*Proc. Natl. Acad. Sci. USA*
Vol. 86, pp. 5517–5521, July 1989
Genetics

# Analysis of mammalian cell genetic regulation *in situ* by using retrovirus-derived "portable exons" carrying the *Escherichia coli lacZ* gene

(promoter/gene fusion/splicing/growth arrest/cell cycle)

DANIEL G. BRENNER*, SUE LIN-CHAO*, AND STANLEY N. COHEN*†

Departments of *Genetics and †Medicine, Stanford University School of Medicine, Stanford, CA 94305-5120

*Contributed by Stanley N. Cohen, May 1, 1989*

ABSTRACT **Self-inactivating derivatives of Moloney murine leukemia retrovirus containing the *Escherichia coli lacZ* gene were used to detect and study the regulation of transcription initiated at chromosomally located promoters in mouse fibroblasts. The introduction of splice acceptor sites in all three translational reading frames relative to *lacZ* and the inclusion of an in-frame ATG translation start codon in one construct allowed synthesis of β-galactosidase fusion proteins upon insertion of retrovirus vectors containing *lacZ* into introns 3′ to either protein-coding or noncoding exons. Selection of *lacZ*-expressing cells by fluorescence-activated cell sorting and the analysis of β-galactosidase production after serum deprivation has yielded lines in which *lacZ* was fused to genes induced by growth arrest in the $G_0$ state.**

Fusions between promoterless reporter genes that encode easily assayable gene products and the controlling elements of other genes have proved to be valuable tools for studying genetic regulation in both prokaryotic and eukaryotic cells. In bacteria, reporter genes can be fused *in vivo* directly to chromosomal DNA sequences (1) by introducing them into bacteriophage Mu and other transposable elements that can integrate into host chromosomes at sites that are oriented uniquely with regard to the transposon but located randomly throughout the chromosome (for reviews, see refs. 2 and 3). In contrast, fusions between reporter genes and regulatory sequences of mammalian cells commonly have been made *in vitro* and then introduced into test cells as extrachromosomal replicons (for example, see refs. 4–10). Since factors such as gene dosage, position effects, chromatin structure, and the presence of regulatory elements acting in cis can influence gene expression, the results obtained in such studies may not be indicative of normal events. Although eukaryotic gene expression *in situ* has been studied previously by using several reporter gene systems (11–14), it has not been practical to analyze conditionally expressed promoter activity on the chromosomes of mammalian cells by the introduction of reporter genes.

We report here the construction and use of a retrovirus-based vector system that enables the direct fusion of the *Escherichia coli lacZ* gene to chromosomally located regulatory elements of mammalian cells *in situ*. The approach depends on the ability of Moloney murine leukemia virus (Mo-MuLV) *lacZ* gene (Mo-MuLV*lac*) constructs to insert in a structurally unique way at multiple chromosomal locations but at a single site in each cell and upon a strategy that enables the detection of transcriptional activity when the constructs are introduced into expressed chromosomal genes. We have identified and cloned three fibroblast-derived cell lines that synthesize differentially regulated β-galactosidase (β-gal) fusion proteins and contain the Mo-MuLV*lac* sequences in different chromosomal locations. In two of these cell lines, β-gal activity is induced by growth arrest in the $G_0$ state.

## MATERIALS AND METHODS

**Construction of Retroviral Vectors and Cell Culture Procedures.** A promoterless *lacZ* gene was placed in a polylinker spacer region 3′ to a fragment of Mo-MuLV [nucleotides 5411–5513, which carry the splice acceptor site for the *env*-encoded mRNA (15)] at a site that had been cleaved with *Sal* I, *Acc* I, or *Hin*cII endonucleases and then made blunt, to provide multiple translational reading frames, by treatment with the Klenow fragment of *E. coli* DNA polymerase I. The resulting 4.1-kilobase DNA fragment, which also included the simian virus 40 (SV40) early region polyadenylylation signal 3′ to *lacZ*, was introduced by means of *Hin*dIII linkers into a corresponding site of the self-inactivating retrovirus vector pHHAM (16) and was oriented such that transcription from the viral long terminal repeat (LTR) is opposite in direction to the orientation of the *lacZ* gene. DNA (10 μg) from the resulting pMo-MuLVlac plasmid was introduced by calcium phosphate precipitation (17) into the ψ2 packaging cell line (18) together with pSV2neo DNA (1 μg) (5). After selection on the antibiotic G418 (800 μg/ml, BRL) for stable cointegrants, transfected cells were cloned at random and screened by polymerase chain reaction (19) analysis for the presence of Mo-MuLV*lac* RNA in the supernatants (D.G.B., S.L.-C., and S.N.C., unpublished results). A clone yielding a titer estimated to be $1 \times 10^5$ virus particles per ml was expanded and treated with mitomycin C (20 μg/ml) for 2 hr in the dark at 37°C to terminate cellular replication, washed extensively with medium, and cocultivated with NIH 3T3 cells for 72 hr in the presence of Polybrene (4 μg/ml). The cocultured NIH 3T3 cells were subsequently grown in Dulbecco's modified Eagle's medium (GIBCO) containing 10% (vol/vol) fetal calf serum (FCS, GIBCO). Production of the packaged virus continued for 72 hr after addition of mitomycin C, whereas no viable helper cells were detected in cultures after 10 days. All cells employed were tested routinely by Hoechst staining for possible mycoplasma contamination.

**Assay of *lacZ* Expression.** β-gal activity in fixed cells was detected by histochemical staining after incubation for 2–24 hr at 37°C with 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) (20). β-gal activity in cell extracts was assayed as described (7, 21). For transient expression assays of β-gal activity, the plasmid pSVL (9) was introduced along with β-gal test plasmids into COS-7 cells by the addition of DEAE-dextran (7, 22). pSVL carries the luciferase (*lux*)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: β-gal, β-galactosidase; X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactoside; FCS, fetal calf serum; SV40, simian virus 40; Mo-MuLV, Moloney murine leukemia virus; LTR, long terminal repeat.

gene, whose expressed product was used to indicate the efficiency of DNA uptake. The substrate fluorescein di-β-D-galactopyranoside (Molecular Probes) was used to identify and select *lacZ*-expressing viable cells by fluorescence-activated cell sorting (23, 24) on a Coulter Epics V dual laser cell sorter with a 525-nm band-pass filter. In reconstruction experiments, clones of expressing cells representing at least 1% of the population, as judged by subsequent staining with X-Gal (which may not detect low levels of *lacZ* expression detectable by the cell sorter), could be selected directly by sorting; populations having a lower frequency of expressing cells were initially sorted to enrich for subpopulations that produced β-gal prior to the cloning of individual cells.

For immunoblots, cell lysates made in 25 mM Tris Cl, pH 7.4/50 mM NaCl/0.5% sodium deoxycholate/2% (vol/vol) Nonidet P-40/0.1% $NaDodSO_4$/1 mM phenylmethylsulfonyl fluoride were subjected to immunoprecipitation with rabbit polyclonal antibody to β-gal (5 Prime to 3 Prime, Inc.) and protein A-Sepharose (Sigma). Precipitates were electrophoresed on $NaDodSO_4$/6% polyacrylamide gels, transferred to nitrocellulose, and tested for immunoreactivity using the ProtoBlot Western blot alkaline phosphatase system (Promega) and mouse monoclonal anti-β-gal antibody (Promega) diluted 1:1500.

**Southern Blot Analysis.** Genomic DNA digested with restriction endonucleases was subjected to electrophoresis in 0.7% agarose gels, blotted onto nitrocellulose paper (25), and hybridized with a $^{32}P$-labeled probe consisting of a 4.1-kilobase *lacZ* DNA fragment that had been radioactively labeled by random priming (26).

## RESULTS

**Design of Mo-MuLV*lac* Retroviral Vectors.** To identify and study the expression of chromosomally located mammalian cell genes active under particular conditions or at specific times, we constructed a series of *lacZ* gene-based cassettes in which the *lac* promoter and other upstream DNA sequences have been removed. As the addition of amino acids at the N-terminal end of a truncated β-gal protein does not affect its enzymatic activity (27), fusion of *lacZ* 3′ to protein-coding segments of other genes is practical, and analysis of translational as well as transcriptional control can be carried out by using the same constructs. Mo-MuLV*lac* retroviruses were made by introducing the promoterless *lacZ* gene into Mo-MuLV self-inactivating viruses, which delete both the enhancer and promoter of the viral LTR prior to integration into the chromosome (16, 28, 29), which facilitates detection of *lacZ* gene expression initiated at host promoters.

In the Mu-d*lac* constructs that served as a paradigm for the present experiments, the *lacZ* reporter gene carries its own translation initiation sequence, thus allowing the synthesis of a polycistronic message upon insertion of Mu-d*lac* anywhere within a gene that is downstream from an active promoter (1). Introns comprise much of the length of mammalian genes, and insertions of *lacZ* gene mRNA sequences into intron-encoded segments of transcripts will be removed by splicing during maturation of mRNA. We therefore introduced a splice acceptor site 5′ to the *lacZ* gene in all three translational reading frames, thus permitting the formation of translational fusions with *lacZ* when Mo-MuLV*lac* "portable exons" are inserted into introns (Fig. 1*A*). The LTR upstream from the splice acceptor site in these vectors contains translation stop codons in all three reading frames. A construct (construct A) that additionally contains an ATG translation start codon in frame with the *lacZ* gene was used to detect *lacZ* expression when Mo-MuLV*lac* integrates into introns that follow nontranslated first exons; earlier observations have suggested that retroviral insertions may occur preferentially at DNase I hypersensitive sites near the 5′ ends of genes (30, 31). The various *lacZ*-containing cassettes were introduced into Mo-MuLV self-inactivating derivatives that in the proviral state do not yield viral transcripts (16). As self-inactivation is sometimes incomplete, cassettes were introduced in an orientation opposite to that of normal retroviral transcription.

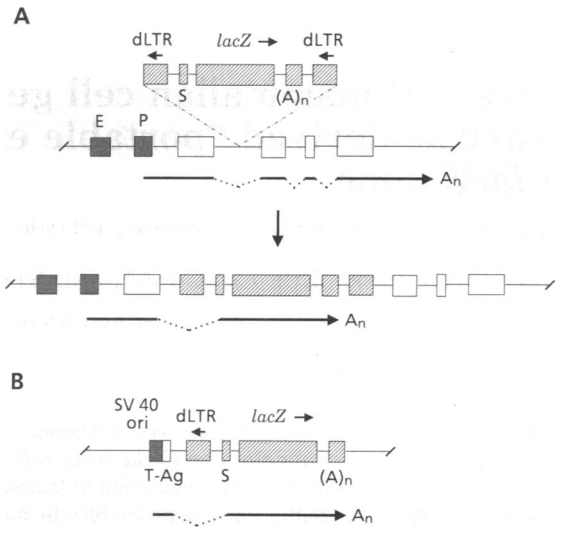


FIG. 1. (*A*) Structure of prototype Mo-MuLV*lac* portable exon. Hatched boxes represent the components of the vector DNA: the *E. coli lacZ* gene (*lacZ*), oriented as indicated, is flanked 5′ by a splice acceptor site (S), which allows *lacZ*-encoded mRNA to be spliced directly to the transcript segment encoded by an upstream exon, and 3′ by a polyadenylylation signal [$(A)_n$] derived from SV40. These three elements are carried by a self-inactivating retroviral vector such that the proviral LTRs (labeled dLTR, to designate that an inactivating deletion has occurred) are oriented opposite to the orientation of *lacZ*. Hypothetical promoter (P) and enhancer (E) elements located in the chromosome are indicated (crosshatched boxes). Open boxes represent hypothetical exonic DNA sequences that flank the insertion site. The site of insertion of the virus into an intron of an actively expressed gene, whose spliced polyadenylylated ($A_n$) RNA is represented by a heavy horizontal line, is shown schematically; the dotted lines represent introns. Below this, the provirus is shown as an integrated sequence, and the resulting transcript, in which the mRNA segment encoded by the first exon of the chromosomal gene is now spliced to the *lacZ* message, is also shown. (*B*) Structure of the region of the plasmid pSAS/Cont used to test functionality of the *lacZ* cassette's splice acceptor site configuration. SV40 sequences (SV40 ori, crosshatched box) that contain promoter and enhancer elements direct transcription of the mRNA (heavy line) as indicated. The adjoining first exon of the SV40 large tumor antigen gene (T-Ag, open box), which carries a splice donor site at its 3′ end, and part of the intron that follows it are 5′ to the *lacZ* cassette. The mRNA segment removed by splicing is shown (dotted line).

**Functional Testing of Design Strategy.** Previous work has shown that separately derived splice donor and splice acceptor sites can interact functionally (32). Plasmids were con-

Table 1. Transient expression of β-gal after transfection of COS-7 cells

| Plasmid DNA | β-gal expression | |
|---|---|---|
| | % cells stained with X-Gal | Enzyme assay, units/μg |
| None | 0.0 | 4 |
| pCont | 0.0 | 6 |
| pSAS/Cont | 4.1 | 175 |

Cells were stained histochemically, and cell extracts were assayed for β-gal activity 48 hr after transfection. The values given are from a representative experiment and have been corrected for variations in transfection efficiencies.

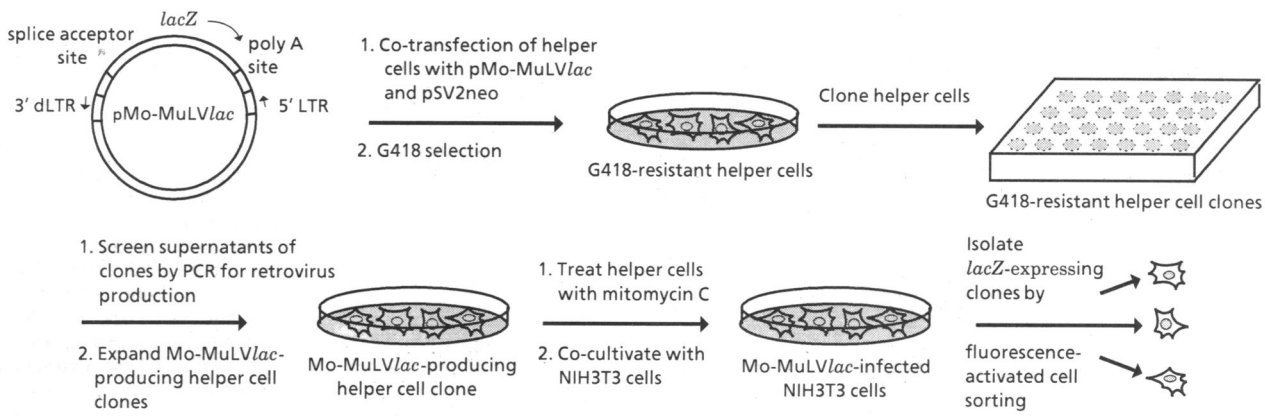


FIG. 2. Experimental design for the isolation of *lacZ*-expressing clones using Mo-MuLV*lac*.

structed to test the functionality of the Mo-MuLV splice acceptor site/*lac* constructs by introducing a *lacZ* cassette downstream from the SV40 virus origin of replication, which contains the large tumor antigen promoter (Fig. 1*B*) (33, 34). The plasmid pSAS/Cont contains a splice acceptor site 5′ to *lacZ*, whereas pCont does not. Plasmids were introduced by transfection into COS-7 cells, and β-gal activity was assessed by histochemical staining and by a transient expression assay (Table 1), both of which showed activation of the inserted Mo-MuLV*lac* cassette and a strict requirement for the splice acceptor site. An S1 analysis (data not shown) confirmed that at least 95% of the *lacZ*-hybridizable mRNA was present in a transcript of the length predicted to result from splicing of the SV40-derived splice donor sequence to the splice acceptor site sequence of the *lacZ* cassette.

**Generation of *lacZ*-Expressing Cell Lines Containing Mo-MuLV*lac* Integrants.** The experimental design is outlined in Fig. 2. DNA from plasmids containing Mo-MuLV*lac* constructs were separately introduced by cotransfection with pSV2neo into Ψ2 cells to rescue the virus in infectious form. After cocultivation of mitomycin C-treated retrovirus-producing helper cell lines with NIH 3T3 cells for several days, histochemical staining with X-Gal showed *lacZ* expression by some of the NIH 3T3 cells (Fig. 3*A*) but not by mock-infected cells. Although *lacZ*-expressing cells were observed with all three reading frame constructs, the frequency of blue-stained cells was 5–10 times greater for construct A, which contained an ATG codon in frame with *lacZ*, and a sample of the infected cells derived from this construct was chosen for further study. Most *lacZ*-expressing cells showed X-Gal staining throughout the cell; however, some cells were stained exclusively in the nucleus (Fig. 3*A*), implying that linkage of the cytoplasmically synthesized β-gal to a peptide sequence directing nuclear localization of the fusion protein had occurred.

Infected NIH 3T3 cells expressing the *lacZ* gene encoded by retroviruses from Mo-MuLV*lac* construct A were viably stained with the β-gal analogue fluorescein di-β-D-galactopyranoside and collected by cell sorting. Initial analysis of the sorted cells showed that 0.1% of the population expressed β-gal; this was increased to 2.5% by growing the sorted cells for a week and resorting the $Lac^+$ population. Following this enrichment step, at least one-third of the cells that survived cloning by cell sorting expressed β-gal, as judged by staining with X-Gal.

**Characterization of *lacZ*-Expressing Clones.** Initial staining with X-Gal of sorted *lacZ*-expressing individual clones showed two distinct patterns: clones in which all cells were stained blue homogeneously and clones showing heterogeneous staining of the progeny of a single cell (Fig. 3 *B* and *C*). A Southern blot (Fig. 4 *Left*) of genomic DNA isolated from clones of these cell types showed the presence of a single integrated copy of Mo-MuLV*lac*. Two clones showing heterogeneous staining (i.e., clones 54.6 and 54.7) and one showing homogeneous staining (i.e., clone 54.2) were selected for further characterization; detailed Southern blot analysis showed a different restriction map at the site of Mo-MuLV*lac* insertion for each clone (Fig. 4 *Right*).

The persistence of staining heterogeneity following multiple serial passages of clones 54.6 and 54.7 suggested that different members of each clonal cell population were expressing *lacZ* to different extents. To investigate the possibility that staining heterogeneity in these clones was the result of synthesis of β-gal during only a certain phase of the cell cycle, these cell lines were subjected to serum deprivation for 2 days to synchronize them in $G_0$ (35, 36). In contrast

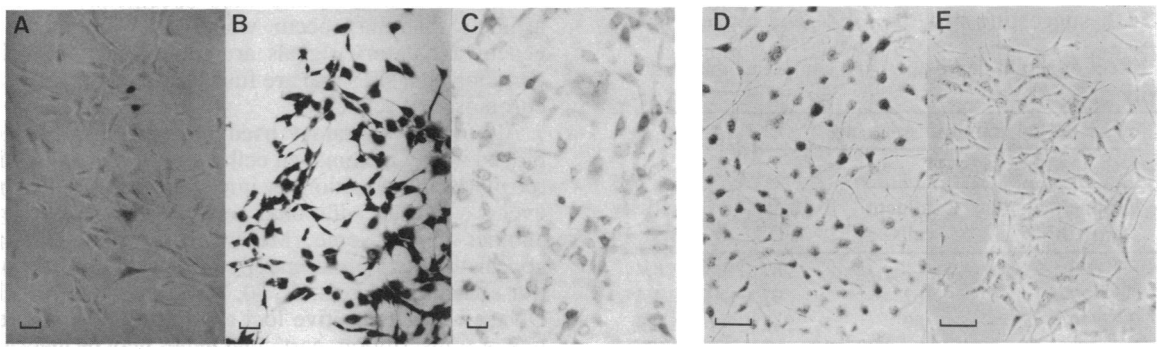


FIG. 3. Histochemical staining of NIH 3T3 cells infected by the Mo-MuLV*lac* construct A. (*A*) After several days of cocultivation, β-gal production was seen only in the nucleus of some cells; the stain is seen throughout in others. Clones were isolated by cell sorting for β-gal expression. Clone 54.2 (*B*) and clone 54.6 (*C*) were maintained in medium containing 10% FCS. (*D*) Clone 54.7 was grown in medium with 0.5% FCS for 48 hr to accomplish growth arrest in $G_0$ (36). (*E*) The growth-arrest block was removed by the addition of 10% FCS, and clone 54.7 was grown for an additional 72 hr. (Bars = 100 μm.)

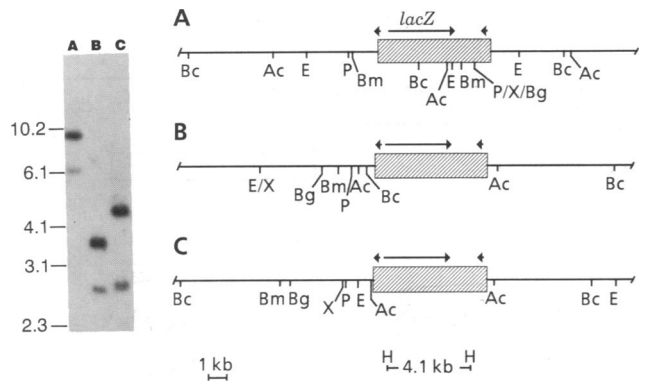


FIG. 4. (*Left*) Southern blot analysis of proviral integration sites. Genomic DNA (4 μg) isolated from cell lines 54.2 (lane A), 54.6 (lane B), and 54.7 (lane C) was treated with *Acc* I and hybridized to a $^{32}$P-labeled 4.1-kilobase (kb) *Hin*dIII fragment from Mo-MuLV*lac* DNA containing the splice acceptor site, *lacZ*, and the polyadenylylation site. The enzyme *Acc* I recognized a unique site within Mo-MuLV*lac*, and each integration site is therefore characterized by the presence of two unique bands. The positions and sizes (in kilobases) of molecular weight standards are indicated. (*Right*) Restriction maps of the chromosomal loci at the integration sites of Mo-MuLV*lac* in cell lines 54.2 (map A), 54.6 (map B), and 54.7 (map C). The data were derived from Southern hybridization analysis with the probe indicated above. Ac, *Acc* I; Bc, *Bcl* I; Bg, *Bgl* II; Bm, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; P, *Pst* I; X, *Xba* I.

to nonsynchronized cells grown in medium containing 10% FCS, most of the $G_0$-synchronized cells now stained blue (Fig. 3 *D* and *E*). Induction of *lacZ* expression by serum deprivation was totally reversible during repeated cycles in 0.5% and 10% FCS. Moreover, *lacZ* expression was increased in clusters of cells that showed the contact inhibition characteristic of high density cell growth. Cell line 54.2, originally found to stain homogeneously with X-Gal after asynchronous growth in medium containing 10% FCS, showed a slight decrease in staining intensity but no change in the frequency of blue staining cells after serum starvation.

The conclusion that the genes fused to *lacZ* in cell lines 54.6 and 54.7 are induced by growth arrest was supported by direct assay of β-gal in extracts obtained from cells synchronized in $G_0$ by serum deprivation, grown to confluence, and maintained in the presence of 10% FCS or allowed to divide actively in media containing 10% FCS (Table 2). In addition, direct assay of the *lacZ* gene product showed that lines 54.6 and 54.7 were induced to different extents by growth arrest, consistent with the differing X-Gal staining intensity observed in these cells following serum starvation. In contrast to the observed induction of *lacZ* expression in lines 54.6 and 54.7, serum starvation resulted in reduced β-gal synthesis in cell line 54.2, suggesting that the *lacZ* gene present in the

Table 2. Effect of serum deprivation and contact inhibition on β-gal activity

| Cell line | β-gal activity,* units/mg | | | |
|---|---|---|---|---|
| | 10% FCS | | | |
| | Exponential growth† | Confluent cells | 0.5% FCS | Activity ratio‡ |
| 54.2 | 698 | — | 287 | 0.4 |
| 54.6 | 22 | 62 | 97 | 4.4 |
| 54.7 | 29 | 89 | 241 | 8.3 |

*The values given are from a representative experiment in which cells were maintained at the indicated FCS concentrations for 48 hr prior to the assay.
†Cells at 20% confluence were used for exponential growth measurements.
‡Ratio of β-gal activity in 0.5% FCS/10% FCS (exponential growth).

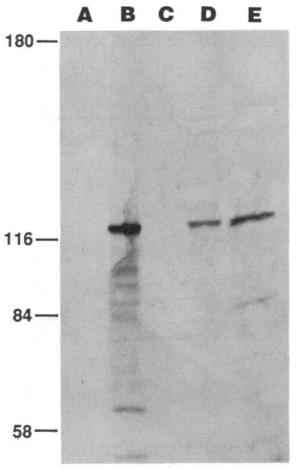


FIG. 5. Western blot analysis of cell extracts made from uninfected NIH 3T3 cells (lane A) or cell line 54.2 (lane B), both cultured in the presence of 10% FCS, and uninfected NIH 3T3 cells (lane C), cell line 54.6 (lane D), or cell line 54.7 (lane E), all maintained for 48 hr in the presence of 0.5% FCS. These extracts contained equal amounts of protein. The positions and sizes (in kDa) of the protein standards used are indicated.

clone is fused to a constitutively expressed gene that shows a decreased level of function when cell growth rate is slowed (36). β-gal synthesis was induced in clones 54.6 and 54.7 by confluent cell growth in 10% FCS (Table 2).

Western blot analysis of protein obtained by immunoprecipitation of extracts from each of the cell lines showed the presence of β-gal protein (Fig. 5). Interestingly, while insertions in both translated and untranslated exons can result in *lacZ* expression with Mo-MuLV*lac* portable exon construct A, all three of the insertions studied yielded β-gal fusion proteins of the predicted length if synthesis was initiated at the in-frame AUG codon located just downstream from the splice acceptor site. This finding is consistent with the notion that Mo-MuLV*lac* portable exon construct A has, in the three cell lines, inserted into an intron immediately 3′ to a nontranslated exon.

## DISCUSSION

The approach reported here is potentially applicable for the identification, mapping, isolation, and study of transcription units of a variety of mammalian genes whose products are not yet known, but whose expression is subject to regulation by specific alterations in growth and/or development or is affected by regulatory genes or chemical agents. The Mo-MuLV*lac* retrovirus constructs, which are mammalian cell analogues of Mu-d*lac* and other transposable elements used previously to study control of gene expression in bacteria, have enabled the identification and cloning of mouse fibroblast cells that contain *lacZ* cassettes inserted in three chromosomally located genes that are differentially affected by growth arrest. The vectors have been designed to generate *lacZ* fusions *in situ* and thus specifically to circumvent artifacts that may occur when DNA fragments carrying genetic regulatory signals are removed from their normal chromosomal sites and are fused to reporter genes on extrachromosomal replicons.

Although the strategy used here can in principle be employed in any mammalian cell line, it assumes that inserts of retrovirus into functionally important genes will be nonlethal events when an additional copy of the gene is present on another chromosome. Moreover, as retroviral insertions appear to occur preferentially at certain chromosomal sites (for example, see refs. 11, 30, 31, 37), some of which may be DNase I hypersensitive loci associated with active replication or transcription, particular genes may be inaccessible for insertions in specific cell types, in certain stages of the cell cycle, or in cells grown under particular experimental conditions. It also remains to be determined whether Mo-MuLV*lac* insertions near chromosomally located promoters affect the activity or regulation of the promoters. The Mo-

MuLV*lac* vectors themselves should be useful in the investigation of these questions, as well as in the detection and study of genes affected by regulated splicing (38).

The choice of β-gal as the reporter gene product allows the formation of enzymatically active fusions with eukaryotic proteins (39), the detection of conditionally or temporally expressed genes (40), and the isolation of viable cells by fluorescence-activated cell sorting (23, 24). Because we anticipated that detection of insertions into functional genes would be optimized by a strategy that allows the expression of *lacZ* cassettes introduced into introns, we incorporated splice acceptor sites 5′ to *lacZ* in all three translational reading frames. Mo-MuLV self-inactivating virus derivatives, which lack transcriptional activity when they are in the proviral state, facilitated the detection of *lacZ* gene transcription originating at chromosomally located promoters.

As described, one of our splice acceptor site constructs (construct A) contained an ATG translation start codon in frame with the *lacZ* gene, which allows detection of *lacZ*-expressing cells following retrovirus insertions into introns immediately downstream from untranslated exons located near transcription start sites. The increased frequency of detection of X-Gal-staining cells observed for this Mo-MuLV*lac* construct may result from the propensity for retroviral insertions at sites near the beginning of genes (30, 31) and the existence of untranslated first exons in many genes (41). All three of the transcriptionally active genes reported in the present studies were identified by using this construct. While construct A also allowed *lacZ* gene fusion to translated protein-coding sequences, the cell-cycle-regulated genes detected in clones 54.6 and 54.7 and the constitutively expressed gene detected in clone 54.2 all appear to involve the initiation of β-gal-containing fusion proteins at an AUG codon in frame with *lacZ* in the Mo-MuLV*lac* cassette, which yields three independently generated proteins of the same size.

Recently, genes specifically expressed during growth arrest of mammalian cells have been described (36). As in the case of growth arrest genes whose regulatory signals have been detected by use of *lacZ*-based portable exons, the growth-arrest-specific (*gas*) genes cloned by differential hybridization of a genomic DNA library with $G_0$-specific cDNA probes are induced by both serum deprivation and density-dependent inhibition of cellular growth (36). It remains to be determined whether the genes induced in cells synchronized in $G_0$ and detected by use of *lacZ*-based portable exons are identical to, or functionally related to, the previously identified *gas* genes.

**Note Added in Proof.** Nonretroviral *lacZ* reporter gene constructs containing splice acceptor sites recently have been employed in mouse embryonic stem cells to detect developmentally regulated genes (42).

D.G.B. acknowledges the technical assistance of, and the open exchange of advice and information with, G. Nolan during the early phases of this project (43). We thank Mary Kovacs for the expert assistance she provided in operating the cell sorter and Rachel Hernandez for excellent general technical assistance. Our investigations were supported by National Institutes of Health Grant GM 27241 to S.N.C. D.G.B. was the recipient of a National Institutes of Health Postdoctoral Fellowship during certain of these studies.

1. Casadaban, M. J. & Cohen, S. N. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4530–4533.
2. Shapiro, J. A. (1983) *Mobile Genetic Elements* (Academic, New York).
3. Berg, C. M., Berg, D. E. & Groisman, E. A. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington, DC), pp. 879–925.
4. Mulligan, R. C. & Berg, P. (1980) *Science* **209**, 1422–1427.
5. Southern, P. J. & Berg, P. (1982) *J. Mol. Appl. Genet.* **1**, 327–341.
6. Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044–1051.
7. Hall, C. V., Jacob, P. E., Ringold, G. M. & Lee, F. (1983) *J. Mol. Appl. Genet.* **2**, 101–109.
8. Selden, R. F., Howie, K. B., Rowe, M. E., Goodman, H. M. & Moore, D. D. (1986) *Mol. Cell. Biol.* **6**, 3173–3179.
9. de Wet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R. & Subramani, S. (1987) *Mol. Cell. Biol.* **7**, 725–737.
10. Shapira, S. K. & Casadaban, M. J. (1987) *Gene* **52**, 83–94.
11. Barklis, E., Mulligan, R. C. & Jaenisch, R. (1986) *Cell* **47**, 391–399.
12. Hamada, H. (1986) *Mol. Cell. Biol.* **6**, 4179–4184.
13. O'Kane, C. J. & Gehring, W. J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 9123–9127.
14. Hiller, S., Hengstler, M., Kunze, M. & Knippers, R. (1988) *Mol. Cell. Biol.* **8**, 3298–3302.
15. Cepko, C. L., Roberts, B. E. & Mulligan, R. C. (1984) *Cell* **37**, 1053–1062.
16. Hawley, R. G., Covarrubias, L., Hawley, T. & Mintz, B. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2406–2410.
17. Graham, F. L. & Van der Eb, A. J. (1973) *Virology* **52**, 456–467.
18. Mann, R., Mulligan, R. C. & Baltimore, D. (1983) *Cell* **33**, 153–159.
19. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
20. Sanes, J. R., Rubenstein, J. L. R. & Nicolas, J.-F. (1986) *EMBO J.* **5**, 3133–3142.
21. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
22. Sompayrac, L. M. & Danna, K. J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7575–7578.
23. Jongkind, J. F., Verkerk, A. & Sernetz, M. (1986) *Cytometry* **7**, 463–466.
24. Nolan, G. P., Fiering, S., Nicolas, J.-F. & Herzenberg, L. A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2603–2607.
25. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
26. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
27. Casadaban, M. J., Chou, J. & Cohen, S. N. (1980) *J. Bacteriol.* **143**, 971–980.
28. Yu, S.-F., von Rüden, T., Kantoff, P. W., Garber, C., Seiberg, M., Rüther, U., Anderson, W. F., Wagner, E. F. & Gilboa, E. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3194–3198.
29. Yee, J.-K., Moores, J. C., Jolly, D. J., Wolff, J. A., Respess, J. G. & Friedmann, T. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5197–5201.
30. Vijaya, S., Steffen, D. L. & Robinson, H. L. (1986) *J. Virol.* **60**, 683–692.
31. Rohdewohld, H., Weiher, H., Reik, W., Jaenisch, R. & Breindl, M. (1987) *J. Virol.* **61**, 336–343.
32. Chu, G. & Sharp, P. A. (1981) *Nature (London)* **289**, 378–382.
33. Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L. & Weissman, S. M. (1978) *Science* **200**, 494–502.
34. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. & Ysebaert, M. (1978) *Nature (London)* **273**, 113–120.
35. Williams, J. G. & Penman, S. (1975) *Cell* **6**, 197–206.
36. Schneider, C., King, R. M. & Philipson, L. (1988) *Cell* **54**, 787–793.
37. Shih, C.-C., Stoye, J. P. & Coffin, J. M. (1988) *Cell* **53**, 531–537.
38. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. & Sharp, P. A. (1986) *Annu. Rev. Biochem.* **55**, 1119–1150.
39. Nielsen, D. A., Chou, J., MacKrell, A. J., Casadaban, M. J. & Steiner, D. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5198–5202.
40. Chou, J., Lemaux, P. G., Casadaban, M. J. & Cohen, S. N. (1979) *Nature (London)* **282**, 801–806.
41. Crick, F. (1979) *Science* **204**, 264–271.
42. Gossler, A., Joyner, A. L., Rossant, J. & Skarnes, W. C. (1989) *Science* **244**, 463–465.
43. Brenner, D. G. & Cohen, S. N. (1988) *J. Cell. Biochem. Suppl.* **12B**, 197.