Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 7317–7321, August 1993
Biochemistry

Human foamy virus genome possesses an internal, Bel-1-dependent and functional promoter

(intragenic cap site/spumaretrovirus/transcription/trans-activation)

MARTIN LÖCHELT, WALTER MURANYI, AND ROLF M. FLÜGEL*

Abteilung Retrovirale Genexpression, Forschungsschwerpunkt Angewandte Tumorvirologie, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 242, 69009 Heidelberg, Federal Republic of Germany

Communicated by Charles Weissmann, May 7, 1993 (received for review December 14, 1992)

ABSTRACT **The human foamy or spumaretrovirus (HSRV) is a complex retrovirus that encodes the three retroviral genes *gag*, *pol*, and *env* and, in addition, at least three *bel* genes. The HSRV Bel-1 protein was identified as a transcriptional trans-activator. HSRV transcription starts in the 5′ long terminal repeat at a defined guanine residue. We report here that a second efficiently utilized start site of transcription is contained within a HSRV *env* DNA sequence upstream of the *bel* genes. Bel-specific transcripts that initiate at the internal transcriptional start site at nt 9196 were identified in HSRV-infected cells by primer extension and S1 nuclease analysis, and the intragenic promoter was shown to be constitutively activated by Bel-1 in the HSRV provirus. In transient expression assays with indicator gene constructs, expression by the HSRV intragenic promoter/enhancer is Bel-1 dependent. The data provide evidence for an intragenic start site of transcription in the genome of a complex, exogenous human retrovirus and are discussed in terms of a model for regulating spumaretroviral gene expression.**

The human foamy or spumaretrovirus (HSRV) is an exogenous retrovirus that was isolated from the lymphoblastoid cells of a nasopharyngeal carcinoma patient (1). The HSRV genome encodes accessory *bel* genes located 3′ of *env* (2, 3). The pronounced complexity of the HSRV genome was revealed in the peculiar HSRV splicing pattern: about a dozen different HSRV transcripts were identified that included singly and multiply spliced mRNA species of the different *bel* genes (4). The *bel-1* gene product was identified as a transcriptional trans-activator for the long terminal repeat (LTR)-directed transcription (5–7) and shown to be absolutely required for viral replication and gene expression (8). The results so far indicated that HSRV gene expression is complex and comparable in structural and functional complexity to other human retroviruses (9). It is still unknown how expression of the various accessory transcripts and gene products in comparison with HSRV gag, pol, and env mRNAs and proteins is regulated. Despite repeated attempts to identify complex retroviral *rev* and *rex* analoga in HSRV, corresponding genes and/or RNA response elements have not been reported for any foamy virus (10).

In this study, defined indicator gene constructs and RNA from HSRV-infected cells were used to identify the internal start site of HSRV transcription located near the 3′ end of the *env* gene. The internal HSRV promoter depends on Bel-1 for efficient gene expression. These results point to, as far as we know, a unique mode of retroviral gene expression.

MATERIALS AND METHODS

Plasmids and Construction of Recombinant Clones. Infectious HSRV DNA clone pHSRV13 (8) served as source of viral DNAs; nucleotide numbering starts at the first base of the 5′ LTR. The eukaryotic expression clone pBCbell has been described (8). Indicator gene plasmids were constructed (see Fig. 1*B*) by inserting the blunt-end 283-bp *Nde* I–*Nhe* I DNA fragment, HSRV nt 8971–9253, into the *Eco*RV site of the promoterless pPLSEAP (5) upstream of the gene for the secreted form of the human alkaline phosphatase (SEAP) (11). Plasmids that contained the HSRV *Nde* I–*Nhe* I fragment in the original 5′ to 3′ sense- (+) or in the antisense-orientation (−) were designated pNNSEAP+ and pNNSEAP−, respectively. Plasmids pNNCAT+ in sense- and pNNCAT− in antisense-orientation that contain the HSRV *Nde* I–*Nhe* I DNA fragment upstream of the chloramphenicol acetyltransferase (CAT) gene were similarly constructed by using the unique *Bam*HI site in vector pBLCAT6 (12). Plasmids containing the HSRV U3 region 5′ of the SEAP gene (pdel 778/10, here designated pHSRV-U3-SEAP) or the HSRV 5′ LTR upstream of the CAT gene (pBC12/HSRV/CAT, here termed pHSRV-LTR-CAT) have been described (5).

Cell Culture, Transfection, and Expression Assays. The cultivation of COS-7, the cultivation of human embryonic lung (HEL) fibroblasts, and virus infections were done as described (8). Transfections by electroporation of 10 μg of DNA were done with equimolar amounts of indicator gene constructs and Bel-1 expression clones or the parental pBC12CMV vector. CAT assays were done as described (13) and normalized to coexpressed β-galactosidase activity. For quantitation and to ascertain that the assays were performed in the linear range, extracts were serially diluted, reaction products were separated by TLC, and the radioactive spots were determined with a TLC scanner. SEAP activity was quantified as described and is expressed in milliunits (5, 11).

RNA Extraction and Primer Extensions. Total RNA was harvested by cell lysis in guanidinium thiocyanate and sedimentation through cesium chloride (14). Primer extensions were done as described with 5′-^{32}P-labeled primer oligonucleotides (13). Primer SEAPan (5′-TTCTCCTCCTCAACTGGGATGATGC-3′) is complementary to the 5′ part of the SEAP gene. Primers 9307a (5′-CAATTCCTTGTAGAGCAGAAGCTGCTGC-3′) and 9382a (5′-GGATAGGCTTTAAGTATCCCAAGAGACT-3′) were derived from the antisense strand of the HSRV provirus 5′ of *bel-1* and designated according to the HSRV DNA sequence.

S1 Nuclease Analysis. S1 nuclease protection assays were done with 30 μg of total RNA as described (15, 16). The hybridization probe was prepared by PCR using the HSRV sense primer 8971s (5′-TATGTTCCTAGCATCGTGACTG-3′), ^{32}P-5′ end-labeled antisense primer 9307a, and pHSRV13 DNA. The probe was hybridized with total RNA at 46°C

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: CAT, chloramphenicol acetyltransferase; HEL, human embryonic lung fibroblasts; HSRV, human spumaretrovirus; LTR, long terminal repeat; SEAP, secreted form of human alkaline phosphatase.

*To whom reprint requests should be addressed.

overnight, and the protected DNA was separated on sequencing gels with dideoxynucleotide sequencing reactions run in parallel.

RESULTS

Detection of an Intragenic HSRV Bel-1-Dependent Promoter/Enhancer. To determine whether HSRV sequences upstream of *bel-1* carry a promoter/enhancer element, the 283-bp *Nde* I–*Nhe* I DNA fragment located ≈200 nt upstream of *bel-1* was inserted into the promoterless pPLSEAP vector in the sense (pNNSEAP+) and antisense orientation (pNNSEAP−) 5′ of the SEAP gene (Fig. 1*B*). The eukaryotic expression clone pBCbel1 that carries *bel-1* under the control of a cytomegalovirus promoter was used as source of Bel-1 transactivator activity in transient expression assays in COS-7 cells. After trans-complementation with pBCbel1, high SEAP activities were detectable in pNNSEAP+ cotransfections but were not detectable in those from pNNSEAP− (Table 1). In the absence of Bel-1, pNNSEAP+ in sense and pNNSEAP− in antisense orientation did not show any activity (Table 1). When analogous expression assays were carried out with CAT reporter gene constructs (Fig. 1*B*), the Bel-1-mediated trans-activation of the HSRV internal promoter of the pNNCAT+ plasmid was ≈200-fold (Table 1). Thus, SEAP and CAT activity was strongly expressed from those constructs that contain the HSRV internal promoter in the sense orientation and when trans-activated by Bel-1 (Table 1).

To compare the activity of the HSRV internal promoter with that of the homologous LTR promoter, plasmids pHSRV-U3-SEAP and pHSRV-LTR-CAT that contain either the U3 region or the 5′ LTR upstream of the indicator genes were analyzed in parallel. The Bel-1-trans-activated internal promoter had ≈17% (pNNSEAP+ versus pHSRV-U3-SEAP) and 80% (pNNCAT+ versus pHSRV-LTR-CAT) activity of that of the Bel-1-trans-activated LTR promoter (Table 1). In the absence of Bel-1, indicator gene expressions were not detectable. The transient expression assays were repeated three times and gave similar values. It follows that the level of activity of the HSRV internal promoter is

Table 1. Bel-1-mediated trans-activation of HSRV LTR- and internal promoter-dependent gene expression

	SEAP,* activity in milliunits		CAT,† activity in cpm	
	− Bel-1‡	+ Bel-1§	− Bel-1‡	+ Bel-1§
pNNSEAP+	<0.5	186		
pNNSEAP−	<0.5	<0.5		
pHSRV-U3-SEAP	<0.5	1084		
pNNCAT+			125	25,700
pNNCAT−			20	190
pHSRV-LTR-CAT			50	32,300

*SEAP activity is expressed in milliunits of SEAP activity (11), where 1.0 milliunit of SEAP hydrolyzes 1.0 pmol of *p*-nitrophenylphosphate per min, which equals an increase of 0.04 *A* units at 405 nm.

†CAT activity is expressed in cpm of monoacetylated chloramphenicol after TLC. Cpm were normalized to undiluted cell extracts used for the CAT assays.

‡Data are from cotransfections with the vector backbone pBC12CMV without *bel-1* insert.

§Data are from cotransfections with the Bel-1 expression vector pBCbel1.

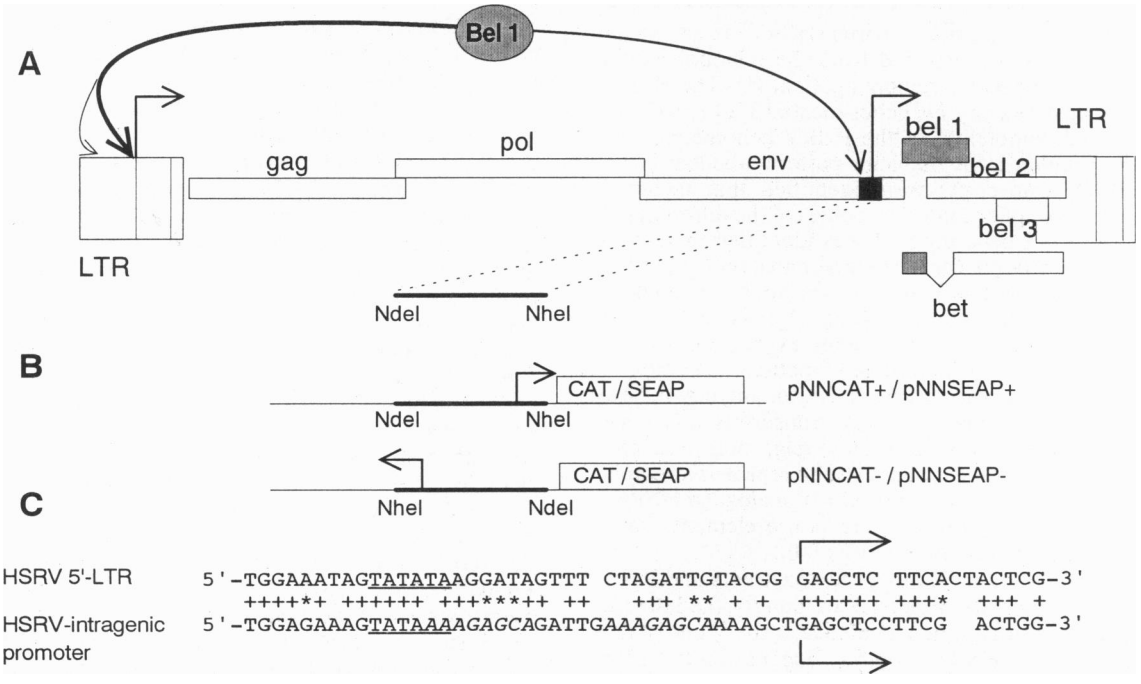

FIG. 1. Genomic organization of HSRV DNA. (*A*) Position of the *Nde* I–*Nhe* I DNA fragment (nt 8971–9253) with the internal promoter/enhancer in *env* is indicated by the black box inside the *env* gene. The start sites and direction of transcription in the 5′ LTR and the internal promoter in *env* upstream of *bel-1* are indicated by arrows above the LTR and *env*, respectively. The *bel-1* gene is shown by a shaded box, and the Bel-1 trans-activator is shown by a shaded ellipse with long arrows, indicating the different Bel-1 DNA target sites identified in the HSRV LTR (4, 5) and the intragenic promoter/enhancer. (*B*) Schematic drawing of CAT and SEAP reporter gene constructs used. The HSRV *Nde* I–*Nhe* I DNA fragment that contains the internal promoter is shown by a thick line. Rectangular arrows indicate direction of transcription; plus and minus signs mark sense and reverse orientations. (*C*) Sequence comparison between the HSRV LTR and intragenic promoters. The HSRV *env* DNA from nt 9156 to 9211 is aligned with the HSRV LTR from nt 740 to 794. A rectangular arrow indicates the start site and direction of transcription; TATA boxes are underlined. A direct octamer repeat (5′-AAAGAGCA-3′) between the TATA box and the cap site of the intragenic promoter is in italics. To maximize the alignment, gaps were introduced; plus signs indicate identical nucleotides, and asterisks mark either purine or pyrimidine nucleotides.

comparable with that of the HSRV LTR or U3 promoter, depending on the gene constructs compared.

Identification of the Internal Transcription Start Site in the HSRV Insert of pNNSEAP+. To determine the SEAP mRNA start site in pNNSEAP+, primer extensions with 5′-labeled SEAPan primer were carried out. RNA was isolated from COS-7 cells transfected with pNNSEAP+ in the presence (Fig. 2, lanes 1 and 2) or absence (lane 3) of a Bel-1 expression plasmid. The extension products were loaded on a sequencing gel next to the reaction products of a dideoxynucleotide sequencing of pNNSEAP+ DNA also primed with SEAPan (lanes G–C). Only RNA from cells cotransfected with *bel-1* directed the synthesis of a specific extension product of discrete size (arrow). The size of the cDNA when aligned to the sequencing ladder is equivalent to the 3′-terminal cytosine in the antisense sequence 5′-AAGGAGCTC-3′ (marked by star in Fig. 2). It follows that the 5′ terminus of the template SEAP mRNA is the guanine at nt position 9196—i.e., the SEAP mRNA started at this guanine (marked as cap site in Fig. 1*C*). This result was confirmed by sequencing the product of primer extensions and S1 nuclease analysis, demonstrating that the cap site of the SEAP mRNA is at HSRV nt 9196 (data not shown).

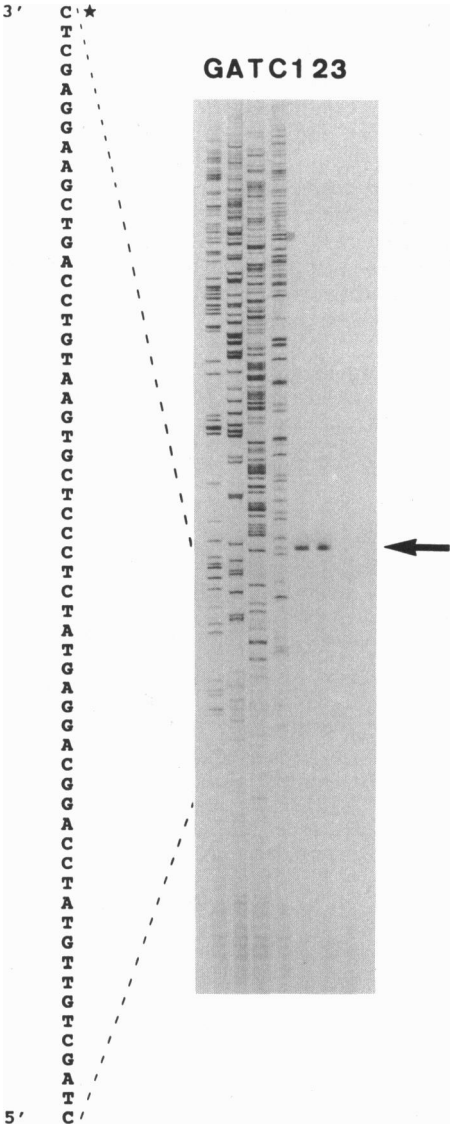

FIG. 2. Determination of the 5′-end of the SEAP mRNA transcribed from the internal HSRV promoter in plasmid pNNSEAP+ by primer extensions. Primer extensions were performed with RNA from transfections with pNNSEAP+ alone (lane 3) or trans-complemented with pBCbell (lane 2) or a pBCbell derivative truncated directly 3′ of *bel-1* (lane 1). The specific extension product (arrow) directed by RNA derived from *bel-1*-cotransfected cells (lanes 1 and 2) was analyzed in parallel with a sequencing reaction of pNNSEAP+ with primer SEAPan (left four lanes marked G, A, T, and C indicate the ddNTP used). Part of the sequence of the antisense strand is given, ending at the cytosine residue (star) that corresponds to the length of the extension product.

Determination of the Intragenic Transcriptional Start Site in HSRV-Infected Cells. To demonstrate that the internal start site of transcription is active in the HSRV provirus, total RNA from HSRV-infected (Fig. 3*A*, lanes 1 and 3) and mock-infected HEL cells (lanes 2 and 4) was subjected to primer extensions. Primers were used that are located between the internal cap site and the start of the *bel-1* coding region. As expected, RNA of HSRV-infected cells and primer 9307a (lanes 1 and 2) directed the synthesis of cDNA of ≈112 nt (solid arrow, lane 1), whereas with primer 9382a (lanes 3 and 4) a major band of 187 nt (open arrow, lane 3) was obtained in addition to bands of lower molecular size. This result confirms the utilization of nt 9196 as internal start site of transcription. Subsequently, S1 nuclease analysis was done according to Weaver and Weissmann (15). RNA from HSRV-infected (Fig. 3*B*, lane 1) or mock-infected HEL cells (lane 2) was hybridized to a HSRV DNA probe from nt 8971 to 9307 that encompasses the internal promoter region. Only

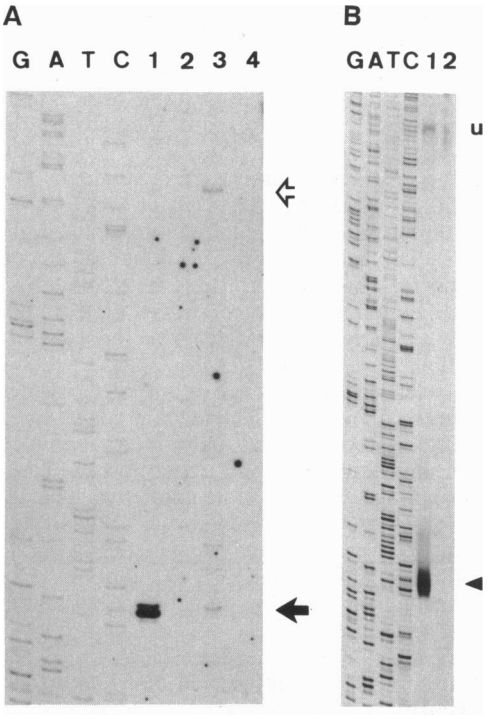

FIG. 3. Identification of the intragenic start site of transcription in HSRV-infected HEL cells by primer extensions and S1 nuclease analysis. (*A*) Primer extension analyses were done with total RNA from HEL cells harvested 72 hr after HSRV− (lanes 1 and 3) or mock-infection (lanes 2 and 4). Primers 9307a (lanes 1 and 2) and 9382a (lanes 3 and 4) are located upstream of *bel-1*. The length of the extension products was found to be 112 nt (primer 9307a, solid arrow) and 187 nt (primer 9382a, open arrow) by using a parallel-run sequencing reaction of pHSRV13 DNA that was started with primer 9307a (lanes G to C). (*B*) S1 nuclease analysis was done with RNA from HSRV− (lane 1) and mock-infected HEL cells (lane 2). RNA from HSRV-infected cells protected DNA of ≈112 nt (arrowhead). u, Undigested hybridization probe (nt 8971–9307). Lanes G to C represent the dideoxynucleotide sequencing ladder of HSRV DNA extended with primer 9307a and run in parallel for length determination.

RNA from HSRV-infected cells protected DNA of 112 ± 2 nt (arrowhead in Fig. 3*B*), which corresponds in length to a protecting mRNA that starts at HSRV position 9196. The results obtained by both independent methods with RNA from virus-infected cells clearly prove that the HSRV intragenic start site of transcription at nt 9196 is efficiently used in HSRV-infected cells.

A typical TATA box (Fig. 1*C*, underlined) is located 26 nt 5′ of the internal transcription start site (rectangular arrow). Sequences at the intragenic promoter display a marked homology to corresponding sequences in the HSRV 5′ LTR (Fig. 1*C*) to sequences upstream of *bel-1* in the genomes of simian foamy viruses types 1 and 3 (17–19) and a similarity to the LTR promoter of the neurovirulent murine leukemia virus, strain Cas-BrE 15-1 (20). This similarity opens the possibility that determinants responsible for the brain-specific lesions observed in HSRV-transgenic mice share common features (21).

DISCUSSION

Primer extensions and S1 nuclease protection assays were used to identify and precisely map the distinctive HSRV internal cap site. RNA sequencing (data not shown) confirmed that the internal start site of transcription is a guanine residue at genomic position 9196 located ≈250 nt upstream of the *bel-1* initiator codon in the HSRV *env* gene (Fig. 1*A*). Transcripts that start at the internal cap site were unambiguously identified in HSRV-infected cells, confirming that this promoter is functionally active in the provirus. It is remarkable that HSRV has the capacity to transcribe mRNAs from an internal cap site because so far it has been reported that genes of complex retroviruses are exclusively expressed from transcripts that start at the LTR cap site. Recently, evidence for an internal activation-dependent, T-lymphocyte-specific promoter was reported in a defective mouse mammary tumor provirus integrated into a mouse thymoma cell line (22). The deletion in the U3 region of this defective provirus shortens the reading frame that codes for the recently identified eukaryotic superantigen (23). An intragenic enhancer, but no internal promoter, was reported for human immunodeficiency virus type 1 (24).

Internal promoters have been reported for widely used retroviral vector constructs. The level of gene expression in retroviral vectors constructed to carry internal promoters can vary and depends on the specific constructs used and on the selection conditions if a selectable marker was used (25, 26). The detection of internally initiated transcripts of the HSRV provirus and comparative analyses of the 5′ LTR to the internal promoter (M.L. and R.M.F., unpublished results) revealed that both promoters are active in the context of a single DNA molecule.

The precise location of the Bel-1 DNA target site(s) for trans-activation within the internal HSRV *Nde* I–*Nhe* I DNA fragment remains to be determined. Sequence comparisons of the internal promoter/enhancer with the known Bel-1 response elements did not show any strong similarity; it seems that different Bel-1 DNA target sites can be used in the U3 region of the HSRV 5′ LTR (5, 6), in the heterologous human immunodeficiency virus LTR (27, 28), and in the internal HSRV promoter/enhancer. Due to the position directly upstream of the *bel* genes and the fact that additional methionine codons are not located in any one of the reading frames between the new cap site and *bel-1*, the internal promoter should direct the expression of *bel-1*, *bet* (Fig. 1*A*), and possibly other *bel* genes (4, 8). Primer extensions started in *bel-1* showed the presence of *bel-1*/*bet*-specific transcripts (unpublished results). The main structural distinction of the internally initiated bel transcripts in comparison with the 5′ LTR-directed bel mRNAs is the presence of a completely different leader mRNA sequence that may influence splicing and/or translation (4).

The presence of an active internal HSRV promoter indicates a distinctive mode of retroviral gene expression for at least the spumavirus group of complex retroviruses. Two active and Bel-1-dependent HSRV promoters open the way for a bimodal way of transactivation of foamy virus gene expression. We postulate that the internal HSRV transcription unit serves to establish and control a balance between the expression of viral structural and regulatory proteins (Fig. 1*A*). The activity of the HSRV internal promoter would replace *rev* and *rex* analogous functions that have not been found to date in foamy viruses (10). The proposed mechanism of regulating gene expression would yield the same net result of balancing the amounts of structural and regulatory gene products as in other complex retroviruses. One consequence of the model is that the Bel-1-dependence of the LTR promoter and of the internal promoter is differentially regulated by unknown positively and/or negatively acting factors. These factors may influence the temporal control of transcriptional activity and, in addition, could determine the range of target cells susceptible to HSRV infection.

Recently, it has been established that internal promoters in retroviral vectors are essential for the long-term expression of genes in primary myoblasts as a somatic tissue for gene therapy (29). Foamy viruses with an exceptionally broad host range of cells and two functional promoters may be particularly suited for such purposes.

We thank Howard Temin for critically reading the manuscript; Angel Alonso, Hajo Delius, and Michael Pawlita for discussions; and Harald zur Hausen for support. This work was supported by grants from the Bundesministerium für Forschung und Technologie (01KI 8901/6) and the Commission of the European Community (TS2-CT187-0186-D).

1. Achong, B. G., Mansell, P. W. A., Epstein, M. A. & Clifford, P. (1971) *J. Natl. Cancer Inst.* **46,** 299–307.
2. Flügel, R. M., Rethwilm, A., Maurer, B. & Darai, G. (1987) *EMBO J.* **6,** 2077–2084.
3. Maurer, B., Bannert, H., Darai, G. & Flügel, R. M. (1988) *J. Virol.* **62,** 1590–1597.
4. Muranyi, W. & Flügel, R. M. (1991) *J. Virol.* **65,** 727–735.
5. Keller, A., Partin, K. M., Löchelt, M., Bannert, H., Flügel, R. M. & Cullen, B. R. (1991) *J. Virol.* **65,** 2589–2594.
6. Venkatesh, L. K., Theodorakis, P. A. & Chinnadurai, G. (1991) *Nucleic Acids Res.* **19,** 3661–3666.
7. Rethwilm, A., Erlwein, O., Baunach, G., Maurer, B. & ter Meulen, V. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 941–945.
8. Löchelt, M., Zentgraf, H. & Flügel, R. M. (1991) *Virology* **184,** 43–54.
9. Flügel, R. M. (1991) *J. Acquired Immune Defic. Syndr.* **4,** 739–750.
10. Flügel, R. M. (1992) in *Human Retroviruses: Frontiers in Molecular Biology*, ed. Cullen, B. R. (Oxford Univ. Press, Oxford), pp. 193–214.
11. Berger, J., Hauber, J., Hauber, R., Geiger, R. & Cullen, B. R. (1988) *Gene* **66,** 1–10.
12. Boshart, M., Klüppel, M., Schmidt, A., Schütz, G. & Luckow, B. (1992) *Gene* **110,** 129–130.
13. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed., pp. 7.79–16.59.
14. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
15. Weaver, R. F. & Weissmann, C. (1979) *Nucleic Acids Res.* **7,** 1175–1193.
16. Davis, L. G., Dibner, M. D. & Battey, J. F. (1986) *Basic Methods in Molecular Biology* (Elsevier, New York), pp. 276–284.
17. Mergia, A., Shaw, K. E. S., Lackner, J. & Luciw, P. (1990) *J. Virol.* **64,** 406–410.
18. Kupiec, J.-J., Kay, A., Hayat, M., Ravier, R., Peries, J. & Galibert, F. (1991) *Gene* **101,** 185–194.

19. Renne, R., Friedl, E., Schweizer, M., Fleps, U., Turek, R. & Neumann-Haefelin, D. (1992) *Virology* **186,** 597–608.
20. Portis, J. L., Perryman, S. & McAtee, F. J. (1991) *J. Virol.* **65,** 1877–1883.
21. Bothe, K., Aguzzi, A., Lassmann, H., Rethwilm, A. & Horak, I. (1991) *Science* **253,** 555–557.
22. Miller, C. L., Garner, R. & Paetkau, V. (1992) *Mol. Cell. Biol.* **12,** 3262–3272.
23. Acha-Orbea, H., Shakov, A. N., Scarpellino, L., Kolb, E., Müller, V., Vessaz-Shaw, A., Fuchs, R., Blöchlinger, K., Rollini, P., Billotte, J., Sarafidou, M., MacDonald, H. R. & Diggelmann, H. (1991) *Nature (London)* **350,** 207–211.
24. Verdin, E., Becker, N., Bex, F., Droogmans, L. & Burny, A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4874–4878.
25. Emerman, M. & Temin, H. M. (1986) *Nucleic Acids Res.* **14,** 9381–9396.
26. Miller, A. D. (1992) *Curr. Top. Microbiol. Immunol.* **158,** 1–24.
27. Keller, A., Garrett, E. D. & Cullen, B. R. (1992) *J. Virol.* **66,** 3946–3949.
28. Lee, A. H., Lee, K. J., Kim, S. & Sung, Y. C. (1992) *J. Virol.* **66,** 3236–3240.
29. Dai, Y., Roman, M., Naviaux, R. K. & Verma, I. M. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10892–10895.