for RNA fingerprinting of DNA clones. This technique should be useful for examining other closely related families of genes, since it is both rapid and much more sensitive than restriction enzyme analysis. For instance, the *env*-LTR region of MCF-247 yields approximately 40 identifiable oligonucleotides (Fig. 2A). Assuming an average size of 12 nucleotides, this corresponds to 480 nucleotides or some 25% of the genome. A similarly detailed analysis using restriction enzymes would require the use of 80 restriction enzymes with a 6-bp recognition sequence. A minor modification of this approach allows ready correlation between specific members of a gene family and expressed RNAs, for example in virions or suitably selected from virus-producing cells.

Recombinational origins of polytropic viruses. The generation of recombinant polytropic viruses remains a fascinating

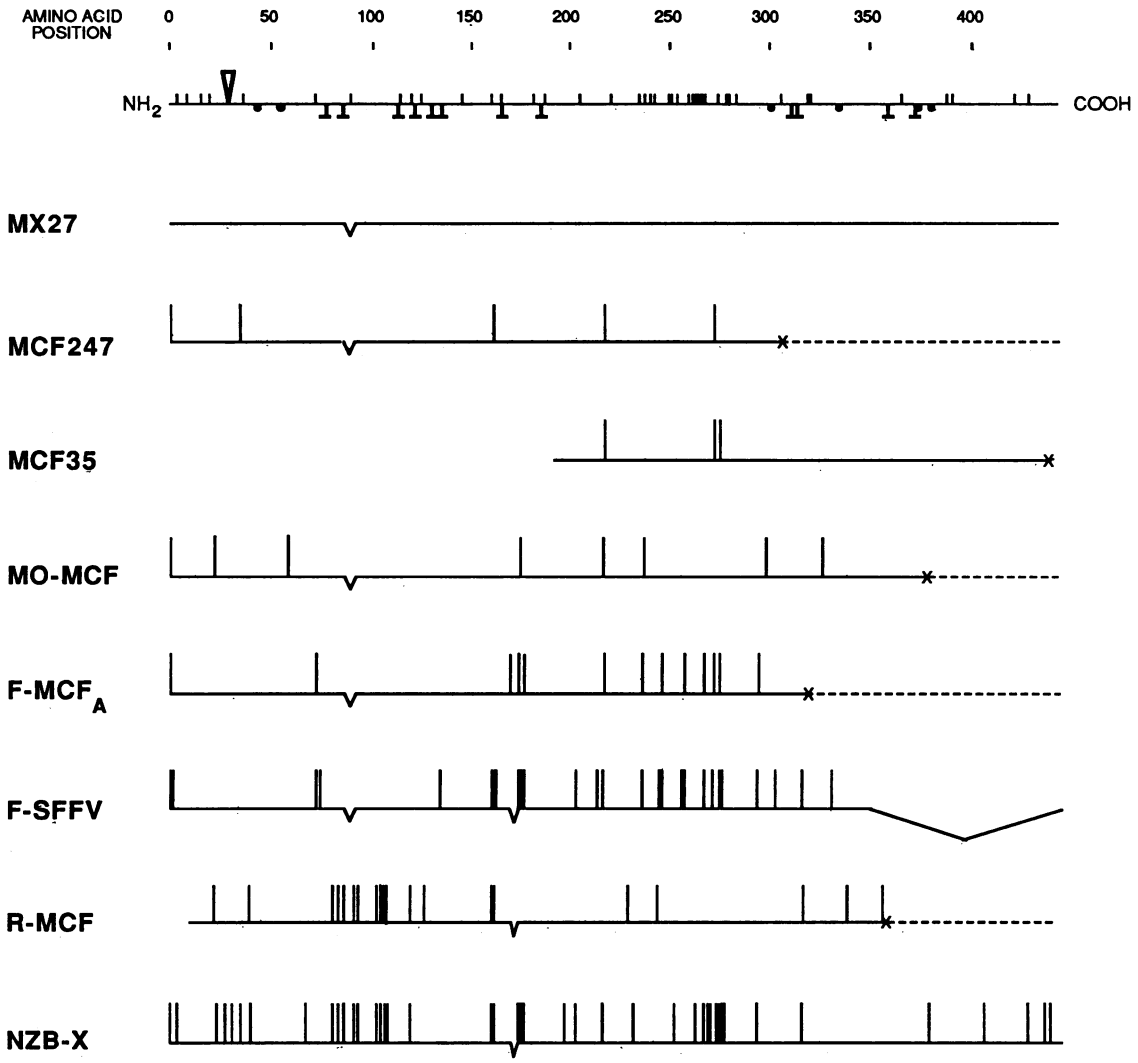

FIG. 5. gp70 of MX27, xenotropic virus, and six recombinant viruses. The predicted gp70 amino acid sequences of MCF-247 (24), MCF35 (48), Moloney MCF (Mo-MCF) (4), Friend MCF (F-MCF$_A$) (1), (R-MCF) (61), (F-SFFV) (9), and NZB-xenotropic (NZB-X) (43) are compared with that of MX27. At the top is a schematic representation of the topography of the molecules, showing (|), proline residues; (⊥), cysteine residues; (●), glycosylation sites; (▽), leader sequence cleavage site; (×), 3′ recombination point; (----), presence of the corresponding ecotropic sequences. Amino acid differences from MX27 (|) and deletions (V) are shown. For the sake of simplicity the actual amino acids are not shown. The complete comparisons are available by request.

puzzle replete with basic issues of virus expression and interaction. We have cloned and sequenced representative members of two families of endogenous proviral genes. Do they play a role in the generation of recombinant viruses? To answer this question, we have compared the deduced amino acid sequences of the *env* genes of MX27 and MX33 with those of a number of sequenced recombinant viruses (Fig. 5 and 6). Unfortunately, no sequence data is available for HRS/J-derived polytropic viruses. However, RNA fingerprints of polytropic viruses derived from HRS/J and AKR mice are virtually identical (19, 58, 59), suggesting that the *env* donors in the two strains are likely to be very similar. The nonecotropic *env* sequences of recombinants derived from either AKV or Moloney MLV (Mo-MLV) are virtually identical to that of MX27 (Fig. 5). Furthermore, MX27 encodes almost all the T_1 oligonucleotides present in a variety of AKV-derived polytropic viruses (35), including 102 which is not encoded by MX33. Thus we can conclude that a provirus of the MX27 type contributed *env* sequences to the AKV- and Mo-MLV-derived polytropic viruses, with the scattered amino acid differences being caused by mutations in the course of viral replication. However, MX27 appears not to be a candidate *env* gene donor since it is presumably translationally inactive due to the G→A transition in the initiator codon. It is of interest to note that an endogenous provirus cloned from RFM/Un mice, which is distinguishable from MX27 by flanking restriction enzyme sites, contains the same mutation (41), suggesting that these two clones may represent a subfamily of polytropic sequences.

The modified polytropic proviruses can serve as *env* donors in some recombinants. Two viruses derived by recombination with the Friend strain of MLV have been sequenced and appear to have acquired *env* genes from different proviral classes. One isolate (1) shows virtual identity of MX27 (Fig. 5), whereas a second (31) differs by

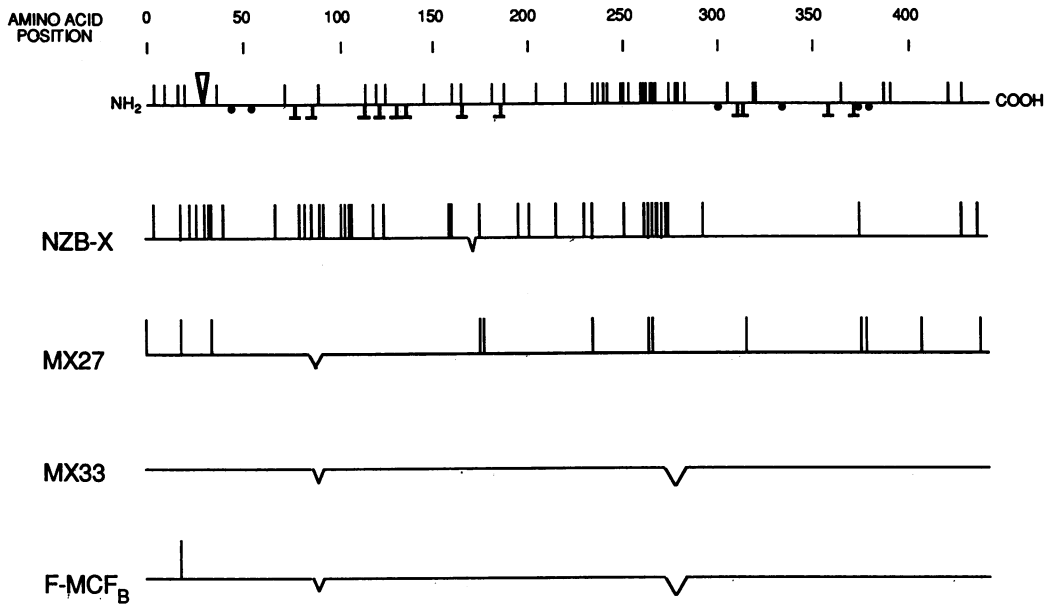

FIG. 6. Comparison of MX33 gp70 predicted sequence with xenotropic, MX27, and Friend MCF sequences. The differences between MX33 and the other viruses are shown (|). Other symbols are defined in the legend to Fig. 5. The Friend MCF (F-MCF$_B$) sequence shown here is from Koch et al. (31).

only one amino acid from MX33 (Fig. 6). It thus appears that different ecotropic viruses can recombine with different endogenous sequences. It is possible that viral sequences themselves directly determine whether recombination can take place. Alternatively, recombination may depend on the tissue tropism of the specific ecotropic viruses since the different LTRs of Friend and Mo-MLV apparently confer distinct specificity for replication in spleen and thymus (6). Endogenous sequences expressed in a differentiation-specific fashion would thus provide distinct opportunities for recombination to occur. This would have important consequences for the generation of AKV-derived polytropic viruses since the recombinational event altering the U_3 region of the LTR also alters the tissue tropism of the virus (23).

Two other viruses, Friend spleen focus-forming virus (F-SFFV) and Rauscher MCF (R-MCF), seemed to have been formed in more complex fashion. Inspection of Fig. 5 shows that the first 150 amino acids of R-MCF are derived from a xenotropic virus, while the remaining 160 nonecotropic amino acids are from an endogenous polytropic virus. Thus, at least two recombinational events were necessary to generate this recombinant gp70. Similarly, amino acids 150 through 200 of F-SFFV may be of xenotropic origin, whereas the remaining nonecotropic sequences come from a polytropic virus. Both these viruses were originally isolated after multiple virus passages through mice (57), a procedure which presumably allowed multiple recombinational events to take place.

Host range determination in murine retroviruses. The viral surface glycoprotein specifically recognizes and binds cell surface receptors. Analysis of superinfection resistance has shown that ecotropic, xenotropic, and polytropic viruses recognize different receptors (49). This interaction is a key event in the replication cycle of any virus since the ability of the *env* gene product to recognize specific receptors determines the host range of a particular virus. It was originally suggested that the polytropic host range of the MCF viruses is created de novo after the recombination event between an ecotropic and a xenotropic virus (21). This is clearly not the case; rather, the host range is an intrinsic property of the *env* gene of the nonecotropic donor. We have not been able to demonstrate this directly since we lack an infectious clone of a polytropic provirus of the MX27 class. However, the demonstration that some polytropic recombinant viruses have an entire gp70 coding region which is nonecotropic in origin (38, 48) strongly implies that the MX27-type provirus has a polytropic host range.

The sequence comparisons in Fig. 5 and 6 show that the *env* genes of polytropic and xenotropic viruses are closely related but carry a number of characteristic differences localized within small regions of the coding sequence. At least some of these differences must be responsible for the differences in host range. The most obvious are the 4-amino-acid deletion (amino acids 49 to 53 of the mature gp70) in the polytropic viruses relative to xenotropic viruses and a variable region at amino acids 263 to 281 (in the region encompassing the MX33 deletion). Host range studies of two different MCF viruses lend support to the notion that these regions are important in determining host range. A role for the former difference is suggested by the finding that R-MCF, which has xenotropic sequences through the first 150 amino acids of gp70 (see above) appears to be xenotropic in host range (61, 62). Our preliminary studies of modified polytropic virus, constructed by replacing the gp70 coding sequences of MCF-247 with those of MX33 indicate that while it is polytropic in host range it replicates poorly on mink cells compared with MCF-247 (Stoye and Coffin, in preparation). This observation provides indirect evidence that the variable region centered around amino acid 270 may also be important for gp70-receptor interactions. An analogous situation exists for the gp85 molecules of the different subgroups of avian leukosis viruses in which we have recently shown that two variable regions determine the avian leukosis virus host range (16, 17).

Evolutionary relationships of murine endogenous viruses. To examine the relationships between the different proviruses present in the mouse germ line we compared the predicted gp70 amino acid sequences of a polytropic, a

TABLE 2. Quantitative *env* gene comparisons[a]

	No. of amino acid differences	No. of insertions or deletions	% Homology
Entire gp70			
Polytropic	0	0	100
Modified polytropic	13	1	96
Xenotropic	39	2	90
Ecotropic	148	13	58
Endogenous feline	186	5	54
N terminus			
Polytropic	0	0	100
Modified polytropic	8	1	96
Xenotropic	33	2	86
Ecotropic	134	13	40
Endogenous feline	134	5	48
C terminus			
Polytropic	0	0	100
Modified polytropic	5	0	97
Xenotropic	6	0	96
Ecotropic	14	0	91
Endogenous feline	52	0	66

[a] The predicted amino acid sequences of a modified polytropic virus (MX33; Fig. 3), a xenotropic virus (NZB-xenotropic virus; 43), an ecotropic virus (AKV; 20), and an endogenous feline virus (enFeLV; J. Mullins, personal communication) were aligned for maximum homology with a polytropic provirus (MX27; Fig. 3). The number of amino acid differences, the number of insertions and deletions necessary to generate optimum alignment, and the percent homology, calculated with an arbitrary gap penalty of 3 are shown for the entire gp70 molecule including the leader sequence (440 amino acids in MX27), the N-terminal two-thirds of gp70 (286 amino acids), and the C-terminal third of gp70 (154 amino acids).

modified polytropic, a xenotropic, and an ecotropic virus (Table 2, Fig. 7). Not suprisingly, in light of their similar host range, the polytropic and modified polytropic proviruses were most closely related to one another, showing a 96% overall amino acid similarity. Hybridization analysis with class-specific oligonucleotide probes showed that both polytropic and modified polytropic proviruses are present in a wide variety of inbred mouse strains (Stoye and Coffin, in preparation), indicating divergence of these two classes of provirus before the generation of inbred strains. The xenotropic virus was also closely related (90% homology to the polytropic sequences) but showed slightly greater divergence that did the modified polytropic from the polytropic class. By contrast, the ecotropic virus was not closely related, with only 58% overall homology, a figure very similar to that of the endogenous feline leukemia provirus.

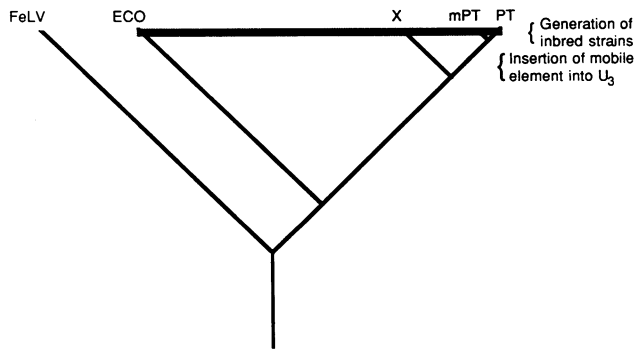

FIG. 7. Evolutionary tree of viral gp70s. The evolutionary relationship between the envelope glycoproteins of various MLVs and feline leukemia viruses (FeLV) is illustrated by plotting the amino acid sequence divergence by using the differences listed in Table 2.

Surprisingly, the relationship between ecotropic and polytropic viruses is very different if we compare separately the N-terminal and C-terminal regions. The N-terminal two-thirds of ecotropic *env* appears more distantly related to polytropic sequences than does feline leukemia virus, whereas the C-terminal third is almost as closely related as are the modified polytropic and xenotropic viruses. In light of this relationship it seems reasonable to speculate that evolution of the ecotropic virus endogenous to inbred strains of mice might have involved a recombination between a nonecotropic virus of the type present in the mouse genome with another virus which has not been fixed in laboratory strains of mice. Thus, the recombinations which occur in highly leukemic strains of mice might in fact be reciprocal to an event that occurred many years ago. In this context it is of interest to note that the minimum polytropic *env* gene substitution in MCF viruses encompasses the whole of the N-terminal region of gp70 (35; see also Fig. 5) and that this region is thought to comprise one functional domain of the gp70 molecule (36, 46).

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant CA 24530 from the National Cancer Institute. J.P.S. was supported by fellowships from the European Molecular Biology Organisation and the American Cancer Society, Massachusetts Division, Inc.

We thank Mary Bostic-Fitzgerald and Susan E. Morrison for assistance in the preparation of this manuscript and numerous colleagues at Tufts for helpful discussions.

LITERATURE CITED

1. **Adachi, A., K. Sakai, N. Kitamura, S. Nakanishi, O. Niwa, M. Matsuyama, and A. Ishimoto.** 1984. Characterization of the *env* gene and long terminal repeat of molecularly cloned Friend mink cell focus-inducing virus DNA. J. Virol. **50**:813–821.
2. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques *in situ*. Science **196**:180–182.
3. **Biggin, M. D., T. J. Gibson, and G. F. Hong.** 1983. Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA **80**:3963–3965.
4. **Bosselman, R. A., F. van Straaten, C. Van Beveren, I. M. Verma, and M. Vogt.** 1982. Analysis of the *env* gene of molecularly cloned and biologically active Moloney mink cell focus-forming proviral DNA. J. Virol. **44**:19–31.
5. **Brownlee, G. G.** 1972. Determination of sequences in RNA. Elsevier/North-Holland Publishing Co., Amsterdam.
6. **Chatis, P. A., C. A. Holland, J. E. Silver, T. N. Fredrickson, N. Hopkins, and J. W. Hartley.** 1984. A 3′ end fragment encompassing the transcriptional enhancers of nondefective Friend virus confers erythroleukemogenicity on Moloney leukemia virus. J. Virol. **52**:248–254.
7. **Chattopadhyay, S. K., M. W. Cloyd, D. R. 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.
8. **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.
9. **Clark, S. P. and T. W. Mak.** 1983. Complete nucleotide sequence of an infectious clone of Friend spleen focus-forming provirus: gp55 is an envelope fusion glycoprotein. Proc. Natl. Acad. Sci. USA **80**:5037–5041.
10. **Coffin, J. M.** 1982. Endogenous retroviruses, p. 1109–1203. *In* R. Weiss, N. Teich, H. E. Varmus, and J. M. Coffin (ed.), Molecular biology of tumor viruses. Part III. RNA tumor viruses, 2nd ed., vol. 1. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

11. **Coffin, J. M., M. Champion, and F. Chabot.** 1978. Nucleotide sequence relationships between the genomes of an endogenous and an exogenous avian tumor virus. J. Virol. **28**:972–991.
12. **Copeland, N. G., H. G. Bedigian, C. Y. Thomas, and N. A. Jenkins.** 1984. DNAs of two molecularly cloned endogenous ecotropic proviruses are poorly infectious in DNA transfection assays. J. Virol. **49**:437–444.
13. **Copeland, N. G., K. W. Hutchinson, and N. A. Jenkins.** 1983. Excision of the DBA ecotropic provirus in dilute coat-color revertants of mice occurs by homologous recombination involving the viral LTRs. Cell **33**:379–387.
14. **Dale, R. M. K., B. A. McClure, and J. P. Houchins.** 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application for sequencing corn mitochondrial 18S rDNA. Plasmid **13**:31–40.
15. **Dolberg, D. S., L. T. Bachelor, and H. Fan.** 1981. Endogenous type C retroviral sequences of mice are organized in a small number of virus-like classes and have been acquired recently. J. Virol. **40**:96–106.
16. **Dorner, A. J., and J. M. Coffin.** 1986. Determinants of receptor interaction and cell killing on the avian retrovirus glycoprotein gp85. Cell **45**:365–374.
17. **Dorner, A. J., J. P. Stoye, and J. M. Coffin.** 1985. Molecular basis of host range variation in avian retroviruses. J. Virol. **53**:32–39.
18. **Etzerodt, M., T. Mikkelver, F. S. Pedersen, N. D. Kjeldergaard, and P. Jorgensen.** 1984. The nucleotide sequence of the AKV murine leukemia virus genome. Virology **134**:196–207.
19. **Green, N., H. Hiai, J. H. Elder, R. S. Schwartz, R. H. Khiroya, C. Y. Thomas, P. N. Tsichlis, and J. M. Coffin.** 1980. Expression of leukemogenic recombinant viruses associated with a recessive gene in HRS/J mice. J. Exp. Med. **152**:249–264.
20. **Herr, W.** 1984. Nucleotide sequence of AKV murine leukemia virus. J. Virol. **49**:471–478.
21. **Hartley, J. W., N. K. Wolford, L. J. Old, and W. P. Rowe.** 1977. A new class of murine leukemia virus associated with the development of spontaneous lymphomas. Proc. Natl. Acad. Sci. USA **74**:789–792.
22. **Hoggan, M. D., R. R. O'Neill, and C. A. Kozak.** 1986. Nonecotropic murine leukemia viruses in BALB/c and NFS/N mice: characterization of the BALB/c *Bxv-1* provirus and the single NFS endogenous xenotrope. J. Virol. **60**:980–986.
23. **Holland, C. A., J. W. Hartley, W. P. Rowe, and N. Hopkins.** 1985. At least four viral genes contribute to the leukemogenicity of murine retrovirus MCF 247 in AKR mice. J. Virol. **53**:158–165.
24. **Holland, C. A., J. Wozney, and N. Hopkins.** 1983. Nucleotide sequence of the gp70 gene of murine retrovirus MCF 247. J. Virol. **47**:413–420.
25. **Horowitz, J. M., and R. Risser.** 1985. Molecular and biological characterization of the endogenous ecotropic provirus of BALB/c mice. J. Virol. **56**:798–806.
26. **Jenkins, N. A., N. G. Copeland, B. A. Taylor, and B. K. Lee.** 1982. Organization, distribution, and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of *Mus musculus*. J. Virol. **43**:26–36.
27. **Kelly, M., C. A. Holland, M. L. Lung, S. K. Chattopadhyay, D. R. Lowy, and N. H. Hopkins.** 1983. Nucleotide sequence of the 3′ end of MCF 247 murine leukemia virus. J. Virol. **45**:291–298.
28. **Khan, A. S.** 1984. Nucleotide sequence analysis establishes the role of endogenous murine leukemia virus DNA segments in formation of recombinant mink cell focus-forming murine leukemia viruses. J. Virol. **50**:864–871.
29. **Khan, A. S., and M. A. Martin.** 1983. Endogenous murine leukemia proviral long terminal repeats contain a unique 190-base-pair insert. Proc. Natl. Acad. Sci. USA **80**:2699–2703.
30. **Khan, A. S., W. P. Rowe, and M. A. Martin.** 1982. Cloning of endogenous murine leukemia virus-related sequences from chromosomal DNA of BALB/c and AKR/J mice: identification of an *env* progenitor of AKR-247 mink cell focus-forming proviral DNA. J. Virol. **44**:625–636.
31. **Koch, W., W. Zimmerman, A. Oliff, and R. Friedrich.** 1984. Molecular analysis of the envelope gene ang long terminal repeat of Friend mink cell focus-inducing virus: implications for the functions of these sequences. J. Virol. **49**:828–840.
32. **Levy, D. E., R. A. Lerner, and M. C. Wilson.** 1985. Normal expression of polymorphic endogenous retroviral RNA containing segments identical to mink cell focus-forming virus. J. Virol. **56**:691–700.
33. **Loenen, W. A. M., and F. R. Blattner.** 1983. Lambda charon vectors (Ch 32, 33, 34, and 35) adapted for DNA cloning in recombination-deficient hosts. Gene **26**:171–179.
34. **Lowy, D. R., E. Rands, S. K. Chattopadhyay, C. F. Garon, and G. L. Hager.** 1980. Molecular cloning of infectious murine leukemia virus DNA from infected mouse cells. Proc. Natl. Acad. Sci. USA **77**:614–618.
35. **Lung, M. L., J. W. Hartley, W. P. Rowe, and N. H. Hopkins.** 1983. Large RNase T_1-resistant oligonucleotides encoding p15E and the U3 region of the long terminal repeat distinguish two biological classes of mink cell focus-forming type C viruses of inbred mice. J. Virol. **45**:275–290.
36. **Machida, C. A., R. K. Bestwick, B. A. Boswell, and D. Kabat.** 1985. Role of a membrane glycoprotein in Friend virus-induced erythroleukemia: studies of mutant and revertant viruses. Virology **144**:158–172.
37. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
38. **Mark, G. E., and U. R. Rapp.** 1984. Envelope gene sequence of two in vitro-generated mink cell focus-forming leukemia viruses which contain the entire gp70 of the endogenous nonecotropic parent. J. Virol. **49**:530–539.
39. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
40. **Messing, J. and J. Vieira.** 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene **19**:269–276.
41. **Nikbakht, K. N., L. R. Boone, P. L. Glover, F. E. Myer, and W. K. Yang.** 1987. Characterization of a molecular clone of RFM/Un mouse chromosomal DNA that contains a full-length endogenous murine leukemia virus-related proviral genome. J. Gen. Virol. **68**:683–693.
42. **Nikbakht, K. N., C.-Y. Ou, L. R. Boone, P. L. Glover, and W. K. Yang.** 1985. Nucleotide sequence analysis of endogenous murine leukemia virus-related proviral clones reveals primer-binding sites of glutamine tRNA. J. Virol. **54**:889–893.
43. **O'Neill, R. R., C. E. Buckler, T. S. Theodore, M. A. Martin, and R. Repaske.** 1985. Envelope and long terminal sequences of a cloned infectious NZB xenotropic murine leukemia virus. J. Virol. **53**:100–106.
44. **O'Neill, R. R., A. S. Khan, M. D. Hoggan, J. W. Hartley, M. A. Martin, and R. Repaske.** 1986. Specific hybridization probes demonstrate fewer xenotropic than mink cell focus-forming murine leukemia virus *env*-related sequences in DNAs from inbred laboratory mice. J. Virol. **58**:359–366.
45. **Ou, C.-Y., L. R. Boone, and W. K. Yang.** 1983. A novel sequence element and the nucleotide structural features in the long terminal repeat of a BALB/c mouse genomic leukemia virus-related DNA clone. Nucleic Acids Res. **11**:5603–5620.
46. **Pinter, A., and W. J. Honnen.** 1984. Characterization of structural and immunological properties of specific domains of Friend ecotropic and dual-tropic murine leukemia virus gp70s. J. Virol. **49**:452–458.
47. **Quint, W., W. Boelens, P. van Wezenbeek, T. Cuypers, E. R. Maandag, G. Selten, and A. Berns.** 1984. Generation of AKR mink cell focus-forming viruses: a conserved single-copy xenotrope-like provirus provides recombinant long terminal repeat sequences. J. Virol. **50**:432–438.
48. **Quint, W., W. Boelens, P. van Wezenbeck, E. R. Maandag, and A. Berns.** 1984. Generation of AKR mink cell focus-forming virus: nucleotide sequence of the 3′ end of somatically acquired

AKR-MCF. Virology **136**:425–434.
49. **Rein, A.** 1982. Interference grouping of murine leukemia viruses: a distinct receptor for MCF-recombinant viruses in mouse cells. Virology **120**:251–257.
50. **Rimm, D. L., D. Horness, J. Kucera, and F. R. Blattner.** 1980. Construction of coliphage lambda Charon vectors with BamHI cloning sites. Gene **12**:301–309.
51. **Risser, R., J. M. Horowitz, and J. McCubrey.** 1983. Endogenous mouse leukemia viruses. Annu. Rev. Genet. **17**:85–121.
52. **Roblin, R., J. M. Young, R. J. Mural, T. E. Bell, and J. N. Ihle.** 1982. Molecular cloning and characterization of murine leukemia virus-related sequences from C3H/HeN mouse DNA. J. Virol. **43**:113–126.
53. **Rommelaere, J., D. V. Faller, and N. Hopkins.** 1978. Characterization and mapping of RNase T_1-resistant oligonucleotides derived from the genomes of AKV and MCF murine leukemia viruses. Proc. Natl. Acad. Sci. USA **75**:67–71.
54. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
55. **Steffan, D. L., R. Mural, D. Cowing, J. Mielcarz, J. Young, and R. Roblin.** 1982. Most of the murine leukemia virus sequences in the DNA of NIH/Swiss mice consist of two closely related proviruses, each repeated several times. J. Virol. **43**:127–135.
56. **Stoye, J. P., and J. M. Coffin.** 1985. Endogenous retroviruses, p. 357–404. *In* R. Weiss, N. Teich, H. E. Varmus, and J. M. Coffin (ed.), Molecular biology of tumor viruses. Part III. RNA tumor viruses, 2nd ed., vol. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
57. **Teich, N., J. Wyke, T. Mak, A. Bernstein, and W. Hardy.** 1982. Pathogenesis of retrovirus-induced disease, p. 785–998. *In* Molecular biology of tumor viruses. Part III. RNA tumor viruses, 2nd ed., vol. 1. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
58. **Thomas, C. Y., and J. M. Coffin.** 1982. Genetic alterations of RNA leukemia viruses associated with the development of spontaneous thymic leukemia in AKR/J mice. J. Virol. **43**:416–426.
59. **Thomas, C. Y., R. Khiroya, R. S. Schwartz, and J. M. Coffin.** 1984. Role of recombinant ecotropic and polytropic viruses in the development of spontaneous thymic lymphomas in HRS/J mice. J. Virol. **50**:397–407.
60. **Van Beveren, C., J. M. Coffin, and S. Hughes.** 1982. Appendixes, p. 559–1221. *In* Molecular biology of RNA tumor viruses. RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
61. **Vogt, M., C. Haggblom, S. Swift, and M. Haas.** 1985. Envelope gene and long terminal repeat determine the different biological properties of Rauscher, Friend, and Moloney mink cell focus-inducing viruses. J. Virol. **55**:184–192.
62. **Vogt, M., C., Haggblom, S. Swift, and M. Haas.** 1986. Specific sequences of the *env* gene determine the host range of two XC-negative viruses of the Rauscher virus complex. Virology **154**:420–424.