MOLECULAR AND CELLULAR BIOLOGY, Dec. 1984, p. 2899–2904
0270-7306/84/122899-06$02.00/0

Vol. 4, No. 12

Dilute-Coat-Color Locus of Mice: Nucleotide Sequence Analysis of the d^{+2J} and d^{+Ha} Revertant Alleles

KEITH W. HUTCHISON,[1]† NEAL G. COPELAND,[2] AND NANCY A. JENKINS[2]*

The Jackson Laboratory, Bar Harbor, Maine 04609,[1] and University of Cincinnati College of Medicine, Department of Microbiology and Molecular Genetics, Cincinnati, Ohio 45267-0524[2]

Received 9 July 1984/Accepted 18 September 1984

The unstable dilute-coat-color mutation (*d*) of DBA/2J mice has been shown to be the result of integration of an ecotropic murine leukemia virus within the mouse genome. Molecular cloning and restriction enzyme analysis of the dilute allele and the viral preintegration site (+ allele), as well as two independent dilute revertants (d^{+2J} and d^{+Ha}), suggested that reversion is due to virus excision occurring by homologous recombination involving the viral long terminal repeats. The DNA sequence has now been determined for the cell-virus junctions of the provirus associated with the *d* mutation, for the viral preintegration site, and for the two revertant sites. These data (i) indicate that the *d* mutation was caused by a normal virus integration, (ii) confirm that virus excision occurs by precise homologous recombination, as exactly one long terminal repeat is present in each revertant site, and (iii) suggest that the virus induced the *d* mutation by integration into a noncoding sequence.

Upon infection of a cell, retroviral RNA is reverse transcribed into DNA, which then integrates into the host genome at many different sites (for a review, see reference 35). The integrated form of the virus, termed the provirus, bears several structural and functional similarities to transposable elements of bacteria, yeast, and *Drosophila* species. Both retroviral proviruses and transposable elements integrate into DNA with little, if any, sequence homology to the element genome, terminate with short inverted repeats, and generate a short duplication of cellular DNA at the site of integration.

Such elements, which are not directed to a specific site of integration, have the potential of acting as insertional mutagens. Retroviruses and transposable elements have been shown to be capable of either activating or inactivating normal gene function. Activation of cellular *onc* genes by the integration of retroviruses has been demonstrated in a number of systems (6, 16, 21, 22–27), and is generally regarded as the causative event for tumor induction by retroviruses. Activation can occur either directly, by utilizing the promoter sequences of the viral long terminal repeat (LTR), or indirectly, by using LTR enhancer sequences. In the latter case, the virus is not required to integrate near the 5′ end of the gene and can integrate some distance away (23–25, 27). Inactivation of a gene by a transposable element is most easily understood when the integration is within the transcriptional unit (8, 13, 17, 28, 31). Nevertheless, inactivation of gene expression has been seen by integration of transposable elements external to the known transcriptional and regulatory boundaries of some genes (19, 20).

Germline integration of retroviruses in mice has caused at least two mutations (3, 10, 11). The integration of a Moloney murine leukemia virus (Mo-MuLV) at the *Mov-13* locus resulted in a recessive lethal mutation (10). In this case, the virus induced the mutation by integrating into the first intron of the α1(I) collagen gene (4), resulting in the inhibition of collagen gene expression. We have shown that the integration of an ecotropic MuLV (designated *Emv-3*, endogenous ecotropic proviral locus 3 [12]) at the dilute locus has caused the dilute (*d*)-coat-color mutation of DBA/2J mice (3). That the dilute mutation also results from the inactivation of a gene is suggested by the fact that it is a recessive mutation. Furthermore, deletions within the dilute locus produce the same coat-color phenotype as the viral integration does (29; unpublished data).

Previously we showed by hybridization and restriction enzyme mapping that revertants of the *d* mutation (d^+) still retain one LTR, suggesting that virus excision occurred by homologous recombination and that the virus was not integrated in a coding sequence (3). In extending this analysis, we have now sequenced the DNA at the site of integration of the provirus associated with the *d* mutation. The results presented here support the proposal that virus integration occurred in noncoding sequences. Furthermore, sequence analysis of two revertant sites showed that revertants of the *d* mutation retain exactly one LTR, confirming that virus excision occurs by homologous recombination.

MATERIALS AND METHODS

Molecular cloning and DNA sequencing. The cloning of an 18-kilobase *Eco*RI fragment containing the complete proviral genome of *Emv-3* into Charon 9 and subcloning in pBR325 was done as previously described (3). For sequencing the cell-virus junctions of *Emv-3*, a *Kpn*I subclone containing all the flanking cellular sequences in the U3-R region of the 5′ LTR and the U5 region of the 3′ LTR was constructed as previously described (3). Subclones were constructed in the M13mp8 and M13mp9 cloning vectors. Similarly, fragments of the 9.0-kilobase-pair *Eco*RI fragment from the DNA of a C57L/J mouse (+/+ at the dilute locus) were subcloned into the M13 vectors. The 9.0-kilobase *Eco*RI fragment had been similarly subcloned into pBR325 (3).

DNA sequences were determined by the dideoxy chain-terminator DNA sequence procedure of Sanger et al. (30). The sequencing strategy is given in Fig. 1. Single-stranded template DNA was purified and the sequence reactions were carried out as described in the Bethesda Research Laboratories Cloning/Sequencing kit instruction manual. DNA se-

* Corresponding author
† Present address: Department of Biochemistry, University of Maine, Orono, ME 04469.

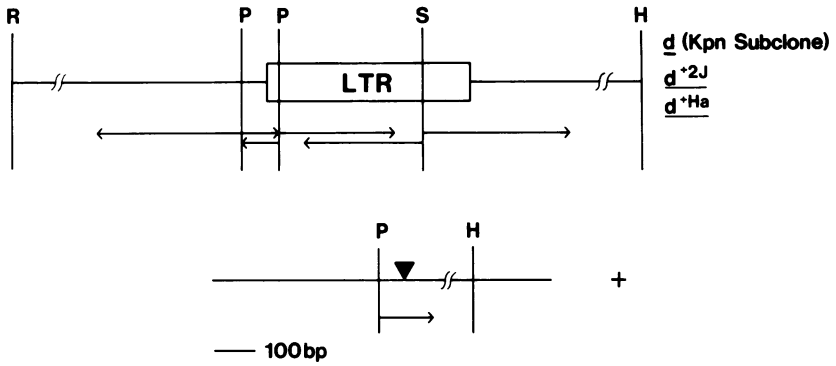

FIG. 1. Derivation of M13 subclones of wild-type (+), *d*, d^{+2J}, and d^{+Ha} DNAs. M13 subclones were isolated after *Eco*RI (R), *Pst*I (P), *Sma*I (S), and *Hin*dIII (H) digestions of DNA from the *Kpn*I subclone of the original *d* clone (3), the d^{+2J} and d^{+Ha} revertant clones, and the clone containing the *Emv-3* preintegration site (+ allele) isolated from C57L/J DNA. The *Kpn*I subclone derived from the *d* clone was produced by deleting part of the 5′ and 3′ *Emv-3* LTRs, as well as all of the internal *Emv-3* sequences, to generate a subclone that contains exactly one LTR (which is a composite LTR containing both 5′ and 3′ LTR sequences), as well as 5′ and 3′ flanking cellular DNA (3). This subclone should be identical to the d^{+2J} and d^{+Ha} clones if in these two reversions the virus was excised by homologous recombination, leaving exactly one LTR behind in revertant DNAs. The directions of the sequencing are shown by the arrows. Sequencing of the preintegration site (+) DNA began at the *Pst*I site 5′ to the LTR and continued through the viral integration site (indicated by a solid triangle).

quences were analyzed on an Apple IIe computer by using the sequencing program of Larson and Messing (18).

RESULTS

DNA sequence of the *Emv-3* LTR. The DNA sequence of the LTR of *Emv-3*, the provirus associated with the dilute mutation, was determined by sequencing several M13 subclones derived as described above (Fig. 1). The *Emv-3* LTR is 523 base pairs (bp) in length and bears a strong homology to the sequence of the AKR-type ecotropic MuLV LTR (7, 34; Fig. 2). There are only five single-base changes between the *Emv-3* LTR and the LTR of AKR(623), a prototype AKR ecotropic provirus. Three of these changes (nucleotides 49, 366, and 510) have been seen previously in the nucleotide sequence of the LTR of BL/Ka(B) virus, a nonleukemogenic B-ecotropic virus isolated from C57BL/Ka mice (15). None of these base changes occur in any of the known viral LTR regulatory sequences. The *Emv-3* LTR is 99 bp shorter than the AKR(623) LTR, missing one of the two 99-bp direct repeats which encompass the enhancer equivalent sequence (nucleotides 109 to 207). The LTR terminates in the normal 11-bp inverted repeat, and the LTR does not contain the 190-bp insert associated with many endogenous MuLV LTR sequences (14).

Revertants of the *d* mutation retain exactly one LTR. Revertants of the *d* mutation are readily detectable by a change in the coat color of the mouse and occur at a high frequency (3, 11). Previously we showed by hybridization and restriction enzyme mapping that revertants have lost the virus but retain one LTR (3). DNA of two independent revertants (d^{+2J} and d^{+Ha}) have now been sequenced. To the nucleotide, exactly one LTR remained in each revertant site (Fig. 3). The five nucleotides which distinguish the *Emv-3* LTR from the AKR(623) LTR were also present in the revertant LTRs; no additional base changes were detected; there was only one copy of the 99-bp repeat; and the cell-LTR junctions were identical to the cell-virus junctions of the intact provirus (Fig. 3, see below).

The sequence of the preintegration site for the *Emv-3* provirus was also determined (Fig. 3). As is typical of retroviral integrations, there was a duplication of cellular DNA upon integration. In the case of the *Emv-3* provirus, four bases of cellular DNA, 5′-GTAC-3′, were duplicated and flank the provirus as direct repeats. These duplicated bases are part of a short inverted repeat, 5′-TGTACA-3′. This sequence resembles the integration site of the ecotropic MuLV AKR(614) where the five-base sequence, 5′-GTGAC-3′, is duplicated upon integration (34). Again, this sequence is part of a short imperfect inverted repeat, 5′-TGTGACA-3′.

***Emv-3* provirus integrated into A-T-rich DNA.** The sequence of 623 bp of cellular DNA flanking the provirus was also determined (Fig. 4). This region is particularly adenine-thymine (A-T) rich (65%), with the 40 nucleotides immediately surrounding the *Emv-3* provirus being 73% A-T. The integration sites for several other retroviruses have also been shown to be A-T rich (32, 33), suggesting that these may be preferred regions for virus integration. Many of the A-T residues surrounding the *Emv-3* provirus are clustered, generating A-T tracts in the DNA sequence (Fig. 4).

***Emv-3* provirus has integrated into a noncoding sequence.** The position of nonsense codons in all six possible reading frames flanking *Emv-3* are shown in Fig. 5. These nonsense codons appear approximately every 50 bases. The longest open reading frame is 234 bp, into which the virus has integrated. It appears, however, that this open reading frame is not utilized for purposes of coding as there are no appropriately placed promoters or splice sequences to access it (see below).

Since it was possible that the virus integrated into an intron or into 5′ or 3′ noncoding sequences, a search was made for potential splice sites and promoter sites. Sequences which resemble the canonical "TATA" box with the associated "CAT" box were found in the flanking cellular DNA. A stability profile was then generated for each of these sites according to the method of Bensimhon et al. (1), who recently reported that eucaryotic promoters are characterized by highly stable, helical regions (high guanine · cytosine) flanking a stability minimum region (the TATA box). None of the potential TATA sequences in the DNA flanking the *Emv-3* provirus fit this pattern, although the promoter within the viral LTR did (data not shown). This suggests that these TATA sequences do not represent useful promoter sites.

Several potential splice sites were also identified in this region and are interspersed among the nonsense codons shown in Fig. 5. However, in the absence of information on

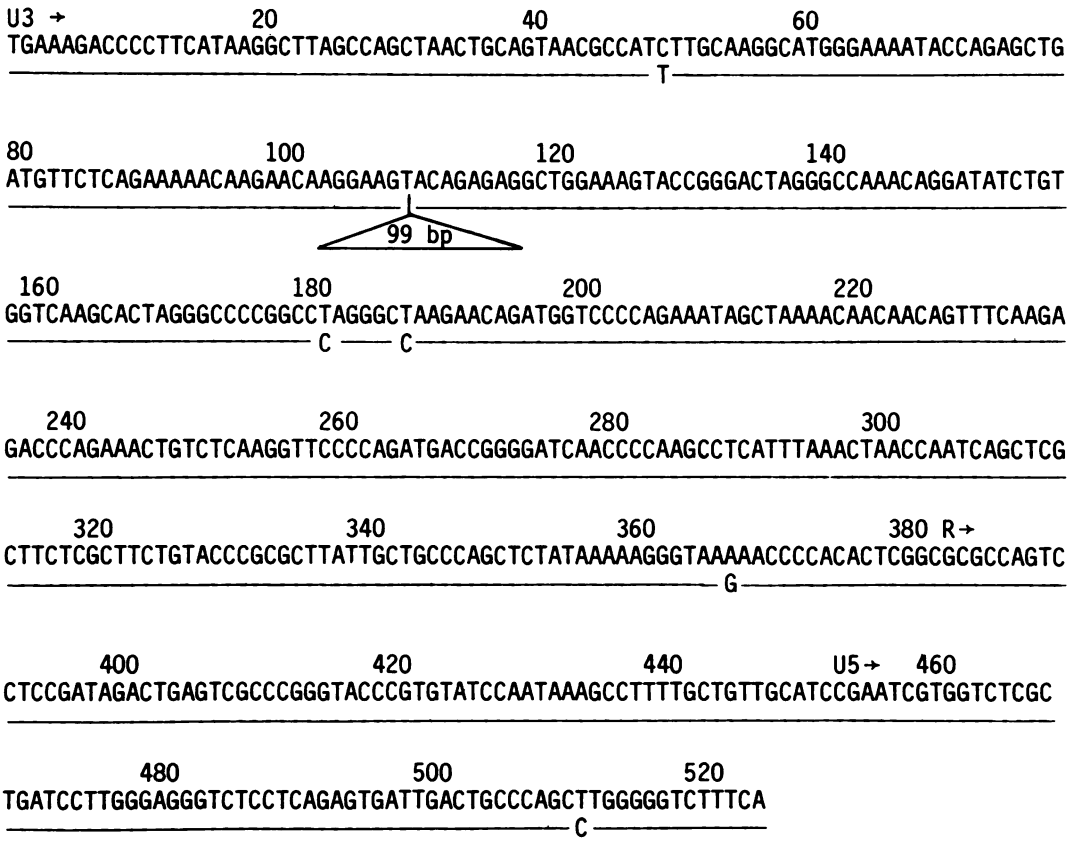

FIG. 2. DNA sequence of the *Emv-3* LTR. The sequence of the *Emv-3* LTR present in the *Kpn*I subclone of the *d* clone was determined as described in the text. This sequence is compared with that of the AKR(623) MuLV LTR (7, 34), with homologous regions shown as a solid line. Five single-base differences were noted (nucleotides 49, 180, 186, 366, and 510) between the two LTRs. In addition, only one copy of the 99-bp sequence that is duplicated in the AKR(623) LTR is present in the *Emv-3* LTR.

RNA transcripts from this segment of DNA it cannot be determined whether any of these splice sites are used.

DISCUSSION

Mice carry a large complement of retroviruses in their genomes. As such, they have the potential for an enormous impact on the expression of the mouse genome. Retroviruses and their structural homologs, the transposable elements, have been shown to act as insertional mutagens on many occasions. These mutations can act as activators or inactivators of gene function. The mutator function of transposable elements has been shown to confer selective advantage in bacteria (2) and may be the reason for the tolerance of a large viral load in the mouse genome.

We have shown that the normal integration of an ecotropic MuLV at the dilute locus of mice has caused the dilute-coat-color mutation (3). The results presented here support several predictions. Within the limitations of hybridization and restriction enzyme mapping, our previous work strongly suggested that the *Emv-3* provirus in dilute-coat-color revertants was excised by homologous recombination between the *Emv-3* LTRs. The DNA sequence of two dilute revertants described here shows that this is the case, with recombination being precise and exactly one LTR being retained in each revertant DNA. In this regard, recombination leading to excision is unlike recombination leading to retroviral integration, which requires little if any homology to the genomic DNA. Loss of a retrovirus or transposable element by homologous recombination between LTRs has

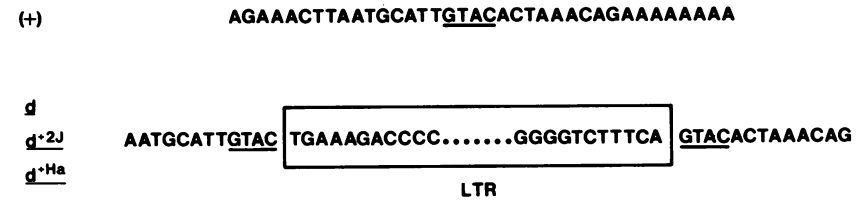

FIG. 3. DNA sequence of the preintegration site (+) and the cell-LTR junctions of the *d* (*Kpn* subclone), the d^{+2J}, and the d^{+Ha} revertants. The 4 bp of cellular DNA duplicated during integration of the *Emv-3* provirus (GTAC) are underlined. The sequence of the LTRs remaining in the d^{+2J} and d^{+Ha} revertants were identical to the sequence (Fig. 2) of the LTR present in the *d Kpn* subclone.

```
                    20                  40                  60
AATATCAGGTAAATAGTGACACTTGGTGTCAGGGGAAATGTTACTCACATTGTAGGAAGATATGTTTATTGCAAGGCT

80                  100                 120                 140
TTGGCCTTCAGCTGCCTAGAGAAGAAATCCATTGCTCAATAGACTGAGTATATCCTTTCATAGTAGAGGTTCAAGATG

160                 180                 200                 220
GCCATTTGAAGATTTGTAGAGGAAGCTACAGTGCTCAATAATTGAAATGTAGCATGAAGAGAAATGAAAATGTATGTG

240                 260                 280                 300
CTTGGAATCCCAGCAGTGGTAACACTACAGCTCTACAAGGGCGCCACAACTTCAAGGATGGGATAAATGGAGACTCGT

320                 340                 360                 380
ACCTGCAGCTCTTAGCCTTACTTCATTTCTGAACAGAAACTTAATGCATTGTACACTAAACAGAAAAAAAACTTACAT

400                 420                 440                 460
GCTGACCATTTTGAGAAAGCATGGCCAGTGATGACCACAGAGGAAAAACAGTTCATAAATACCCACAAATTAAATAT

480                 500                 520                 540
ATACACCTTAGGAAACATTAGCTTATTCTTTATTTACAGGCAAAAGTACAGATCTTATTTAAGTATGTGGTTAAGTGT

560                 580                 600                 620
AAATTTAATTTATTGAATTAATTGTAACTTTTATACTTAAGAGTATCTTATTATAGAAAGGGAATTGTGTCCATCCAT
```

FIG. 4. Sequence of 623 bp of cellular DNA flanking the *Emv-3* provirus. The sequence was determined as described in the text with the M13 subclones isolated as described in the legend to Fig. 1. The integration site of the *Emv-3* provirus (GTAC) is shown as a solid box.

been seen previously and seems to be the common mode of element excision (19, 28, 36). We cannot, however, determine at this point whether the recombination event was intrachromosonal, producing a chromosome containing a single LTR as well as a circular virus, also with a single LTR, or whether the recombination event was due to unequal crossing over during meiosis between the 5′ LTR on one chromosome and the 3′ LTR on the other chromosome. Such an event would produce a chromosome with one LTR and a chromosome with a tandemly duplicated provirus. The latter allele would presumably still have the mutant phenotype and would not be detected in our mouse colony.

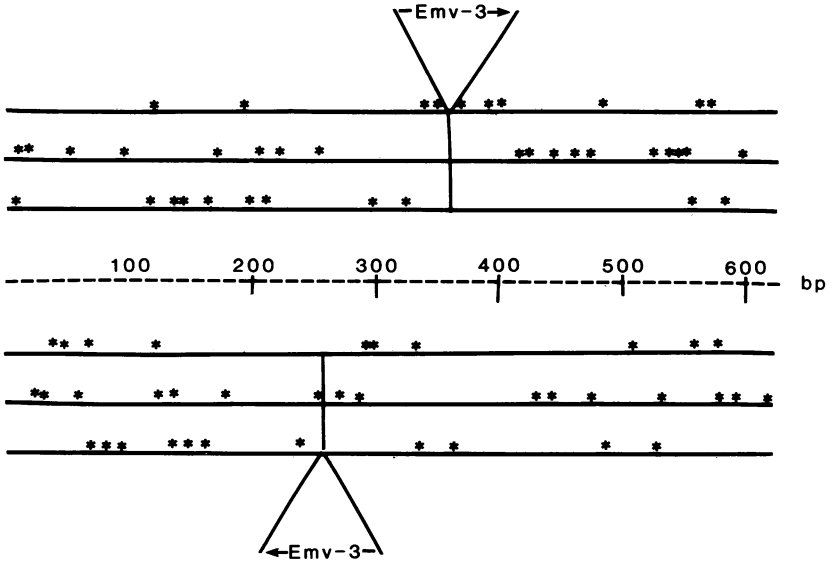

FIG. 5. Location of nonsense codons in all six possible reading frames in the cellular DNA sequences immediately flanking *Emv-3*. The presence of nonsense codons are indicated by an asterisk. The orientation of the *Emv-3* provirus 5′ to 3′ is shown by an arrow.

The integration site of the *Emv-3* provirus, like the integration sites of other retroviruses (32, 33), is A-T rich. Whether this is an important feature of the integration sites of ecotropic MuLVs is unknown. They could represent regions of the DNA which would be easily denatured. In this regard, Isberg and Syvanen (9) suggested that transposition in bacteria involves single-stranded regions of DNA.

The retention of the LTR in the revertant DNA suggests that the virus did not integrate into a coding sequence to cause the dilute-coat-color mutation. This is also consistent with sequencing data. All six possible reading frames surrounding the *Emv-3* provirus are blocked with multiple nonsense codons (Fig. 5), the longest open reading frame being 234 bp in length. It is possible, however, that the virus integrated into 5′ noncoding sequences. In fact, several promoter-like sequences appear in the flanking cellular DNA. However, when these were analyzed for helical stability, they did not resemble other eucaryotic promoters.

A second alternative is that the virus integrated into an intron. This result would be similar to that seen by Varmus et al. (36) for inactivation of v-*src* by a Mo-MuLV integration into a Rous sarcoma virus carried in a transformed rat cell line and by Wolf and Rotter (37) for inactivation of the p53 cellular-encoded tumor antigen by integration of Mo-MuLV into the first p53 intron. Hawley et al. (5) found repetitive DNA elements related to intracisternal A particles within the intervening sequences of κ-light-chain genes in two mutant hybridoma cell lines defective in κ-light-chain synthesis. Furthermore, integration of Mo-MuLV within an intron of the α1(I) collagen gene has also been shown to induce a recessive lethal mutation (4). Consistent with this model, there are a number of potential splice sites which might serve as boundaries for introns near the viral integration site at the dilute locus. Work is currently in progress to determine whether RNA is transcribed from this segment of the dilute locus and whether any of these splice sites are used.

A third alternative is that the integration of the virus is some distance from the dilute locus and is acting in some unknown manner, inactivating genes from a distance. Such a model has been proposed by Modolell et al. (20) for the inactivation of genes in *D. melanogaster* by a 7.3-kilobase mobile element.

The sequence we determined for the *Emv-3* LTR was deduced by sequencing parts of the 5′ and 3′ viral LTRs. Therefore, it is conceivable that the two LTRs diverged and differ in the regions not sequenced. However, both LTRs retained by the revertants are identical to the composite LTR, suggesting that this is not the case. Furthermore, the major LTR functions, including promoter, enhancer and polyadenylate addition signals, are in the regions sequenced, and all appear normal. Therefore, this LTR has the potential of being transcriptionally active. Whether the LTR remaining in revertant DNAs is functionally active remains to be determined.

The present work supports the model that the *Emv-3* provirus caused the dilute mutation by integrating into a noncoding region of the mouse genome. It further confirms that revertants of the dilute locus result from precise homologous recombination between viral LTRs. However, the location of the dilute gene with respect to the *Emv-3* provirus is still unknown, as is the nature of the dilute gene product. Experiments in progress with several X-ray-induced deletion mutations that have been generated at the dilute locus (29) will hopefully answer these questions.

ACKNOWLEDGMENTS

We thank J. Barker and H. G. Bedigian (The Jackson Laboratory, Bar Harbor, Maine) for critical reading of this manuscript and Barbara K. Lee and Marie Menzietti for expert technical assistance.

This work was supported by Public Health Service grant CA 37283 from the National Cancer Institute and American Cancer Society grant MV-124A.

LITERATURE CITED

1. **Bensimhon, M., J. Gabarro-Arpa, R. Ehrlich, and C. Reiss.** 1983. Physical characteristics in eucaryotic promoters. Nucleic Acids Res. **11**:4521–4540.
2. **Chao, L., C. Vargas, B. B. Spear, and E. C. Cox.** 1983. Transposable elements as mutator genes in evolution. Nature (London) **303**:633–635.
3. **Copeland, N. G., K. W. Hutchison, and N. A. Jenkins.** 1983. Excision of the DBA ecotropic provirus in dilute coat-color revertants of mice occurs by homologous recombination involving the viral LTRs. Cell **33**:379–387.
4. **Harbers, K., M. Kuehn, H. Delius, and R. Jaenisch.** 1984. Insertion of retrovirus into the first intron of α1(I) collagen gene leads to embryonic lethal mutation in mice. Proc. Natl. Acad. Sci. U.S.A. **81**:1504–1508.
5. **Hawley, R. G., M. J. Shulman, H. Murialdo, D. M. Gibson, and N. Hozumi.** 1982. Mutant immunoglobulin genes have repetitive DNA elements inserted into their intervening sequences. Proc. Natl. Acad. Sci. U.S.A. **79**:7425–7429.
6. **Hayward, W. G., B. E. Neel, and S. M. Astrin.** 1981. Activation of a cellular oncogene by promoter insertion in ALV induced lymphoid leukosis. Nature (London) **290**:475–480.
7. **Herr, W.** 1984. Nucleotide sequence of AKV murine leukemia virus. J. Virol. **49**:471–478.
8. **Ikenaga, H., and K. Saigo.** 1982. Insertion of a movable genetic element, 297, into the TATA box for the H3 histonegene in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. **79**:4143–4147.
9. **Isberg, R. R., and M. Syvanen.** 1982. DNA gyrase is a host factor required for transposition of Tn*5*. Cell **30**:9–18.
10. **Jaenish, R., K. Harbers, A. Schnieke, J. Lohler, I. Chumakov, D. Jahner, D. Grotkopp, and E. Hoffman.** 1983. Germline integration of Moloney murine leukemia virus at the *Mov-13* locus leads to recessive lethal mutation and early embryonic death. Cell **32**:209–216.
11. **Jenkins, N. A., N. G. Copeland, B. A. Taylor, and B. K. Lee.** 1981. Dilute (*d*) coat color mutation of DBA/2J mice is associated with the site of integration of an ecotropic MuLV genome. Nature (London) **293**:370–374.
12. **Jenkins, N. A., N. G. Copeland, B. A. Taylor, and B. K. Lee.** 1982. Organization, distribution, and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of *Mus musculus*. J. Virol. **43**:26–36.
13. **Johnsrud, L., M. P. Calos, and J. H. Miller.** 1978. The transposon Tn*9* generates a 9 bp repeated sequence during integration. Cell **15**:1209–1219.
14. **Kahn, A. S., and M. A. Martin.** 1983. Endogenous murine leukemia proviral long terminal repeats contain a unique 190-base-pair insert. Proc. Natl. Acad. Sci. U.S.A. **80**:2699–2703.
15. **Kim, J. P., H. S. Kaplan, and K. E. Fry.** 1982. Characterization of an infective molecular clone of the B-tropic, ecotropic BL/Ka(B) murine retrovirus genome. J. Virol. **44**:217–225.
16. **Kuff, E., A. Feenstra, K. Lueders, G. Rechavi, D. Givol, and E. Canaani.** 1983. Homology between an endogenous viral LTR and sequences inserted in an activated cellular oncogene. Nature (London) **302**:547–548.
17. **Kuhn, S., H.-J. Fritz, and P. Starlinger.** 1979. Close vicinity of IS1 integration sites in the leader sequence of the *gal* operon of *E. coli*. Mol. Gen. Genet. **167**:235–241.
18. **Larson, R., and J. Messing.** 1983. Apple II computer software for DNA and protein sequence data. DNA **2**:31–35.
19. **Levis, R., and G. M. Rubin.** 1982. The unstable WDZL mutation of *Drosophila* is caused by a 13 kilobase insertion that is imprecisely excised in phenotypic revertants. Cell **30**:543–550.

20. **Modelell, J., W. Bender, and M. Meselson.** 1983. *Drosophila melanogaster* mutation suppressible by the suppressor of Hairy-wing are insertions of a 7.3-kilobase mobile element. Proc. Natl. Acad. Sci. U.S.A. **80**:1678–1682.
21. **Neel, B. G., W. S. Hayward, H. L. Robinson, J. Fang, and S. M. Astrin.** 1981. Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete viral RNAs: oncogenesis by promoter insertion. Cell **23**:323–334.
22. **Noori-Daloii, M. R., R. A. Swift, H.-J. Kung, L. B. Crittenden, and R. L. Witter.** 1981. Specific integration of REV proviruses in avian bursal lymphomas. Nature (London) **294**:574–576.
23. **Nusse, R., A. van Ooyen, D. Cox, Y. K. T. Fung, and H. Varmus.** 1984. Mode of proviral activation of a putative mammary oncogene (*int-1*) on mouse chromosome 15. Nature (London) **307**:131–136.
24. **Nusse, R., and H. E. Varmus.** 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell **31**:99–109.
25. **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**:209–213.
26. **Payne, G. S., S. A. Courtneidge, L. B. Crittenden, A. M. Fadly, J. M. Bishop, and H. E. Varmus.** 1981. Analysis of avian leukosis virus DNA and RNA in bursal tumors: viral gene expression is not required for maintenance of the tumor state. Cell **23**:311–322.
27. **Peters, G., S. Brookes, R. Smith and C. Dickson.** 1983. Tumorigenesis by mouse mammary tumor virus: evidence for a common region for provirus integration in mammary tumors. Cell **33**:369–377.
28. **Roeder, G. S., P. J. Farabaugh, D. T. Chaleff, and G. R. Fink.** 1980. The origins of genetic instability in yeast. Science **209**:1375–1380.
29. **Russell, L. B.** 1971. Definition of functional units in a small chromosomal sequence of the mouse and its use in interpreting the nature of radiation-induced mutations. Mutat. Res. **11**: 107–123.
30. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5468.
31. **Schaller, H.** 1978. The intergenic region and the origins for filamentous phage DNA replication. Cold Spring Harbor Symp. Quant. Biol. **43**:401–408.
32. **Shimotohno, K., S. Mitzutani, and H. M. Temin.** 1980. Sequence of retrovirus provirus resembles that of bacterial transposable elements. Nature (London) **285**:550–554.
33. **Tamura, T., and T. Takano.** 1982. Long terminal repeat (LTR)-derived recombination of retroviral DNA: sequence analyses of an aberrant clone of baboon endogenous virus DNA which carries an inversion from the LTR to the *gag* region. Nucleic Acids Res. **10**:5333–5343.
34. **van Beveren, C., E. Rands, S. K. Chattopadhyay, D. R. Lowy, and I. M. Verma.** 1982. Long terminal repeat of murine retroviral DNAs: sequence analysis, host-proviral junctions, and preintegration site. J. Virol. **41**:542–556.
35. **Varmus, H. E.** 1982. Form and function of retroviral proviruses. Science **216**:812–820.
36. **Varmus, H. E., N. Quintrell, and S. Ortiz.** 1981. Retrovirus as mutagens: insertion and excision of a nontransforming provirus alter expression of a resident transforming provirus. Cell **25**:23–36.
37. **Wolf, D., and V. Rotter.** 1984. Inactivation of p53 gene expression by an insertion of Moloney murine leukemia virus-like DNA sequences. Mol. Cell. Biol. **4**:1402–1410.