Instability of Large Direct Repeats in Retrovirus Vectors

BRETT W. RHODE, MICHAEL EMERMAN,[†] AND HOWARD M. TEMIN*

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

Received 18 August 1986/Accepted 21 November 1986

Retrovirus vectors were constructed with large (0.85- to 1.3-kilobase-pair) direct repeats in their genomes. Deletions involving the direct repeats occurred at a high frequency. Deletions occurred both when the direct repeats were in tandem and when they were separated by additional sequences. These deletions occurred during virus replication.

There is evidence that some deletions in retroviral genomes occur between direct repeats (10, 14, 17), while others occur without apparent specificity (4, 19, 20). Because we have been studying parameters that govern the stability and expression of genes in retrovirus vectors (2, 5-8a, 16), we wished to determine the stability of retrovirus vectors that contain large direct repeats. We found that we could not maintain large direct repeats, either in tandem or separated by 1.0 kilobase pair (kbp), in a retrovirus vector constructed from the genome of spleen necrosis virus.

ME111 is a spleen necrosis virus-based vector (6, 7) that contains one copy of the *neo* gene, one copy of the *tk* gene, and the *cis*-acting regions needed for virus replication (Fig. 1A and C). pME112 is the same as pME111, except that pME112 contains three copies of the *neo* gene in tandem direct repeats 5' to the *tk* gene. pME118 contains three copies of the *neo* gene in tandem direct repeats 3' to the *tk* gene, while pME116 contains only one copy of the *neo* gene (Fig. 1A and C).

To analyze the structure of the viral genomes produced from plasmids containing proviral DNA, chicken embryo fibroblasts (CEF) were transfected with 5 µg of each plasmid DNA and 0.1 µg of helper virus (Rev-A) DNA by the dimethyl sulfoxide-polybrene method (12), and virus was harvested 5 days posttransfection. The structures of the resultant viral genomes were analyzed by Southern transfer and hybridization of the Hirt supernatant fraction of DNA (9) isolated from cells infected with the harvested viruses. We found that the linear unintegrated ME112 viral DNA was the same size as that produced by ME111 (Fig. 1B). Similarly, the linear unintegrated DNA produced by ME118 was the same size as that produced by ME116 (Fig. 1B). Restriction enzyme mapping of the unintegrated viral DNAs revealed that viruses recovered from pME112 and pME118 contained only one copy of the *neo* gene (data not shown). This result indicated that large tandem direct repeats are not stable in these retrovirus vectors.

To analyze the effect on the stability of the viral genome of large direct repeats separated by additional sequences, we constructed pME126 (Fig. 2B). pME126 contains two copies of the 0.85-kbp promoter from the herpes simplex virus type 1 *tk* gene (*tk pro*) separated by the *neo* gene. Virus recovered by transfection of CEF with pME126 and helper virus DNA could transform TK⁻ cells to a TK⁺ phenotype but could not

transform cells to a Neo^R phenotype (data not shown), suggesting that the *neo* gene between the direct repeats had been deleted. This hypothesis was tested directly by analysis of unintegrated ME126 viral DNA. Full-length unintegrated viral DNA produced by ME126 virus should be 6.8 kbp. However, we found that the viral genome had undergone a deletion and was 5.0 kbp (Fig. 2A, lanes 1 and 4). Digestion of the unintegrated viral DNA with *Bgl*II revealed that the internal 1.8-kbp *Bgl*II fragment was deleted (Fig. 2A, compare lane 2 with lane 3 for pME126). This result, and other restriction endonuclease mapping of the unintegrated viral DNA, showed that the major deletion occurred between the direct repeats in ME126, thus deleting the *neo* gene.

Close examination with a *neo*-specific probe of the virus DNA produced after transfection with pME126 and helper virus DNA showed that no full-length (6.8-kbp) ME126 virus could be detected, although viruses still containing at least part of the *neo* gene were present (Fig. 2A, lane 5). This result indicated that deletions and other rearrangements that did not involve the direct repeat occurred in the virus stock, although the most common (or fastest-replicating) virus had a deletion involving the large direct repeats.

Stocks of retrovirus vectors are usually produced by transfection of recombinant plasmids into the appropriate cells. Because transfections are very mutagenic (3, 5, 15), we wished to determine whether the deletions involving the direct repeats occurred during virus replication or only during transfection. Our experimental strategy was to transfect cells with pME112, to select a cell clone in which the integrated plasmid DNA still contained tandem copies of the *neo* gene, and then to superinfect the cell clone with helper virus. The DNA from the transfected cell clone and the DNA produced by virus released from this cell clone after superinfection could then be analyzed separately.

D17 dog cells were transfected with pME112 (in the absence of helper virus DNA) and selected for Neo^R, and a cell clone was picked and analyzed. pME111 cut with *Bgl*II gives a *neo*-specific band of 1.85 kbp, while pME112 cut with *Bgl*II gives *neo*-specific bands of 1.85 and 1.3 kbp (Fig. 3). The genomic DNA of the cell clones transfected with pME112 contained more than one contiguous copy of the *neo* gene, indicated by the presence of both the 1.85- and the 1.3-kbp *Bgl*II bands (Fig. 3A, lane 5). *Bam*HI digestion of pME111 and pME112 gave *neo*-specific bands of 3.9 and 6.5 kbp, respectively (Fig. 3A, lanes 2 and 4). The *Bam*HI digestion of DNA from the Neo^R D17 cell clone analyzed (Fig. 3A, lane 6) is consistent with the integration of three copies of plasmid DNA. Of these three plasmid integrants in

* Corresponding author.

[†] Present address: Institut Pasteur, Unité d'Oncologie Virale, 75724 Paris, France.

FIG. 1. Analysis of retrovirus vectors containing tandem direct repeats. (A) Restriction enzyme analysis of pME111, pME112, pME116, and pME118. The presence of p before the DNA type indicates plasmid DNA. The absence of p indicates unintegrated viral DNA. To separate proviruses from the pBR322 vectors, 1 μ g of each plasmid was digested with *Sal*I and was run on a 1.0% agarose gel. The common band in each lane represents the pBR322 sequences of all plasmids, while the larger band in each lane represents the size of the viral DNA. (B) Southern blot of ME111, ME112, ME116, and ME118 linear unintegrated viral DNAs. Unintegrated viral DNA was produced by cotransfecting CEF with 5 μ g of plasmid DNA and 0.1 μ g of helper virus (Rev-A) DNA. At 5 days after transfection, virus was recovered and used to infect CEF. At 3 days postinfection, Hirt supernatant DNA was extracted, separated on a 1.0% agarose gel, transferred to nitrocellulose, and hybridized with a *neo*-specific probe. Longer exposures did not reveal additional bands. The titers of ME111 and ME112 were both over 10^5 NEO TU/ml. (C) DNA genomes of the cloned form of the retroviruses analyzed (pBR322 sequences not shown). Symbols: $\square$, spleen necrosis virus long terminal repeats; —, spleen necrosis virus sequences; ||||| , ||||| , as labeled in the figure. pME112 is identical to pME111, except that pME111 contains one copy of the *neo* gene, while pME112 contains three copies of the *neo* gene. Similarly, pME118 is identical to pME116, except that pME116 contains one copy of the *neo* gene, while pME118 contains three copies of the *neo* gene.

the Neo^R D17 cell clone (Fig. 3A, lane 6), one contained at least five copies of the *neo* gene (the >9-kbp band), one contained two copies of the *neo* gene (the 5.2-kbp band), and one contained less than one copy of the *neo* gene (the 3.5-kbp band). These non-genome-sized integrants must have been generated during the transfection.

We then superinfected the D17 cell clone with Rev-A, collected virus 5 days postinfection, and passed the resultant virus stock on CEF or D17 cells. The Hirt supernatant fraction of DNA was extracted 3 days later, and the structure of the viral genome was analyzed and compared with the structure of the genomic DNA in the original transfected D17 cell clone. In contrast to the plasmid integrants in the genomic DNA, the resultant virus obtained from this cell

clone after superinfection with Rev-A contained exactly one copy of the *neo* gene. That is, the structure was identical to that of ME111 (Fig. 3A, lanes 7 to 9). This result indicated that the deletions occur rapidly during virus replication.

In our experiments, duplication of the large direct repeats in virus replication would not have been observed, because these duplications make the viral genome too large to be packaged efficiently (8a). However, both deletions and duplications involving small direct repeats are major sources of variation of replication-competent retroviruses. Many isolates of the same virus have deletions or insertion of a single copy of small tandem direct repeats in the U3 region of the long terminal repeat of retroviruses (17, 18), and in structural genes (1).

On the other hand, analysis of two deletions in a defective variant that became prominent in a stock of Rous sarcoma virus that had been extensively passaged showed that neither deletion took place at a direct repeat, although a direct repeat was near one of the deletions (19). We also detected deletions at a low level that did not involve the direct repeats (Fig. 2A, lane 5). We do not know whether, after extended passage, one of these deletions that did not occur at the direct repeats would overgrow the rest of the population.

FIG. 2. Analysis of retrovirus vectors containing nontandem direct repeats. (A) Southern blot of linear unintegrated viral DNA extracted from ME126-infected CEF. The Hirt supernatant DNA was prepared in the same manner as described in the legend to Fig. 1. Lanes: 1, ME126 uncut; 2, ME126 digested with *Bgl*II; 3, pME126 digested with *Bgl*II; 4 and 5, ME126 uncut. Lanes 1 to 3 were hybridized with probes specific for the *tk* promoter and gene. Lane 4 was hybridized with a probe specific for the *tk* promoter. Lane 5 was hybridized with a *neo*-specific probe. The 1-kbp ladder from Bethesda Research Laboratories, Gaithersburg, Md., was used as a marker. (B) DNA genome of the cloned form of ME126 (pBR322 sequences not shown). Symbols are the same as in Fig. 1C. ME126 contains two copies of the *tk* promoter sequence separated by the 1.0-kbp *neo* gene. The *tk* gene is present at the 3' end of the virus.

FIG. 3. Analysis of a cell clone transfected with pME112. (A) Genomic DNA was obtained from a D17 cell clone that had been transfected with pME112 and selected for Neo^R. The same Neo^R D17 clone used for genomic DNA isolation was superinfected with Rev-A, and the resultant virus was collected 5 days later. Unintegrated viral DNA was isolated from a Hirt supernatant fraction 3 days after infection of CEF with the recovered virus. Lanes: 1 and 2, pME111 digested with *Bgl*II and *Bam*HI, respectively; 3 and 4, pME112 digested with *Bgl*II and *Bam*HI, respectively; 5 and 6, genomic DNA of a ME112-transfected Neo^R clone digested with *Bgl*II and *Bam*HI, respectively; 7 to 9, linear unintegrated ME112 viral DNA digested with *Bgl*II, digested with *Bam*HI, and uncut, respectively. The blot was hybridized with the *neo*-specific probe. There are large nonspecific spots on the film between lanes 3 and 4 at about 4 kbp and between lanes 6 and 7 at about 9 kbp. (B) DNA genome of the cloned form of ME112 (pBR322 sequences not depicted), with the three copies of the *neo* gene present and the region used as a hybridization probe. Symbols are the same as in Fig. 1C; —, probe.

In summary, we were unable to recover full-length virus from plasmids that contained large direct repeats. We have shown that the deletions can occur during virus replication. In addition to large direct repeats, other sources of instability in retrovirus vectors are genomes too large to be packaged efficiently (8a), the presence of internal 3' polyadenylation sites (11, 16), selection for expression of some internal promoters (6), and intron sequences that affect splicing (5, 13).

We thank J. Couch, S. Hellenbrand, J. Rein, and A. Troup for technical assistance and J. Dougherty, J. Embretson, C. Miller, and W. Sugden for helpful comments on the manuscript.

This work was supported by Public Health Service grants CA-2243 and CA-07215 from the National Institutes of Health. M.E. was supported by Public Health Service training grant CA-09135-10. H.M.T. is an American Cancer Society research professor.

LITERATURE CITED

1. Alizon, M., S. Wain-Hobson, L. Montagnier, and P. Songio. 1986. Genetic variability in the AIDS virus: nucleotide sequence analysis of two isolates from African patients. *Cell* 46:63-74.
2. Bandyopadhyay, P. K., and H. M. Temin. 1984. Expression of the complete chicken thymidine kinase gene inserted into a retrovirus vector. *Mol. Cell. Biol.* 4:749-754.
3. Calos, M., J. Lebokowski, and M. Botchan. 1983. High mutation frequency in DNA transfected into mammalian cells. *Proc. Natl. Acad. Sci. USA* 80:3015-3019.
4. Darlix, J.-L., and P.-F. Spahr. 1984. High spontaneous mutation rate of Rous sarcoma virus demonstrated by direct sequencing of the RNA genome. *Nucleic Acids Res.* 11:5953-5968.
5. Dougherty, J. P., and H. M. Temin. 1986. High mutation rate of a spleen necrosis virus-based retrovirus vector. *Mol. Cell. Biol.* 6:4387-4395.
6. Emerman, M., and H. M. Temin. 1984. High-frequency deletions in retrovirus vectors containing exogenous DNA with promoters. *J. Virol.* 50:42-49.
7. Emerman, M., and H. M. Temin. 1984. Genes with promoters in retrovirus vectors can be independently suppressed by an epigenetic mechanism. *Cell* 39:459-467.
8. Emerman, M., and H. M. Temin. 1986. Quantitative analysis of gene suppression in integrated retrovirus vectors. *Mol. Cell. Biol.* 6:792-800.
- 8a. Gelinis, C., and H. M. Temin. 1986. Nondefective spleen necrosis virus-derived vectors define the upper size limit for packaging reticuloendotheliosis viruses. *Proc. Natl. Acad. Sci. USA* 83:9211-9215.
9. Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* 26:265-269.
10. Hughes, S., and E. Kosik. 1984. Mutagenesis of the region between *env* and *src* of the SR-A strain of Rous sarcoma virus for the purpose of constructing helper-independent vectors. *Virology* 136:89-99.
11. Joyner, A. L., and A. Bernstein. 1983. Retrovirus transduction: generation of infectious retroviruses expressing dominant and selectable genes is associated with in vivo recombination and deletion events. *Mol. Cell. Biol.* 3:2180-2190.
12. Kawai, S., and M. Nishizawa. 1984. A new procedure for DNA transfection with polycation and dimethyl sulfoxide. *Mol. Cell. Biol.* 4:1172-1174.
13. Miller, C. K., and H. M. Temin. 1986. Insertion of several different DNAs in reticuloendotheliosis virus strain T suppresses transformation by reducing the amount of subgenomic mRNA. *J. Virol.* 58:75-80.
14. Omer, C. A., K. Pogue-Geile, R. Guntaka, K. A. Staskus, and A. J. Faras. 1983. Involvement of directly repeated sequences in the generation of deletions of the avian sarcoma virus *src* gene. *J. Virol.* 47:380-382.
15. Razzaque, A., H. Mizusawa, and M. Seidman. 1983. Rearrangement and mutagenesis of a shuttle vector plasmid after passage in mammalian cells. *Proc. Natl. Acad. Sci. USA* 80:3010-3014.
16. Shimotohno, K., and H. M. Temin. 1981. Formation of infectious progeny virus after insertion of herpes simplex thymidine kinase gene into DNA of an avian retrovirus. *Cell* 26:67-77.
17. Shimotohno, K., and H. M. Temin. 1982. Spontaneous variation and synthesis in the U3 region of the long terminal repeat of an avian retrovirus. *J. Virol.* 41:163-171.
18. Van Beveren, C., E. Rands, S. K. Chattopadhyay, D. R. Lowy, and I. M. Verma. 1982. Long terminal repeat of murine retroviral DNAs: sequence analysis, host-proviral junctions, and preintegration site. *J. Virol.* 41:542-556.
19. Voynow, S. L., and J. M. Coffin. 1985. Evolutionary variants of Rous sarcoma virus: large deletion mutants do not result from homologous recombination. *J. Virol.* 55:67-78.
20. Wang, L.-H., B. Edelstein, and B. J. Mayer. 1984. Induction of tumors and generation of recovered sarcoma viruses by, and mapping in, two molecularly cloned *src* deletion mutants. *J. Virol.* 50:904-913.