JOURNAL OF VIROLOGY, Mar. 1985, p. 998–1000
0022-538X/85/030998-03$02.00/0

Vol. 53, No. 3

Characterization of an Endogenous Retrovirus-Repetitive DNA Chimera in the Mouse Genome

THOMAS G. FANNING,[1]†* DAVID W. MORRIS,[1] ROBERT D. CARDIFF,[1] AND HARVEY D. BRADSHAW, JR.[2]‡

Department of Pathology, School of Medicine, University of California, Davis, California 95616,[1] and Department of Biochemistry, Louisiana State University Medical Center, New Orleans, Louisiana 70112[2]

Received 17 September 1984/Accepted 6 November 1984

We found that an endogenous mouse mammary tumor provirus, GR-MTV-8, is embedded within a member of the BAM HI family of long interspersed repetitive mouse DNAs. GR-MTV-8 appears to be transcriptionally silent at its normal chromosomal position in the mouse genome. The provirus is transcriptionally active, however, when cloned and transfected into mouse cells (Kennedy et al., Nature (London) 295:622–624, 1982). We propose that the transcriptional inactivity of GR-MTV-8 in situ is due to an inhibitory effect, possibly involving DNA methylation, attributable to the flanking BAM HI element.

All inbred laboratory mice and many feral mice contain genetically inherited proviral copies of mouse mammary tumor virus (MuMTV) (6). In several inbred mouse strains (e.g., BALB/c and C57BL), these endogenous MuMTV proviruses are inactive (nontumorigenic), and female mice from these strains rarely develop mammary tumors (10). Other inbred strains (e.g., GR/A and C3H) harbor active (tumorigenic) MuMTV proviruses in their genomes, and female mice of these strains often develop mammary tumors (10).

Mammary tumors from high mammary tumor-incidence mouse strains exhibit increased numbers of MuMTV proviruses (5, 25), which are invariably hypomethylated (4, 13). Although the exact mode of MuMTV proviral amplification is at present unknown, a plausible mechanism is the transcription of an endogenous provirus followed by reverse transcription and integration of the newly synthesized viral DNA (34). Occasionally a virus is integrated near a specific locus which by an as-yet-unknown mechanism appears to activate a cellular oncogene (27, 28).

Of the many endogenous MuMTV proviruses that have been characterized (6, 11, 24, 33), only a few appear to be tumorigenic (11, 24). The benign character of many proviruses may be the result of their transcriptional inactivity, and indeed, in low mammary tumor-incidence mouse strains, little or no MuMTV-specific RNA can be detected in the mammary gland cells of female mice (35). In addition, high mammary tumor-incidence mouse strains appear to contain MuMTV RNA transcribed from the tumorigenic proviruses, but not from other MuMTV proviruses present in the same cell (12, 16).

How might the transcriptional silencing of MuMTV proviruses be achieved? The most obvious mechanism is mutation of the transcriptional regulatory signals. A second possibility is that host sequences flanking the provirus exert an inhibitory effect on proviral transcription (3, 7). Inhibitory effects might result from the influence of the flanking DNA on the methylation or conformation (or both) of the provirus (3, 20).

In the present report we demonstrate that one of the MuMTV proviruses present in the GR/A mouse strain, GR-MTV-8, is embedded within a member of the BAM HI family of interspersed repetitive mouse sequences (the major family of interspersed repetitive mouse DNAs has numerous aliases; BAM HI [9, 32], L1Md [36], and MIF-1 [1] are the most common). We discuss this finding with respect to the possible role of BAM HI sequences in inhibiting GR-MTV-8 proviral transcription. In addition, we note that since BAM HI elements appear to be (or in the past were) mobile elements themselves (9), transcription and amplification of one element may result in concomitant transcription and amplification of the other member in a chimeric structure.

We prepared a genomic library from GR/A mouse tumor DNA in the vector λ L47 (23) and identified and isolated 24 MuMTV-positive clones (D. Morris, H. Bradshaw, and R. Cardiff, manuscript in preparation). We then screened the MuMTV-positive clones with probes representing various portions of the BAM HI family of repetitive mouse DNAs. Three clones hybridized to plasmid pMXB6 (9), which contained a 2.7-kilobase (kb) insert of BAM HI DNA (Fig. 1). By restriction mapping, Southern blotting, and sequencing of selected subclones (not shown), we were able to demonstrate that all three lambda clones contained the same MuMTV provirus. We chose one of the clones, λ GRT-25, for further study. This clone contained a complete MuMTV provirus plus ca. 0.6 kb of host DNA flanking the 3′ end of the provirus and at least 2 kb of host DNA flanking the 5′ end of the provirus.

We established the identity of the cloned MuMTV by sequencing ca. 600 base pairs (bp) of the long terminal repeat region of the provirus (not shown). Our sequence matched exactly with that reported by Kennedy et al. (21) for GR-MTV-8 (GR-MTV-8 is referred to by its clone designation, GR-40, in many publications [18, 19, 21]). In addition, our sequences of the neighboring host DNA matched perfectly with the host sequences reported by Kennedy et al. (21) (Fig. 1). Thus, we have no doubt that clone λ GRT-25 contains a complete copy of GR-MTV-8.

To determine the exact location of the provirus in the BAM HI element, we sequenced a large portion of the 2.7-kb insert in pMXB6. Our results indicated that the GR-MTV-8 provirus was located ca. 1.6 kb from one end of the BAM HI element (Fig. 1). We sequenced several hundred bp of GR-MTV-8 flanking DNA, and a comparison with the cor-

* Corresponding author.
† Present address: Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20205.
‡ Present address: Department of Biochemistry, University of Washington, Seattle, WA 98195.

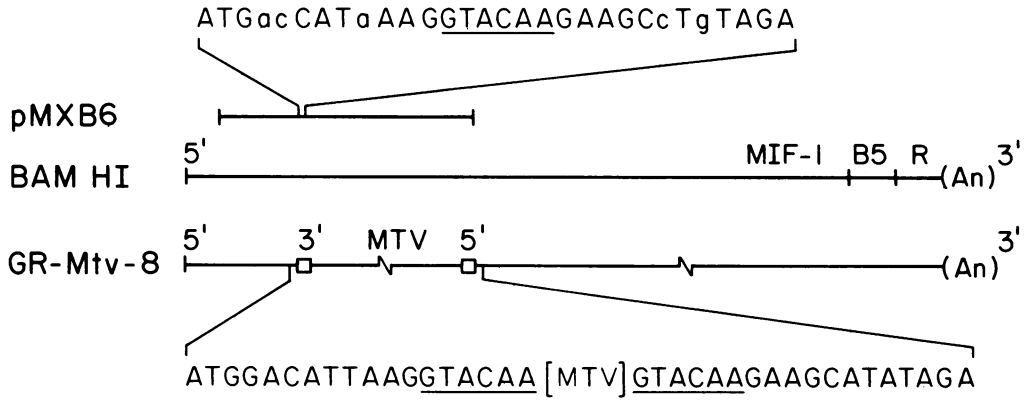

FIG. 1. Location of the GR-MTV-8 provirus in a BAM HI family member. The pMXB6 plasmid contains a 2.7 kb insert of BAM HI DNA, which maps near the 5′ end of the element (9). BAM HI elements, although often truncated at the 5′ end, are 7 kb in length (9). Subcomponents have been named R (14), BAM 5 (B5) (8), and MIF-1 (1). All BAM HI elements characterized to date contain polydeoxyadenylate tracks (An) at the 3′ end (9, 14, 36, 37). The GR-MTV-8 provirus (also known as GR-40 [18, 19, 21]) is located ca. 1.6 kb from the consensus 5′ end of the BAM HI element, making the entire chimeric structure about 17 kb in size (MuMTV proviruses are about 10 kb). The top and bottom lines show portions of the sequences flanking the GR-MTV-8 provirus (bottom) and the corresponding sequence from pMXB6 (top). The 6-bp sequence duplicated by insertion of the provirus is underlined, and nonmatching bases are given in lowercase letters (top line). Only a short stretch of sequence is presented in the figure; several hundred bp of DNA flanking the provirus have been sequenced and show, overall, 85% homology with the pMXB6 sequence.

responding region of pMXB6 demonstrated that the two BAM HI family members share about 85% homology. One conserved region in the two sequences was the 6-bp sequence duplicated during the insertion of the provirus (Fig. 1, underlined sequences).

MuMTV mRNA transcripts are readily detectable in GR/A mammary glands, solid mammary tumors, and tumor cell lines (29, 35), and results from several laboratories suggest that these RNAs originate from only two of the endogenous GR-MuMTV proviruses, GR-MTV-2 and GR-MTV-3 (12, 16, 26). No evidence for in situ transcription of GR-MTV-8 was obtained (12, 16, 26). Nevertheless, when cloned and transfected into mouse cells, GR-MTV-8 was transcribed, and transcription was under hormonal control (18, 19, 21). This indicates that GR-MTV-8 possesses all of the regulatory signals necessary for transcription, and yet the provirus appears to be transcriptionally silent in both normal and malignant mouse cells. Could the Bam HI sequences flanking GR-MTV-8 play a role in transcriptional silencing of this provirus?

The inhibitory effect of flanking DNA on proviral transcription has been noted by others (3, 7). The exact mechanism by which inhibition occurs is presently unknown, but it may involve the state of proviral DNA methylation (3, 20), although other mechanisms have been proposed (7). Certainly the proviral methylation pattern, as well as the location of the provirus in the host genome, are important and may be correlated with one another and with transcriptional activity (20). Because virtually all BAM HI elements are highly methylated (2, 30) and poorly transcribed, we propose the following scenario. Upon integration into the BAM HI element, the GR-MTV-8 provirus assumed a methylation pattern similar to that of the BAM HI element, rendering the provirus transcriptionally silent. When these methylation constraints are removed by cloning, proviral transcriptional activity is restored (18, 19, 21).

MuMTV proviruses exhibit various extents of hypomethylation in different mouse tissues. For example, in liver, brain, and kidney, all proviruses appear to be extensively methylated. In other tissues, such as spleen, testes, and placenta, various MuMTV proviruses are hypomethylated (17; T. Fanning, W.-S. Hu, and R. Cardiff, manuscript in preparation). Although one or more proviruses have been found to be hypomethylated in many nonmammary tissues, no MuMTV-specific mRNAs have been detected in these same tissues (17). To explain this observation, we (17) and others (15) have proposed that MuMTV proviruses adopt the methylation patterns of the surrounding sequences ("hitchhiker" effect [17]) and that these flanking sequences are hypomethylated in a tissue-specific fashion (15, 17). Our results with GR-MTV-8 certainly support such a model, especially in view of the following facts: (i) GR-MTV-8 appears to be highly methylated in all adult mouse tissues (15), and (ii) virtually all BAM HI elements are highly methylated in all adult mouse tissues (2, 30).

Finally, we note that not all BAM HI elements are transcriptionally silent and that BAM HI RNA is detectable in some mouse cells (8, 32). A provirus which integrated into an active BAM HI element might, therefore, be transcribed along with the element itself. Alternatively, the provirus might cause transcription of a nontranscribed BAM HI element. To our knowledge, no events of this nature have been reported, either for MuMTV or for any other proviruses or provirus-like sequences. However, because the mouse genome contains such a vast array of repetitive elements (31), proviruses (34), and provirus-like elements (22), it is conceivable that transcriptional activation of one class of elements by members of another class may occur under certain circumstances. Since BAM HI elements appear to be propagated in a manner analogous to proviruses, via reverse transcription of RNA (9), such events could lead to coamplification of repetitive DNA-provirus chimeric structures in the mouse genome.

Supported by Public Health Service grants 5R01CA21454 and 1R01 CA 31702-01 from the National Institutes of Health.

LITERATURE CITED

1. **Brown, S. D., and G. Dover.** 1981. Organization and evolutionary progress of a dispersed repetitive family of sequences in widely separated rodent genomes. J. Mol. Biol. **150**:441–466.
2. **Chapman, V., L. Forrester, J. Sanford, N. Hastie, and J. Rossant.** 1984. Cell lineage-specific undermethylation of mouse

repetitive DNA. Nature (London) **307**:284–286.
3. **Chumakov, I., J. Stuhlmann, K. Harbers, and R. Jaenisch.** 1982. Cloning of two genetically transmitted Moloney leukemia proviral genomes: correlation between biological activity of the cloned DNA and viral genome activation in the animal. J. Virol. **42**:1088–1098.
4. **Cohen, J. C.** 1980. Methylation of milk-borne and genetically transmitted mouse mammary tumor virus proviral DNA. Cell **19**:653–662.
5. **Cohen, J. C., P. R. Shank, V. L. Morris, R. D. Cardiff, and H. E. Varmus.** 1979. Integration of the DNA of mouse mammary tumor virus in virus-infected normal and neoplastic tissue of the mouse. Cell **16**:333–345.
6. **Cohen, J. C., and H. E. Varmus.** 1979. Endogenous mammary tumor virus DNA varies among wild mice and segregates during inbreeding. Nature (London) **278**:418–423.
7. **Cooper, G. M., and L. Silverman.** 1978. Linkage of the endogenous avian leukosis virus genome of virus-producing chicken cells to inhibitory cellular DNA sequences. Cell **15**:573–577.
8. **Fanning, T. G.** 1982. Characterization of a highly repetitive family of DNA sequences in the mouse. Nucleic Acids Res. **10**:5003–5013.
9. **Fanning, T. G.** 1983. Size and structure of the highly repetitive BAM HI element in mice. Nucleic Acids Res. **11**:5073–5091.
10. **Fanning, T. G., and R. G. Cardiff.** 1984. Alterations of mouse mammary tumor virus DNA during mammary tumorigenesis, p. 71–94. *In* G. Klein (ed.), Advances in viral oncology, vol. 4. Raven Press, New York.
11. **Fanning, T. G., J. P. Puma, and R. D. Cardiff.** 1980. Selective amplification of mouse mammary tumor virus in mammary tumors of GR mice. J. Virol. **36**:109–114.
12. **Fanning, T. G., J. P. Puma, and R. D. Cardiff.** 1980. Identification and partial characterization of an endogenous form of mouse mammary tumor virus that is transcribed into the virion-associated RNA genome. Nucleic Acids Res. **8**:5715–5723.
13. **Fanning, T. G., A. B. Vassos, and R. D. Cardiff.** 1982. Methylation and amplification of mouse mammary tumor virus DNA in normal, premalignant, and malignant cells of GR/A mice. J. Virol. **41**:1007–1013.
14. **Gebhard, W., T. Meitinger, and J. Höchtl, and H. G. Zachau.** 1982. A new family of interspersed repetitive DNA sequences in the mouse genome. J. Mol. Biol. **157**:453–471.
15. **Günzburg, W. H., and B. Groner.** 1984. The chromosomal integration site determines the tissue-specific methylation of mouse mammary tumor virus proviral genes. EMBO J. **3**:1129–1135.
16. **Herrlich, P., N. E. Hynes, H. Ponta, U. Rahmsdorf, N. Kennedy, and B. Groner.** 1981. The endogenous proviral mouse mammary tumor virus genes of the GR mouse are not identical and only one corresponds to the exogenous virus. Nucleic Acids Res. **9**:4981–4995.
17. **Hu, W.-S., T. G. Fanning, and R. D. Cardiff.** 1984. Mouse mammary tumor virus: specific methylation patterns of proviral DNA in normal mouse tissues. J. Virol. **49**:66–71.
18. **Hynes, N. E., N. Kennedy, U. Rahmsdorf, and B. Groner.** 1981. Hormone-responsive expression of an endogenous proviral gene of mouse mammary tumor virus after molecular cloning and gene transfer into cultured cells. Proc. Natl. Acad. Sci. U.S.A. **78**:2038–2042.
19. **Hynes, N. E., U. Rahmsdorf, N. Kennedy, L. Fabiani, R. Michalides, R. Nusse, and B. Groner.** 1981. Structure, stability, methylation, expression and glucocorticoid induction of endogenous and transfected proviral genes of mouse mammary tumor virus in mouse fibroblasts. Gene **15**:307–317.
20. **Jaenisch, J., and D. Jähner.** 1984. Methylation, expression and chromosomal position of genes in mammals. Biochim. Biophys. Acta **782**:1–9.
21. **Kennedy, N., G. Knedlitschek, B. Groner, N. E. Hynes, P. Herrlich, R. Michalides, and A. J. J. van Ooyen.** 1982. Long terminal repeats of endogenous mouse mammary tumor virus contain a long reading frame which extends into adjacent sequences. Nature (London) **295**:622–624.
22. **Kuff, E. L., L. A. Smith, and K. K. Lueders.** 1981. Intracisternal A-particle genes in *Mus musculus*: a conserved family of retrovirus-like elements. Mol. Cell. Biol. **1**:216–227.
23. **Loenen, W. A. M., and W. J. Brammar.** 1980. A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes. Gene **10**:249–259.
24. **Michalides, R., R. Van Nie, R. Nusse, N. E. Hynes, and B. Groner.** 1981. Mammary tumor induction loci in GR and DBAf mice contain one provirus of the mouse mammary tumor virus. Cell **23**:165–173.
25. **Morris, V. L., E. Mederios, G. M. Ringold, J. M. Bishop, and H. E. Varmus.** 1977. Comparison of mouse mammary tumor virus-specific DNA in inbred, wild and asian mice, and in tumors and normal organs from inbred mice. J. Mol. Biol. **114**:73–91.
26. **Nusse, R., J. deMoes, J. Hilkens, and R. van Nie.** 1980. Localization of a gene for expression of mouse mammary tumor virus antigens in the GR-MTV-2− mouse strain. J. Exp. Med. **157**:712–719.
27. **Nusse, R., and H. E. Varmus.** 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell **31**:99–109.
28. **Peters, G., S. Brookes, R. Smith, and C. Dickson.** 1983. Tumorigenesis by mouse mammary tumor virus: evidence for a common region for provirus integration in mammary tumors. Cell **33**:369–377.
29. **Ringold, G. M., K. R. Yamamoto, J. M. Bishop, and H. E. Varmus.** 1977. Glucocorticoid-stimulated accumulation of mouse mammary tumor virus RNA: increased rate of syntheses of viral RNA. Proc. Natl. Acad. Sci. U.S.A. **74**:2879–2883.
30. **Sanford, J., L. Forrester, and V. Chapman.** 1984. Methylation patterns of repetitive DNA sequences in germ cells of *Mus musculus*. Nucleic Acids Res. **12**:2823–2836.
31. **Singer, M. F.** 1982. Highly repeated sequences in mammalian genomes. Int. Rev. Cytol. **76**:67–112.
32. **Soriano, P., M. Meunier-Rotival, and G. Bernardi.** 1983. The distribution of interspersed repeats is nonuniform and conserved in the mouse and human genomes. Proc. Natl. Acad. Sci. U.S.A. **80**:1816–1820.
33. **Traina, V. L., B. A. Taylor, and J. C. Cohen.** 1981. Genetic mapping of endogenous mouse mammary tumor viruses: locus characterization, segregation, and chromosomal distribution. J. Virol. **40**:735–744.
34. **Varmus, H. E.** 1982. Form and function of retroviral proviruses. Science **216**:812–820.
35. **Varmus, H. E., N. Quintrell, E. Medeiros, J. M. Bishop, R. C. Nowinski, and N. H. Sarkar.** 1973. Transcription of mouse mammary tumor virus genes in tissues from high and low tumor incidence mouse strains. J. Mol. Biol. **79**:663–679.
36. **Voliva, C. F., C. L. Jahn, M. B. Comer, C. A. Hutchison, and M. H. Edgell.** 1983. The L1Md long interspersed repeat family in the mouse: almost all examples are truncated at one end. Nucleic Acids Res. **11**:8847–8859.
37. **Wilson, R., and U. Storb.** 1983. Association of two different repetitive DNA elements near immunoglobulin light chain genes. Nucleic Acids Res. **11**:1803–1817.