JOURNAL OF VIROLOGY, June 1982, p. 1108–1113
0022-538X/82/061108-06$02.00/0

Vol. 42, No. 3

NOTES

Nucleotide Sequence Analysis of the Long Terminal Repeat of Integrated Simian Sarcoma Virus: Evolutionary Relationship with Other Mammalian Retroviral Long Terminal Repeats

SUSHILKUMAR G. DEVARE,* E. PREMKUMAR REDDY, J. DORIA LAW, AND STUART A. AARONSON

Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, Maryland 20205

Received 13 November 1981/Accepted 1 February 1982

Nucleotide sequence analysis of the long terminal repeat (LTR) of the integrated simian sarcoma virus showed that the simian sarcoma virus LTR comprised 504 nucleotides with an inverted repeat of seven bases at its 5′ and 3′ termini. At the site of simian sarcoma virus integration, cellular flanking sequences adjacent to the proviral LTR contained a direct repeat of four bases. A 13-base sequence after the 5′ LTR was found to be complementary to prolyl tRNA, suggesting that $tRNA^{Pro}$ may serve as the primer for reverse transcription of simian sarcoma virus RNA. The U_5 and R regions, derived respectively from the 5′ end and terminally redundant sequences of the viral RNA, were found to have similar organization and sequence homology close to that of Moloney murine sarcoma virus or Moloney murine leukemia virus. These results indicate that regions within LTRs with known functionally important sequences have been most well conserved during retrovirus evolution.

An important structural feature of the retroviral genome is the appearance of two long terminal repeats (LTRs) at both 5′ and 3′ ends of the proviral genome (11, 16, 17, 36). These LTRs bear striking similarities to the terminal repeats of procaryotic transposable elements and contain promoter-like sequences as well as mRNA capping and polyadenylation signals (7, 31, 37, 40, 41, 44). The recent demonstration that molecularly cloned LTR sequences have the capacity to initiate transcription in vitro of avian and murine type C retroviruses has provided direct evidence that transcription promotion signals in the LTR have the capacity to act as efficient promoters (27, 45). Moreover, the promoter signals in the LTRs have been shown to regulate and initiate transcription when present either at 5′ or 3′ end of the functional gene (4, 14) and even in opposite orientations (H. E. Varmus, R. Swanstrom, N. Quintrell, G. Payne, R. Parker, S. Oritz, H. Oppermann, G. Mardon, J. Majors, P. Luciw, S. Hughes, W. DeLorbe, M. Capecchi, and J. M. Bishop, Int. Congr. Virol. Abstr. Vth, Strasbourg, France, p. 12, 1981).

Detailed analysis of the structure of the LTR has been limited to a few avian and mammalian retroviruses. Their sequence complexity ranges from around 300 bases for avian viruses to around 1,200 bases for mouse mammary tumor virus. The LTRs of mammalian type C viruses appear to comprise around 600 bases. Among mammalian type C retroviruses, complete nucleotide sequence data is available only for Moloney murine leukemia virus (38) and Moloney murine sarcoma virus (MSV) (30), which derived its LTR from Moloney murine leukemia virus.

In view of the known evolutionary relationships among mammalian type C retroviruses as well as the important functional role of the LTR in the virus life cycle, it was of interest to analyze and compare by primary sequence analysis different mammalian type C retroviral LTRs. Simian sarcoma virus (SSV) has recently been molecularly cloned and shown to comprise some sequences of a mammalian type C helper virus, designated simian sarcoma associated virus, and a unique sequence cell-derived gene (10, 33) of primate origin (34). In the present report, we undertook primary sequence analysis of the SSV LTR and have compared its sequence organization with that of other type C viral LTRs.

Localization of LTR sequences within integrated SSV DNA and sequencing strategy. The isolation of a molecular clone of integrated SSV DNA from SSV-11 nonproductively transformed normal rat kidney cells has been described previously (33). Charon 16A λ phage containing the 5.8 kilobase pair of the SSV-11 provirus, designated λ-SSV-11 Cl 1, was propagated in *Escherichia coli* K-12 DP50 *supF*. λ-SSV-11 Cl 1 DNA was excised with *Eco*RI to

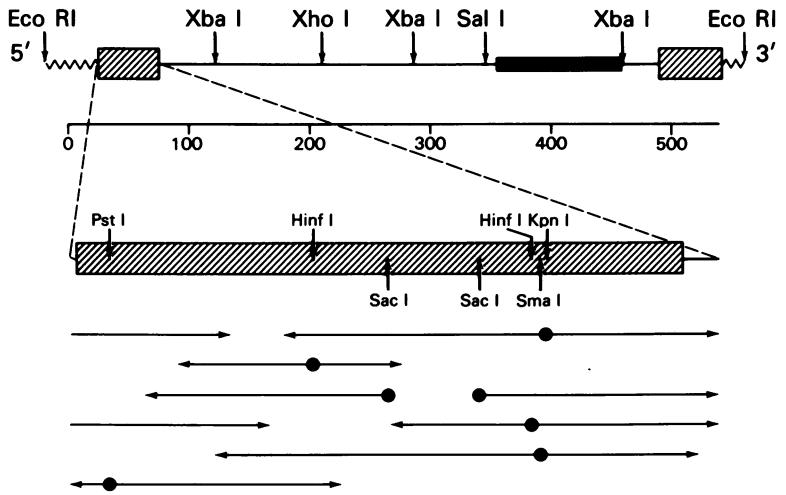

FIG. 1. Restriction enzyme map and strategy for sequencing SSV LTR. The 5' LTR sequence was derived by using the restriction sites on the diagrammatic map. The 5' ends were labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase (24). The 3' ends were labeled with [^{32}P]cordycepin triphosphate and terminal transferase (35). The 5' or 3' end-labeled DNA fragments were digested with appropriate restriction endonucleases and isolated by agarose or polyacrylamide gel electrophoresis. The nucleotide sequence was determined by the procedure of Maxam and Gilbert (24). ●, Labeled end of each fragment. The extent and direction of sequencing are indicated by the arrows.

obtain the viral insert, purified by agarose gel electrophoresis, and ligated to *Eco*RI-digested pBR322. *E. coli* C600 was transfected (42) with SSV DNA ligated to pBR322 and plated on agar containing NZ-amine type A (Hum-ko-Sheffield, Lynnhurst, N.J.) and yeast extract (Difco Laboratories, Detroit, Mich.). Colonies containing SSV DNA were identified by in situ hybridization (2) with nick-translated SSV probe. Plasmid DNA isolated from chloramphenicol-treated cells (3) was used for sequence analysis.

The restriction map of proviral SSV DNA was constructed based on the restriction endonuclease cleavage products as visualized on agarose gels after staining with ethidium bromide (33) and later confirmed by the partial restriction mapping technique of Smith and Birnstiel (39). The location of LTR sequences within the SSV genome was defined by the presence of an identical constellation of restriction sites near either end of SSV DNA. Thus, the presence of *Pst*I, *Hin*fI, *Sac*I, and *Kpn*I sites at either end of the DNA roughly defined the location of the LTR in the SSV integrated genome (33). Figure 1 shows the location of these restriction sites within the 5' LTR. The 5' to 3' orientation of the molecule was defined as described previously (33). We analyzed both 5' and 3' LTR as well as host flanking sequences. However, for the sake of simplicity, the strategy for sequencing only the 5' LTR is shown in Fig. 1. The sequence of both strands of DNA was determined, and all restriction cleavage sites were confirmed by sequence analysis.

Sequence organization of the SSV LTR and its adjacent regions. The nucleotide sequence of 5' and 3' LTRs as well as cellular flanking regions is shown in Fig. 2. The SSV LTR comprised 504 bases and possessed an imperfect inverted repeat of 7 bases with a sequence 5'-T-G-A-A-G-G-A-3' at its 5' and 3' terminus (at position 1-7 and 498-504). A four-base host sequence (T-A-A-T), apparently repeated at the site of virus integration, was detected. In addition, putative regulatory signals within the LTR can be summarized as follows (Fig. 3). (i) The adenine-thymine-rich eucaryotic transcription initiation signal, commonly known as the Hogness box, was detected at 331-337 with the sequence 5'-T-A-T-A-A-A-A-3'. Some of the characteristic sequences around this Hogness box promoter consist of a pentamer 5'-C-C-A-A-T-3' (Cat box) around 40 bases preceding the promoter sequence (9). Thus, the presence of a Cat box at 290-294 indicates that the sequence at 331-337 may be the promoter for transcription of viral RNA. (ii) Another important feature is the capping site for the viral genomic RNA synthesis. It is known that RNA transcription is initiated around 22 to 25 bases from the promoter sequence (7, 31, 40, 44). Thus, the G-C-G sequence at position 359-361 is the most likely site for initiation of viral RNA synthesis. (iii) The hexanucleotide AATAAA precedes the polyadenylation site in several eucaryotic genes (29). Such a polyadenylation signal was found at position 407-412 in the SSV LTR. At 16 bases downstream from AATAAA, a dinucleotide C-A (at 428-429) may represent the preferred site for polyadenylation (29). Based on the putative signals for initiation and termination of transcription, the SSV LTR is composed of a track of 358

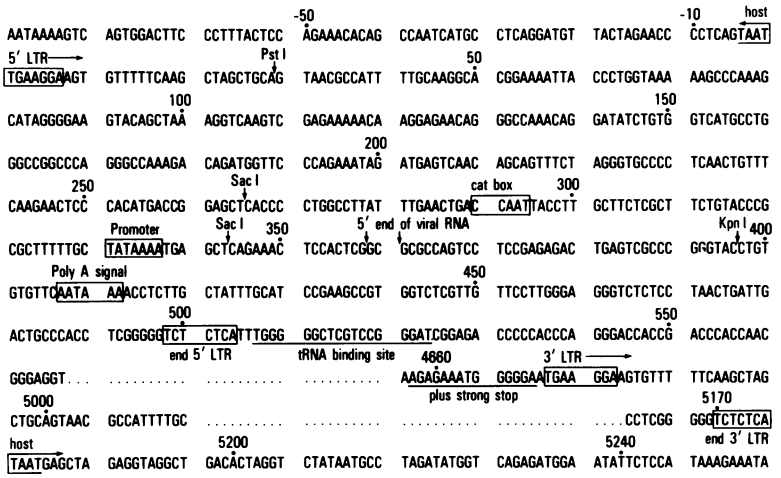

FIG. 2. Sequence of SSV LTR and its adjacent regions of integrated proviral genome. The sequence of LTR and its flanking regions proceeding in the 5′ to 3′ direction with the same polarity as that of the SSV genomic RNA is shown. The major structural features of the LTR are indicated.

nucleotides derived from the 3′ end of viral RNA, directly followed to the right by a stretch of about 146 nucleotides derived from the 5′ end of the viral genome. It is known that retroviral genomic RNA contains a direct terminally repeated sequence of 50 to 60 nucleotides, termed R or *trs* (5, 6). Thus, if the C-A signal at position 428-429 corresponds to the 3′ end of SSV RNA, the sequence of 69 to 71 nucleotides between this CA and the GCG triplet representing the 5′ end of the viral RNA should constitute the R region of SSV (Fig. 2 and 3). (iv) Retroviral genomes contain a tRNA binding site immediately following the LTR (13, 43). This site serves as a primer for synthesis of viral DNA. Sequence analysis of the SSV genome showed a stretch of 18 bases at position 507-524 (5′-T-G-G-G-G-G-C-T-C-G-T-C-C-G-G-G-A-T-3′) which was complementary to the 3′ sequence of prolyl tRNA (28). These results suggest that, as with Moloney MSV (7, 31), prolyl tRNA may serve as primer for reverse transcription of SSV. It is of interest to note that Moloney MSV (7, 30), a mammalian type C virus, and spleen necrosis virus (37), a retrovirus of avian origin belonging to the reticuloendotheliosis virus group, both prefer $tRNA^{Pro}$ as a primer for initiation of plus-strong-stop DNA. In contrast, typical type C avian viruses utilize $tRNA^{Trp}$ primer (43). Previous studies have demonstrated immunological cross-reaction of the reverse transcriptase (25) and major structural proteins (1, 18) of the reticuloendotheliosis virus group with the analogous proteins of mammalian type C retroviruses. Thus, $tRNA^{Pro}$ binding site may be another marker for evolutionary linkage among these retroviruses.

Preceding the 3′ LTR, we detected a purine-rich sequence of 16 bases (at position 4656 to 4671). Similar purine-rich sequences have been found at analogous positions in murine mammary tumor virus (8), avian sarcoma virus (41),

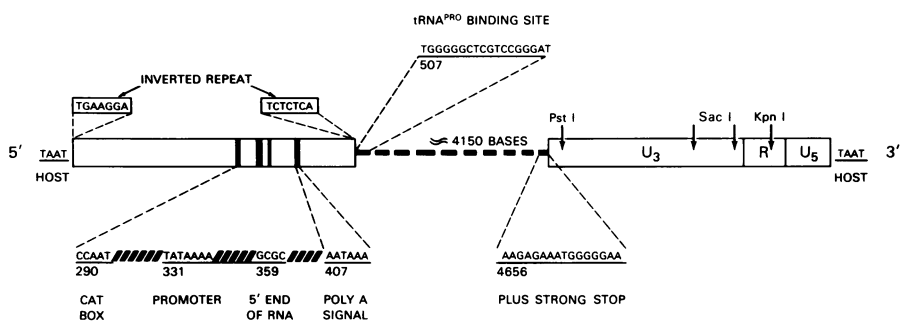

FIG. 3. Summary of the major features of SSV LTR and its adjacent sequences. Important features of the SSV LTR sequence, including putative signals for promoter and polyadenylation; 5′ terminus of SSV genomic RNA, $tRNA^{Pro}$ binding site, and plus-strong-stop DNA; U_3 sequences unique to the 3′ end of genomic RNA; terminally redundant sequences of genomic RNA (R); sequences unique to 5′ end of genomic RNA (U_5); and position of an imperfect inverted repeat have been illustrated diagrammatically.

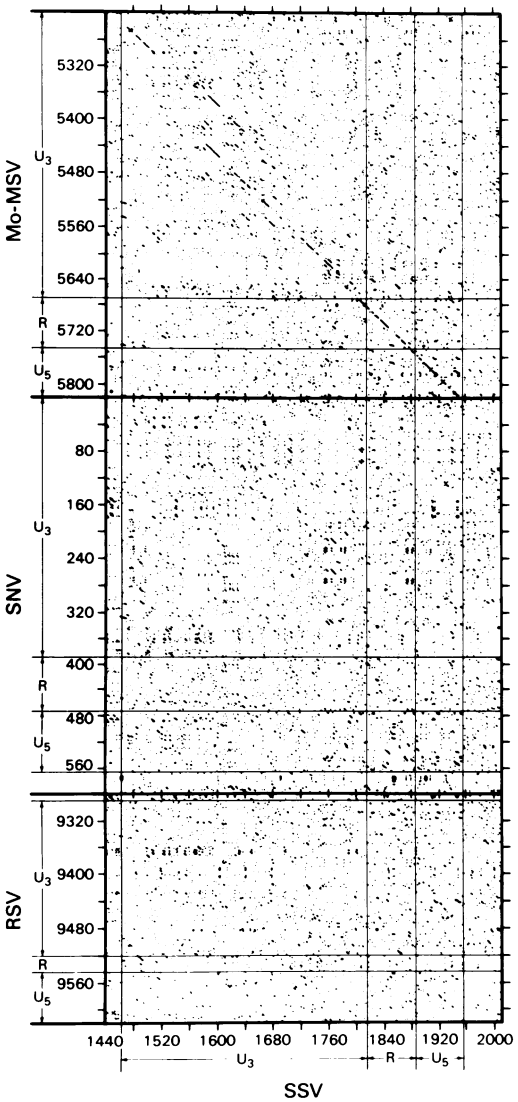

FIG. 4. Homology between the LTRs of Moloney MSV, spleen necrosis virus, Rous sarcoma virus, and SSV. The sequence homology between LTRs of Moloney MSV (7, 30), spleen necrosis virus (38), and Rous sarcoma virus (Schwartz et al., in press) was determined by a two-dimensional dot matrix homology comparison computer program developed by J. V. Maizel, Jr. (19, 23). Each dot represents the center of a three-base homology between two LTRs. The grid allows comparison of the homology represented by a line at −45° to the reference to U_3, R, and U_5 region of LTR. The numbers in parentheses represent the positions of nucleotide bases in the 5′ or 3′ LTR.

Moloney MSV (7), Moloney murine leukemia virus (37), avian endogenous virus *ev*1 (15), and avian myeloblastosis virus (Rushlow et al., personal communication). Though not conclusively proven, it has been proposed that this purine-rich sequence may serve as a primer for plus-strong-stop DNA (41).

Comparison of LTR sequences of Moloney MSV and SSV. We next compared the SSV LTR sequence with that of Moloney MSV, another mammalian type C virus. To do so, we utilized the two-dimensional dot matrix homology comparison computer program developed by J. V. Maizel, Jr. (19, 23). In this program, two sequences are compared with each other base by base. Thus, each row or column represents the comparison of a single base of one sequence with the corresponding base of the other. A dot represents a three-base homology, and homology of a long stretch of sequence is observed as a line of dots extending at a −45° angle. Duplication regions are observed as a second line parallel to the first, whereas deletion or insertions are observed as a line parallel to the major line but shifted up or down by the number of bases not found in the other gene.

Considerable homology was observed between R and U_5 regions of SSV and Moloney MSV LTR sequences (Fig. 4, top panel). These results indicate that the DNA sequences coding for the 5′ ends of their RNAs possess significant sequence homology and support the contention that these sequences are well conserved (22). Another finding revealed by this analysis was the deletion of around 72 base pairs in the U_3 region of the SSV LTR compared with the analogous region of Moloney MSV. This was detected by the presence of a line parallel to the major line but shifted up by around 70 bases (Fig. 4, top panel). The Moloney MSV LTR contains an almost perfect duplication of 72 base pairs in the U_3 region (7, 31). Although the function of these sequences in the life cycle of MSV is not yet known, it is interesting to note that similar tandem repeats have been shown to occur on the right side of the origin for DNA replication of simian virus 40 (32). These sequences appear to play an important role in promotion of the early transcription of the simian virus 40 genome (12, 26). The absence of such duplication in the SSV LTR does not impair its function, as cloned SSV DNA transforms cells in tissue culture at a very high efficiency. Moreover, the SSV genome can be rescued from such transformants by superinfection with a type C helper virus (33).

In contrast to the conservation of U_5 and R regions, considerable divergence was observed between the U_3 regions of SSV and Moloney MSV (Fig. 4). In fact, sequence analysis of independent DNA clones of Moloney MSV M1 (7) and MSV-124 (30, 31), as well as Moloney MuLV (38, 40, 44), from which the MSV LTRs were derived, has revealed as much as 5 to 10% divergence in their respective U_3 sequences with no appreciable differences in U_5 or R regions. Moreover, comparison of the sequences of a molecular clone of simian sarcoma-associated virus with that of SSV revealed a similar degree of divergence in U_3 sequences as compared with

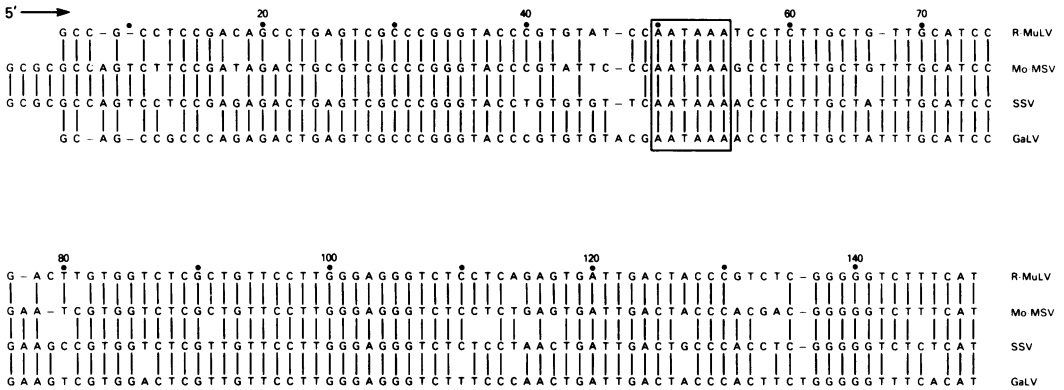

FIG. 5. Comparison of DNA sequences corresponding to the 5′ terminal region for SSV, Moloney MSV (7, 30), Rauscher murine leukemia virus, and gibbon ape leukemia virus (21). The sequences have been aligned to maximize homology. The putative signal for polyadenylation is indicated by the box.

U_5 and R regions. These findings suggest that the U_5 and R regions have functionally important sequences as compared with U_3.

We also utilized dot matrix analysis to compare the SSV LTR sequence with that reported for avian spleen necrosis virus (37), a virus immunologically related to mammalian type C retroviruses (1, 18), and Rous sarcoma virus (D. Schwartz, R. Tizard, and W. Gilbert, Cell, in press), a prototype avian type C virus. Neither of these viral LTRs showed appreciable sequence homology with the LTR sequence of SSV (Fig. 4).

Conservation of 5′-terminal nucleotide sequences in mammalian retroviruses. Although complete LTR sequence analysis has been performed for only a few mammalian retroviruses, sequence analysis has been reported for the short stretch of nucleotides localized to the 5′ end of viral RNA designated as strong-stop DNA (13, 43). Comparison of DNA sequences corresponding to the 5′ end of SSV RNA with those of Moloney MSV (7, 31), Rauscher murine leukemia virus, and gibbon ape leukemia virus (21) showed that these sequences are conserved among these mammalian retroviruses (Fig. 5). These results are consistent with evidence that SSV and gibbon ape leukemia virus have an ancestral rodent origin (20). Sequence analysis of strong-stop DNA of other mammalian retroviruses of rodent, feline, and primate origin have also indicated significant nucleic acid homology in their 5′ terminal regions (21, 22). In contrast, avian retroviruses do not share sequence homology with mammalian retroviruses in this limited region of the LTR (22).

In the present studies, a comparison of 5′ terminal sequences of SSV RNA with those of other mammalian and avian retroviruses (21, 22) did not show statistically significant relationships (data not shown). These results, in combination with the known lack of detectable immunological cross-reaction among the structural proteins of avian and mammalian type C retroviruses, argue that these major type C virus groups have diverged very markedly from a common progenitor or were generated through different evolutionary pathways.

We thank K. C. Robbins, whose SSV DNA clone we utilized, and S. R. Tronick for help and advice during the course of these studies. The dot matrix computer program was generously provided by J. V. Maizel, Jr.

ADDENDUM IN PROOF

Van Beveren et al. have recently published the LTR sequence of an additional Moloney MSV-124 DNA clone (C. Van Beveren, F. van Straaten, J. A. Galleshaw, and I. M. Verma, Cell **27**:97–108, 1981). The LTR sequence of this clone is almost identical, even in its U_3 region, with that of one of our previously published Moloney MSV-124 clones (31).

LITERATURE CITED

1. **Barbacid, M., E. Hunter, and S. A. Aaronson.** 1979. Avian reticuloendotheliosis viruses: evolutionary linkage with mammalian type C retroviruses. J. Virol. **30**:508–514.
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. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
4. **Blair, D. G., W. L. McClements, M. K. Oskarsson, P. J. Fischinger, and G. F. Vande Woude.** 1980. Biological activity of cloned Moloney sarcoma virus DNA: terminally redundant sequences may enhance transformation efficiency. Proc. Natl. Acad. Sci. U.S.A. **77**:3504–3508.
5. **Coffin, J. M., T. C. Hageman, A. M. Maxam, and W. A. Haseltine.** 1978. Structure of the genome of Moloney murine leukemia virus: a terminally redundant sequence. Cell **13**:761–773.
6. **Coffin, J. M., and W. A. Haseltine.** 1977. Terminal redundancy and origin of replication of Rous sarcoma virus RNA. Proc. Natl. Acad. Sci. U.S.A. **74**:1908–1912.
7. **Dhar, R., W. L. McClements, L. W. Enquist, and G. F. Vande Woude.** 1980. Nucleotide sequence of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. Proc. Natl. Acad. Sci. U.S.A.

77:3937–3941.
8. **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.
9. **Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. K. DeRiel, B. G. Forget, S. M. Weissman, J. L. Slightom, A. E. Blechl, O. Smithies, F. E. Baralle, C. C. Shoulders, and N. J. Proudfoot.** 1980. The structure and evolution of the human β-globin gene family. Cell **21**:653–668.
10. **Gelmann, E. P., F. Wong-Staal, R. A. Kramer, and R. C. Gallo.** 1981. Molecular cloning and comparative analyses of the genomes of simian sarcoma virus and its associated helper virus. Proc. Natl. Acad. Sci. U.S.A. **78**:3373–3377.
11. **Gilboa, E., S. Goff, A. Shields, F. Yoshimura, S. Mitra, and D. Baltimore.** 1979. In vitro synthesis of a 9 kbp terminally redundant DNA carrying the infectivity of Moloney murine leukemia virus. Cell **16**:863–874.
12. **Gruss, P., R. Dhar, and G. Khoury.** 1981. Simian virus 40 tandem repeated sequences as an element of the early promoter. Proc. Natl. Acad. Sci. U.S.A. **78**:943–947.
13. **Haseltine, W. A., D. G. Kleid, A. Panet, E. Rothenberg, and D. Baltimore.** 1976. Ordered transcription of RNA tumor virus genomes. J. Mol. Biol. **196**:109–131.
14. **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) **290**:475–480.
15. **Hishinuma, F., P. J. DeBona, S. Astrin, and A. M. Skalka.** 1981. Nucleotide sequence of acceptor site and termini of integrated avian endogenous provirus evl: integration creates a 6 bp repeat of host DNA. Cell **23**:155–164.
16. **Hsu, T. W., J. L. Sabran, G. E. Mark, R. V. Guntaka, and J. M. Taylor.** 1978. Analysis of unintegrated avian RNA tumor virus double-stranded DNA intermediates. J. Virol. **28**:810–818.
17. **Hughes, S. H., P. R. Shank, D. H. Spector, H. J. Kung, J. M. Bishop, H. E. Varmus, P. K. Vogt, and M. L. Brietman.** 1978. Proviruses of avian sarcoma virus are terminally redundant, co-extensive with unintegrated linear DNA and integrated at many sites. Cell **15**:1397–1410.
18. **Hunter, E., A. S. Bhown, and J. C. Bennett.** 1978. Amino-terminal amino acid sequence of the major structural polypeptides of avian retroviruses: sequence homology between reticuloendotheliosis virus p30 and p30s of mammalian retroviruses. Proc. Natl. Acad. Sci. U.S.A. **75**:2708–2712.
19. **Konkel, D. A., J. V. Maizel, Jr., and P. Leder.** 1979. The evolution and sequence comparison of two recently diverged mouse chromosomal β-globin genes. Cell **18**:865–873.
20. **Lieber, M. M., C. J. Sherr, G. J. Todaro, R. E. Benveniste, R. Callahan, and H. G. Coon.** 1975. Isolation from the Asian mouse *Mus caroli* of an endogenous type C virus related to infectious primate type C virus. Proc. Natl. Acad. Sci. U.S.A. **72**:2315–2319.
21. **Lovinger, G. G., and G. Schochetman.** 1979. 5'-terminal nucleotide sequences of the Rauscher leukemia virus and gibbon ape leukemia virus genomes exhibit a high degree of correspondence. J. Virol. **32**:803–811.
22. **Lovinger, G. G., and G. Schochetman.** 1980. 5'-terminal nucleotide sequences of type C retroviruses: features common to noncoding sequences of eucaryotic messenger RNAs. Cell **20**:441–449.
23. **Maizel, J. V., Jr., and R. P. Lenk.** 1981. Enhanced graphic matrix analysis of nucleic acid and protein sequences. Proc. Natl. Acad. Sci. U.S.A. **78**:7665–7669.
24. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.
25. **Moelling, K., M. Gelderblom, G. Pauli, R. Friis, and H. Bauer.** 1975. A comparative study of the avian reticuloendotheliosis virus: relationship to murine leukemia virus and viruses of the avian sarcoma-leukosis complex. Virology **65**:546–557.
26. **Moreau, P., R. Hen, B. Wasylyk, R. Everett, M. P. Gaub, and P. Chambon.** 1981. The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Res. **9**:6047–6068.
27. **Ostrowski, M. C., D. Berard, and G. L. Hager.** 1981. Specific transcriptional initiation in vitro on murine type C retrovirus promoters. Proc. Natl. Acad. Sci. U.S.A. **78**:4485–4489.
28. **Peters, G., and J. E. Dahlberg.** 1979. RNA-directed DNA synthesis in Moloney murine leukemia virus: interaction between the primer tRNA and the genome RNA. J. Virol. **31**:398–407.
29. **Proudfoot, N. J., and G. G. Brownlee.** 1974. Sequence at the 3' end of the globin mRNA shows homology with immunoglobulin light chain mRNA. Nature (London) **252**:359–362.
30. **Reddy, E. P., M. J. Smith, and S. A. Aaronson.** 1981. Complete nucleotide sequence and organization of the Moloney murine sarcoma virus genome. Science **214**:445–450.
31. **Reddy, E. P., M. J. Smith, E. Canaani, K. C. Robbins, S. R. Tronick, S. Zain, and S. A. Aaronson.** 1980. Nucleotide sequence analysis of the transforming region and large terminal redundancies of Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **77**:5234–5238.
32. **Reddy, V. B., B. Thimmappaya, R. Dhar, K. N. Subramanian, B. S. Zain, J. Pan, C. L. Celma, and S. M. Weissman.** 1978. The genome of simian virus 40. Science **200**:494–502.
33. **Robbins, K. C., S. G. Devare, and S. A. Aaronson.** 1981. Molecular cloning of integrated simian sarcoma virus: genome organization of infectious DNA clones. Proc. Natl. Acad. Sci. U.S.A. **78**:2918–2922.
34. **Robbins, K. C., R. L. Hill, and S. A. Aaronson.** 1982. Primate origin of the cell-derived sequences of simian sacroma virus. J. Virol. **41**:721–725.
35. **Roychoudhury, R., and R. Wu.** 1980. Terminal transferase-catalyzed addition of nucleotides to the 3' termini of DNA. Methods Enzymol. **65**:43–62.
36. **Shank, P. R., S. H. Hughes, H. J. Kung, J. E. Majors, N. Quintrell, R. V. Guntaka, J. M. Bishop, and H. E. Varmus.** 1978. Mapping unintegrated avian sarcoma virus DNA: termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. Cell **15**:1383–1395.
37. **Shimotohno, K., S. Mizutani, and H. M. Temin.** 1980. Sequence of retrovirus provirus resembles that of bacterial transposable elements. Nature (London) **285**:550–554.
38. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) **293**:543–548.
39. **Smith, H. O., and M. L. Birnstiel.** 1976. A simple method for DNA restriction site mapping. Nucleic Acids Res. **3**:2387–2398.
40. **Sutcliffe, J. G., T. M. Shinnick, I. M. Verma, and R. A. Lerner.** 1980. Nucleotide sequence of Moloney leukemia virus: 3' end reveals details of replication, analogy to bacterial transposons, and an unexpected gene. Proc. Natl. Acad. Sci. U.S.A. **77**:3302–3306.
41. **Swanstrom, R., W. J. DeLorbe, J. M. Bishop, and H. E. Varmus.** 1981. Nucleotide sequence of cloned unintegrated avian sarcoma virus DNA: viral DNA contains direct and inverted repeats similar to those in transposable elements. Proc. Natl. Acad. Sci. U.S.A. **78**:124–128.
42. **Taketo, A.** 1972. Sensitivity of *Escherichia coli* to viral nucleic acids. V. Competence of calcium-treated cells. J. Biochem. **72**:973–979.
43. **Taylor, J. M.** 1977. An analysis of the role of tRNA species as primers for the transcription into DNA of RNA tumor virus genomes. Biochem. Biophys. Acta **473**:57–71.
44. **Van Beveren, C., J. C. Goddard, A. Berns, and I. M. Vema.** 1980. Structure of Moloney murine leukemia viral DNA: nucleotide sequence of the 5' long terminal repeat and adjacent cellular sequences. Proc. Natl. Acad. Sci. U.S.A. **77**:3307–3311.
45. **Yamamoto, T., B. deCrombrugghe, and I. Pastan.** 1980. Identification of a functional promoter in the long terminal repeat of Rous sarcoma virus. Cell **22**:787–797.