JOURNAL OF VIROLOGY, June 1986, p. 983–987
0022-538X/86/060983-05$02.00/0
Copyright © 1986, American Society for Microbiology
Vol. 58, No. 3

Nucleotide Sequence Analysis of Squirrel Monkey Retrovirus Reveals a Novel Primer-Binding Site for $tRNA_{1,2}^{Lys}$

ING-MING CHIU†* AND SUZANNE F. SKUNTZ

Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, Maryland 20892

Received 23 August 1985/Accepted 3 January 1986

Nucleotide sequences of a DNA fragment containing the long terminal repeat (LTR) of squirrel monkey retrovirus (SMRV) were determined. Sequence analysis showed that the SMRV LTR is 456 base pairs (bp) long and is bounded by 2-bp inverted repeats. Within the U3 region, there are two 43-bp repeats and two 42-bp repeats which are homologous to each other. These repeats are likely to provide enhancer activities commonly observed in other enhancer sequences. Following the repeats are transcriptional regulatory sequences including a CAT box, a Goldberg-Hogness box, and a polyadenylation signal, all positioned within the U3 region of SMRV LTR. A 22-nucleotide sequence immediately downstream from the LTR was found to be complementary to $tRNA_{1,2}^{Lys}$, suggesting that $tRNA_{1,2}^{Lys}$ serves as the primer for the reverse transcription of SMRV viral RNA.

Squirrel monkey retrovirus (SMRV) was first isolated from squirrel monkey lung tissue by cocultivation with canine cells (17). It is a type D retrovirus present in multiple proviral copies in the genomes of squirrel monkeys (3, 9). Recently, three independent isolates of type D retroviruses have been isolated from macaques suffering from an acquired immune deficiency syndrome (AIDS) that resembles human AIDS (11, 24, 44). This raised the possibility of using animal models for studying the pathogenesis of this devastating human disease. However, in contrast to the well-characterized genome structures of type A, B, and C retroviruses (50), very little has been documented on the structure and organization of the type D retrovirus genome.

Previous studies have demonstrated limited antigenic homology between the *gag* gene products of the type D retroviruses and those of the type B and mammalian type C viruses (1). In an effort to extend these studies, we have molecularly cloned the closed circular form of the proviral SMRV DNA in a bacteriophage vector (3). Using Southern hybridization in conjunction with nucleotide sequence analysis, we have shown that type A, type B, avian type C, and type D viruses share extensive sequence homology in the *pol* gene regions (4, 5). Furthermore, by Southern hybridization and heteroduplex analysis, we (4) and others (8) have shown conservation between the *env* genes of type D and mammalian type C viruses. In contrast, we did not detect any sequence homology between the SMRV long terminal repeat (LTR) and other retroviral LTRs by filter hybridization (4).

The retroviral LTRs contain all the regulatory sequences required for transcription of the viral genome and are important for integration of the proviral DNA into the host genome (47). The LTR may also lead to the activation of host genes adjacent to the viral integration site which may cause a variety of neoplasias in animals (16, 20, 27, 30). To elucidate the structure and the interrelationship of type D retroviral LTR sequences with those of other genera of the retroviruses, we determined the nucleotide sequence of SMRV LTR. Here we report that the SMRV LTR possesses all the sequences required for retroviral transcription, including a CAT box, a Goldberg-Hogness box, and a polyadenylation signal. We also detected a presumptive primer-binding site (PBS) unlike that of any other known retroviruses which is complementary to the last 22 nucleotides of $tRNA_{1,2}^{Lys}$.

SMRV DNA cloned in λgtWES · λB (3) was subsequently cloned into pBR322, and the resultant plasmid was designated pSMRV. On the basis of the difference in the restriction maps between the linear proviral DNA and pSMRV, we deduced that pSMRV was derived from the closed circular proviral DNA with one copy of the LTR (3). By using subgenomic DNA fragments of pSMRV as probes, we concluded that the LTR of SMRV may reside between the *Sph*I site and the *Eco*RI site (7.7 to 8.4 kilobase pairs [kbp]) of the pSMRV restriction map (3) (Fig. 1). We therefore subcloned the 1.1-kbp *Bam*HI-*Eco*RI fragment of pSMRV into plasmid pAT153 (46). The resultant clone, pSS4.01, was used for sequence analysis by the method of Maxam and Gilbert (25). The sequencing strategy is shown in Fig. 1. Sequences from both strands were independently determined and mutually confirmed. To ensure that there were no restriction sites too close to be detected by restriction enzyme mapping, every restriction site shown was sequenced through from another restriction site. The complete nucleotide sequence of the 1.1-kbp *Bam*HI-*Eco*RI DNA which contains the SMRV LTR is shown in Fig. 2.

Since the SMRV clone we characterized had only a single LTR, it is presumed that this LTR should be bound by a polypurine-rich region on its 5′ side and a PBS on its 3′ side. We therefore examined these reference sequences along with other common structural features (Fig. 2). We first searched for sequences in the SMRV LTR complementary to the known primer tRNAs, $tRNA^{Pro}$ (2), $tRNA^{Trp}$ (2), $tRNA_3^{Lys}$ (2), and $tRNA^{Phe}$ (7, 15, 29), using the SRCHN program of Wilbur and Lipman (51). We did not find any such sequences in the SMRV LTR. To test whether type D retroviruses use a previously unidentified PBS, the sequences complementary to the ones presented in Fig. 2 were used to search for homology with all the nucleotide sequences in the National Institutes of Health database, using the GNSRCH computer program (51). The nucleotide sequence 863 to 884 was found to be complementary to the 3′ end of the $tRNA_{1,2}^{Lys}$ of various species including rabbit,

* Corresponding author.
† Present address: Department of Medicine and Comprehensive Cancer Center, Ohio State University, Columbus, OH 43210.

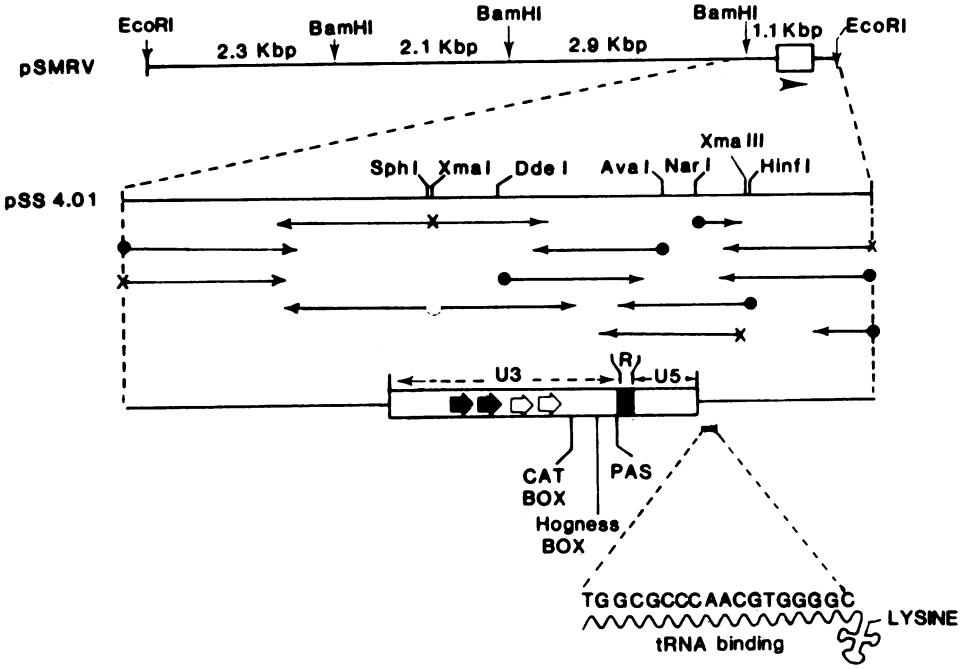

FIG. 1. Sequencing strategies of SMRV LTR and its adjacent regions. The open boxes represent LTRs. The arrowhead under the LTR represents the direction of transcription. The 1.1-kbp *Bam*HI-*Eco*RI fragment derived from pSMRV was subcloned into pAT153 and designated pSS4.01. The open and closed circles represent end labeling at the 3′ end by terminal transferase or Klenow fragment, respectively. Crosses represent end labeling at the 5′ end by polynucleotide kinase. The U3, R, and U5 boundaries are noted as are the potential CAT box, promoter (Goldberg-Hogness box), polyadenylation signal (PAS), and PBS for $tRNA_{1,2}^{Lys}$. The closed arrowheads within the U3 box are two identical 43-bp repeats. The open arrowheads are two imperfect 42-bp repeats which are homologous to the 43-bp repeats.

mouse, and *Drosophila* (31, 41). We note that $tRNA_{1,2}^{Lys}$ (31, 41) and $tRNA_3^{Lys}$ (31) differ by 6 nucleotides in this 22-nucleotide region. Since the sequence of the $tRNA_{1,2}^{Lys}$ corresponding to the 19th base of SMRV PBS is methylated (31) and cannot pair with any base, the real PBS is likely to be only 18 bp like that of other known retroviruses. The tRNA-binding sequence serves conveniently as a landmark for the 3′ end of the LTR. Two nucleotides upstream from

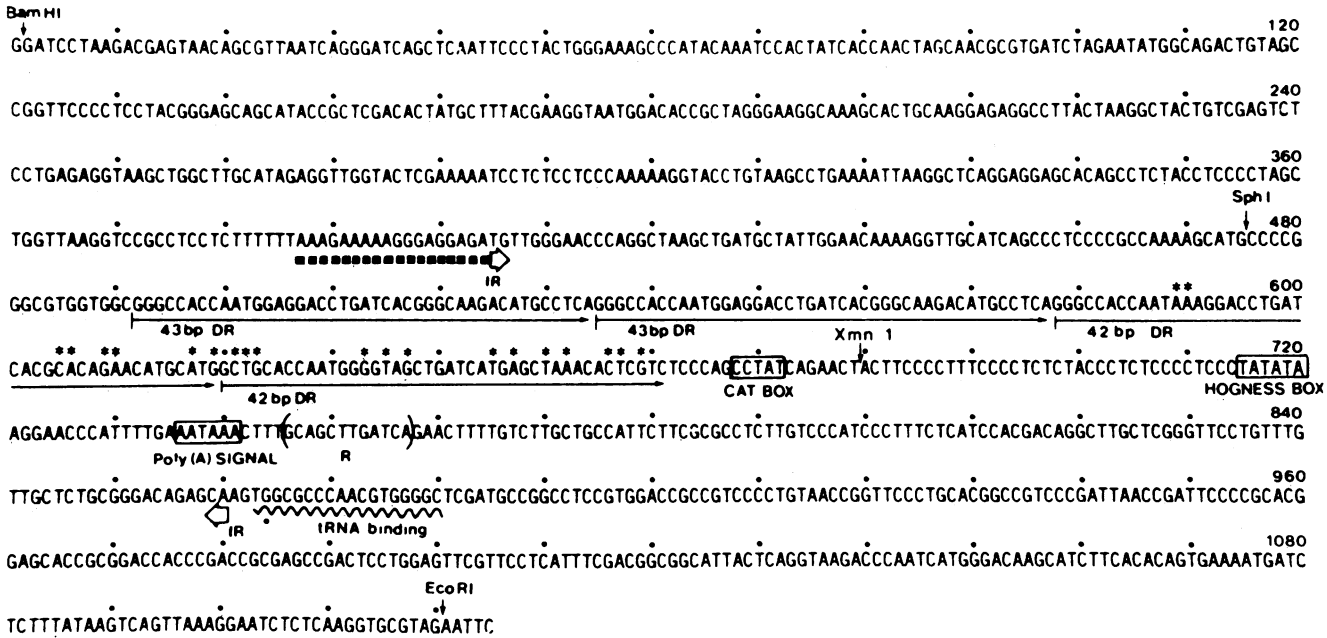

FIG. 2. Nucleotide sequence of SMRV LTR. The restriction sites for *Bam*HI, *Sph*I, *Xmn*I, and *Eco*I are indicated. The 43-bp direct repeats (DR) and the 42-bp direct repeats are underlined. Asterisks indicate the sequences in the 42-bp direct repeats which are different from the 43-bp direct repeats. The inverted repeats (IR) which demarcate the LTR are indicated by open arrows. The tRNA-binding sequence is indicated by a wavy line. The transcriptional signals including a CAT box, a Goldberg-Hogness (TATATA) box, and a polyadenylation signal are noted. The terminal repeat (R) in the retroviral RNA is bound by parentheses.

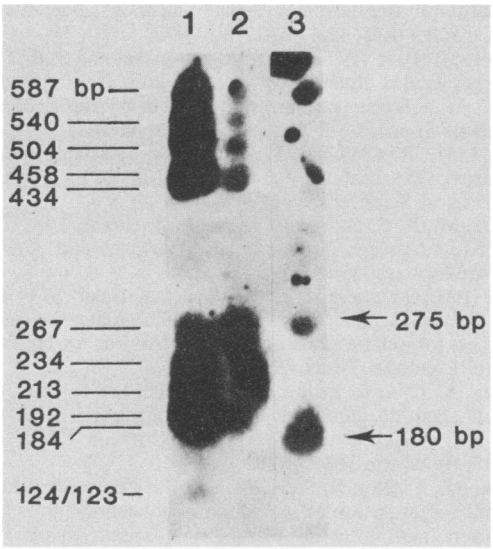

FIG. 3. Boundary of the 5′ end of SMRV LTR. Linear SMRV DNA was digested with *Xmn*I (lane 3), separated on a 2.5% agarose gel, and transferred to GeneScreen Plus nylon membrane. The membrane was then hybridized to an SMRV LTR probe (1.1-kbp *Bam*HI-*Eco*RI DNA fragment) which was nick translated (34) to a specific activity of 2×10^8 cpm/μg. The marker used was pBR322 DNA digested with *Hae*III and end labeled with polynucleotide kinase and [γ-^{32}P]ATP (lane 1, 5,000 cpm; lane 2, 1,000 cpm). The upper and lower arrows indicate DNA fragments derived from the left end and the right end of the SMRV LTR, respectively.

the PBS is a dinucleotide CA (nucleotides 859 to 860) which is also found invariably in the 3′ end of other LTRs and can, therefore, serve as a second reference point.

A comparison of restriction enzyme maps of linear SMRV and cloned pSMRV DNAs has suggested that the SMRV LTR is about 400 bp long (3). Approximately 400 bp upstream from the tRNA-binding sequence are two polypurine-rich regions (nucleotides 387 to 404 and 436 to 446) followed by the sequence TG which forms an inverted repeat with the 3′ end of the LTR. To determine which of the two polypurine-rich regions defines the 5′ boundary of LTR, we undertook detailed restriction mapping of the linear SMRV proviral DNA. Since the cloned pSMRV DNA has a unique *Xmn*I site which is located in the LTR region (data not shown), the linear SMRV DNA when digested with *Xmn*I should yield an 8.4-kbp DNA fragment, a 181-bp DNA fragment from the right-end LTR, and a 275- or 232-bp DNA fragment from the left-end LTR. The size of the last DNA fragment is dependent on which of the two polypurine-rich regions defines the 5′ end of the left-end LTR. The linear SMRV proviral DNA was isolated from acutely infected canine fetal thymus cells (Fcf2Th) as previously described (3). A 275-bp and a 180-bp DNA fragment were detected when *Xmn*I-digested linear SMRV DNA was fractionated on a 2.5% agarose gel, transferred to GeneScreen Plus nylon membrane (New England Nuclear Corp., Boston, Mass.), and hybridized to an SMRV LTR probe (43) (Fig. 3). These data suggested that the polypurine-rich sequence located at nucleotides 387 to 404 defines the 5′ end of the SMRV LTR.

At 88 bp inside the LTR is a pair of 43-bp repeats followed by a pair of 42-bp repeats. The two 43-bp direct repeats are identical. The sequence homologies between each of the 42-bp repeats and the 43-bp repeat are 81.4 and 70.0%, respectively. At 7 bp downstream is a variant of the CAT box, CCTAT (nucleotide 668 to 672), followed by the Goldberg-Hogness box, TATATA (nucleotides 715 to 720). We tentatively assigned the transcription initiation site to the sequence GC (nucleotides 746 to 747) which is 78 and 31 bp downstream from the CAT box and the TATA box, respectively. The polyadenylation signal, AATAAA, is observed

TABLE 1. Length (in base pairs) of retroviral LTR regions

Classification[a]	Retrovirus[b]	LTR[c]	U3	R	U5	PBS[d]	tRNA[e]	PAS[f]	Reference
A	IAP	350	210	85	55	18	AT-Phe	R	7, 15, 21, 29
B	MMTV	1,328	1,195	15	118	18	GC-Lys_3	U3	13, 14, 19
aC	RSV	335	234	21	80	18	TT-Trp	U3	37
mC	M-MuLV	594	449	68	77	23	TT-Pro	R	40
	SSV	508	362	69	77	18	TT-Pro	R	12
	BaEV	554	421	66	67	23	TT-Pro	R	45
D	SMRV	456	341	12	103	22	AG-$Lys_{1,2}$	U3	This work
Unclassified	BLV	530	201	238	91	18	AT-Pro	U3	10, 35
	HTLV-I	755	354	228	173	18	GT-Pro	U3	38
	HTLV-II	763	315	247	201	18	AT-Pro	U3	39, 42
	HTLV-III	634	453	98	83	18	G-Lys_3	R	26, 32, 36, 48
Human endogenous	Clone 4-1	495	NA[g]	NA	NA	18 (2)	TTTCT-Glu	NA	33
	ERV3	593	494	63	36	18 (1)	TTT-Arg	R	28
	RTVL-H	415	NA	NA	NA	18 (1)	TT-His	NA	23

[a] aC, Avian type C viruses; mC, mammalian type C viruses.
[b] IAP, intracisternal A particle; MMTV, mouse mammary tumor virus; RSV, Rous sarcoma virus; M-MuLV, Moloney murine leukemia virus; SSV, simian sarcoma virus; BaEV, baboon endogenous virus; BLV, bovine leukemia virus; HTLV, human T-cell leukemia virus.
[c] The convention was to consider the terminal TT bases near the PBS of type C viruses as part of the LTR. However, these two nucleotides do not form part of the inverted repeat in any of the LTRs of type A, type B, type D, and other unclassified retroviruses including HTLV-1, HTLV-II, HTLV-III, and BLV. Therefore, a more general nomenclature was adopted here to consider these two bases as positioned between the LTR and the PBS.
[d] The number of mismatches, if there are any, is shown in parentheses.
[e] The sequences between the 5′ LTR and the PBS are also listed.
[f] PAS shows the location of the polyadenylation signal.
[g] NA, data not available.

at nucleotides 736 to 741. With the identification of the AATAAA sequence, it was possible to assign the polyadenylation site CA at nucleotide 757 which is 16 bp downstream from the polyadenylation signal. These distances agree very well with the previously established distances for such signals (47). Thus, we were able to establish that the SMRV LTR is 456 bp long and is followed by a $tRNA_{1,2}^{Lys}$ PBS. The size of the LTR in type D retroviruses is different from those of other known retroviruses. The size of the terminal repetitive region (R) of SMRV viral RNA (12 bp) is the smallest among all the sequenced retroviral LTRs.

The U3 region of retroviral LTRs usually contains repetitive elements distal from the CAT and TATA promoter sequences. These repetitive sequences act to enhance RNA transcription. In this communication, we showed that the SMRV LTR contains 43-bp repeats which are repeated four times with different degrees of perfection (Fig. 2). We did not observe the enhancer core sequence $TGG\overset{AAA}{TTT}$ (22, 49) in our LTR direct repeats. However, the proximity of the tandem repeats to the canonical promoter elements strongly suggests that they function as enhancer sequences. Both the Moloney murine leukemia virus LTR and simian virus 40 early-gene promoter contain 72-bp repeats. Although the primary nucleotide sequences of the enhancer elements are not homologous, they are interchangeable to provide some enhancer function with different host-cell specificity (22). Like SMRV LTR, both the intracisternal A particle LTR (7) and the visna virus LTR (18) contain 43-bp direct repeats. However, their primary sequences are not noticeably homologous to the SMRV LTR direct repeats. Other than the size of the direct repeats, we did not detect any similarity between SMRV LTR and any other retroviral LTRs or any other published sequences. It will be of interest to test the host-cell specificity of the 43-bp repeats among these retroviral LTRs. Such an experiment will further our understanding toward the biological roles of the direct repeats in enhancer promotion.

The sequences flanking the 5′ LTR of the retroviral genome contain the binding site for a specific tRNA which serves as a primer for reverse transcription. Different genera of retroviruses use different tRNAs as primers (Table 1). Our data indicate that SMRV uses a novel PBS complementary to $tRNA_{1,2}^{Lys}$. The fact that type B and type D oncoviruses and human AIDS retroviruses share similar isoaccepting tRNA-binding sites may suggest a closer genetic interaction among these three genera of retroviruses. We have very recently established that human AIDS retrovirus is a member of the lentivirus subfamily (6). Consequently, it will be of interest to find out which Lysine-tRNA is used by the type D retroviruses associated with simian AIDS (11, 24, 44) and by the other lentiviruses as a primer for reverse transcription.

We thank S. Aaronson and S. Tronick for their support and N. Sarver for critical review of the manuscript. We are grateful to P. Luciw and colleagues for providing us their human AIDS-associated retrovirus sequence before publication. Plasmid clone pAT153 was provided to us by Rosita Gol. The excellent secretarial work of Nancy Mode and the photographic work of Karen Laufer are also highly appreciated.

LITERATURE CITED

1. **Barbacid, M., L. K. Long, and S. A. Aaronson.** 1980. Major structural proteins of type B, type C, and type D oncoviruses share interspecies antigenic determinants. Proc. Natl. Acad. Sci. USA **77**:72–76.
2. **Chen, H. R., and W. C. Barker.** 1984. Nucleotide sequences of the retroviral long terminal repeats and their adjacent regions. Nucleic Acids Res. **12**:1767–1778.
3. **Chiu, I.-M., P. R. Anderson, S. A. Aaronson, and S. R. Tronick.** 1983. Molecular cloning of the unintegrated squirrel monkey retrovirus genome: organization and distribution of related sequences in primate DNAs. J. Virol. **47**:434–441.
4. **Chiu, I.-M., R. Callahan, S. R. Tronick, J. Schlom, and S. A. Aaronson.** 1984. Major *pol* gene progenitors in the evolution of oncoviruses. Science **223**:364–370.
5. **Chiu, I.-M., R. C. C. Huang, and S. A. Aaronson.** 1985. Genetic relatedness between intracisternal A particles and other major oncovirus genera. Virus Res. **3**:1–11.
6. **Chiu, I.-M., A. Yaniv, J. E. Dahlberg, A. Gazit, S. F. Skuntz, S. R. Tronick, and S. A. Aaronson.** 1985. Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses. Nature (London) **371**:366–368.
7. **Christy, R. J., A. R. Brown, B. B. Gourlie, and R. C. C. Huang.** 1985. Nucleotide sequences of murine intracisternal A-particle gene LTRs have extensive variability within the R region. Nucleic Acids Res. **13**:289–302.
8. **Cohen, M., N. Rice, R. Stephens, and C. O'Connell.** 1982. DNA sequence relationship of the baboon endogenous virus genome to the genomes of other type C and type D retroviruses. J. Virol. **41**:801–812.
9. **Colcher, D., R. L. Heberling, S. S. Kalter, and J. Schlom.** 1977. Squirrel monkey retrovirus: an endogenous virus of a New World primate. J. Virol. **17**:294–301.
10. **Couez, D., J. Deschamps, R. Kettmann, R. M. Stephens, R. V. Gilden, and A. Burny.** 1984. Nucleotide sequence analysis of the long terminal repeat of integrated bovine leukemia provirus DNA and of adjacent viral and host sequences. J. Virol. **49**:615–620.
11. **Daniel, M. D., N. W. King, N. L. Letvin, R. D. Hunt, P. K. Sehgal, and R. C. Desrosiers.** 1984. A new type D retrovirus isolated from macaques with an immunodeficiency syndrome. Science **223**:602–605.
12. **Devare, S. G., E. P. Reddy, J. D. Law, and S. A. Aaronson.** 1982. Nucleotide sequence analysis of the long terminal repeat of integrated simian sarcoma virus: evolutionary relationship with other mammalian retroviral long terminal repeats. J. Virol. **42**:1108–1113.
13. **Donehower, L. A., A. L. Huang, and G. L. Hager.** 1981. Regulatory and coding potential of the mouse mammary tumor virus long terminal redundancy. J. Virol. **37**:226–238.
14. **Fasel, N., K. Pearson, E. Buetti, and H. Diggelmann.** 1982. The region of mouse mammary tumor virus DNA containing the long terminal repeat includes a long coding sequence and signals for hormonally regulated transcription. EMBO J. **1**:3–7.
15. **Hawley, R. G., M. J. Shulman, and N. Hozumi.** 1984. Transposition of two different intracisternal A particle elements into an immunoglobulin kappa-chain gene. Mol. Cell. Biol. **4**: 2565–2572.
16. **Hayward, W. S., B. G. Neel, and S. M. Astrin.** 1981. Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis. Nature (London) **311**:414–416.
17. **Heberling, R. L., S. T. Barker, S. S. Kalter, G. C. Smith, and R. J. Helmke.** 1977. Oncornavirus: isolation from a squirrel monkey (*Saimiri sciureus*) lung culture. Science **195**:289–292.
18. **Hess, J. L., J. F. Clements, and O. Narayan.** 1985. *Cis*- and *trans*-acting transcriptional regulation of visna virus. Science **229**:482–485.
19. **Kennedy, N., G. Knedlitschek, B. Groner, N. E. Hynes, P. Herrlich, R. Michalides, and A. J. J. Van Ooyen.** 1982. Long terminal repeats of endogenous mouse mammary tumor virus contain a long open reading frame which extends into adjacent sequences. Nature (London) **295**:622–624.
20. **Kuff, E. L., A. Feenstra, K. Lueders, G. Rechavi, D. Givol, and E. Canaani.** 1983. Homology between endogenous viral LTR and sequence inserted in an activated cellular oncogene. Nature (London) **302**:547–548.
21. **Kuff, E. L., A. Feenstra, K. Lueders, L. Smith, R. Hawley, N. Hozumi, and M. Shulman.** 1983. Intracisternal A-particle genes as movable elements in the mouse genome. Proc. Natl. Acad.

Sci. USA 80:1992–1996.
22. **Laimins, L. A., G. Khoury, C. Gorman, B. Howard, and P. Gruss.** 1982. Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA 79:6453–6457.
23. **Mager, D. L., and P. S. Henthorn.** 1984. Identification of a retrovirus-like repetitive element in human DNA. Proc. Natl. Acad. Sci USA 81:7510–7514.
24. **Marx, P. A., D. H. Maul, K. G. Osborn, N. W. Lerche, P. Moody, L. J. Lowenstine, R. V. Henrickson, L. O. Arthur, R. V. Gilden, M. Gravell, W. T. London, J. L. Sever, J. A. Levy, R. J. Munn, and M. B. Gardner.** 1984. Simians AIDS: isolation of a type D retrovirus and transmission of the disease. Science 223:1083–1086.
25. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499–560.
26. **Muesing, M. A., D. H. Smith, C. D. Cabradilla, C. V. Benton, L. A. Lasky, and D. J. Capon.** 1985. Nucleic acid structure and expression of human AIDS/lymphadenopathy retrovirus. Nature (London) 313:450–458.
27. **Nilsen, T. W., P. A. Maroney, R. G. Goodwin, F. M. Rottman, L. B. Crittenden, M. A. Raines, and H.-J. Kung.** 1985. c-*erb*B activation in ALV-induced erythroblastosis: novel RNA processing and promoter insertion result in expression of an amino-truncated EGF receptor. Cell 41:719–726.
28. **O'Connell, C. D., and M. Cohen.** 1984. The long terminal repeat sequences of a novel human endogenous retrovirus. Science 26:1204–1206.
29. **Ono, M., and H. Ohishi.** 1983. Long terminal repeat sequences of intracisternal A particle genes in the Syrian hamster genome: identification of $tRNA^{Phe}$ as a putative primer tRNA. Nucleic Acids Res. 11:7169–7179.
30. **Payne, G. S., J. M. Bishop, and H. E. Varmus.** 1982. Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. Nature (London) 295:209–214.
31. **Raba, M., K. Limburg, M. Burghagen, J. R. Katze, M. Simsek, J. E. Heckman, U. L. Rajbhandary, and H. J. Gross.** 1979. Nucleotide sequence of three iso-accepting lysine tRNAs from rabbit liver and SV40 transformed mouse fibroblasts. Eur. J. Biochem. 97:305–318.
32. **Ratner, L., W. Haseltine, R. Patarca, K. J. Livak, B. Starcich, S. F. Josephs, E. R. Doran, J. A. Rafalski, E. A. Whitehorn, K. Baumeister, L. Ivanof, S. R. Petteway, Jr., M. L. Pearson, J. A. Lautenberger, T. S. Papas, J. Ghrayeb, N. T. Chang, R. C. Gallo, and F. Wong-Staal.** 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature (London) 313:277–284.
33. **Repaske, R., P. E. Steele, R. R. O'Neill, A. B. Rabson, and M. A. Martin.** 1985. Nucleotide sequence of a full-length human endogenous retroviral segment. J. Virol. 54:764–772.
34. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. 113:237–251.
35. **Sagata, N., T. Yasunaga, Y. Ogawa, J. Tsuzuku-Kawamura, and Y. Ikawa.** 1984. Bovine leukemia virus: unique structural features of its long-terminal repeats and its evolutionary relationship to human T-cell leukemia virus. Proc. Natl. Acad Sci. USA 81:4741–4745.
36. **Sanchez-Pescador, R., M. D. Power, P. J. Barr, K. S. Steimer, M. M. Stempien, S. L. Brown-Shimer, W. W. Gee, A. Renard, A. Randolph, J. A. Levy, D. Dina, and P. A. Luciw.** 1985. Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2). Science 227:484–492.
37. **Schwartz, D., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell 32:853–869.
38. **Seiki, M., S. Hattori, Y. Hirayama, and M. Yoshida.** 1983. Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. Proc. Natl. Acad. Sci USA 80:3618–3622.
39. **Shimotohno, K., D. W. Golde, M. Miwa, T. Sugimura, and I. S. Y. Chen.** 1984. Nucleotide sequence analysis of the long terminal repeat of human T-cell leukemia virus type II. Proc. Natl. Acad. Sci. USA 81:1079–1083.
40. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) 293:543–548.
41. **Silverman, S., I. C. Gillam, G. M. Tener, and D. Soell.** 1979. The nucleotide sequence of lysine $tRNA_2$ from *Drosophila*. Nucleic Acids Res. 6:435–442.
42. **Sodroski, J., M. Trus, D. Perkins, R. Patarca, F. Wong-Staal, E. Gelmann, R. Gallo, and W. A. Haseltine.** 1984. Repetitive structure in the long-terminal repeat element of a type II human T-cell leukemia virus. Proc. Natl. Acad. Sci. USA 81: 4617–4621.
43. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503–517.
44. **Stromberg, K., R. E. Benveniste, L. O. Arthur, H. Rabin, W. E. Giddens, Jr., H. D. Ochs, W. R. Morton, and C.-C. Tsai.** 1984. Characterization of exogenous type D retrovirus from a fibroma of a macaque with simian AIDS and fibromatosis. Science 224:289–292.
45. **Tamura, T., M. Noda, and T. Takano.** 1981. Structure of the baboon endogenous virus genome: nucleotide sequences of the long terminal repeat. Nucleic Acids Res. 9:6615–6626.
46. **Twigg, A. J., and D. Sherrat.** 1980. *Trans*-complementable copy-number of mutants of plasmid ColE1. Nature (London) 283:216–218.
47. **Varmus, H. E.** 1982. Form and function of retroviral proviruses. Science 216:812–820.
48. **Wain-Hobson, S., P. Sonigo, O. Panos, S. Cole, and M. Alizon.** 1985. Nucleotide sequence of the AIDS virus, LAV. Cell 40:9–17.
49. **Weiher, H., M. Konig, and P. Gruss.** 1983. Multiple point mutations affecting the simian virus 40 enhancer. Science 219:626–631.
50. **Weiss, R., N. Teich, H. Varmus, and J. Coffin.** 1982. Molecular biology of tumor viruses (2nd ed.): RNA tumor viruses, p. 25–207. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
51. **Wilbur, W. J., and D. J. Lipman.** 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA 80:726–730.