JOURNAL OF VIROLOGY, Apr. 1990, p. 1675–1682
0022-538X/90/041675-08$02.00/0

Vol. 64, No. 4

Structural Diversity and Nuclear Protein Binding Sites in the Long Terminal Repeats of Feline Leukemia Virus

RUTH FULTON,[1] MARK PLUMB,[1] LESLEY SHIELD,[2] AND JAMES C. NEIL[1,2]*

Beatson Institute for Cancer Research, Bearsden, Glasgow G61 1BD,[1] and Department of Veterinary Pathology, University of Glasgow, Bearsden, Glasgow G61 1QH,[2] Scotland

Received 26 July 1989/Accepted 18 December 1989

The long terminal repeat U3 sequences were determined for multiple feline leukemia virus proviruses isolated from naturally occurring T-cell tumors. Heterogeneity was evident, even among proviruses cloned from individual tumors. Proviruses with one, two, or three repeats of the long terminal repeat enhancer sequences coexisted in one tumor, while two proviruses with distinct direct repeats were found in another. The enhancer repeats are characteristic of retrovirus variants with accelerated leukemogenic potential and occur between −155 and −244 base pairs relative to the RNA cap site. The termini of the repeats occur at or near sequence features which have been recognized at other retrovirus recombinational junctions. In vitro footprint analysis of the feline leukemia virus enhancer revealed three major nuclear protein binding sites, located at consensus sequences for the simian virus 40 core enhancer, the nuclear factor 1 binding site, and an indirect repeat which is homologous to the PEA2 binding site in the polyomavirus enhancer. Only the simian virus 40 core enhancer sequence is present in all of the enhancer repeats. Cell type differences in binding activities to the three motifs may underlie the selective process which leads to outgrowth of viruses with specific sequence duplications.

Retrovirus long terminal repeats (LTRs), like the more extensively characterized DNA tumor virus enhancers, are modular structures encompassing binding sites for multiple cellular transcription factors (16, 34, 43). It is becoming clear that the modular enhancer elements interact with general, as well as tissue-specific, factors which presumably regulate the function of retrovirus LTRs in different tissues and thereby exert a controlling influence on their oncogenic activity (17, 27, 50). The identification of protein binding sites in the LTR and of the factors which interact with them will therefore be essential to understand LTR activity in normal and neoplastic tissues.

The LTRs of feline leukemia virus (FeLV) play an important role in leukemogenesis by subverting regulation of the expression of host cell genes. In cases in which FeLV has transduced the coding sequences of the c-*myc* gene, expression of this host gene appears to be exclusively from the viral LTR promoter and the normal c-*myc* alleles are transcriptionally silent (10). In other tumors of similar phenotype, FeLV proviruses are found integrated immediately upstream and in the opposite orientation to c-*myc*, where they appear to activate transcription from the normal c-*myc* promoters by enhancer insertion (10, 29, 31). Although a number of feline leukemia and feline sarcoma virus LTRs have been sequenced (9, 13, 14, 29, 40, 46) and a limited functional study has been done (51), no systematic study has been undertaken to discover whether LTR sequence changes are responsible for increased pathogenicity or tissue-specific oncogenic potential. We have begun a detailed structural and functional analysis of tumor-derived FeLV LTRs to examine these issues.

In the present study, we examined sequence variation in the LTRs of primary tumor isolates and nuclear factor binding sites within the FeLV LTR enhancer. We report that proviruses cloned directly from T-cell tumor tissue frequently have sequence duplications and clustered point mutations and that the direct repeats duplicate complete functional binding sites for a set of nuclear factors.

* Corresponding author.

MATERIALS AND METHODS

FeLV proviral clones. The origins of most of the proviral clones analyzed in this study have been published previously. The structures are shown in diagrammatic form in Fig. 1. Five clones from thymic lymphosarcoma T17 were analyzed. The structures of three (T17T-22, T17T-31, and T17T-34) have been published (11). Two other clones from the same tumor were studied in detail, one a recombinant containing a v-*myc* gene (T17M-1) and the other an apparently full-length FeLV provirus (T17H-1). The CT4 and CT8 clones from thymic lymphosarcoma 84793 have been described previously (32), while the pBam8 clone was kindly provided by J. Casey (Cornell University, Ithaca, N.Y.) (21).

Cell lines. The feline cell lines used to generate nuclear extracts were AH927, an immortal cell line of fibroblastic origin (39), and T3, a lymphoid tumor cell line established from a naturally occurring thymic lymphosarcoma (32). The latter cell line appears to be derived from a mature thymocyte and expresses functionally rearranged T-cell antigen receptor genes (35).

Sequencing strategy. Conserved *Pst*I and *Kpn*I sites within the FeLV LTRs were utilized for cloning and subsequent sequencing of the eight U3 regions. Following digestion with both enzymes, the DNA fragments were cloned into homologous restriction sites in bacteriophages m13mp18 and mp19 (33). Both strands were sequenced, using universal sequencing primers and synthetic primers based on the determined sequence by the Sanger dideoxy chain termination method (42, 47) or by double-stranded sequencing, using synthetic primers from within the LTR or the noncoding sequences 3′ to the FeLV *env* gene.

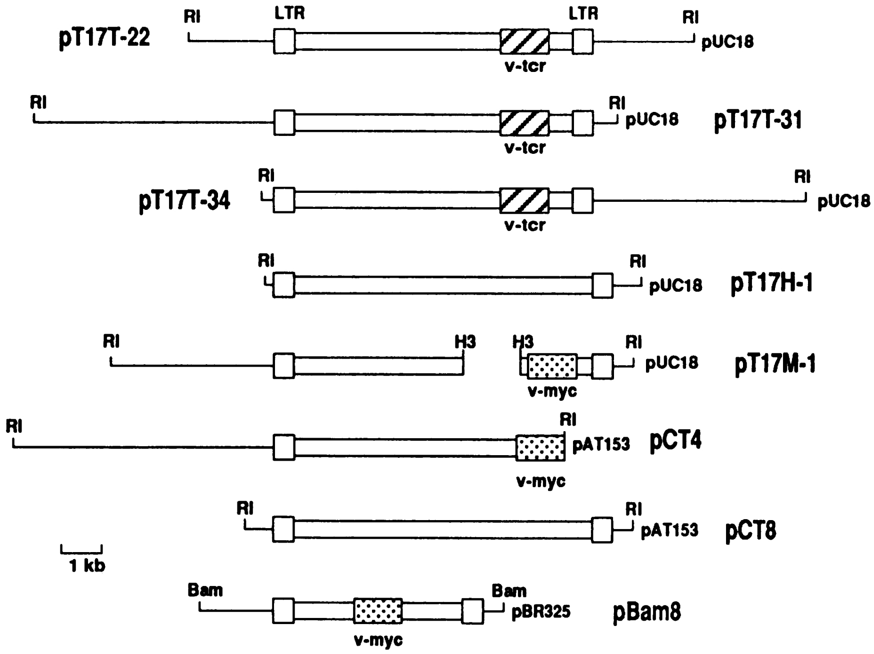

FIG. 1. Structures of molecularly cloned proviruses analyzed in this study. Five of the clones (T17-22, T17-31, T17-34, T17M-1, and T17H-1) came from tumor T17 (11). The T17-22, T17-31, and T17-34 proviruses carry the transduced T-cell antigen receptor β-chain gene, and although they are similar in internal proviral structure, differences in flanking sequences indicate that they are independent integrants. T17M-1 is a v-*myc*-containing virus, and T17H-1 is an apparently full-length FeLV provirus. A partial clone of a v-*myc*-containing provirus (CT4) and a full-length helper FeLV genome (CT8) were obtained from another field case tumor (32). A further v-*myc*-containing primary tumor isolate (pBam8) was provided by J. Casey. Detailed restriction maps are available in the original publications describing these clones. Viral sequences are indicated by the boxes. Larger boxes denote the LTRs, and hatching or stippling indicates host-derived inserts, as marked.

PCR. The Perkin Elmer-Cetus polymerase chain reaction (PCR) kit was employed with recombinant Taq polymerase. The amplimer sequences were 5′-TTACTCAAGTATGTT CCCATG-3′ and 5′-CTGGGGAGCCTGGAGACTGCT-3′ (Fig. 2, asterisks). Reaction mixtures contained 1 μg of genomic DNA or 1 ng of plasmid DNA, 1 μg of each primer (1 μg/μl), 10 μl of 10× reaction buffer, 500 μM deoxynucleotides, and water to a volume of 95 μl. After a 10-min incubation at 95°C, 0.5 U (5 μl) of Taq polymerase was added and amplification was achieved with 30 cycles of denaturation at 92°C, annealing at 37°C, and polymerization at 72°C. The products of the PCR were resolved on 6% nondenaturing acrylamide gels and analyzed by Southern blotting or by subcloning into m13 and sequencing of both strands.

Footprint analysis. Crude nuclear protein extracts were prepared essentially as described previously (37), but all the buffers used contained 50 fM sodium orthovanadate, 10 mM glycerophosphate, and 2 mM Levamisole. Typically, nuclei containing 100 to 150 mg of DNA yielded 5 ml (5 to 10 mg of protein per ml) of crude nuclear extract.

*Kpn*I-*Pst*I restriction fragments containing FeLV LTR sequences derived from both pT17T-31 and pFGA5 were subcloned into pIC20R (25). The plasmid was linearized by digestion with *Hin*dIII, treated with calf intestinal phosphatase, and 5′ end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase, and the insert was isolated after secondary digestion with *Eco*RI (38).

Footprint protection assays were performed essentially by published procedures (38). Briefly, assay mixtures at a final volume of 100 μl contained 80 μl of nuclear extract, 5 ng of end-labeled restriction fragment, and 1 μg of poly(dI-dC), with or without 100 ng of double-stranded competitor oligonucleotide DNA. After limited DNase I digestion, nucleic acids were purified and resolved by denaturing polyacrylamide gel electrophoresis and autoradiography. Markers were prepared from the same fragment by the chemical sequencing method of Maxam and Gilbert (26), using the G+A-specific reaction.

Oligonucleotides. Complementary single-stranded oligonucleotides were synthesized (by Oswell DNA Service, Edinburgh University, Edinburgh, Scotland) and annealed to give *Bam*HI and *Bgl*II restriction sites at the 5′ and 3′ ends, respectively. The sequences used were (i) the nuclear factor 1 (NF1) binding site (30) from the adenovirus origin of replication (GATCTTATTTGGCTTGAAGCCAATATG; a single base-pair change was introduced to make a perfect inverted repeat) and (ii) the simian virus 40 (SV40) core enhancer sequence (GATCCATCTGTGGTTAAGCACCTG GA; see Fig. 4 and Table 1).

RESULTS

Exogenous FeLV U3 sequences are highly conserved. The FeLV proviruses analyzed are shown in Fig. 1. The sequences presented in Fig. 2 are the complete U3 and adjacent sequences up to the conserved *Kpn*I site in the R region of the LTR; the sequences are compared with that of FeLV-A/Glasgow-1, a field isolate of relatively low leukemogenicity (46). Apart from the heterogeneous direct repeats detected here, the U3 sequences of exogenous FeLV are highly conserved. Comparison of the sequences presented here and other published sequences (9, 13, 40, 45) showed an overall U3 sequence conservation of at least 95%.

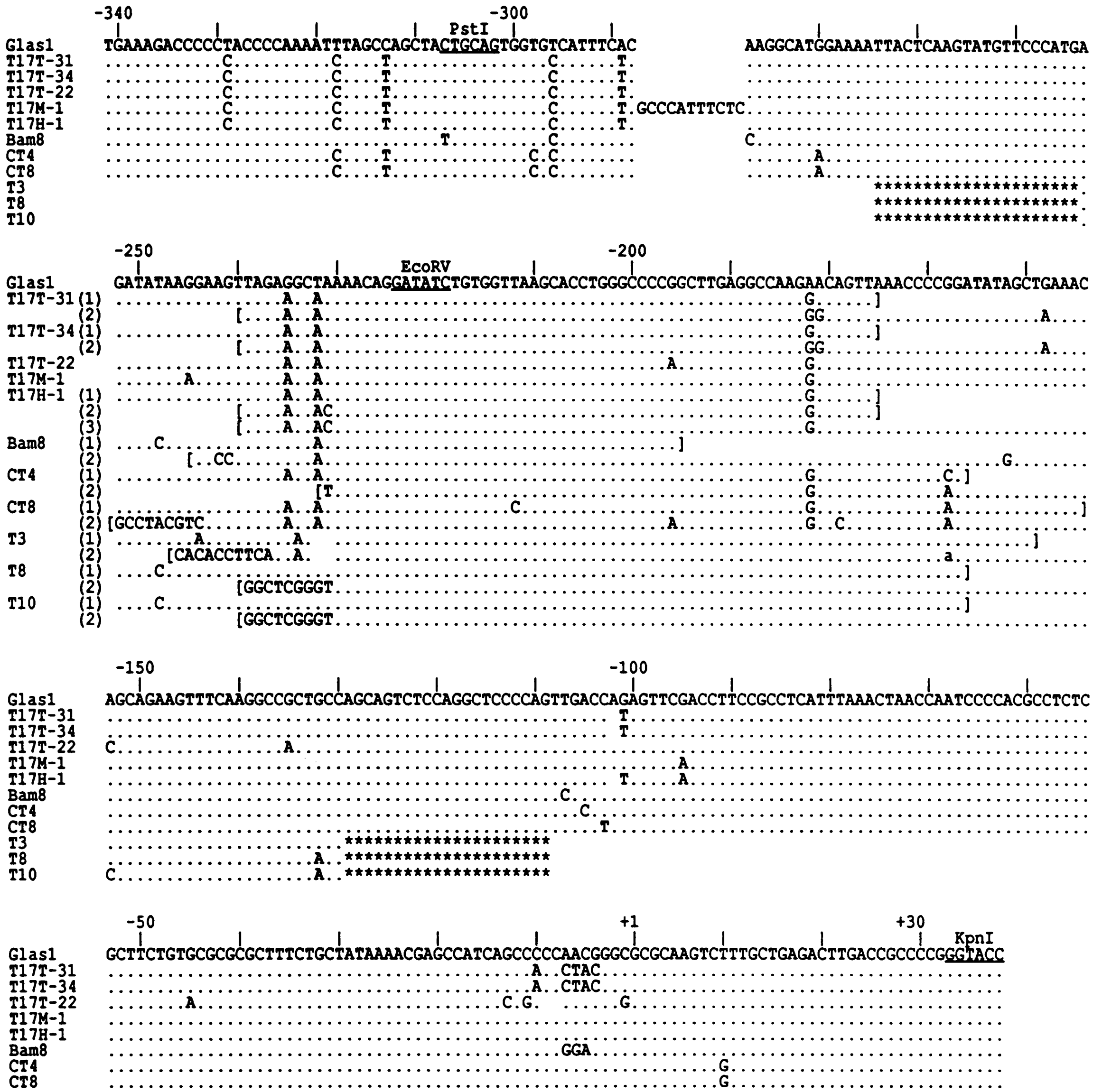

FIG. 2. U3 sequences of T-cell tumor-derived FeLV proviruses. Complete sequences were determined for the cloned proviruses whose genetic structures are shown diagrammatically in Fig. 1. Additional partial LTR sequences came from cloning of PCR products from tumors T3, T8, and T10 (Fig. 3). The amplimers used for the PCR are shown as asterisks. A lowercase letter in the T3 sequence denotes a polymorphic site in independent PCR clones. Direct repeats are indicated by brackets, and the sequence is resumed on the line below. Dots indicate identity with the previously published FeLV-A/Glasgow-1 (Glas1) sequence (46).

This figure includes multiple isolates from the United Kingdom and the United States. A recent Japanese FeLV proviral isolate, FeLV-FT1 (29), is more divergent (85% homology with pFGA-5), suggesting that some U3 sequence drift may be occurring in geographically separated cat populations.

Primary tumor-derived FeLV LTRs have heterogeneous direct repeats. Direct repeats of 49 to 90 base pairs were found in tumor-derived proviral LTRs. The termini of the repeats range from positions −155 to −244 from the presumptive RNA cap site in the FeLV-A/Glasgow-1 LTR, but no two isolates have 5′ and 3′ coterminal repeats unless they were obtained from the same tumor or had a common passage history. Of the proviruses cloned from tumor T17, T17T-31 and T17T-34 have identical sequences, including a 64-base-pair direct repeat, whereas the other provirus containing the v-*tcr* gene (T17T-22) is closely related, sharing numerous point mutational differences from the prototype sequence, but has only a single copy of the direct repeat. The apparently full-length (though replication-defective) provirus T17H-1 shares point mutations with the v-*tcr* viruses but has three copies of the direct repeat which is coterminal 5′ and 3′ with the repeat in T17T-31 and -34. Further divergence was seen in the FeLV-*myc* recombinant provirus (T17M-1) cloned from the same tumor. This LTR lacks the enhancer repeats but has a unique 11-base-pair insert at −286. The insert is an imperfect repeat of an upstream conserved region (UCR) motif identified as a *cis*-acting negative regulatory element in murine leukemia viruses (A. Khan, personal communication).

The CT4 and CT8 proviruses were cloned from a single tumor (no. 84793) and represent two of the three single-copy

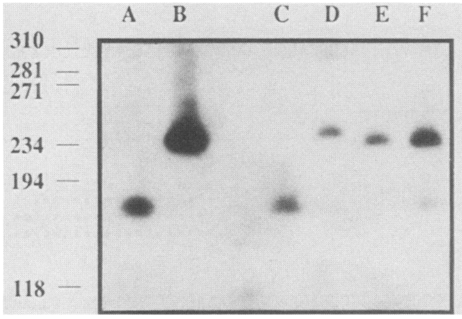

FIG. 3. Amplification of FeLV LTR sequences by PCR. Using two primers from conserved domains of the LTR (Fig. 2), we selectively amplified the enhancer sequences. The PCR products were separated on a 6% acrylamide gel, blotted, and hybridized with a probe consisting of the 64-base-pair repeat of the T17T-31 proviral LTR. The probe DNA was obtained by digestion of an LTR subclone (*Pst*I-*Kpn*I) with *Eco*RV and purification of the 64-base-pair fragment from acrylamide gels. The DNA was self-ligated prior to being labeled with [^{32}P]dCTP to a specific activity of at least 10^8 cpm/μg by the random priming method. Amplified plasmid DNAs are shown in lanes A (pFGA-5) and B (pT17T-31), while amplified cellular DNAs from established cell lines or primary tumors are shown in lanes C (F422 cell line), D (T3 cell lien), E (T10 tumor), and F (T8 tumor). Marker sizes (in base pairs) are indicated at the left.

proviruses detected in this tumor by Southern blot analysis (32), but their LTRs contain quite distinct direct repeats. The CT4 repeat is 65 base pairs, spanning positions −167 to −231, while that of CT8 spans positions −155 to −243 and the repeated sequences in CT8 are interrupted by a short sequence (GCCTACGTC) of unknown origin. A further unique structure was found in the L115 (pBam8) FeLV-*myc* isolate (21), which contained the shortest direct repeat, 49 base pairs.

Three further partial LTR sequences were obtained after PCR amplification of tumor DNA. Figure 3 shows a Southern blot analysis of the PCR products amplified from genomic DNAs, using the LTR primers indicated by asterisks in Fig. 2. The unit length LTR core enhancer (165-base-pair PCR fragment) is exemplified by the cloned FeLV-A/Glasgow-1 (Fig. 3, lane A). The predicted size of the duplicated T17T-31 enhancer fragment is 229 base pairs, which agrees well with the observed result (lane B). Duplicated LTR structures were the predominant forms detected in DNA from tumors T3, T8, and T10 (lanes D, E, and F). DNA from the F422 cell line yielded only the unit length LTR core enhancer (lane C), in agreement with the LTR structure of the FeLV-*myc* provirus cloned from this line (7). Subcloning and sequencing of the PCR fragments were performed; the results are presented along with the other sequences in Fig. 2. Each presented sequence is based on at least three independent subclones. The T3 tumor contains a primary FeLV-*myc* isolate (32, 35) and has a unique 75-base-pair repeat which is interrupted by a short sequence of unknown origin (CACACCTTCA).

Tumors T8 and T10 (32) were induced by inoculation with a crude tumor extract of the highly leukemogenic Rickard strain of FeLV (which was generated by four rounds of in vivo passage of a thymic tumor extract) (41). The PCR-amplified sequences from these tumors showed that in contrast to the other LTRs, which came from independently derived (naturally occurring) tumors, the LTRs from proviruses in T8 and T10 had identical 64-base-pair direct repeats, including a short insertion (GGCTCGGGT) between the repeats. These results strongly suggest that the direct repeat

TABLE 1. LTR factor binding sites

LTR binding site[a]	Sequence	Reference
SV40 core enhancer		
Consensus	TGTGGAAAG	16, 34
SL3-3	TGTGGTTAA	20, 48
FeLV	TGTGGTTAA	Fig. 2
FeLV CT8(1)	TGTGGTCAA	Fig. 2
Akv	TGTGGTCAA	20
MoMuLV	TGTGGTATG	43
NF1		
Consensus	TGG(N)6–7GCCA	12
FeLV	CCCGGCTTGAGGCCAAGGA	Fig. 2
MoMuLV	CCCGGCTCAGGGCCAAGAA	43
MMTV	TTTGGAATTTATCCAAATC	18
Adenovirus	TTTGGATTGAAGCCAATAT	30
PEA-2		
FeLV	CGCTGCCAGCAG	Fig. 2
Polyomavirus	AACTGACCGCAG	17
MoMuLV	GAGAACCATCAG	43
GRE		
Consensus	AGAACAnnnTGTACC	1
MoMSV	TCAGGAACAGAGAGACAG	28
	CCAAGAACAGATGGAACA	28
	CCAAGAACAGATGGTCCC	28
MMTV	TTAAGAACAGTTTGTAAC	28
	ATCAGAACATTTGATACC	28
	AATAGAACACTAAGAGCT	28
	AAAAGAACATAGGAAATA	28
FeLV G1	CCAAGAACAGTTAAACCC	Fig. 2
FeLV T17T-31	CCAAGGACAGTTAAACCC	Fig. 2
LVb		
FeLV	AAACAGGATATCT	Fig. 2
MoMuLV	AAACAGGATATCT	43

[a] MMTV, Mouse mammary tumor virus; FeLV G1, FeLV-A/Glasgow 1.

was present in the FeLV-Rickard inoculum. Since tumor T8 had an FeLV proviral insertion several kilobases upstream of c-*myc* in the opposite transcriptional orientation (10, 32), it appears that the repeats are associated with viruses of high leukemogenic potential, whether these are oncogene-containing viruses (T3, L115, and T17) or viruses which activate genes by proviral insertion.

Interestingly, two clusters of mutations also distinguish the majority of the tumor-derived viruses from previously analyzed FeLV LTRs, typified by the FeLV-A/Glasgow-1 sequence. The mutation clusters occur at or near the boundaries of the direct repeats (Fig. 2), suggesting that they might be a consequence of the duplication event. However, one cluster at the 3′ end is an A → G transition in a glucocorticoid response element (GRE) consensus sequence (Table 1). This sequence overlaps with a NF1 protein binding site (see below). The other mutational cluster at the 5′ end of the direct repeats is not in any recognized motif but flanks the core enhancer binding site, as described below.

DNase I footprinting of repeated enhancer domain reveals binding to several consensus sequences. Nuclear factor binding sites in the enhancer domain of the LTR were determined by in vitro footprint protection assays. 5′-End-labeled restriction fragments of the pT17T-31 LTR (Fig. 4A) or the pFGA-5 (FeLV-A/Glasgow-1) LTR (Fig. 4B) were bound by crude nuclear extracts prepared from either a T-cell tumor line (T3) or a feline fibroblast cell line (AH927). Multiple

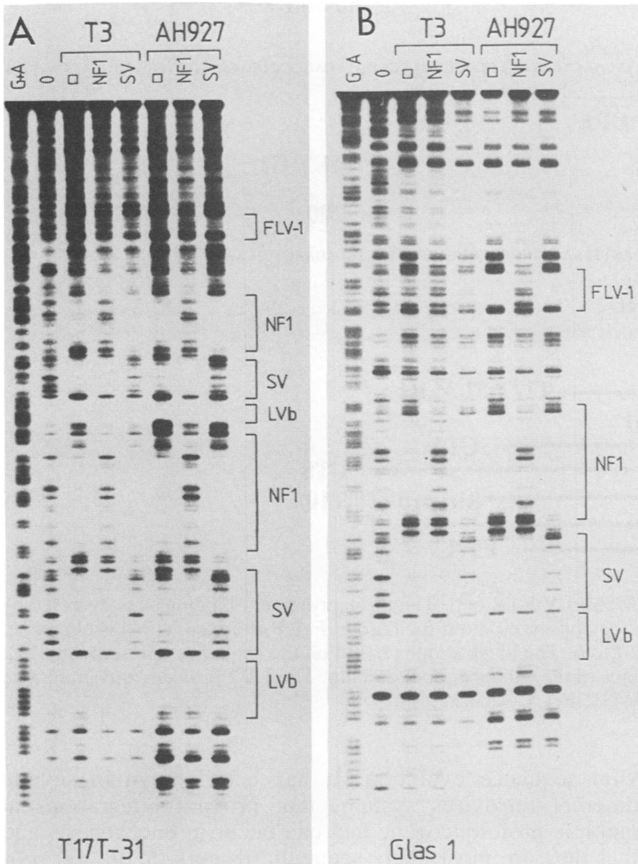

FIG. 4. In vitro DNase I footprint analysis of the pT17T-31 (A) and pFGA-5 (B) FeLV LTRs. Crude nuclear extracts from either T3 or AH927 cells were bound to a 5′-end-labeled probe in the presence or absence of excess competitor oligonucleotides based on the adenovirus NF1 (NF1) or SV40 enhancer core (SV) binding site. After digestion with DNase I, nucleic acids were resolved on denaturing polyacrylamide gels (see Materials and Methods). Symbols: □, nuclear extract without inhibitor; 0, control with no extract added; G+A, chemical sequencing lane. Glas 1, FeLV-A/Glasgow-1.

protein binding sites were detected in the duplicated region of the T17T-31 LTR; these are shown in Fig. 5, in which they are compared with the Moloney murine leukemia virus (MoMuLV) LTR binding sites identified by Speck and Baltimore (43).

The 5′-most protected sequence shown in Fig. 4 and 5 is homologous to the SV40 core enhancer binding site consensus sequence (Table 1) and is present in all of the direct repeats of the FeLV enhancer. Multiple proteins which bind to the SV40 core enhancer have been described (3, 19, 27, 48). It is not yet clear which if any of these bind to the FeLV LTR. Preliminary evidence from gel shift assays suggests that there may be cell type-specific differences in the proteins binding to this site (unpublished results).

Another footprint occurs over an inverted repeat sequence which is similar to the NF1 binding site consensus sequence (Fig. 4). Homologous NF1 binding sites are also found in the MoMuLV (43) and mouse mammary tumor virus (18) LTRs (Table 1). The DNase I protection pattern was abolished (Fig. 4) by an excess of cold double-stranded oligonucleotides containing the NF1 binding site of the adenovirus origin of replication (see Table 1 and Materials and Methods), and gel shift assays showed that the protein-DNA complexes formed with the adenovirus and FeLV NF1 binding sites were indistinguishable (data not shown). The T3 and AH927 nuclear extracts both protected the NF1 binding sites, although other studies indicate that the levels of NF1 in T3 cells are significantly lower than those in AH927 cells. The 3′ end of the NF1 binding site in the FeLV LTR contains a sequence which is homologous to the GRE in the LTR of the Moloney murine sarcoma virus (MoMSV) and mouse mammary tumor virus enhancers (Fig. 5 and Table 1). The FeLV (and MuLV) GRE consensus sequences are within the NF1 footprint, and in the case of the FeLV LTR, no protection was observed over that provided by the GRE after NF1 was eliminated by competition with adenovirus NF1 oligonucleotide (Fig. 4A and B). Binding to the MoMSV GRE has been recorded, but only with purified glucocorticoid receptor (28). The lack of GRE binding activity after competition with NF1 in our assays may have been due to a low concentration or affinity of the receptor in our extracts or may have reflected a cooperative binding interaction which was abolished by NF1 competition.

Sequence alignment of the FeLV and MoMuLV LTRs reveals additional areas of conservation and divergence (Fig. 5). Several binding motifs identified in the murine virus are not present in FeLV. These include LVa, LVc, and a proximal NF1 site, where alignment with MoMuLV requires a 15- to 20-base-pair deletion in FeLV. There is conservation over a sequence (CAGGAT) 5′ to the SV40 core-related binding site, which has been shown to bind a general transcription factor, LVb, in the MoMuLV LTR (43). The LVb motif is competely conserved in all the FeLV sequences, but no consistent DNase I protection pattern was seen at this site in the footprint assays (Fig. 4A and B) with crude nuclear extracts. LVb was previously identified by gel shift and methylation interference assays which used partially purified nuclear extracts rather than by footprinting (43).

A novel FeLV-specific binding site (FLV-1) is located 3′ to the enhancer duplications. A third protected sequence occurs outside the duplicated region of the T17T-31 LTR and was seen most clearly in the footprint protection pattern observed with the FeLV LTR pFGA5 (Fig. 4B). The sequence bound contains an indirect repeat which is similar to the PEA2 binding site of the polyomavirus A enhancer (Table 1). Binding was detected very weakly or not at all with T3 nuclear extracts. The footprint was abolished by competition with the adenovirus NF1-binding oligonucleotide, suggesting that the site represents either a weak NF1 binding site which was not observed in T3 nuclear extracts because of low NF1 levels or a binding site for a distinct protein which is eliminated fortuitously by the 1,000-fold molar excess of competitor NF1 oligonucleotide, which is weakly homologous to the site (Table 1). Until there is further definition of the factors binding to this site, we refer to the site as FLV-1.

Although the FLV-1 binding site is in a region of high sequence conservation in FeLV, there is only weak homology in the equivalent location in MoMuLV. However, direct analysis of the MoMuLV sequence for binding activity will be required to confirm the apparent lack of the FLV-1 site.

DISCUSSION

In this study we analyzed sequence variation in the LTR sequences of FeLV in naturally occurring lymphomas. We found that, unlike FeLV LTRs from other sources, those cloned directly from tumor DNA frequently have enhancer

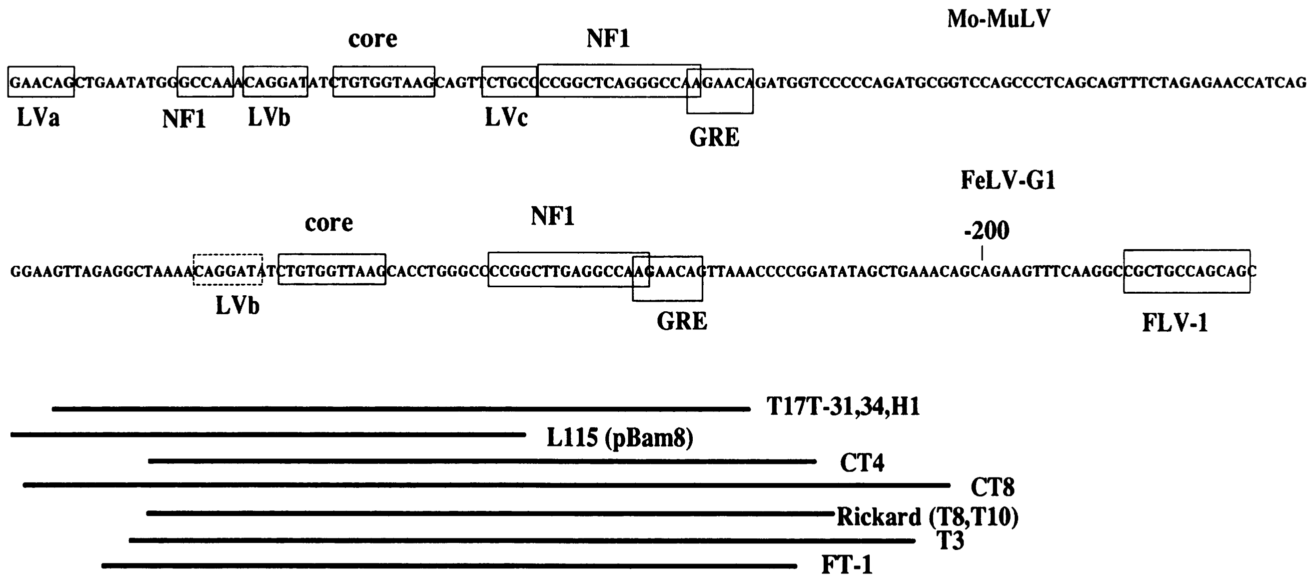

FIG. 5. Binding motifs in the FeLV LTR compared with those in the MoMuLV LTR (43). The three prominent binding sites detected by DNase I footprinting are shown as boxes on the FeLV LTR. In addition, the conserved motif for LVb in FeLV is boxed with a broken line. The extent of LTR sequence duplications is indicated by the lines at the bottom. The illustration is based on the sequences presented in Fig. 2 and an additional sequence (FT-1) reported recently which had three copies of the enhancer core domain. The FT-1 provirus was integrated upstream of c-*myc* in a naturally occurring thymic lymphoma (29). FeLV-G1, FeLV-A/Glasgow-1.

sequence duplications. The duplications are analogous to those found in murine leukemia viruses with accelerated disease potential (4, 5, 15, 22). A similar correlation appears to hold for FeLV; we found enhancer duplications in viruses from the majority of naturally occurring lymphoid tumors but not in viruses isolated from healthy animals. The duplications are unique to each independent isolate but encompass a series of binding sites for feline nuclear factors, including in every case a site closely related to the SV40 core enhancer sequence. These results suggest that this site is important for the T-cell leukemogenic properties of FeLV.

Despite the very close relationship between murine and feline retroviruses, as shown by the conservation of sequence, location, and spatial organization of nuclear factor binding sites in the enhancer (Fig. 5), there are notable differences in their patterns of sequence variation. The weakly pathogenic murine leukemia virus Akv differs from the highly leukemogenic variant SL3-3 by a single base change in the SV40-like core sequence (20). This mutation is particularly deleterious for enhancer function in T-lymphoma cells and appears to be a major determinant of the reduced leukemogenic potential of Akv (3, 23, 48). Almost all of the FeLV isolates are identical to SL3-3 within this recognition sequence, regardless of their degree of pathogenicity. Strong conservation of the SL3-3-like core sequence in FeLV may reflect a requirement for efficient enhancer function in T cells to allow viral establishment and persistence. FeLV is an efficiently transmitted infectious agent, while Akv is transmitted as a germ line insert in AKR mice (24). The core sequence mutation of Akv may have attenuated its pathogenic function, allowing survival of mice bearing the germ-line provirus.

Heterogeneity of LTR direct repeats in a single tumor (e.g. CT4 and CT8) might arise from concomitant infection with divergent viral genomes due to phenotypic mixing in vivo. Alternatively, the diversity may arise from a process of sequential proviral integrations with intervening periods of viral sequence evolution. It has been shown in several different retrovirus systems that proviral integrations at multiple proto-oncogene loci can occur in oncogenesis and that they are most likely sequential events (6, 36, 49). For tumor T17, analysis of the host-virus sequence junctions of recombinant proviruses carrying v-*tcr* and v-*myc* (Fig. 1) showed no evidence of any common intermediate in transduction of the two cellular gene loci (R. Fulton and J. C. Neil, unpublished results). On the basis of the LTR sequence divergence of the proviral populations, we suggest further that the transductions of v-*myc* and v-*tcr* may have occurred at different stages of the leukemogenic process.

Previous models for the generation of the retrovirus enhancer repeats have been based on template jumping by reverse transcriptase, involving predicted stem-loop structures in viral RNA (5, 44). The termini of the FeLV LTR direct repeats can in a number of cases be aligned as short or imperfect direct repeats (Fig. 6), which might favor a template jumping process. However, the termini also display some sequence features in common with other retrovirus recombination sites. Marked purine/pyrimidine strand bias has been noted as a feature of such recombination sites (2, 7, 31) and is seen particularly at the 5′ junctions of the repeats. A motif found at several 5′ FeLV-*onc* junctions (ACCCC) (7) is present in the center of the clustered 3′ junctions of the LTR repeats. A hitherto unrecorded feature is that this motif is contained within the terminal inverted repeat of the FeLV LTR. Furthermore, short motifs homologous to the termini of the LTR inverted repeats are present at or close to the ends of the duplications. In the T17M-1 LTR, for example, there are short matches to both ends of the LTR inverted repeats close to the site of an 11-base-pair insertion (Fig. 6). It is conceivable, therefore, that the viral integrase protein may play a role in generating the enhancer repeats.

It is interesting to note the close similarity of the FeLV LTR terminal inverted repeat to a motif within the LTR of a mouse retrotransposon which is recognized by an uncharac-

```
     *                                    *
GGAAGTTAGAGGC    T17T-31     CAAAAACAGGATATC     CT4
   ||||| |  |    T17T-34      |||  | ||||||      Rickard (T8, T10)
AACAGTTAAACCC    T17H-1      TAAACCCCGGATATA
      *                               *

       *
GTTAGAGACAAAACA
| || ||   |||||    T3
GATATAGCTGAAACA
        *

                  *
    **  *  *      *
ATAAGGAAGTTAGAGGCTAAAACAGGAT                     all 5' junctions
               -230

                                    *
 *                *                 *     *    *
CGGCTTGAGGCCAAGAACAGTTAAACCCCGGATATAGCTGAAACAGCA   all 3' junctions
                                              -150

      *
CATTTCACAAGGCATGGAAAATTACTC                      T17M-1 insert
          -280

(CACACCTTCA) AGACAAACAGG                         T3 junction
             -230

TGGAAATCCCCC                                     ds DNA binding site
                                                 LTR-IS

TGAAAGACCCCC                                     FeLV LTR
GGGGGTCTTTCA                                     inverted repeats
```

FIG. 6. Sequences at the junctions of LTR direct repeats. The upper part of the figure shows sequence matches at the 5′ and 3′ junctions of LTR duplications for individual isolates. Asterisks denote the terminal nucleotides of the direct repeats. In the lower portion, motifs are underlined which match previously observed sequences at FeLV recombinational junctions (ACCCC) (7, 31) or the inverted repeats at the ends of the LTR. A match is also noted to a double-stranded DNA (ds DNA) binding site for a cellular protein which interacts with the LTR of a murine transposable element (8). The termini of direct repeats or points of sequence insertion are marked by asterisks above the sequence. Parentheses in the T3 sequence indicate an insert between the repeats which does not match the prototype LTR of FeLV-A/Glasgow-1 (Fig. 2).

terized double-stranded DNA binding protein. Binding at the motif has been proposed to play a role in suppressing recombination at related sites in chromosomal DNA (8). A corollary of this proposal is that newly synthesized and unprotected proviral DNA might then act as a sensitive substrate for recombination.

Whatever the mechanism(s) by which the repeats are formed, efficient enhancers must result if the structures are to achieve wide dispersal, particularly if they arise de novo in an infected host. In the modular arrays of binding sites for transcription factors found in viral enhancers, the number and interrelationships of binding sites are critical for activity (34). If a random recombination process generates the duplicated structures, many of these will not be detected because impaired enhancer function precludes their propagation. The LTR enhancer duplications in FeLV proviruses in primary tumors of T-cell origin invariably include the SV40-like enhancer core site and an adjacent motif for LVb (43). This common central domain of the enhancer duplications may be the basic functional enhancer unit (34), containing binding site(s) for crucial regulators of transcription, some of which may be expressed in a tissue-specific fashion.

We identified a novel binding site (FLV-1) closely 3′ to the enhancer duplications. The binding site is homologous (8 of 12 base positions) to PEA-2, a negative regulatory site in the polyomavirus enhancer (17). An analogous role may be postulated for the FLV-1 site in FeLV, based on the observations that the FLV-1 site is invariably excluded from the enhancer duplications and binding activity in a T-cell tumor (T3) nuclear extract was low or absent, while it was abundant in fibroblast nuclear extracts. We do not yet know whether this is a feature of the T-cell lineage, the differentiation stage of the T3 cell line, or the tumor phenotype. NF1 activity was also relatively low in T3 lymphoma cells compared with that in the AH927 fibroblast line (M. Plumb, unpublished results), and the principal binding activity in T3 was to the SV40-like core enhancer sequence. Viruses with duplications of this binding site may therefore be selectively adapted for replication and enhancer function in the target cell for T-cell leukemogenesis.

ACKNOWLEDGMENTS

We thank Robert McFarlane for technical assistance.
This work was supported by the Cancer Research Campaign.

LITERATURE CITED

1. **Beato, M.** 1989. Gene regulation by steroid hormones. Cell **56**:335–344.
2. **Besmer, P., J. E. Murphy, P. C. George, F. Qui, P. J. Bergold, L. Lederman, H. W. Snyder, Jr., D. Brodeur, E. E. Zuckerman, and W. D. Hardy.** 1986. A new acute transforming feline retrovirus and relationship of its oncogene v-kit with the protein kinase gene family. Nature (London) **320**:415–421.
3. **Boral, A. L., S. A. Okenquist, and J. Lenz.** 1989. Identification of the SL3-3 virus enhancer core as a T-lymphoma cell-specific element. J. Virol. **63**:76–84.
4. **Chattopadhyay, S. K., B. M. Baroudy, K. L. Holmes, T. N. Frederickson, M. L. Lander, H. C. Morse III, and J. W. Hartley.** 1989. Biologic and molecular genetic characteristics of a unique MCF virus that is highly leukemogenic in ecotropic virus-negative mice. Virology **168**:90–100.
5. **Clark, S. P., R. Kaufhold, A. Chan, and T. W. Mak.** 1985. Comparison of the transcriptional properties of the Friend and Moloney retrovirus long terminal repeats: importance of tandem duplications and of the core enhancer sequence. Virology **144**:481–494.
6. **Cuypers, H. T., G. Selten, M. Zjilstra, R. de Goede, C. Melief, and A. Berns.** 1986. Tumor progression in murine leukemia virus-induced T-cell lymphomas: monitoring clonal selections with viral and cellular probes. J. Virol. **60**:230–231.
7. **Doggett, D. L., A. L. Drake, V. Hirsch, M. E. Rowe, V. Stallard, and J. I. Mullins.** 1989. Structure, origin, and transforming activity of the feline leukemia virus-*myc* recombinant provirus FTT. J. Virol. **63**:2108–2117.
8. **Edelmann, W., B. Kroger, M. Goller, and I. Horak.** 1989. A recombinational hotspot in the LTR of a mouse retrotransposon identified in an in vitro system. Cell **57**:937–946.
9. **Elder, J. H., and J. I. Mullins.** 1983. Nucleotide sequence of the envelope gene of Gardner-Arnstein feline leukemia virus B reveals unique sequence homologies with a murine mink cell focus-forming virus. J. Virol. **46**:871–880.
10. **Forrest, D., D. Onions, G. Lees, and J. C. Neil.** 1987. Altered structure and expression of c-myc in feline T-cell tumours. Virology **158**:194–205.
11. **Fulton, R., D. Forrest, R. McFarlane, D. Onions, and J. C. Neil.** 1987. Retroviral transduction of T-cell antigen receptor β-chain and myc genes. Nature (London) **326**:190–194.
12. **Gronostajski, R. M.** 1986. Analysis of nuclear factor I binding to DNA using degenerate oligonucleotides. Nucleic Acids Res. **14**:9117–9132.
13. **Guilhot, S., A. Hampe, L. D'Auriol, and F. Galibert.** 1987. Nucleotide sequence analysis of the LTR's and env genes of SM-FeSV and GA-FeSV. Virology **161**:252–258.
14. **Hampe, A., M. Gobet, J. Even, C. J. Sherr, and F. Galibert.** 1983. Nucleotide sequences of feline sarcoma virus long terminal repeats and 5′ leaders show extensive homology to those of other mammalian retroviruses. J. Virol. **45**:466–472.
15. **Holland, C. A., C. Y. Thomas, S. K. Chattopadhyay, C. Koehne, and P. V. O'Donnell.** 1989. Influence of enhancer sequences on

thymotropism and leukemogenicity of mink cell focus-forming viruses. J. Virol. **63**:1284–1292.

16. **Jones, N. C., P. W. J. Rigby, and E. B. Ziff.** 1988. Trans-acting protein factors and the regulation of eukaryotic transcription. Genes Dev. **2**:267–281.
17. **Kryszke, M.-H., J. Piette, and M. Yaniv.** 1987. Induction of a factor that binds to the polyoma virus A enhancer on differentiation of embryonal carcinoma cells. Nature (London) **328**: 254–256.
18. **Kuo, W.-L., L. R. Vilander, M. Huang, and D. O. Peterson.** 1988. A transcriptionally defective long terminal repeat within an endogenous copy of mouse mammary tumor virus proviral DNA. J. Virol. **62**:2394–2402.
19. **Landschultz, W. H., P. F. Johnson, E. Y. Adashi, B. J. Graves, and S. L. McKnight.** 1988. Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. **2**:786–800.
20. **Lenz, J., D. Celander, R. L. Crowther, R. Patarca, D. W. Perkins, and W. A. Haseltine.** 1984. Determination of the leukaemogenicity of a murine retrovirus by sequences within the long terminal repeat. Nature (London) **308**:467–470.
21. **Levy, L. S., M. B. Gardner, and J. W. Casey.** 1984. Isolation of a feline leukaemia provirus containing the oncogene myc from a feline lymphosarcoma. Nature (London) **308**:853–856.
22. **Li, Y., E. Golemis, J. W. Hartley, and N. Hopkins.** 1987. Disease specificity of nondefective Friend and Moloney murine leukemia viruses is controlled by a small number of nucleotides. J. Virol. **61**:693–700.
23. **LoSardo, J. E., L. A. Cupelli, M. K. Short, J. W. Berman, and J. Lenz.** 1989. Differences in activities of murine retroviral long terminal repeats in cytotoxic T lymphocytes and T-lymphoma cells. J. Virol. **63**:1087–1094.
24. **Lowy, D. R., E. Rands, S. K. Chattopadhyay, C. F. Garon, and G. L. Hager.** 1980. Molecular cloning of infectious integrated murine leukemia virus DNA from infected mouse cells. Proc. Natl. Acad. Sci. USA **77**:614–618.
25. **Marsh, J. L., M. Erfle, and E. J. Wykes.** 1984. The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene **32**:481–485.
26. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavage. Methods Enzymol. **65**: 499–560.
27. **Mercurio, F., and M. Karin.** 1989. Transcription factors AP-3 and AP-2 interact with the SV40 enhancer in a mutually exclusive manner. EMBO J. **8**:1455–1460.
28. **Miksicek, R., A. Heber, W. Schmid, U. Danesch, G. Posseckert, M. Beato, and G. Schutz.** 1986. Glucocorticoid responsiveness of the transcriptional enhancer of Moloney murine sarcoma virus. Cell **46**:283–290.
29. **Miura, T., M. Shibuya, H. Tsujimoto, M. Fukusawa, and M. Hayami.** 1989. Molecular cloning of a feline leukemia provirus integrated adjacent to the c-myc gene in a feline T-cell leukemia cell line and the unique structure of its long terminal repeat. Virology **169**:458–461.
30. **Nagata, K., R. A. Guggenheimer, and J. Hurwitz.** 1983. Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. Proc. Natl. Acad. Sci. USA **80**:6177–6181.
31. **Neil, J. C., D. Forrest, D. L. Doggett, and J. I. Mullins.** 1987. The role of feline leukaemia virus in naturally-occurring leukaemias. Cancer Surveys **6**:117–137.
32. **Neil, J. C., D. Hughes, R. McFarlane, N. M. Wilkie, D. E. Onions, G. Lees, and O. Jarrett.** 1984. Transduction and rearrangement of the myc gene by feline leukaemia virus in naturally occurring T-cell leukaemias. Nature (London) **308**:814–820.
33. **Norrander, J., T. Kempe, and J. Messing.** 1983. Improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **26**: 101–106.
34. **Ondek, B., L. Gloss, and W. Herr.** 1988. The SV40 enhancer contains two distinct levels of organization. Nature (London) **333**:40–45.
35. **Onions, D., G. Lees, D. Forrest, and J. Neil.** 1987. Recombinant feline viruses containing the myc gene rapidly produce clonal tumours expressing T-cell antigen receptor gene transcripts. Int. J. Cancer **40**:40–45.
36. **Peters, G., A. E. Lee, and C. Dickson.** 1986. Concerted activation of two potential proto-oncogenes in carcinomas induced by mouse mammary tumour virus. Nature (London) **320**:628–631.
37. **Plumb, M., J. Frampton, H. Wainwright, M. Walker, K. Macleod, G. Goodwin, and P. Harrison.** 1989. GATAAG; a cis-control region binding an erythroid-specific nuclear factor with a role in globin and non-globin gene expression. Nucleic Acids Res. **17**:73–92.
38. **Plumb, M. A., and G. H. Goodwin.** 1988. Detection of sequence-specific DNA-protein interactions by the DNA footprinting technique. Methods Mol. Biol. **4**:139–164.
39. **Rasheed, S., and M. B. Gardner.** 1980. Characterization of cat cell cultures for expression of retrovirus, FOCMA and endogenous sarc genes, p. 393–400. *In* M. Essex and A. J. McClelland (ed.), Proceedings of the Third International Feline Leukemia Virus Meeting. Elsevier/North-Holland Publishing Co., Amsterdam.
40. **Riedel, N., E. A. Hoover, P. W. Gasper, M. O. Nicolson, and J. I. Mullins.** 1986. Molecular analysis and pathogenesis of the feline aplastic anemia retrovirus, feline leukemia virus. C-SARMA. J. Virol. **60**:242–250.
41. **Rojko, J. L., E. A. Hoover, L. E. Mathes, R. G. Olsen, and J. P. Schaller.** 1979. Pathogenesis of experimental feline leukemia virus infection. J. Natl. Cancer Inst. **63**:759–768.
42. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
43. **Speck, N. A., and D. Baltimore.** 1987. Six distinct nuclear factors interact with the 75-base-pair repeat of the Moloney leukemia virus enhancer. Mol. Cell. Biol. **7**:1101–1110.
44. **Spiro, C., J.-P. Li, R. K. Bestwick, and D. Kabat.** 1988. An enhancer instability that diversifies the cell repertoire for expression of a murine leukemia virus. Virology **164**:350–361.
45. **Stewart, M. A., D. Forrest, R. McFarlane, D. Onions, N. Wilkie, and J. C. Neil.** 1986. Conservation of the c-myc coding sequence in transduced feline v-myc genes. Virology **154**:121–134.
46. **Stewart, M. A., M. Warnock, A. Wheeler, N. Wilkie, J. I. Mullins, D. E. Onions, and J. C. Neil.** 1986. Nucleotide sequences of a feline leukemia virus subgroup A envelope gene and long terminal repeat and evidence for the recombinational origin of subgroup B viruses. J. Virol. **58**:825–834.
47. **Tabor, S., and C. C. Richardson.** 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84**:4676–4771.
48. **Thornell, A., B. Hallberg, and T. Grundstrom.** 1988. Differential protein binding in lymphocytes to a sequence in the enhancer of the mouse retrovirus SL3-3. Mol. Cell. Biol. **8**:1625–1637.
49. **Tsichlis, P. N., P. Gunter-Strauss, and M. A. Lohse.** 1985. Concerted DNA rearrangements in Moloney leukemia virus-induced thymomas: a potential synergistic relationship in oncogenesis. J. Virol. **56**:258–267.
50. **Wingender, E.** 1988. Compilation of transcription regulating proteins. Nucleic Acids Res. **16**:1879–1902.
51. **Wong, T. C., R. S. Goodenow, B. T. Sher, and N. Davidson.** 1985. The promoter of the long terminal repeat of fellne leukemia virus is effective for expression of a mouse H-2 histocompatibility gene in mouse and human cells. Gene **34**:27–38.