JOURNAL OF VIROLOGY, Apr. 1984, p. 42–49
0022-538X/84/040042-08$02.00/0

Vol. 50, No. 1

# High-Frequency Deletion in Recovered Retrovirus Vectors Containing Exogenous DNA with Promoters

MICHAEL EMERMAN AND HOWARD M. TEMIN*

*McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, Wisconsin 53706*

Received 10 September 1983/Accepted 8 December 1983

We previously described infectious retrovirus vectors constructed from spleen necrosis virus which contain the herpes simplex virus thymidine kinase gene and the mouse α-globin gene (K. Shimotohno and H. M. Temin, Nature [London] **299**:255–268, 1982). In the present study we report that when $TK^-$ chicken cells infected with a virus containing the mouse α-globin promoter and other 5′ noncoding sequences in addition to the α-globin coding sequences were selected for thymidine kinase (TK) activity, all virus-producing $TK^+$ cell clones shed virus with a deletion. These deletions were of different sizes and included the mouse α-globin coding sequences and the mouse α-globin transcriptional promoter. One of the deleted viruses was molecularly cloned. DNA sequencing showed that the deleted sequences are flanked by a short direct repeat. This deleted virus was also shown to have an advantage over the nondeleted parent both in multiplication and in its specific TK-transforming unit titer. In contrast to the results described above, $TK^+$ cell clones established with viruses that contained only the coding sequences from the mouse α-globin gene did not delete and were stable over many cell passages. The implications of the high-frequency deletion of the viruses with internal promoters are discussed in terms of the evolution of retroviruses and the construction of retrovirus vectors.

In eucaryotic cells deletions occur in chromosomal DNA (5, 6), in viral DNAs during passage (10, 28), and in DNA introduced into cells by transfection (2, 18, 23). In some cases the deletions are associated with the loss of function of genes (22, 28), whereas in other cases they result in a gain of function (25, 28). We are interested in determining some of the parameters that influence the rate and nature of these DNA rearrangements in vertebrate cells and their viruses.

Viral vectors containing a gene whose product can be selected in vivo are good substrates for studying DNA rearrangements. These viral genomes are easily recovered from infected cells, the viral vectors can be manipulated in vitro to study the effects of different sequences on the rearrangements, and the selectable marker ensures that the sequences under selection will be maintained and that any rearrangements in the nonselected DNA can be monitored.

We have been constructing vectors from the genome of an avian retrovirus, spleen necrosis virus (SNV). Retroviruses have an additional advantage over DNA viral vectors, since in their natural life cycle retroviruses are efficiently integrated into the genomes of infected cells and maintained as stable genetic elements. Moreover, the use of recombinant retroviruses to study genetic changes will elicit information relevant to the evolution of highly oncogenic retroviruses.

Previous work of Shimotohno and Temin (20) with a recombinant retrovirus containing the mouse α-globin gene inserted 5′ to the herpes simplex virus thymidine kinase (*tk*) gene demonstrated that the introns in the mouse α-globin gene are spliced out of the viral vector as the virus is passaged in chicken embryo fibroblasts. In the present study we determined the effect of placing the cells infected with this recombinant retrovirus under selection for thymidine kinase (TK) activity. We found that all $TK^+$ cell clones established with a virus that includes the mouse α-globin promoter and other 5′ noncoding sequences in addition to the α-globin coding sequences (+promoter; Fig. 1A) contain a provirus which has a deletion. The deletion includes the mouse α-globin promoter and the entire coding sequences for the α-globin gene. One of the deleted viruses was molecularly cloned from unintegrated viral DNA and was shown to have a selective advantage over the nondeleted parent virus both in replication and in the efficiency of transforming $TK^-$ cells to a $TK^+$ phenotype. On the other hand, an analogous recombinant retrovirus that contains only the coding sequences of the mouse α-globin gene (−promoter; Fig. 1B) did not undergo deletions and indeed was stable through several months of cell passage.

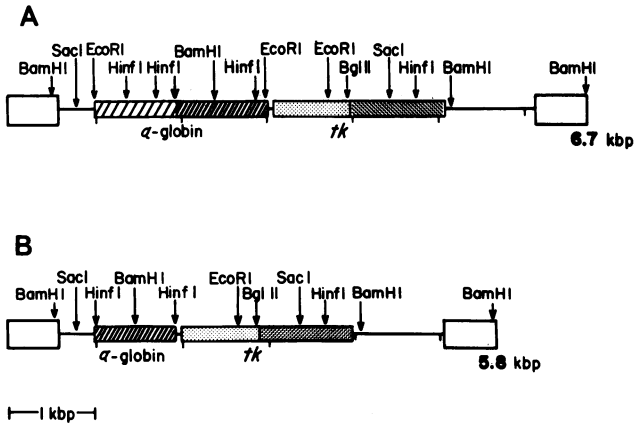


FIG. 1. Restriction enzyme cleavage site maps of the SNV–α-globin–*tk* insert in pBR322. (A), +Promoter; (B), −promoter. The open boxes represent the viral LTRs. The single-slashed boxes represent the α-globin promoter and 5′ noncoding region. The double-slashed boxes represent the α-globin coding sequences and 3′ noncoding region. The single-dotted boxes represent the *tk* promoter and 5′ noncoding region. The double-dotted boxes represent the *tk* coding sequences and 3′ noncoding region. The line represents SNV sequences.

* Corresponding author.

## MATERIALS AND METHODS

**Cells.** Chicken embryo fibroblasts (from SPAFAS chickens) were grown in Temin-modified Eagle medium with 20% tryptose phosphate broth, 2% calf serum, and 2% fetal bovine serum. Chicken hepatoma cells lacking endogenous TK activity ($CHTK^-$) were a gift of Michael Wigler and were grown in Eagle medium with 5% calf serum, 1% heat-inactivated chicken serum, and 100 μg of bromodeoxyuridine per ml. Selection of $TK^+$ cells was carried out in Eagle medium with 10% calf serum, 1% heat-inactivated chicken serum, $10^{-4}$ M hypoxanthine, $5 \times 10^{-7}$ M methotrexate, and $3 \times 10^{-5}$ M thymidine (HAT medium). $CHTK^-$ cells were plated in medium without bromodeoxyuridine the day before infection. One day after infection, the medium was replaced with HAT medium. Colonies surviving HAT selection were picked in cloning cylinders ca. 1 or 2 weeks after infection and were maintained thereafter in HAT medium. Buffalo rat liver (BRL) $TK^-$ cells were grown in Eagle medium with 7% calf serum.

**Recovery of virus from plasmids.** Chicken embryo fibroblasts in 60-mm dishes were transfected by the calcium phosphate coprecipitation method. Typically, 1 μg of a recombinant plasmid was mixed with 0.1 μg of reticuloendotheliosis virus strain A (REV-A) helper viral DNA (pSW253) and 5 μg of salmon sperm DNA. A 0.5-ml volume of the mixture in HEPES buffer (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid) was used for each dish. pME107 (see below) was cut with *Sal*I and ligated to form concatemers before transfection. Virus was collected 3 or 5 days after transfection, and titers for TK-transforming activity were determined on BRL $TK^-$ cells or $CHTK^-$ cell as previously described (19). Virus stocks gave equivalent results whether titers were determined on BRL $TK^-$ or $CHTK^-$ cells.

**Molecular cloning of unintegrated viral DNA.** Chicken embryo fibroblasts (80 100-mm plates) were infected with medium from $CHTK^+$ cell clone J (see Fig. 3) at 0.5 ml per plate (titer of ca. $10^5$ TK-transforming units [TKTU] per ml) with 0.5 ml of Polybrene (20 μg/ml; Aldrich Chemical Co.) in Eagle medium. Three days after infection, unintegrated viral DNA was isolated by Hirt fractionation (9). The Hirt supernatant fraction was treated with proteinase K (20 μg/ml; Beckman Instruments, Inc.) for 2 h at 37°C, extracted successively with phenol and chloroform, precipitated with 2 volumes of ethanol, and suspended in 10 ml of distilled water. Ten grams of CsCl and 0.8 ml of a 10-mg/ml solution of ethidium bromide were added, and the solution was centrifuged for 60 h in a Beckman 50 Ti rotor at 35,000 rpm. Sixty fractions were collected from the bottom of the centrifuge tube, and 5 μl of each fraction was spotted on nitrocellulose. The DNA was denatured with 0.1 M NaOH, and the filter was neutralized with 1 M Tris-hydrochloride, baked for 2 h at 80°C, and hybridized to [$^{32}$P]α-globin DNA. Fractions containing the closed circular unintegrated viral DNA were pooled, extracted with isopropanol, and precipitated with ethanol. The pooled fraction, which contained 1 ng of viral DNA in 7.5 μg of total DNA, was digested with *Sal*I. Charon 21A DNA (1) was digested with *Sal*I, treated with alkaline phosphatase, and ligated to an equimolar amount of insert DNA. The ligated DNA was packaged, and phage containing α-globin sequences were purified by serial dilutions of the plaques containing the phage. All phage with the α-globin-containing unintegrated DNA were identical except for the presence of one or two copies of the viral long terminal repeats (LTRs). They are called Charon 21A-ME107 and Charon 21A-ME108, respectively.

**Plasmids and nucleic acid analyses.** SNV–α-globin ΔterR-*tk* ΔterR (called +promoter in this paper; Fig. 1A) and SNV–α-globin Δter*Hin*fI-*tk* ΔterR (called −promoter in this paper; Fig. 1B) were previously described (20). pSW253 is a subclone of an REV-A provirus in pBR322. The insert in Charon 21A-ME107 was cut out of the vector with *Sal*I and subcloned into the *Sal*I site of pBR322. This plasmid is called pME107.

Unintegrated linear viral DNA was prepared as described above for supercoiled DNA except that it was not centrifuged in a CsCl gradient and was treated with RNase (50 μg/ml). The unintegrated viral DNA was electrophoresed in an agarose gel either uncut or cut with restriction enzymes. Procedures for Southern transfer and hybridization were previously described (3). Tracing of bands in the autoradiograms by densitometry was done on a Beckman DU-8 scanning spectrophotometer. Several exposures of the film were made to ensure that the linear range of the film was used. Enzymes were purchased from New England Bio-Labs, Inc., or P-L Biochemicals, Inc., and were used according to the recommendations of the manufacturer. DNA sequencing was done by the method of Maxam and Gilbert (13).

## RESULTS

**High-frequency deletion in a retrovirus vector.** $CHTK^-$ cells were infected with virus recovered after cotransfection of chicken embryo fibroblasts with pSNV–α-globin–*tk* and helper virus (REV-A) DNAs. The SNV–α-globin–*tk* plasmids (20) contain, in addition to all *cis*-acting regions necessary for retrovirus propagation, the herpes simplex virus *tk* gene and the mouse α-globin gene (both with their polyadenylation sites removed [Fig. 1]). The *tk* gene is expressed from its own promoter (19; P. Bandyopadhyay and H. M. Temin, Mol. Cell. Biol., in press). Cell clones were established from cells that survived HAT selection and were cultured in HAT medium. Fourteen cell clones were analyzed from cells that had been infected with the +promoter virus, and eight clones were analyzed that had been established with the −promoter virus.

The virus shed by the $TK^+$ cell clones reflects the resident proviruses in the clones. To screen the proviral structure of a large number of cell clones, we determined the structure of the viral genomes in the viruses shed by these clones. Therefore, cell-free media were collected and were used to infect chicken embryo fibroblasts. Three days after infection, unintegrated DNA was isolated and analyzed by Southern transfer analysis. If no gross rearrangements of the proviruses had occurred, we expected the +promoter viral DNA to be either 6.7 or 6.4 kilobase pairs (kbp), depending on whether the intervening sequences in the α-globin gene had been spliced out. The −promoter viral DNA, on the other hand, was expected to be either 5.8 or 5.5 kbp.

Surprisingly, however, we found that the virus shed by −promoter-infected cell clones was always larger than the virus shed by +promoter cell clones (Fig. 2). In all 11 cases in which the +promoter cell clones were producing α-globin-containing SNV, at least one of the viruses shed by the cell clone had a deletion which resulted in a virus of 5.7 to 4.6 kbp (Table 1). In 8 of these 11 cases, the only virus produced by the cell clones established with the +promoter virus was a deleted virus. All of the −promoter cell clones producing α-globin-containing SNV, on the other hand, shed virus of the expected size (5.5 kbp) when the intervening sequences were spliced out. Moreover, we passaged these cells for over 5 months with no deletions appearing in the virus shed.

In addition to the smaller viruses seen with the +promoter stock, we have also observed in our CHTK$^+$ cell clones viral variants that are larger than the expected size (Table 1 and Fig. 2, lane C). These larger viruses are in clones established from both +promoter and −promoter viruses. Some of these viruses are up to 13.2 kbp, over 50% larger than the wild-type SNV from which the vectors were derived (19). Restriction endonuclease mapping of the unintegrated DNA from these viruses showed that they are not simple duplications of viral DNA (data not shown). About one in five cell clones contained a larger virus as well as one of the expected size (for the −promoter virus) or smaller size (for +promoter virus).

**Mapping the deletions in virus shed by two cell clones.** Figure 3 shows the results of mapping the deletions in virus from two +promoter cell clones by restriction enzyme digestions of the linear unintegrated viral DNA, followed by Southern blotting. (The restriction enzyme cleavage map for virus shed by clone J was confirmed by molecular cloning described below.) In both cases the entire coding sequences of the α-globin gene were deleted along with portions of the 5′ noncoding sequences. Virus from CHTK$^+$ cell clone O had, additionally, lost the *Eco*RI site at the 5′ end of the α-globin insert. Both 5′ and 3′ endpoints of the large deletion in clone O are different from the 5′ and 3′ deletion endpoints in clone J (Fig. 3). This result indicates that the deletion was not due to cryptic splice sites which might have caused a deletion during the RNA stage of the virus life cycle. DNA

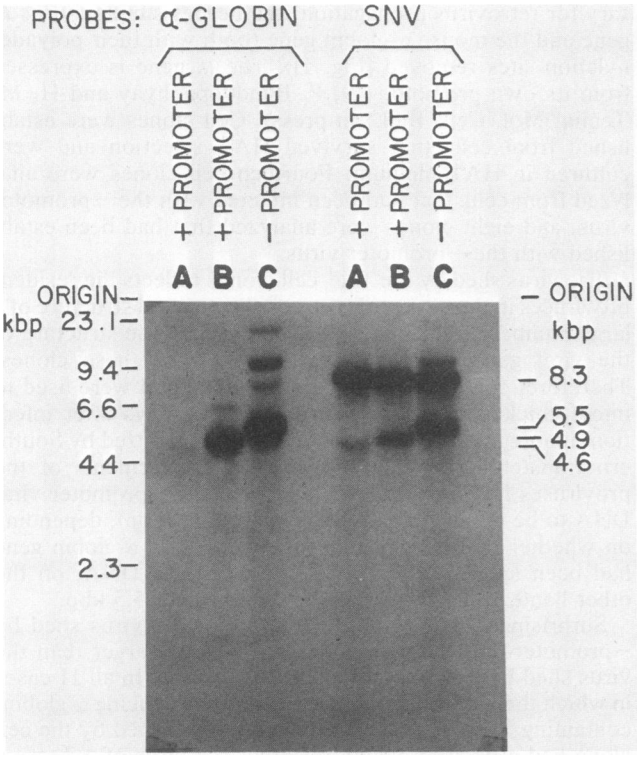


FIG. 2. Southern blot of unintegrated linear viral DNA recovered from chicken embryo fibroblasts infected with virus from CHTK$^+$ cell clones and probed with either an α-globin- or an SNV-specific probe. DNAs recovered from two 100-mm plates of cells were loaded in each lane. The 8.3-kbp band in the SNV probe panel is the REV-A helper virus. Lane A, Virus from +promoter clone O; lane B, virus from +promoter clone K; lane C, virus from −promoter clone E.

TABLE 1. Characterization of virus harvested from CHTK$^+$ cell clones

| Cell clone[a] | Size of α-globin-containing viral DNA (kbp)[b] |
|---|---|
| −Promoter | |
| A | 5.5 |
| B | 5.5 |
| C | 8.5, 5.5 |
| D | 5.5 |
| E | 12.0, 10.5, 7.9, 5.5 |
| F | 5.5 |
| G | 13.2, 10.6, 5.5 |
| H | Nonproducer |
| +Promoter | |
| I | 5.2 |
| J | 4.9 |
| K | 6.5, 4.9 |
| L | 11.0, 8.6, 7.4, 5.2 |
| M | 5.3 |
| N | 5.2 |
| O | 4.6 |
| P | 5.7 |
| Q | 6.7, 5.4 |
| R | 5.5 |
| S | 5.4 |
| T, U, V | Nonproducers |

[a] Cell clones A to H were established by infection with the −promoter virus and selection for TK$^+$ colonies; cell clones I to V were established by infection with the +promoter virus and selection for TK$^+$ colonies. Virus used to establish the cell clones was obtained by transfecting chicken embryo fibroblasts with the −promoter or +promoter plasmid and pSW253 (see text). Virus was collected from the chicken embryo fibroblasts 4 to 7 days after transfection. Cell clones A to O and V were passaged for at least 2 months (i.e., from infection to medium collection) before virus was analyzed. Cell clones P to U were passaged for only 2 weeks before virus was analyzed (see text).

[b] The size of the viral DNA was assayed by infecting chicken embryo fibroblasts with virus from the cell clones and isolating the unintegrated viral DNA 3 days after infection. Southern blots of the uncut DNA were probed separately with an α-globin probe, an SNV probe, and sometimes a *tk* probe. All probes gave identical results except that the SNV probe also hybridized to the helper REV-A viral DNA.

sequencing of the deletion junction for one virus (see below) confirmed this hypothesis because the DNA sequence at the deletion junction bears no resemblance to the consensus sequences for RNA splicing sites (14).

To better understand the nature of the deletion, we molecularly cloned the unintegrated DNA of one deleted virus (recovered from CHTK$^+$ cell clone J, shown in Fig. 3) in the lambda vector Charon 21A. This virus, called ME107, is 4.9 kbp. By fine-structure restriction enzyme mapping and by DNA sequencing across the junction of the deletion, we determined that this virus has a deletion of 1.8 kbp extending from 110 bases 5′ to the TATAA box of the α-globin gene (16) to 15 bases 5′ to the start of the coding sequences of the *tk* gene (29). Thus, in this virus the normal RNA cap sites of both inserted genes as well as the promoters for both genes have been deleted. Therefore, the RNA transcript from which TK is expressed (see below) must be initiated in the 5′ LTR in this clone. Examination of the deletion junction showed that a 6- out of 7-base-pair homology (GAAXCTC) exists at the site of the deletion (Fig. 4). Additionally, there is a 1-base-pair insertion (T/A) near the site of the deletion.

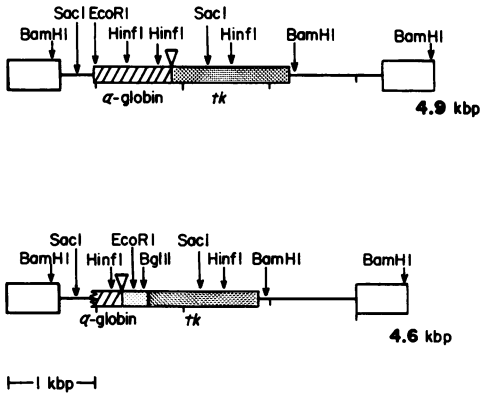


FIG. 3. Restriction enzyme cleavage site maps of linear unintegrated viral DNA recovered from two CHTK$^+$ cell clones, J (top) and O (bottom) (see Table 1), which had been infected with +promoter virus. Symbols are the same as in Fig. 1. A jagged line at the 5′ end of the α-globin sequences in the restriction enzyme cleavage site map of virus from clone O represents a mutation in the DNA that eliminated at least the 5′ *Eco*RI site but has not been mapped in more detail. The ▽ represents the site of the large deletion in each virus.

Therefore, at least two changes—a large deletion and a small insertion—occurred in this virus sometime between the original transfection and the time the virus from the established cell clone was analyzed.

**Selective advantage of the deleted virus compared with the +promoter parent.** To determine what, if any, selective advantage the deleted virus had over the +promoter parent, we transfected chicken embryo fibroblasts with the cloned deleted virus, ME107, and determined the titers of the virus on TK$^-$ cells. ME107 had been circularly permuted about the *Sal*I site in pBR322 and was not infectious in this form. To make virus from the ME107 plasmid, the viral sequences were cut out with *Sal*I and ligated to form concatemers before cotransfections of chicken embryo fibroblasts with helper virus DNA were performed. The virus which resulted from this procedure was capable of transforming TK$^-$ cells to a TK$^+$ phenotype. Moreover, by parallel transfection of the ME107 and +promoter plasmids and subsequent infections with the resultant viruses, we found that ME107 virus was over 100-fold more active at transforming cells to the TK$^+$ phenotype than was the parental +promoter virus (Table 2). The deleted virus also has a growth advantage over the parental +promoter virus (Table 2). However, when the TK titers of the two virus stocks were normalized for the amount of virus production, the ME107 virus still had a 10-fold-higher specific TK titer.

To confirm the finding of increased replication of the deleted virus, we transfected chicken embryo fibroblasts with various dilutions of *Sal*I-cut and -ligated pME107 and the +promoter plasmid in parallel (all in the presence of helper virus DNA). Five days after transfection, we collected virus and then infected fresh chicken embryo fibroblasts and collected the unintegrated DNAs for a hybridization analysis of the viruses present in the stocks. Our results indicate (Fig. 5) that at a 1:10 dilution of pME107 in the +promoter plasmid, the deleted virus still overgrows the +promoter parent virus. Therefore, the viruses which have undergone a deletion have an advantage both in replication and in their ability to transform TK$^-$ cells to a TK$^+$ phenotype.

**Origin of the deleted virus.** In our initial screen of cell clones (clones A to O and V), the clones had been passaged for at least 2 months before analyses of the shed viruses. Therefore, we were unable to determine whether the deletions occurred in the chromosomes during passaging of the cells or at an earlier stage. For this reason, new cell clones were established with the +promoter virus, and virus from these clones was analyzed as soon as possible after infection. If the deletion occurred during cell passage, then we would have expected the cell clones analyzed early after infection to contain only the full-size provirus or to be a mosaic population of cells containing either full-size or deleted proviruses. The mosaic cell population would be represented in the shed viruses by both full-size and deleted viral genomes. At the earliest times assayed (2 weeks after

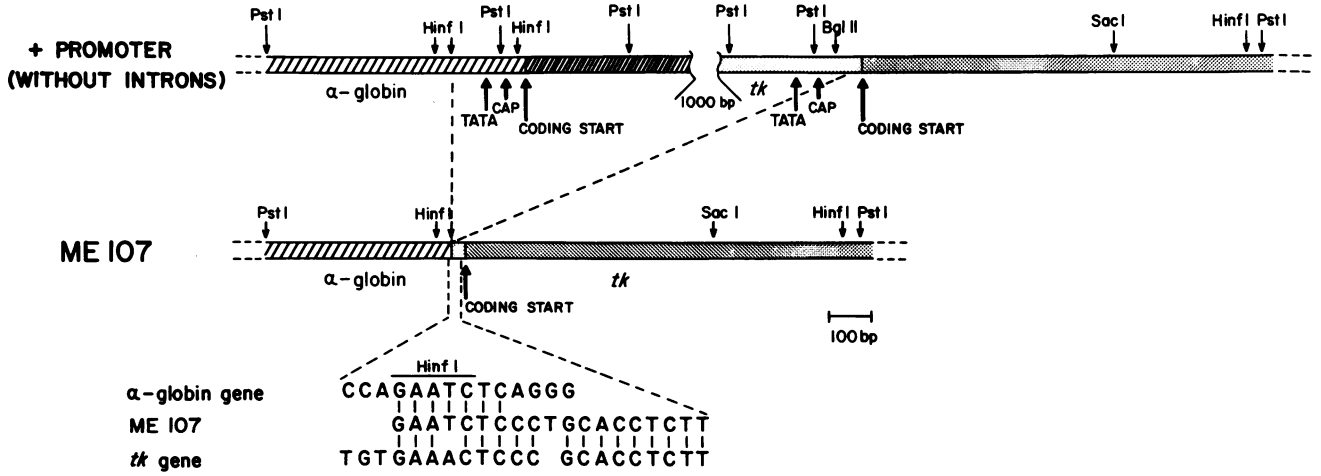


FIG. 4. Detailed map of the junction between the α-globin and *tk* genes in the +promoter virus and in ME107. ME107 is a molecular clone of unintegrated DNA from virus shed by +promoter cell clone J (Fig. 3). The dashed lines show the endpoints of the deletion as determined by fine-structure restriction enzyme mapping and DNA sequencing. DNA sequencing was done by isolating the 1.3-kbp *Pst*I fragment of ME107 shown above, labeling the *Hin*fI sites with polynucleotide kinase, cutting with *Sac*I, and isolating the 550-bp *Hin*fI-*Sac*I fragment. The sequence of pME107 was compared with the published sequences for the α-globin (16) and *tk* (29) genes. The *tk* portion of the +promoter plasmid was also sequenced from the *Bgl*II site to confirm the published sequence. The +promoter virus is represented with the introns spliced out of the gene (21). Symbols are the same as in Fig. 1.

infection), all cell clones (clones P, Q, R, and S; Table 1) contained a deleted virus. Moreover, three of four of these cell clones (clones P, R, and S) did not shed any detectable full-size virus (1% levels of full-size virus would be detected; data not shown). Thus, it is unlikely that the deletion occurred in the chromosomes of these cells during extended passage. One clone (clone Q, Table 1) contained a full-length virus in addition to the deleted virus. However, when this cell clone was passaged for an additional month, the ratios between the amount of the two viruses did not change. Thus, the larger parental virus was not spontaneously deleting in the chromosome. Nonetheless, it is possible that the deletions occurred in the cell clones infected with +promoter virus within the first several cell doublings of the $TK^+$ cell clones.

We considered the possibility that there was a minor population of virus in our +promoter stock which had undergone deletion before infection of the $CHTK^-$ cells. Figure 6 shows a Southern blot of unintegrated viral DNA from the +promoter stock passaged on chicken embryo fibroblasts without selection for the *tk* gene. No bands which correspond to a minor population of virus with a specific deletion are visible. (The same result was found when virus was passaged in $CHTK^-$ cells without selection for TK activity [data not shown].) However, in Figure 6 deleted viral genomes are seen on a long exposure of the gel as a background "smear" beneath the full-length +promoter band. Because the deletion size is variable (Table 1), this would be the expected result if the virus stock contained a minor population of viruses with deleted genomes before the cell clones under TK selection were established.

To determine whether the +promoter plasmid has the intrinsic ability to transform $TK^-$ cells to a $TK^+$ phenotype or whether the deletion is necessary to allow TK expression, we transfected the +promoter plasmid, −promoter plasmid, and parental construction pSW227 into BRL $TK^-$ cells and selected for $TK^+$ cells. All three plasmids gave $TK^+$ transformants at ca. 1.5 transformants per μg of DNA.

## DISCUSSION

All of the virus-producing cell clones (11 of 11) established with +promoter virus shed virus bearing deletions of 0.8 to 1.4 kbp after selection for the *tk* gene. However, none (seven of seven) of the $TK^+$ cell clones established with −promoter virus were smaller than the size predicted. These observations have implications for the evolution of highly oncogenic retroviruses, the nature of deletions in eucaryotic DNA, and the construction and use of viral vectors.

TABLE 2. Specific TKTU[a] of the cloned deleted virus ME107 relative to the parent +promoter virus

| Virus[b] | TK titer (TKTU/ml) | Relative virus production[c] | Relative specific TKTU[d] |
|---|---|---|---|
| +Promoter | $1.0 \times 10^3$ | 1 | 1 |
| ME107 | $1.1 \times 10^5$ | 11 | 10 |

[a] TKTU were assayed by diluting the virus stocks and infecting BRL $TK^-$ cells. TKTU per milliliter are the dilution times the number of $TK^+$ colonies per plate.

[b] Virus stocks were prepared by cotransfecting chicken embryo fibroblasts with 1 μg of the appropriate plasmid and 0.1 μg of helper virus DNA. Medium was collected 5 days after transfection.

[c] Relative amount of recovered virus DNA = recombinant virus/helper virus. Chicken embryo fibroblasts were infected with each virus, and unintegrated viral DNA was extracted 3 days later. The DNA was electrophoresed on a 0.8% agarose gel and blotted to nitrocellulose. The filter was hybridized to the 1.3-kbp *PstI* junction fragment of pME107 (Fig. 4), and the bands were analyzed by densitometry. The filter was then hybridized to an SNV probe to quantify the amount of helper virus DNA in each lane.

[d] Ratio of TKTU titers/relative virus production.

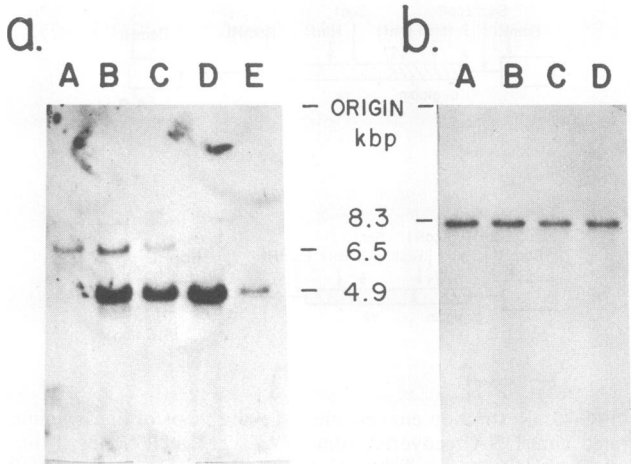


FIG. 5. Southern blot of the unintegrated viral DNAs recovered from chicken embryo fibroblasts infected with +promoter or ME107 or both viral stocks. The viral stocks were made by transfecting chicken embryo fibroblasts on 60-mm dishes with 5 μg of sheared salmon sperm DNA, 0.05 μg of helper REV-A pSW253 DNA, and the following amounts of +promoter and pME107 (cut with *SalI* and ligated to concatemers) DNAs: lane A, 0.5 μg of +promoter; lane B, 0.5 μg of +promoter and 0.05 μg of ME107; lane C, 0.25 μg of +promoter and 0.25 μg of ME107; lane D, 0.5 μg of ME107; lane E, 1/10 the amount of DNA loaded in lane D. The 6.5-kbp band corresponding to the +promoter virus, the 4.9-kbp band corresponding to the ME107 virus, and the 8.3-kbp band corresponding to the helper virus were quantified by densitometry. (a) was probed with the 1.3-kbp *PstI* junction fragment from ME107 (Fig. 4). (b) was an identical filter hybridized to an SNV probe.

**Evolution of retroviruses.** Highly oncogenic retroviruses have evolved from replication-competent retroviruses by insertion of cell DNA (proto-oncogene) into their genome. In all cases except that of Rous sarcoma virus, this insertion is accompanied by a deletion of viral sequences which renders the highly oncogenic retrovirus defective for replication in the absence of a helper virus (4). Chen and Temin (3) showed that in REV-T the insertion of sequences from the replication-competent component of REV-T suppresses the oncogenicity of REV-T. Some variants, however, had regained their oncogenic potential. In most cases these recovered revertant mutants of REV-T had deletions in the inserted sequences. Thus, the formation of a highly oncogenic retrovirus may often require deletions in the viral genome to allow expression of the v-*onc* sequences.

As with the oncogene in REV-T which provides a means for selection by the outgrowth of transformed cells, the insertion of the *tk* gene into a retrovirus allows selection for infected cells because uninfected cells die quickly when HAT medium is applied. We have found that under HAT selection the expression of the *tk* is also accompanied by a deletion of other sequences present in the virus. Our observation that the deleted viruses are a very minor portion of the unselected population of virus but are the major virus species present when cells are selected with HAT medium illustrates how efficiently spontaneously arising replication-defective retrovirus mutants with a selective advantage can overgrow another replication-defective virus.

In both cases in which the deleted viruses were mapped, a

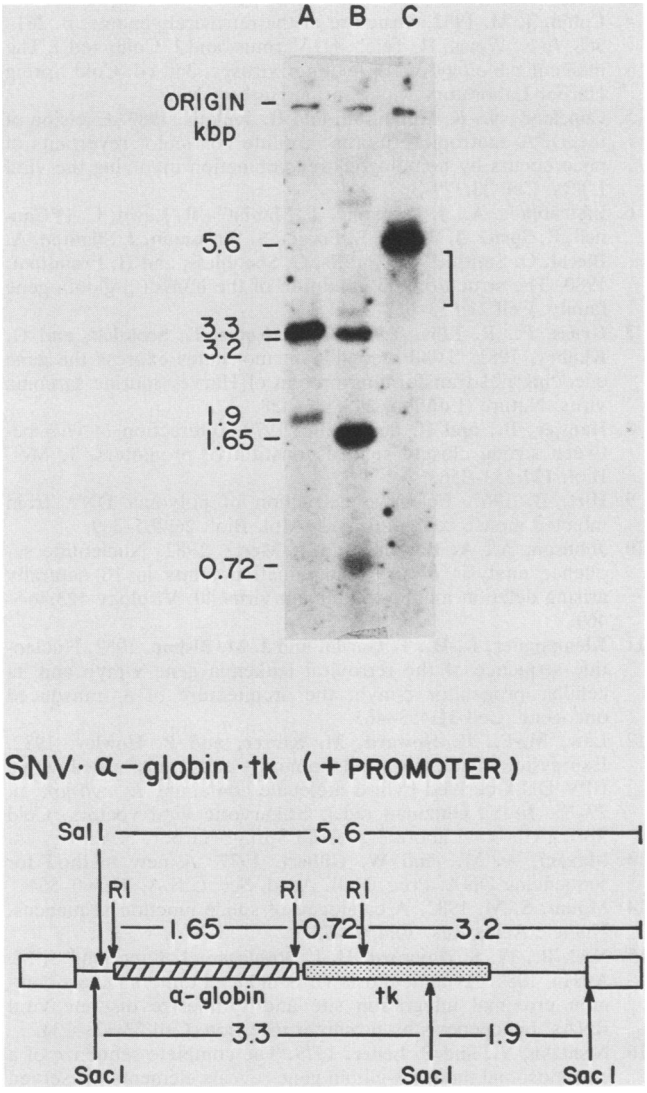


FIG. 6. Structure of the unintegrated viral DNA of the +promoter virus after passage in chicken embryo fibroblasts for 8 days. Each lane is Hirt supernatant DNA from four 100-mm plates of infected cells. Lane A, DNA cut with *Sac*I; lane B, DNA cut with *Eco*RI; lane C, DNA cut with *Sal*I. The filter was hybridized to a probe that consisted of the α-globin and *tk* genes. The area in brackets shows the smear of deleted viral molecules in the range expected after *Sal*I digestion. The restriction enzyme cleavage site map below the autoradiogram shows the expected fragment sizes for a nondeleted +promoter virus with no intervening sequences.

small mutation (either a 1-base-pair insertion or the loss of a restriction enzyme recognition site) was found in addition to the deletion. This result indicates that a high level of mutation occurs in these viruses.

Most, if not all, of the natural isolates of cellular proto-oncogenes that have been transduced into retroviruses do not include their own promoter (one exception may be Harvey murine sarcoma virus [7]). Rather, the transcription of these genes relies on the promoter present in the LTR of the retrovirus. Likewise, c-*myc* activation in chicken bursal lymphomas by the 5′ insertion of an avian leukosis virus is often accompanied by the deletion of one transcription promoter, the 5′ LTR (15). The only DNA sequences missing both in the two deletions we have mapped (Fig. 3) and in the in vitro construction of the −promoter virus, which has never been seen to undergo gross deletions, is a 160-base-pair fragment which includes the α-globin promoter. This sequence extends from 110 base pairs 5′ to the α-globin TATAA box to 20 base pairs 3′ to the normal cap site of the α-globin gene (Fig. 4). Thus, our data suggest that the deletion of a promoter (e.g., the α-globin promoter) near other transcriptional units (e.g., the *tk* promoter) can be selected for by the expression of genes which give the virus or cell an advantage. This result may indicate an inhibitory effect of the 5′ promoter on the activity of the 3′ promoter. This phenomenon is known to occur in *Escherichia coli* and depends on the relative strengths of the promoters (8). It should be noted that a deletion of the 5′ LTR would not be detected in this study because of the requirement for virus production.

**Nature of the deletions.** There appear to be at least two mechanisms for forming deletions in eucaryotic DNA. Naturally occurring deletions in simian virus 40 (10) and deletions in transfected DNA (24) can be formed in the absence of homologous recombination. Likewise, the junction of v-*onc* sequences to viral genes that may be indicative of recombination between retroviruses and cellular oncogenes seems to require very little or no base-pair homology (11, 26). On the other hand, several types of deletions found in chromosomal DNA do occur at regions of homology. These include the excision of a retrovirus from the chromosome by homologous recombination between the LTRs (5, 27) and the deletions found in the globin-like genes where 2- to 8-base-pair homology occurs at the deletion junctions (6). The deletions that we have observed may be similar to those that have been mapped in the human globin-like genes. Although there does not appear to be a "hot spot" for recombination because most of the deletions in our +promoter virus are of different sizes, the sequence at one deletion junction revealed a 6- out of 7-base-pair homology at the deletion junction. Spontaneous deletions in the U3 region of SNV (21) and in the *src* gene of avian sarcoma virus (17) also occurred at regions of short homology.

We do not know precisely when the deletions occurred in the +promoter viruses. They are not an artifact of a single virus preparation because we established clones on three separate occasions, each time with an independently derived virus stock which was made by transfection with a new plasmid preparation.

Several reports have indicated that a high level of mutation occurs in DNA that has been transfected into vertebrate cells (2, 22). These mutations include both deletions and insertions. It is also possible that the deletions seen here occurred by the reverse transcriptase "jumping" templates asymmetrically during virus replication. Although less likely, it is possible that the deletions occurred within the first several cell doublings of the infected cells under selection for TK activity. If the cells which contained the deleted copy of the virus grew faster than cells which contained only the full-length virus, then a high-frequency deletion early after infection would give a result similar to that of the deleted viruses being present in the population before infection of the TK-selected cells.

We have observed that the +promoter virus has a TKTU titer that is equal to or greater than the −promoter stocks which do not delete (data not shown). This result may be reconciled with the observation that nearly all of the population of the +promoter virus is full length under nonselective conditions by the fact that the deleted viruses are 10 times more efficient at transforming cells from $TK^-$ to $TK^+$ than

are their parental viruses (specific TKTU, Table 2). Thus, a minor population of 10% could account for the entire transforming potential of the +promoter stock. Because the deleted virus also has an advantage in replication, deletion events are amplified in the population. Although the +promoter plasmid has the ability to transform BRL $TK^-$ cells to a $TK^+$ phenotype after transfection (Table 2), our inability to find any $CHTK^+$ cells with only a full-size +promoter provirus may be because we did not screen enough clones. Alternatively, the mechanism of TK transformation by transfection of BRL cells with plasmids may be quite different than that by infection of chicken cells with viruses.

**Retrovirus vectors.** Our observations are also important in the construction and use of retrovirus vectors. First, we have found that in some cases (i.e., with the −promoter virus) the vector is stable through many cell generations and there has been no apparent change in the nonselected foreign DNA (the α-globin coding sequences) in this time. This result illustrates the utility that retrovirus vectors may have in the stable insertion of foreign DNA into the genome of cells. On the other hand, the introduction of particular sequences will lead to instability of the vector when selection either for virus replication or for the expression of inserted genes is applied. Shimotohno and Temin (20) have already shown that in some cases it is also necessary to remove the polyadenylation site of the inserted gene, and Bandyopadhyay and Temin (P. Bandyopadhyay and H. M. Temin, Mol. Cell. Biol., in press) have shown that a strong promoter in the orientation opposite to that of viral transcription inhibits virus production. Our report indicates that some DNA sequences in the vector will lead to selection for viruses with rearrangements when expression of internal genes is required for maintenance of the vector. Reports of rearrangements in vectors constructed from DNA viruses have also been published recently (2, 12, 18, 25).

Finally, we have found that some of our cell lines shed virus that forms viral DNA much larger than that of the parent virus. This result implies that retroviruses are able to package genomes much larger (>50%) than that of wild-type virus. This observation indicates that it may be possible to construct recombinant retroviruses with large insertions of foreign DNA.

## ACKNOWLEDGMENTS

We thank S. Hellenbrand and A. Joy for technical assistance, K. Wilhelmsen for help in cloning in λ, and J. Mertz, R. Risser, B. Sugden, T. Schedl, and W. G. Tarpley for helpful comments on the manuscript.

This work was supported by Public Health Service grants CA-22443 and CA-07175 from the National Institutes of Health. M.E. is supported by Public Health Service training grant GM-07215 from the National Institutes of Health. H.M.T. is an American Cancer Society Professor.

## LITERATURE CITED

1. **Blattner, F. R., B. G. Williams, A. E. Blechl, K. Denniston-Thompson, H. E. Faber, L. A. Furlons, D. J. Grunwald, D. O. Kiefer, P. D. Moore, J. W. Schumm, E. I. Sheldon, and O. Smithies.** 1977. Charon phages: safer derivatives of bacteriophage lambda for DNA cloning. Science **196**:161–169.
2. **Calos, M., J. Lebokowski, and M. Botchan.** 1983. High mutation frequency in DNA transfected into mammalian cells. Proc. Natl. Acad. Sci. U.S.A. **80**:3015–3019.
3. **Chen, I. S. Y., and H. M. Temin.** 1982. Substitution of 5′ helper virus sequences into non-rel portion of reticuloendotheliosis virus strain T suppresses transformation of chicken spleen cells. Cell **31**:111–120.
4. **Coffin, J. M.** 1982. Structure of the retroviral genome, p. 261–368. *In* R. Weiss, H. Teich, H. Varmus, and J. Coffin (ed.), The molecular biology of RNA tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. **Copeland, N., K. Hutchison, and H. 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.
6. **Efstratiadis, A., J. Posakony, T. Maniatis, R. Lawn, C. O'Connell, R. Spritz, J. DeRiel, B. Forget, S. Weissman, J. Slighton, A. Blechl, O. Smithies, F. Boralle, C. Shouhless, and H. Proudfoot.** 1980. The structure and evolution of the human β-globin gene family. Cell **21**:653–658.
7. **Gruss, P., R. Ellis, T. Shih, M. Konis, E. Scolnick, and G. Khoury.** 1981. SV40 recombinant molecules express the gene encoding p21 transforming protein of Harvey murine sarcoma virus. Nature (London) **293**:486–488.
8. **Hausler, B., and R. Somerville.** 1979. Interaction *in vivo* between strong closely spaced constitutive promoters. J. Mol. Biol. **127**:353–356.
9. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**:265–269.
10. **Johnson, A., A. Barkan, and J. Mertz.** 1982. Nucleotide sequence analysis of the recombination joints in 16 naturally arising deletion mutants of simian virus 40. Virology **123**:464–469.
11. **Klempnauer, K.-H., T. Gorda, and J. M. Bishop.** 1982. Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. Cell **31**:453–463.
12. **Law, M.-F., B. Howard, M. Sarver, and P. Howley.** 1982. Expression of selective traits in mouse cells transformed with a BPV DNA-derived hybrid molecule containing *E. coli* gpt, p. 79–85. *In* Y. Gluzman (ed.), Eukaryotic viral vectors. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.
14. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
15. **Neel, B., W. S. Hayward, H. L. Robinson, J. Fang, and S. M. Astrin.** 1981. Avian leukosis virus-induced tumors have a common proximal integration site and synthesize discrete viral RNAs: oncogenesis by promoter insertion. Cell **23**:323–334.
16. **Nishioka, Y., and P. Leder.** 1979. The complete sequence of a chromosomal mouse α-globin gene reveals elements conserved throughout vertebrate evolution. Cell **18**:875–887.
17. **Omer, C. A., K. Pogue-Geile, R. Guntaka, K. A. Staskus, and A. J. Faras.** 1983. Involvement of directly repeated sequences in the generation of deletions of the avian sarcoma virus *src* gene. J. Virol. **47**:380–382.
18. **Razzaque, A., H. Mizusawa, and M. Seidman.** 1983. Rearrangement and mutagenesis of a shuttle vector plasmid after passage in mammalian cells. Proc. Natl. Acad. Sci. U.S.A. **80**:3010–3014.
19. **Shimotohno, K., and H. M. Temin.** 1981. Formation of infectious progeny virus after insertion of herpes simplex thymidine kinase gene into DNA of an avian retrovirus. Cell **26**:67–77.
20. **Shimotohno, K., and H. M. Temin.** 1982. Loss of intervening sequences in genomic mouse α-globin DNA inserted in an infectious retrovirus vector. Nature (London) **299**:265–268.
21. **Shimotohno, K., and H. M. Temin.** 1982. Spontaneous variation and synthesis in the U3 region of the long terminal repeat of an avian retrovirus. J. Virol. **41**:163–171.
22. **Simon, A., and M. Taylor.** 1983. High-frequency mutation at the adenine phosphoribosyltransferase locus in Chinese hamster ovary cells due to deletion of the gene. Proc. Natl. Acad. Sci. U.S.A. **80**:810–814.
23. **Subramani, S., and P. Berg.** 1983. Homologous and nonhomologous recombination in monkey cells. Mol. Cell. Biol. **3**:1040–1052.
24. **Tonagawa, S.** 1983. Somatic generation of antibody diversity. Nature (London) **302**:575–581.
25. **Tsui, L.-C., M. L. Breitman, L. Smiminovitch, and M. Buch-**

**wald.** 1982. Persistence of freely replicating SV40 molecules carrying a selectable marker in permissive simian cells. Cell **30**:499–508.

26. **Van Beveren, C., J. A. Galleshaw, V. Jonas, A. Berns, R. F. Doolittle, D. J. Donoghue, and J. M. Verma.** 1981. Nucleotide sequence and formation of the transforming gene of a mouse sarcoma virus. Nature (London) **289**:258–262.
27. **Varmus, H. E., H. Quintrell, and S. Oritz.** 1981. Retroviruses as mutagens: insertion and excision of a non-transforming provirus after expression of a resident transforming provirus. Cell **25**:23–36.
28. **Vogt, P. K.** 1971. Spontaneous segregation of non-transforming viruses from cloned sarcoma viruses. Virology **46**:947–952.
29. **Wagner, M., J. Sharp, and W. Summers.** 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex type 1. Proc. Natl. Acad. Sci. U.S.A. **78**:1441–1445.