Environmental Health Perspectives
Vol. 88, pp. 141–148, 1990

Retroviruses As Genetic Tools to Isolate Transcriptionally Active Chromosomal Regions

by Harald von Melchner* and H. Earl Ruley*

By exploiting the ability of retroviruses to move genes into random sites of mammalian genomes and by exploiting some features of their replication, retrovirus vectors have been developed that select for instances in which the virus integrates into expressed genes. Since integrated proviruses tag transcriptionally active sites, the vectors provide a means to identify and isolate promoters active in different cell types. Furthermore, the viruses may be useful as insertional mutagens, since they select for instances in which integration occurs into expressed sites. This reduces the number of integrants needed to screen for loss of gene function and may enable genes controlling phenotypes in mammalian cells to be isolated.

Introduction

Like transposons in bacteria or retrotransposons in lower eukaryotes (*1–6*), retroviruses are genetic elements capable of moving genes into the genomes of mammalian cells. Upon integration, retroviruses may cause either a recessive loss of gene function due to gene disruption or a dominant gain of function due to transcriptional activation of adjacent genes. Integrations near cellular promoters or enhancers may also activate viral gene expression.

By exploiting the ability of retroviruses to move genes into random sites of mammalian genomes and by taking advantage of some features of their replication, retrovirus vectors have been developed that select for instances in which the virus integrates into expressed genes (*7,8*). After briefly reviewing some aspects of retroviral replication, the present article discusses the development and uses of retrovirus vectors to tag transcriptionally active chromosomal regions.

Replication of Retroviruses

Retroviruses are RNA viruses that replicate through a DNA intermediate. Flanking the ends of the viral genome are short sequence repeats (R) and unique sequences (U5 and U3) that control DNA synthesis, integration, transcription, and RNA processing. Between the control regions are coding sequences for the major structural proteins of the virus particle (*gag* and *env*) and for enzymes (*pol*) found in particles (protease, reverse transcriptase, and integrase) (*2,9–11*).

Shortly after infection, viral RNA is converted into DNA by reverse transcriptase. Prior to integration, terminal sequences of the viral genome join and are duplicated such that the retroviral genome is flanked by long terminal repeats (LTRs), each containing the U3, R, and U5 regions. Circular forms of the viral genome, containing either one or two copies of the LTR, are also found in virus-infected cells. There is evidence that formation of circular molecules with two tandem LTRs creates *cis*-acting recognition sequences for the enzymes catalyzing integration. Thus, sequences containing a spleen necrosis virus LTR-LTR junction when inserted into DNA at an internal site were sufficient to promote integration. This suggests that circular genomes containing two LTRs can serve as integration precursors (*12,13*). However, several investigators have shown that, at least *in vitro*, linear viral DNA can integrate directly without forming a circularized intermediate (*14,15*). The integration steps appear similar to those observed in *E.coli* during phage Mu transposition, involving cleavage at the integration site; joining of the retroviral 3′ ends to protruding 5′ ends of the target DNA; and repair of remaining single-stranded gaps (*14*). Whether linear DNA also serves as an integration precursor *in vivo* remains to be established.

LTR sequences are maintained in the integrated retrovirus (also termed provirus) except that two nucleotides (nt) are lost from each end (*2,9,10*). Cellular DNA sequences are also unaltered except that upon integration, 4 to 6 nt are duplicated such that the provirus is flanked at each end by 4 to 6 bp repeats (*8,10,13*).

*Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, 40 Ames Street, Cambridge, MA 02139.

Address reprint requests to H. von Melchner, Center for Cancer Research, Massachusetts Institute of Technology, 40 Ames Street, Cambridge, MA 02139.

As a provirus, the retroviral genome is replicated with cellular DNA and transcribed as a cellular gene. Provirus transcription is controlled by promoter/enhancer sequences located in the U3 region of the 5′ LTR. Transcripts initiate at the junction between U3 and R (cap site) in the 5′ LTR and terminate at the R-U5 junction in the 3′ LTR. RNA is synthesized by cellular RNA polymerase II and processed by the cellular enzymes. Full-length (genomic) RNA is transported from the nucleus to the cytoplasm and either packaged into virus particles that bud from the cell or translated to yield precursors of the *gag* and *pol* proteins. A fraction of the RNA is spliced to yield mRNA encoding *env*.

It is possible to adapt retroviruses to transduce genes into mammalian genomes. Provided that certain control sequences within the LTRs remain unaltered (*16,17*), much of the retroviral genome can be deleted without impairing its ability to replicate in cells that express proteins necessary for reverse transcription, integration, and particle formation. To do this, vector DNA is transfected into cell lines that contain complete retroviral genomes. The helper virus cannot assemble into particles due to a small deletion that removes sequences (termed Ψ) required for assembly of viral RNA into infectious virions. Since vector DNA does not contain the Ψ deletion, recombinant transcripts are packaged and expelled from the cells as virus particles (Fig. 1). In addition to Ψ, *gag* sequences also enhance the ability of the vectors to be packaged (*16,18–23*).

Retroviruses appear to integrate randomly throughout the genome, although about one-fifth of all integrations involve highly preferred sites (*24*). Integration sometimes results in mutations that either inactivate or augment expression of genes in the vicinity of the provirus (*9,25–27*). Gene inactivation may be caused by insertions into exons (*28–30*) that interrupt open reading frames or introns (*31,32*) and alter normal splicing patterns. Activation of genes adjacent to the provirus involves transcriptional enhancement either by upstream U3 promoters or nearby U3 enhancers. By some yet unknown mechanism, the activity of 3′ LTRs is inhibited by 5′ LTRs (*33*). Thus, in most cases, where downstream cellular genes are activated by 3′ LTR promoters, 5′ LTRs are either rearranged or lost (*34–40*).

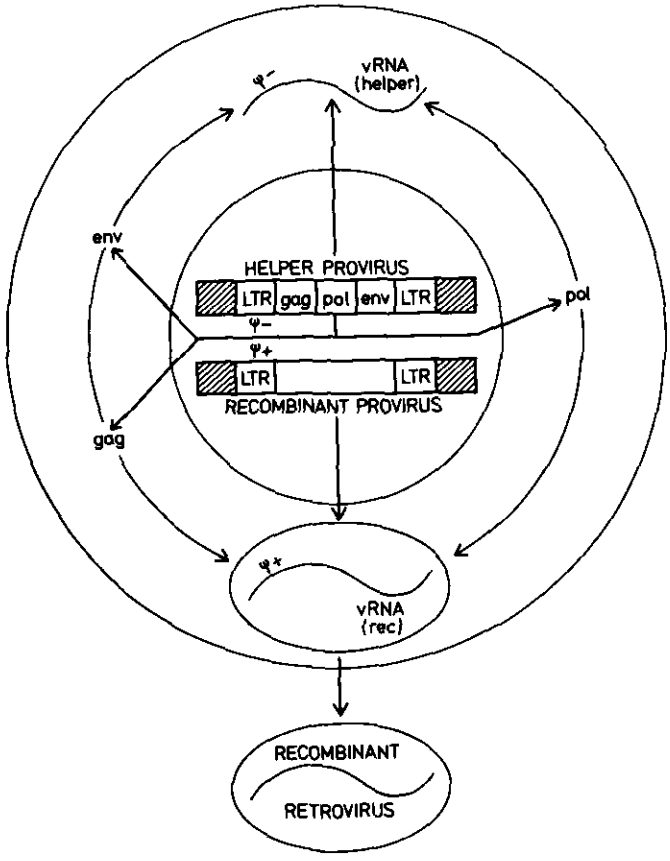

FIGURE 1. Schematic representation of a retrovirus packaging cell. Shaded areas represent chromosomal flanking regions. v, virion; rec, recombinant. Reprinted with permission (*80*).

Development of a Retroviral Promoter Trap

Transcriptional activation of proviral genes has been observed in cells in which the LTR is transcriptionally inactive. In some instances activation was due to insertions near cellular promoters or enhancers (*41–44*). Thus, in principle it is possible to tag transcriptionally active chromosomal regions by infecting cells with retrovirus vectors that contain a selectable marker (e.g., antibiotic resistance) or reporter gene (e.g., *lacZ*) and selecting for recombinants in which the genes are expressed. However, several factors have undermined the practical use of retroviruses as probes for transcriptionally active chromosomal regions. First, enhancers in the LTRs may influence the expression of adjacent genes and thus interfere with the detection of cellular sequences that regulate transcription in a tissue-specific manner (*17,44–47*). Second, 3′ RNA processing signals and AUG codons within the lefthand LTR may interfere with translation of proviral transcripts initiated by nearby cellular promoters, since it has been shown that AUG or termination codons upstream of the protein synthesis initiating codon often inhibit efficient translation and lead to aberrant translation products (*48–52*). For these reasons, a retroviral vector aimed at detecting cellular promoters should lack both enhancers and viral sequences between proviral genes and flanking DNA.

The strategy we have used (*7,8*) involved inserting a selectable marker, histidinol dehydrogenase (*hisD*) (*53*) into the 3′ LTR of an enhancerless Moloney murine leukemia virus (MoMuLV) (Fig. 2). The extra sequence did not interfere with the ability of the virus to be passaged and *his* sequences duplicated normally. Coding sequences of *hisD* in the 5′ LTR were placed just 30 nt from the flanking cellular DNA (Fig. 3A). This removed intervening viral sequences that could interfere with transcriptional activation of provirus genes by cellular

promoters. The virus also contained a second selectable marker, neomycin-phosphotransferase (*neo*), expressed from an internal herpes simplex virus thymidine kinase (*tk*) promoter, to provide an independent measure of virus infectivity (Fig. 2).

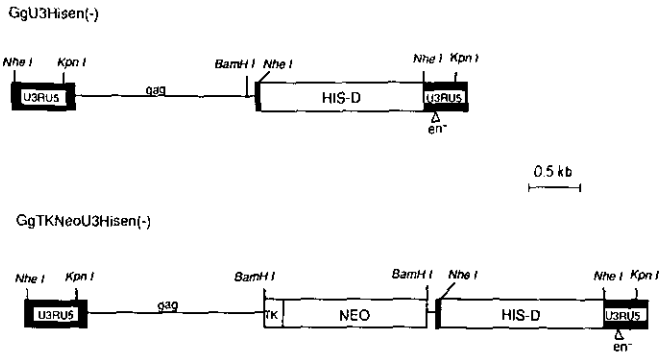

FIGURE 2. Structures of recombinant retrovirus vectors. Gene sequences present in each virus are labeled as follows: HIS-D, *Salmonella typhimurium* histidinol-dehydrogenase; NEO, neomycin-phosphotransferase; gag, truncated gag region of MoMLV; and TK, HSV thymidine kinase promotor. (en−) Designates a 178 nt deletion in the 3' LTR encompassing the viral enhancer. Reprinted with permission (*7*).

Virus-producing cell lines were generated by transfecting U3His vectors into NIH 3T3 cell lines expressing packaging-defective ecotropic (Ψ2) and amphotropic (PA317) helper viruses. Viruses recovered from cloned producer lines were titered on NIH 3T3 cells, selecting either with G418 (neomycin derivative) or L-histidinol. Neo-titers of U3His viruses were high and similar to what we and others have obtained with other MoMuLV vectors (*19,22,54–57*), indicating that insertion of *his* sequences into the LTR did not significantly interfere with virus replication or integration. However, comparing the ratio of neomycin and histidinol titers suggested that provirus integration was 2500-fold less likely to convert cells to a histidinol-resistant (His^r) phenotype than to a neomycin resistant (Neo^r) phenotype.

In principle, the potential to express histidinol resistance could be an intrinsic but inefficient property of each provirus, or alternatively may require secondary events such as mutations or transcriptional activation by adjacent cellular sequences. Several experiments (*7*) suggested that the capacity to transduce histidinol resistance is neither an intrinsic property of the provirus nor the consequence of mutations: α) clones initially selected in G418 did not survive when transferred to me-

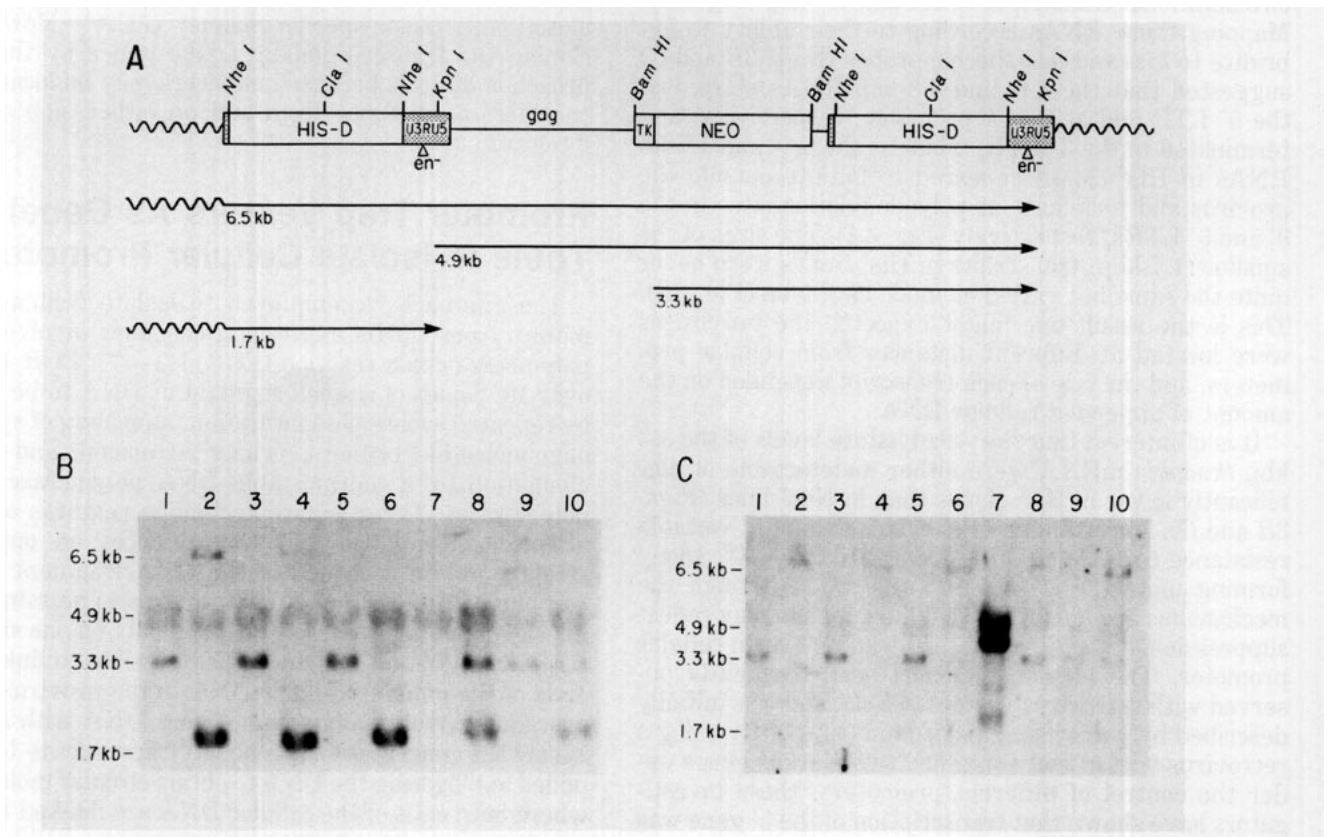

FIGURE 3. Analysis of proviral transcripts. (*A*) GgTKNeoU3Hisen(−) provirus structure, predicted transcripts, and sites of transcription initiation in Neo^R and His^R clones. Total cellular RNA (10 μg/lane) was extracted from Neo^R and His^R clones, fractionated on formaldehyde/agarose gels, transferred to nylon membranes, and hybridized to *his* (*B*) or *neo* (*C*) specific probes as follows: 1, 3, 5, 7, 9 are clones selected in G418; 2, 4, 6, 8, 10 are clones selected in L-histidinol. The sizes of major RNA species as indicated were estimated from RNA standard (BRL). Reprinted with permission (*7*).

dium containing L-histidinol, indicating that most proviruses did not confer histidinol resistance; *b*) the number of doubly resistant colonies produced after plating U3His infected cells in medium containing both G418 and L-histidinol was similar to the number of colonies obtained after selection in L-histidinol alone, implying that only a subset of the proviruses conferring neomycin resistance was capable of expressing histidinol resistance; *c*) virus titers were similar for U3His vectors with or without *tk* promoters, indicating that the ability to passage histidinol resistance did not require the *tk* promotor; *d*) proviruses in His^r clones lacked gross sequence rearrangements as judged by Southern blot analysis; and *e*) proviruses rescued from His^r clones after superinfection with wild-type MoMuLV did not transduce *his* any more efficiently (as compared with *neo*) than did the original U3His vectors.

To further examine why only certain proviruses expressed *his*, transcription of provirus sequences in His^r and Neo^r clones was analyzed by Northern blot hybridization (Figs. 3*A*–*C*). The results indicated that the histidinol-resistant phenotype is conferred by transcripts initiating in the nearby provirus flanking region. Both Neo^r and His^r clones expressed 4.9- and 3.3-kbp provirus transcripts, whereas only His^r clones expressed two additional transcripts of 6.5 and 1.7 kbp (Fig. 3*B*). Mapping these RNAs according to their ability to hybridize to *his*- and *neo*-specific probes (Figs. 3*B* and *C*) suggested that the 4.9- and 3.3-kbp RNAs started at the 5′ LTR and at the *tk* promoter, respectively, and terminated in the 3′ LTR, whereas the 6.5- and 1.7-kb RNAs in His^r clones appeared to initiate outside the provirus and terminate at polyadenylation sites in the 3′ and 5′ LTRs, respectively (Fig. 3*A*). The sizes of the smaller (1.7 kbp) transcripts in His^r clones were never quite the same but varied as much as 100 bp (Fig. 3*B*). This is the result one might expect if the proviruses were located at different distances from cellular promoters, and the size of each transcript depended on the amount of appended cellular RNA.

It is of interest that the steady-state levels of the 3.3 kbp *tkneohis* mRNA were either undetectable or significantly lower in His^r clones than in Neo^r lines (Figs. 3*B* and *C*). Accordingly, His^r clones exhibited variable resistance to G418 ranging from 2 to 2×10^3 colony forming units (CFU) per 10^5 cells (*7*). Although the mechanisms are unknown, the phenomenon may reflect suppression of the *tk* promoter by an upstream cellular promoter. The phenomenon has been frequently observed with retrovirus vectors (*58–63*) and was initially described by Emermann and Temin (*64,65*). By using a retrovirus vector that contained two marker genes under the control of different promoters, these investigators have shown that transcription of the 5′ gene was suppressed when there was selection for the 3′ gene and vice-versa, whereby the suppressed genes produced about 10 to 50% less product than when they were selected (*65*).

Ribonuclease protection assays confirmed that transcripts conferring histidinol resistance are initiated in the nearby flanking region. Accordingly, labeled RNA probes complementary to the provirus coding strand protected fragments of exactly the size expected for transcripts colinear with provirus sequences extending from the *Cla*I site to the 5′ end of the LTR (Fig. 3*A*). Thus, the ability of U3His viruses to convert cells to a His^r phenotype requires that the provirus acquire a promoter from the flanking cellular DNA.

Several investigators have isolated cellular promoters or enhancers by linking random DNA fragments to the coding sequence of a selectable marker, introducing the DNA into recipient cells and selecting for clones that result if the gene is expressed (*66–70*). However, the strategy has several limitations that the U3His vectors avoid. First, DNA-mediated gene transfer is less efficient than retrovirus transduction. Second, introduced gens are frequently amplified in cells surviving selection, thus increasing background and necessitating screening of multiple clones or performing secondary transfections in order to identify clones containing only one gene copy. Third, potential promoter-enhancer elements identified after DNA transfer are not expressed in their normal chromosomal locations.

Transfected enhancerless genes have been used to identify transcriptionally active chromosomal regions (*71–73*). In some cases expression appeared to be regulated in a tissue-specific manner (*71,73*). However, cloning the transcriptionally active genes by this approach is difficult because enhancers may be located at considerable distance from and on either side of the integrants (*10*).

Promoter Trap Vectors As Genetic Tools to Isolate Cellular Promoters

The approach we employed to isolate cellular promoters from U3His expressing cell lines involved the polymerase chain reaction (PCR) (*74–77*). PCR allows over 10^9 copies of a small segment of DNA to be made by repeated cycles of denaturation, annealing of specific oligonucleotide primers, primer extension, and chain elongation by a thermostable DNA polymerase (*Taq* polymerase). The standard technique requires oligonucleotides complementary to sequences on opposite strands and on each end of the DNA fragment to be amplified. Although in U3His provirus expressing cell lines genomic sequences are flanked only on one side by a region of known sequence, it is possible to link both ends of the flanking cellular DNA to the provirus. The procedure involves digesting cellular DNA with an enzyme that generates small enough fragments to be amplified and ligating the DNA to obtain circular molecules where both ends of the cellular DNA are flanked by the provirus (Fig. 4*A*) (*78,79*).

Synthetic oligonucleotide primers complementary to sequences on both sides of the *Pvu*II site at position 72 of the U3His provirus were constructed that would permit extension reactions proceeding away from each other (Fig. 4*A*). To generate small restriction fragments

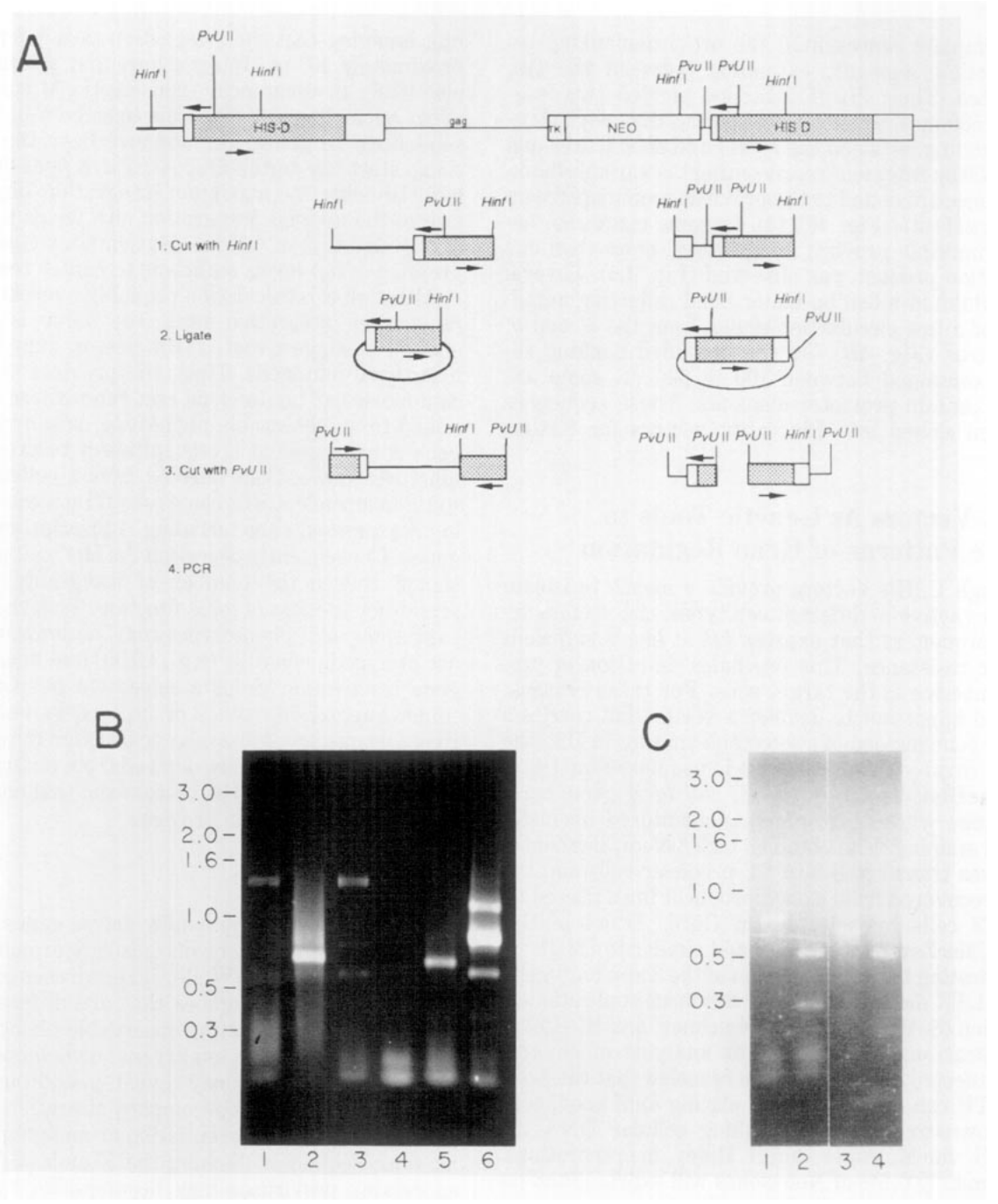

FIGURE 4. Amplification of sequences flanking U3His proviruses. (*A*) Strategy used to generate templates to be amplified by the polymerase chain reaction. (*B*) Ethidium bromide-stained gel showing the PCR products of six His^R cell lines. (*C*) Ethidium bromide-stained gel showing PCR products from clones 2 and 5 (lanes 1 and 3, respectively) digested with *Hinf*I (lanes 3 and 4, respectively). The sizes of the PCR products as indicated were estimated from 1 kbp DNA ladder standard (BRL). Adapted from von Melchner et al. (*8*).

that amplify more efficiently, genomic DNA was digested to completion with *Hinf*I, yielding an average fragment length of 800 bp. Religating *Hinf*I protruding ends positioned 5′ proviral flanking fragments between the proviral priming sites. To avoid background PCR products originating from circles formed at the 3′ end of the provirus, the DNA was digested with *Pvu*II. As shown in Figure 4*A*, this technique separates the priming sites onto different DNA fragments, thus precluding their geometric amplification. Such separation is less likely to occur at the left (5′) end because *Pvu*II sites are an order magnitude less frequent than *Hinf*I sites in mammalian DNA.

PCR products from histidinol-resistant lines were different in size, as one might expect if the proviruses are located at different distances from a flanking *Hinf*I site and the size of each segment depends on the amount of appended cellular DNA (Fig. 4*B*). In each case the am-

plified template represented the provirus-flanking region junction segment, extending between the two *Hinf*I sites. Thus, *Hinf*I digestion yielded two fragments: a 560 bp fragment representing the proviral region extending between the *Pvu*II and *Hinf*I sites and a 100 to 500 bp fragment representing the variable flanking sequence appended to the provirus region upstream of the *Pvu*II site (Fig. 4*C*) (*8*). In some instances, because of multiple proviral integrations, more than one amplification product was observed (Fig. 4*A*). Several clones exhibited a 680 bp minor band reflecting amplification of a few circular molecules from the 3′ end of the provirus (Fig. 4*B*). As the amplified flanking sequences contained between 100 to 500 nt, some are likely to contain promoter elements. These sequences have been cloned into Bluescript vectors for further analysis.

U3lacZ Vectors As Genetic Tools to Analyze Patterns of Gene Regulation

Although U3His vectors provide a means to isolate promoters active in different cell types, the vectors select for promoters that express *his* at levels sufficient to confer resistance. This precludes detection of promoters inactive in the target cells. For these reasons, it seemed important to develop a vector that carries a reporter gene instead of a selectable marker in U3. The strategy involved replacing *his*-D sequences in U3 of GgTKNeoHisU3en(–) by the *E. coli lacZ* gene, a reporter gene whose expression is monitored by histochemical staining. The obtained GgTKNeoU3lacZen(–) vector was transfected into Ψ2 producer cells and the viruses recovered from cloned producer lines titered on NIH 3T3 cells by selecting in G418. Titers of GgTKNeoU3lacZen(–) were high and similar to U3His titers, indicating that the insertion of the 3 kbp *lacZ* gene into the LTR did not interfere with virus replication or integration (S. Reddy, H. von Melchner, and H. Ruley, in preparation). Southern blot analysis of several U3lacZ infected NIH 3T3 clones revealed that the elongated LTR can duplicate, thus placing *lacZ* sequences 30 nt downstream of the flanking cellular DNA (S. Reddy, H. von Melchner, and H. Ruley, in preparation). Approximately 0.5% of Neo clones will express *lacZ* on transcripts initiating in flanking cellular sequences, suggesting that U3lacZ vectors may be used to isolate promoters regulated during different stages of cellular development.

U3His Vectors As Insertional Mutagens

Because selection for U3His expression reduces the size of the integration target to genomic sequences immediately downstream of cellular promoters, U3His vectors may be effective insertional mutagens. It is possible to estimate the maximum number of integration sites that enable NIH 3T3 cells to express histidinol resistance. The total integration target for *neo* transduction cannot exceed the size of the genome (3×10^9 bp), implying that the integration target for *his* is approximately 10^6 nt. Integrations that activate *his* are also likely to occur near transcriptional start sites to avoid appending AUG codons upstream of *his* coding sequences. In genes, the distance from the transcriptional start site to the first AUG averages 50 to 100 nt (*50*). Dividing the maximum integration target by the size of the average integration site yields a maximum of 1×10^4 to 2×10^4 integration sites capable of expressing *his* at levels sufficient to confer resistance.

Although this calculation probably overestimates the number of integration sites, the value is still small enough to suggest that U3His vectors may be used as insertional mutagens. Thus, the provirus would serve as a molecular tag to clone any gene whose function is linked to an observable phenotype, assuming that the gene is expressed at levels sufficient to promote histidinol resistance. Cells derived from a collection of 10^4 to 10^5 independent His^r clones should contain proviruses in all expressed sites including transcriptionally active genes. Consequently, selection for His^r cell populations should reduce the number of integrants needed to screen for the loss of gene function. Isolation of clones expressing null phentotypes may be improved by the use of hypodiploid cells (e.g., CHO) and in cases where gene inactivation yields a selectable phenotype (e.g., tumor suppressing genes or antioncogenes). Alternatively, homozygous loss of gene function may be accomplished by breeding mice derived from embryonal stem cells infected with U3His vectors and that contain germ line integrations of the provirus.

Conclusions

By tagging transcriptionally active genes, promoter trap vectors provide a means to isolate promoters active in different cell types. While U3His viruses may be used to identify and clone genes that are differentially expressed in cells generating observable phenotypes during development (e.g., embryonal or hemopoietic stem cells), U3lacZ vectors may reveal, in addition, patterns of gene regulation that accompany alterations in cellular phenotypes (e.g., differentiation or malignant transformation). Sequences flanking the 5′ ends of *his* or *lacZ* expressing proviruses may be amplified by the polymerase chain reaction cloned, and used as probes to isolate cDNA clones derived from transcripts of genes expressed during distinct stages of cellular development.

Finally, U3His vectors may make effective insertional mutagens because they select for instances in which integration occurs into expressed sites. This reduces the number of integrants needed to screen for loss of gene function and may enable genes controlling phenotypes in mammalian cells to be isolated.

This work was supported by a grant from the National Cancer Institute (PO1 CA42063) to H.E.R. and partially supported by a Cancer Center Core Grant (14051) to P.A. Sharp. H.v.M. was supported by the Deutsche Forschungsgemeinschaft, Bonn-Bad Godes-

berg, West-Germany. H.E.R. is a Rita Allen Foundation scholar and Latham Professor of Biology.

REFERENCES

1. Movable Genetic Elements. Cold Spring Harbor Symp. Quant. Biol. 45 (1980).
2. Baltimore, D. Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. Cell 40: 481–482 (1985).
3. Bellofatto, V., Shapiro, L., and Hodgson, D. A. Generation of Tn5 promoter probe and its use in the study of gene expression in *Caulobacter crescentus*. Proc. Natl. Acad. Sci. USA 81: 1035–1039 (1984).
4. de Bruijn, F. J., and Lupski, J. The use of transposon Tn5 mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids—a review. Gene 27: 131–149 (1984).
5. Cooley, L., Kelley, R., and Spradling, A. Insertional mutagenesis of the *Drosophila* genome with single P elements. Science 239: 1121–1128 (1988).
6. Shapiro, J. A. Mobile genetic elements. In: Academic Press, New York, 1983.
7. von Melchner, H., and Ruley, H. E. Identification of cellular promoters by using a retrovirus promoter trap. J. Virol. 63: 3227–3233 (1989).
8. von Melchner, H., Reddy, S., and Ruley, H. E. Isolation of cellular promoters using a retrovirus promoter trap. Proc. Natl. Acad. Sci. USA, in press.
9. Varmus, H. Retroviruses. Science. 240: 1427–1435 (1988).
10. Varmus, H., and Swanstrom, R. Replication of retroviruses. In: RNA Tumor Viruses, Vol. 2 (R. Weiss, Ed.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1984, pp. 75–134.
11. Panganiban, A. T., and Fiore, D. Ordered interstrand and intrastrand DNA transfer during reverse transcription. Science 241: 1064–1069 (1988).
12. Panganiban, A. T., and Temin, H. M. Circle with two tandem LTRs are precursors to integrated retrovirus DNA. Cell 36: 673–679 (1984).
13. Panganiban, A. T. Retroviral DNA integration. Cell 42: 5–6 (1985).
14. Fujiwara, T., and Mizuuchi, K. Retroviral DNA integration: structure of an integration intermediate. Cell 54: 497–504 (1988).
15. Brown, P. O., Bowerman, B., Varmus, H. E., and Bishop, J. M. Correct integration of retroviral DNA *in vitro*. Cell 49: 347–356 (1987).
16. Murphy, J. E., and Goff, S. P. Construction and analysis of deletion mutations in the U5 regions of Moloney murine leukemia virus: effects on RNA packaging and reverse transcription. J. Virol. 63: 319–327 (1989).
17. Dougherty, J. P., and Temin, H. Promotorless retroviral vector indicates that there are sequences in U3 required for 3′ RNA processing. Proc. Natl. Acad. Sci. USA 84: 1197–1201 (1987).
18. Cone, R. D., and Mulligan, R. C. High efficiency gene transfer into mammalian cells: generation of helper free recombinant retroviruses with broad host range. Proc. Natl. Acad. Sci. USA 81: 6349–6353 (1984).
19. Danos, O., and Mulligan, R. C. Safe and efficient generation of recombinant retroviruses with amphotropic and ecotrophic host ranges. Proc. Natl. Acad. Sci. USA 85: 6460–6464 (1988).
20. Gorelick, R. J., Henderson, L. E., Hanser, J. P., and Rein, A. Point mutants of Moloney murine leukemia virus that fail to package viral RNA: evidence for specific RNA recognition by a "zinc finger-like" protein sequence. Proc. Natl. Acad. Sci. USA 85: 8420–8424 (1988).
21. Mann, R., Mulligan, R. C., and Baltimore, D. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell 33: 153–159 (1983).
22. Miller, A. D., and Buttimore, C. Redesign of retrovirus packaging cell line to avoid recombination leading to helper virus production. Mol. Cell. Biol. 6: 2895–2902 (1986).
23. Miller, A. D., Bender, M. A., Harris, E. A., Kaleko, M, and Gelinas, R. E. Design of retrovirus vectors for transfer and expression of the human beta-globin gene. J. Virol. 62: 4337–4345 (1988).
24. Shih, C. C., Stoye, J. P., and Coffin, J. M. Highly preferred targets for retrovirus integration. Cell 53: 531–537 (1988).
25. Jaenisch, R., Breindl, M., Harbers, K., Jahner, D., and Lohler, J. Retroviruses and insertional mutagenesis. Cold Spring Harb. Symp. Quant. Biol. 50: 439–445 (1985).
26. Jaenisch, R., and Soriano, P. Retroviruses as tools for mammalian development. Symp. Fundam. Cancer Res. 39: 59–65 (1986).
27. Temin, H. M. Evolution of cancer genes as a mutation-driven process. Cancer Res. 48: 1697–1701 (1988).
28. Hartung, S., Jaenisch, R., and Breindl, M. Retrovirus insertion inactivates mouse alpha 1(I) collagen gene by blocking initiation of transcription. Nature. 320: 365–367 (1986).
29. King, W., Patel, M., Lober, L., Goff, S., and Nguyen-Huu, M. Insertion mutagenesis of embryonal carcinoma cells by retroviruses. Science 228: 554–558 (1985).
30. Weinstein, Y., Ihle, J. N., Lavu, S., and Reddy, E. P. Truncation of the c-myb gene by a retroviral integration in an interleukin 3-dependent myeloid leukemia cell line. Proc. Natl. Acad. Sci. USA 83: 5010–5014 (1986).
31. Harbers, K., Kuehn, M., Delius, H., and Jaenisch, R. Insertion of retrovirus into the first intron of alpha 1(I) collagen gene to embryonic lethal mutation in mice. Proc. Natl. Acad. Sci. USA 81: 1504–1508 (1984).
32. Rosson, D., Dugan, D., and Reddy, E. P. Aberrant splicing events that are induced by proviral integration: implications for *myb* oncogene activation. Proc. Natl. Acad. Sci. USA 84: 3171–3175 (1987).
33. Cullen, B. R., Lomedico, P. T., and Ju, G. Transcriptional interference in avian retroviruses—implications for the promoter insertion model of leukaemogenesis. Nature 307: 241–245 (1984).
34. Adler, H. T., Reynolds, P. J., Kelley, C. M., and Sefton, B. M. Transcriptional activation of *lck* by retrovirus promoter insertion between two lymphoid-specific promoters. J. Virol. 62: 4113–4122 (1988).
35. Ikawa, S., Hagino, Y. K., Kawai, S., Yamamoto, T., and Toyoshima, K. Activation of the cellular *src* gene by transducing retrovirus. Mol. Cell. Biol. 6: 2420–2428 (1986).
36. Isfort, R., Witter, L. R., and Kung, H. J. C-*myc* activation in an unusual retrovirus-induced avian T-lymphoma resembling Marek's disease: proviral insertion 5′ of exon one enhances the expression of an intron promoter. Oncogene Res. 2: 81–94 (1987).
37. Li, Y., Holland, C. A., Hartley, J. W., and Hopkins, N. Viral integration near c-*myc* in 10–20% of *mcf* 247-induced AKR lymphomas. Proc. Natl. Acad. Sci. USA 81: 6808–6811 (1984).
38. Morishita, K., Parker, D. S., Mucenski, M. L., Jenkins, N. A., Copeland, N. G., and Ihle, J. N. Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid leukemia cell lines. Cell 54: 831–840 (1988).
39. Payne, G. S., Bishop, J. M., and Varmus, H. E. Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. Nature 295: 209–213 (1982).
40. Westaway, D., Papkoff, J., Moscovici, C., and Varmus, H. E. Identification of a provirally activated *c-Ha-ras* oncogene in an avian nephroblastoma via a novel procedure: cDNA cloning of a chimeric viral-host transcript. *EMBO* J. 5: 301–309 (1986).
41. Barklis, E., Mulligan, R. C., and Jaenisch, R. Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell 47: 391–399 (1986).
42. Jaenisch, R., Jaehner, D., Nobis, P., Simon, I., Loehler, J., Harbers, K., and Grotopp, D. Chromosomal position and activation of retroviral genomes inserted into the germ line of mice. Cell 24: 519–529 (1981).
43. Sorge, J., Cutting, A. E., Erdman, V. D., and Gautsch, J. W. Integration-specific retrovirus expression in embryonal carcinoma cells. Proc. Natl. Acad. Sci. USA 81: 6627–6631 (1984).
44. Taketo, M., Gilboa, E., and Sherman, M. Isolation of embryonal carcinoma cell lines that express integrated recombinant genes flanked by the Moloney murine leukemia virus long terminal repeat. Proc. Natl. Acad. Sci. USA 82: 2422–2426 (1985).
45. Fung, Y. K. T., Fadly, A. M., Crittenden, L. B., and Kung, H.

J. On the mechanism of retrovirus-induced avian lymphoid leukosis: Deletion and integration of the proviruses. Proc. Natl. Acad. Sci. USA 78: 3418–3422 (1981).
46. Fung, Y. K. T., Lewis, W. G., Kung, K. J., and Crittenden, L. B. Activation of the cellular *c-erbB* by LTR insertion: molecular basis for induction of erythroblastosis by avian leukosis virus. Cell 33: 357–368 (1983).
47. Hayward, W. S., Neel, B. G., and Astrin, S. M. Activation of a cellular *onc*-gene by promotor insertion in ALV-induced lymphoid leukosis. Nature 290: 475–480 (1981).
48. Cigan, A. M., Pabich, E. K., and Donahue, T. F. Sequence and structural features associated with translational initiator regions in yeast—a review. Gene 59: 1–18 (1987).
49. Donahue, T. F., and Cigan, A. M. Genetic selection for mutations that reduce or abolish ribosomal recognition of the *HIS4* translational initiator region. Mol. Cell. Biol. 8: 2955–2963 (1988).
50. Kozak, M. Evaluation of the "scanning model" for initiation of protein synthesis in eucaryotes. Cell 22: 7–8 (1980).
51. Kozak, M. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs, Nucl. Acids Res. 15: 8125–8148 (1987).
52. Kozak, M. Effects of intercistronic length on the efficiency of reinitiation by eurcaryotic ribosomes. Mol. Cell. Biol. 7: 3438–3445 (1987).
53. Hartmann, S. C., and Mulligan, R. C. Two dominant-acting selectable markers for gene transfer studies. Proc. Natl. Acad. Sci. USA 85: 8047–8051 (1988).
54. Bender, M. A., Palmer, T. D., Gelinas, R. E., and Miller, D. Evidence that the packaging signal of Moloney murine leukemia virus extends into the gag region. J. Virol. 61: 1639–1646 (1987).
55. Cone, R. D., Reilly, E. B., Eisen, H. N., and Mulligan, R. C. Tissue-specific expression of functionally rearranged lambda 1 Ig gene through a retrovirus vector. Science 236: 954–957 (1987).
56. Guild, B. C., Finer, M. H., Housman, D. E., and Mulligan, R. C. Development of retrovirus vectors useful for expression of genes in cultured murine embryonal cells and hematopoietic cells *in vivo*. J. Virol. 62: 3795–3801 (1988).
57. Mann, R., and Baltimore, D. Varying the position of a retrovirus packaging sequence results in the encapsidation of both unspliced and spliced RNAs. J. Virol. 54: 401–407 (1985).
58. Armentano, D., Yu, S. F., Kantoff, P. W., von Ruder, T., Anderson, W. F., and Gilboa, E. Effect of internal viral sequences on the utility of retroviral vectors. J. Virol. 61: 1647–1650 (1987).
59. Bowtell, D. D., Cory, S., Johnson, G. R., and Gonda, T. J. Comparison of expression in hemopoietic cells by retroviral vectors carrying two genes. J. Virol. 62: 2464–2473 (1988).
60. Garver, R. I. J., Chytil, A., Karlsson, S., Fells, G. A., Brantly, M. L., Courtney, M., Kantoff, P. W., Nienhuis, A. W., Anderson, W. F., and Crystal, R. G. Production of glycosylated physiologically "normal" human alpha 1-antitrypsin by mouse fibroblasts modified by insertion of a human alpha 1-antitrypsin cDNA using a retroviral vector. Proc. Natl. Acad. Sci. USA 84: 1050–1054 (1987).
61. McIvor, R. S., Goddard, J. M., Simmonsen, C. C., and Martin, D. W. J. Expression of a cDNA sequence encoding human purine nucleotide phosphorylase in rodent and human cells. Mol. Cell. Biol. 5: 1349–1357 (1985).
62. Williams, D. A., Orkin, S. H., and Mulligan, R. C. Retrovirus-mediated transfer of human adenosine deaminase gene sequences into cells in culture and into murine hematopoietic cells in vivo. Proc. Natl. Acad. Sci. USA 83: 2566–2570 (1986).
63. Williams, D. A., and Orkin, S. H. Somatic gene therapy. Current status and future prospects. J. Clin. Invest. 77: 1053–1056 (1986).
64. Emerman, M., and Temin, H. M. Genes with promoters in retrovirus vectors can be independently suppressed by an epigenetic mechanism. Cell 39: 459–467 (1984).
65. Emerman, M., and Temin, H. M. Quantitative analysis of gene suppression in integrated retrovirus vectors. Mol. Cell. Biol. 6: 792–800 (1986).
66. Fried, M., Griffith, M., Davies, B., Bjursell, G., Mantia, G. L., and Lania, L. Isolation of cellular DNA sequences that allow expression of adjacent genes. Proc. Natl. Acad. Sci. USA 80: 2177–2121 (1983).
67. Hiller, S., Hengstler, M., Kunze, M., and Knippers, R. Insertional activation of a promotorless thymidine kinase gene. Mol. Cell. Biol. 8: 3298–3302 (1988).
68. Santangelo, G. M., Tornow, J., and Moldave, K. Cloning of open reading frames and promotors from the Saccharomyces cerevisiae genome: construction of genomic libraries of random small fragments. Gene 46: 181–186 (1986).
69. Santangelo, G. M., Tornow, J., McLauglin, C. S., and Moldave, K. Properties of promotors cloned randomly from the Saccharomyces cerevisiae genome. Mol. Cell. Biol. 8: 4217–4224 (1988).
70. Weber, F., Villiers, J., and Schaffner, W. An SV40 "enhancer trap" incorporates exogenous enhancers or generates enhancers from its own sequences. Cell 36: 983–992 (1984).
71. Bhat, K., McBurney, M., and Hamada, H. Functional cloning of mouse chromosomal loci specifically active in embryonal carcinoma stem cells. Mol. Cell. Biol. 8: 3251–3259 (1988).
72. Hamada, H. Activation of an enhancerless gene by chromosomal integration. Mol. Cell. Biol. 6: 4179–4184 (1986).
73. Gossler, A., Joyner, A. L., Rossant, J., and Skarnes, W. C. Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science 244: 463–465 (1989).
74. Faloona, F., and Mullis, K. B. Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. Methods Enzymol. 155: 335–350 (1987).
75. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350–1354 (1985).
76. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487–491 (1988).
77. Scharf, S. J., Horn, G. T., and Erlich, H. A. Direct cloning and sequence analysis of enzymatically amplified genomic sequences. Science 233: 1076–1078 (1986).
78. Collins, F. S., and Weissman, S. M. Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. Proc. Natl. Acad. Sci. USA 81: 6812–6816 (1984).
79. Silver, J., and Keerikatte, V. Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus. J. Virol. 63: 1923–1928 (1989).
80. von Melchner, H. and Höffken, K. Retrovirus-mediated gene transfer into hemopoietic cells. Blut 57: 1–5 (1988).