Volume 13 Number 11 1985 Nucleic Acids Research

Pattern of undermethylation of the major satellite DNA of mouse sperm

Sheldon I.Feinstein[1,4], Vincent R.Racaniello[2,4], Melanie Ehrlich[5], Charles W.Gehrke[6], Dorothy A.Miller[1,4] and Orlando J.Miller[1,3,4]

Departments of [1]Human Genetics and Development, [2]Microbiology, [3]Obstetrics and Gynecology, and [4]The Cancer Center, Columbia University, New York, NY 10032, [5]Department of Biochemistry, Tulane University, New Orleans, LA 70112, and [6]Department of Biochemistry, Experiment Station Chemical Laboratories, University of Missouri, Columbia, MO 65201, USA

Received 8 March 1985; Accepted 30 April 1985

ABSTRACT

Enzymatic hydrolysis and base analysis by high performance liquid chromatography showed that mouse satellite DNA had 30-50% less 5-methylcytosine in sperm than in somatic tissue (1.59 mols % vs 2.40-3.11 mols %). Maxam-Gilbert sequencing and analysis of the intensity of the cytosine bands indicated that the level of methylation of the eight CpGs of the consensus sequence in sperm satellite DNA ranged from 0 to about 50%, considerably lower than the levels reported in somatic tissues. The MnlI site containing one of these CpGs was cut much more extensively in satellite DNA from sperm than from liver, confirming the undermethylation of this site in sperm DNA.

INTRODUCTION

Satellite DNAs are highly repetitive sequences which are found in the genomes of most eukaryotic organisms. In the mouse, about 8% of the genome consists of a light satellite DNA made up of more than a million tandemly repeated copies of a slightly variable 234 base-pair unit whose consensus sequence has been determined by Maxam-Gilbert sequencing of the satellite fraction from somatic cells (1,2).

Satellite DNAs from somatic cells of a number of organisms are more highly methylated than the rest of the DNA. This has been shown in the human (3,4), higher primates (5,6), bovidae (7) and other vertebrates, including the mouse. The consensus sequence of mouse somatic satellite DNA contains eight CpG dinucleotides, seven of which are estimated, from Maxam-Gilbert sequencing gels, to be methylated in 80-90% of the repeat units, while the eighth, at position 205 in the sequence, is methylated in about one-half of the repeats (1,2). No sites of methylation other than CpG are found. This provides an estimate of 2.4-2.9 mol percent 5-methylcytosine, which agrees well with the values

(2.6-3 mol percent) obtained from somatic DNAs by different chemical methods (8-9). These levels are approximately threefold higher than those found in the rest of the mouse DNA from various types of somatic murine cell populations (10).

Although the consensus sequence of the mouse major satellite does not contain a HpaII or MspI site, MspI does cleave mouse satellite DNA into an orderly ladder of multimeric repeat units (11). These occasional MspI sites are presumably due to mutation from A to C at position 33 of the consensus sequence, producing a CCGG. Most of the MspI sites in somatic cell satellite DNA are not cut by HpaII, because they are methylated at the internal cytosine (11). HpaII cleaves about 33% more sites in satellite DNA from sperm and other cells of the male germ line than it does in that from somatic cells, suggesting that mouse satellite DNA is undermethylated in the male germ line (12,13).

We have confirmed and extended these findings using three different approaches. First, we have examined the extent of cleavage with the restriction endonuclease MnlI, which has a single CpG methylation-sensitive site in most repeat units (1). Unlike the HpaII site, this site is present in the consensus sequence, not just in a small percentage of clustered repeats. It is therefore more representative of the majority of the satellite DNA repeats. Second, we have used high performance liquid chromatography (HPLC) to extend our observations to the overall methylation of satellite DNA. Finally, we have used Maxam-Gilbert sequencing (14) to determine the amount of undermethylation at the eight CpG sites in the consensus sequence of sperm satellite DNA. Our results allow us to quantitate more accurately the levels of undermethylation of the satellite DNA in male germ cells in the mouse and to examine the pattern of undermethylation at different sites within the consensus sequence.

MATERIALS AND METHODS

Isolation of DNA

DNA was isolated from various tissues of randomly bred Swiss male mice essentially as described (15). Sperm were isolated by mincing epididymides and allowing the sperm to swim

out. The sperm fraction was estimated microscopically to be over 95% pure. Treatment with 1% Sarkosyl was used to eliminate the remaining somatic cells (16). The sperm cells were collected as a pellet and their DNA isolated by treatment with dithiothreitol, Pronase and phenol (17).

Purification of Satellite DNA

Satellite DNA was purified from main band DNA by isopycnic centrifugation using Hoechst 33258 and a cesium chloride density gradient (18). As expected (19), sperm and testis satellite DNA did not separate as well from the main band DNA as the highly methylated satellite of somatic tissues did. Therefore, two sequential purifications on Hoechst 33258-cesium chloride gradients were used routinely.

Restriction and Sequencing of DNA

Satellite DNA was digested with Sau96I or MnlI (New England BioLabs) as recommended by the supplier. The DNA fragments were fractionated on a 6% acrylamide gel. The monomer size Sau96I fragments were recovered from the gel as described (20) and purified on an Elutip column (Schleicher and Schuell). Eluted DNA was treated with calf intestinal alkaline phosphatase (Boehringer-Mannheim) followed by T4 polynucleotide kinase (Boehringer-Mannheim) labelling with γ^{32}P-ATP (Amersham). Radiolabeled monomers were loaded on a strand separation 6% polyacrylamide gel. Separated strands were sequenced according to the protocols of Maxam and Gilbert (14). Since 5-methyl cytosine is not registered by this method, the extent of cytosine methylation could be estimated visually from the reduced intensity of the relevant cytosine bands in the autoradiographs. In addition the autoradiographs were scanned using a Joyce-Loebl integrating microdensitometer.

High Performance Liquid Chromatography Analysis

Methods of enzyme hydrolysis of DNA, analysis of base composition by high performance liquid chromatography and quantitation of the results were as described (21-23). For the satellite DNAs, 1.5-5.0 ug of digested DNA was chromatographed. Given the much lower 5-methylcytosine content of the main band DNA as well as its greater availability, 3-14 ug of each of these DNAs was used for analysis.

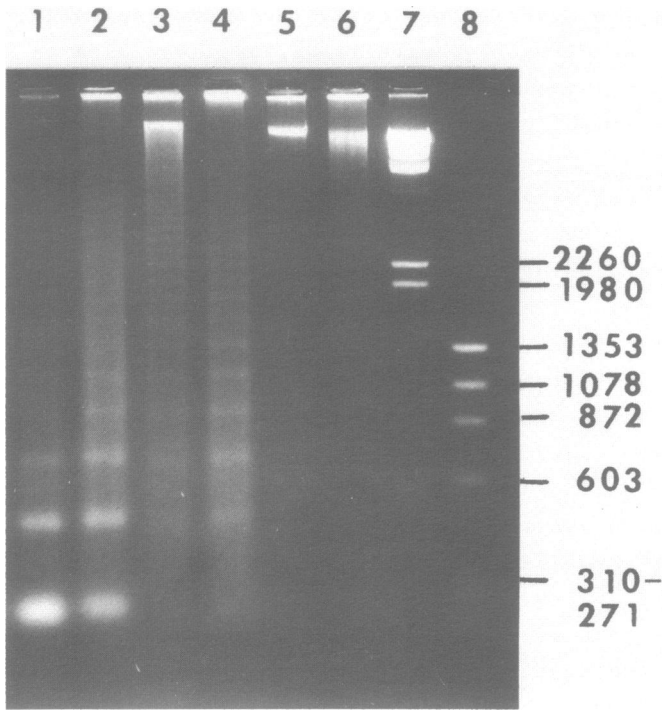

Figure 1. Satellite DNA from mouse liver (lanes 1,3,5) or mouse sperm (2,4,6) was digested for 1.5 hours at 37° with Sau96I 4 units/lane (1,2), with MnlI 2.5 units/lane (3,4) or not digested (5,6). Markers (sizes expressed as base pairs): λ DNA-HindIII fragments (7) and ϕX 174 RF DNA-HaeIII fragments (8). 6% polyacrylamide gel, ethidium bromide stain.

RESULTS

Sau96I digestion of satellite DNA from mouse sperm or liver produced primarily 234 bp monomers and decreasing amounts of dimers, trimers and multimers (Figure 1, lanes 1,2). MnlI digestion of sperm satellite DNA also produced a regular ladder of fragments, with fewer monomers and more higher multimers (Figure 1, lane 4). This suggests that unmethylated MnlI sites are distributed at random throughout most of the satellite sequences. Dimers, trimers and tetramers were quite abundant, indicating that the MnlI site is unmethylated in a moderate percentage of the repeating units. Satellite DNA from liver was cut much less extensively by MnlI (Figure 1, lane 3). The same sizes of fragments were produced, but most of the fragments were

Table 1. Base composition of mouse DNA fractions

	Deoxynucleoside* (mol %)				
Satellite DNA	m^5dCyd	dCyd	dGuo	dThd	dAdo
Sperm	1.59	18.1	14.8	32.9	32.6
Lung	3.11	15.7	16.2	33.2	31.8
Brain	2.91	17.6	16.6	31.8	31.0
Heart	2.41	16.4	17.9	32.4	30.9
Liver	2.40	16.6	17.2	33.5	31.6
Mean value	--	16.9	16.5	32.8	31.6
Main band DNA					
Sperm	0.82	20.5	19.0	31.6	30.4
Lung	0.85	20.9	20.4	29.2	28.7
Brain	0.88	19.9	19.6	30.0	29.5
Heart	0.73	20.0	19.7	29.7	29.9
Liver	0.69	19.4	20.5	29.6	29.8
Mean value	--	20.1	19.8	30.0	29.7

*Determined by high performance liquid chromatography of DNA digested to the nucleoside level. Abbreviations: m^5dCyd, 5-methyldeoxycytidine; dCyd, deoxycytidine; dGuo, deoxyguanosine; dThd, deoxythymidine; dAdo, deoxyadenosine.

in the very high molecular weight classes. Thus, the MnlI site in satellite DNA is methylated in significantly more of the repeating units in liver than in sperm.

The estimate of overall cytosine methylation derived from high performance liquid chromatography analysis of the mouse satellite DNA from various tissues is shown in Table 1. Complete methylation of all eight CpG sites in the consensus sequence of the 234 base pair monomers of satellite DNA would give 3.4 mols percent of 5-methylcytosine (8/234). With 1.59 mol percent 5-methylcytosine, satellite DNA from sperm was slightly less than 50% methylated whereas satellite DNA from somatic cells ranged from 70% methylated (liver and heart) to 90% methylated (lung). These data indicate that the level of overall methylation of satellite DNA is 30-50% lower in mouse sperm than in somatic tissues. As expected, the level of cytosine methylation was higher in the satellite DNA than it was in the main band DNA from the same tissue (Table 1). The 5-methylcytosine content of the main band DNA in different tissues ranged from 0.69 mol percent to 0.88 mol percent. Satellite DNA also had a higher adenine plus thymine content, with HPLC and consensus sequence data yielding closely

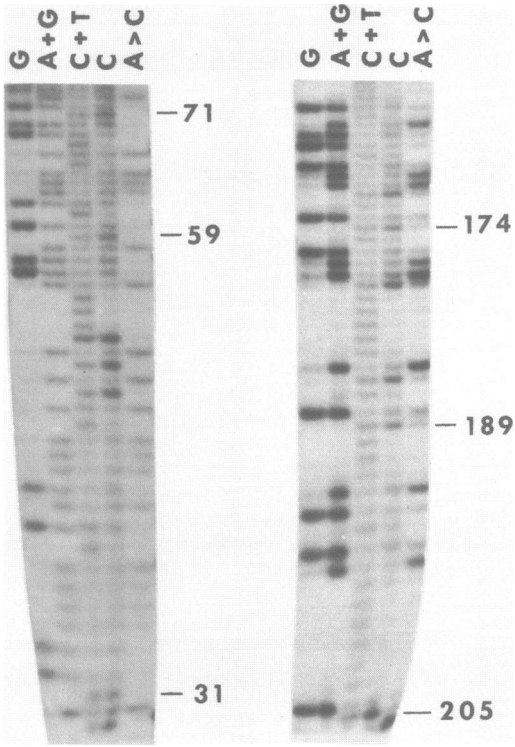

Figure 2. 3'labeled fast and slow strands sequenced after cleavage with Sau96I. Numbered arrows indicate potentially methylated cytosine positions in the sequence.

comparable values. These are in good agreement with previous results (1,2).

Maxam-Gilbert sequencing of satellite DNA confirmed the consensus sequence reported earlier (1,2). The intensity of bands representing cytosines located in CpG dinucleotides was determined from integrated microdensitometric plots produced by scanning autoradiographs of the gels. The intensity of the bands representing cytosine in each of the consensus CpG dinucleotides was compared to that of nearby cytosines to obtain an estimate of the level of methylation. This is possible because the intensity of the bands varies inversely with the extent of methylation. The level of methylation in sperm satellite DNA varied from 0 at position 205 to about 50% ($\pm$ 10%)

Table 2. Level of methylation of specific cytosines in Sau96I monomers from mouse sperm satellite DNA

Position	Methylation* Fast Strand	Slow Strand	Both Strands
13	38 ± 18 (2)	48 ± 6 (2)	43 ± 13 (4)
31	33 ± 12 (5)	- 3 ± 26 (2)	23 ± 23 (7)
59	46 ± 12 (4)	39 ± 18 (2)	44 ± 13 (6)
71	26 ± 20 (4)	44 ± 13 (2)	32 ± 19 (6)
119	52 ± 12 (2)	50 ± 12 (2)	50 ± 10 (4)
174	15 ± 23 (3)	23 ± 19 (4)	20 ± 20 (7)
189	- 3 ± 18 (3)	28 ± 11 (4)	15 ± 21 (7)
205	-46 ± 16 (3)	12 ± 13 (4)	-13 ± 34 (7)

*Mean values in % ± S.D. (number of samples)

Source of variation	d.f.	Mean square
Between positions (Fast)	7	1014.3
Within positions (Fast)	18	277.6
F ratio = 3.65, p<.02		
Between positions (Slow)	7	348.5
Within positions (Slow)	14	235.9
F ratio = 1.5, p>.05		
Between positions (F+S)	7	410.9
Within positions (F+S)	40	456.2
F ratio = 0.9, p>.05		

for the cytosine at position 119 (Table 2). The average level of methylation at the eight CpG sites in sperm estimated in this way was about 30%, which is somewhat lower than that based on HPLC data. An analysis of variance indicates that there may be significant differences between the levels of methylation at different positions in the repeating sequences ($p<.02$ for the fast strand), but the data are too variable for the slow strand or between strands to establish this point or to confirm that there is equal methylation of both strands at each site, as expected from the general symmetry of DNA methylation.

DISCUSSION

Our findings indicate that mouse satellite DNA is considerably less methylated in sperm than it is in somatic cells. These results confirm and extend those of earlier workers (12,13), who found in sperm an approximately 33% increase in the number of CCGG sites that could be cut by HpaII;

however, CCGG sites are found in only a small percentage of clustered repeats. On the other hand, a CGAGG site is found in most of the tandomly repetitive sequences of satellite DNA and can be cut by MnlI unless the C is methylated. This MnlI site is considerably less methylated in sperm than in liver satellite DNA, as shown by more extensive cutting by MnlI. In addition, we have found that the overall level of methylation in sperm satellite DNA is considerably lower than that in the various somatic tissues examined by HPLC. Furthermore, the results of Maxam-Gilbert sequence analysis indicate a generally low level of methylation of the eight CpGs in the consensus sequence of mouse sperm satellite DNA. The data are not sufficiently precise to show whether there are significant differences in the fraction of methylated CpGs among all the eight sites, although these range from approximately 50% of the CpGs methylated at position 119 to few or none methylated at position 205. The 8-nucleotide sequence surrounding position 119 contains three more G residues than that surrounding position 205. These results are consistent with the idea that changes from AT to GC base pairs outside the canonical methylase recognition sites are associated with an increased probability of methylation. Studies on bovine satellite I DNA indicate that four base pair recognition sequences containing only GC base pairs, e.g., CCGG, are more highly methylated than those containing two GC base pairs, e.g., TCGA (24).

Undermethylation of both clustered and dispersed repetitive sequences in the mouse is widespread, though not universal, in cell lineages which give rise to either extraembryonic structures or to the male or female germ line. The CCGG site found in a minority of satellite DNA repeats is almost as undermethylated in testis as in sperm DNA (12) and so is the MnlI site in the consensus repeating sequence, based on analysis of restriction enzyme and Maxam-Gilbert sequence data (unpublished). In addition, the mouse has a minor satellite which has an MspI site in most of the 116 base pair tandemly repeated units (25). This site is fully methylated in liver DNA but is unmethylated not only in sperm but also in oocyte DNA, as shown by an increase in the number of HpaII-digested sites (13).

"R" type or MIF elements of the highly repetitive L1Md or BamH1 family of interspersed repeats are not undermethylated in sperm by this criterion (12, 13) but they are in oocytes (13). The centromeric minor satellite sequences and the dispersed MIF family sequences are also undermethylated at HpaII sites in the trophoectoderm and primitive endoderm, i.e., cell lineages that differentiate early in embryogenesis, but are fully methylated at 7.5 days of development in the primitive ectoderm which gives rise to the embryo (26).

Satellite DNA is undermethylated in germ cells of mammals other than the mouse. Bovine satellite I DNA in sperm is methylated to only one-tenth the level found in calf thymus (27). The highly repetitive (cot $\leq$ 0.05) DNA which makes up 15% of the total human DNA is less methylated in sperm than in somatic tissue (21,28) and a Y-specific fraction of human satellite I DNA is less methylated in sperm than in somatic cells (29). It is still not known whether gamete-associated hypomethylation of satellite sequences reflects a special role for satellite DNA in the germ line and early embryonic tissue or is simply an epiphenomenon. Transcription of the major satellite sequences has not been observed (12), but it has not been ruled out conclusively.

ACKNOWLEDGMENTS

This work was supported by grants from the March of Dimes Birth Defects Foundation (1-361 to D.A.M. and 1-853 to M.E.), the National Cancer Institute (CA27655 to O.J.M. and CA19942 to M.E. and C.W.G.), and the National Institute of General Medical Sciences (GM25193 to O.J.M.).

REFERENCES

1. Hörz, W. and Altenburger, W. (1981). Nucl. Acids Res. 9, 683-696.
2. Manuelidis, L. (1981). FEBS Lett. 129, 25-28.
3. Miller, O.J., Schnedl, W., Allen, J. and Erlanger, B.F. (1974). Nature 251, 636-637.
4. Manuelidis, L. and Wu, J.D. (1978). Chromosoma 66, 1-22.
5. Kahana, Z.E., Miller, O.J., and Erlanger, B.F. (1978). Cold Spr. Harb. Symp. Quant. Biol. 42, 397-400.
6. Schnedl, W., Dev, V.G., Tantravahi, R., Miller, D.A., Erlanger, B.F. and Miller, O.J. (1975). Chromosoma 52, 59-66.

7. Schnedl, W., Erlanger, B.F. and Miller, O.J. (1976). Hum. Genet. 31, 21-26.
8. Salomon, R., Kaye, A.M. and Herzberg, M. (1969). J. Mol. Biol. 43, 581-592.
9. Gruenbaum, Y., Stein, R., Cedar, H. and Razin A. (1981). FEBS Lett. 124, 67-71.
10. Gama-Sosa, M.A., Midgett, R.M., Slagel, V., Githens, S., Kuo, K.C., Gehrke, C.W. and Ehrlich, M. (1983). Biochim. Biophys. Acta 740, 212-219.
11. Reilly, J.G., Thomas, C.A.,Jr. and Sen, A. (1982). Biochim. Biophys. Acta 697, 53-59.
12. Ponzetto-Zimmerman, and Wolgemuth, D. (1984). Nucl. Acids Res. 12, 2807-2822.
13. Sanford, J., Forrester, L., Chapman, V., Chandley, A. and Hastie, N. (1984). Nucl. Acids Res. 12, 2823-2836.
14. Maxam, A. and Gilbert, W. (1979). Methods Enzymol. 65, 504-561.
15. Wigler, M., Sweet, R., Sim, G.K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S. and Axel, R., (1979). Cell 16, 777-785.
16. Venolia, L. and Gartler, S.M. (1983). Somat. Cell Genet. 9, 617-627.
17. Shiurba, R. and Nandi, S. (1979). Proc. Natl. Acad. Sci. USA 76, 3947-3951.
18. Manuelidis, L. (1977). Anal. Biochem. 78, 651-568.
19. Ehrlich, M., Ehrlich, K. and Mayo, J.A. (1975). Biochim. Biophys. Acta 395, 109-119.
20. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. Ehrlich, M., Gama-Sosa, M.A., Huang, L.-H., Midgett, R.M., Kuo, K.C., McCune, R.A. and Gehrke, C. (1982). Nucl. Acids Res. 10, 2709-2721.
22. Kuo, K.C., McCune, R.A., Gehrke, C.W., Midgett, R. and Ehrlich, M. (1980). Nucl Acids Res. 8, 4763-4776.
23. Gehrke, C.W., McCune, R.A., Gama-Sosa, M.A., Ehrlich, M. and Kuo, K.C. (1984). J. Chromatog. 301, 199-219.
24. Sano, H. and Sager R. (1982). Proc. Natl. Acad. Sci. USA 79, 3584-3588.
25. Pietras, D.F., Bennett, K.L., Siracusa, L.D., Woodworth-Gutai, M., Chapman, V.M., Gross, K.W., Kane-Haas, C. and Hastie N.D. (1983). Nucl. Acids Res. 11, 6965-6983.
26. Chapman, V., Forrester, L., Sanford, J., Hastie, N. and Rossant, J. (1984). Nature 307, 284-286.
27. Sturm, K.S. and Taylor, J.H. (1981). Nucl. Acids Res. 9, 4537-4546.
28. Gama-Sosa, M.A., Wang, R.Y.-H., Kuo, K.C., Gehrke, C.W. and Ehrlich, M. (1983). Nucl. Acids Res. 11, 3087-3095.
29. Cooke, H.J., Schmidtke, J. and Gosden, J. (1982). Chromosoma 87, 491-502.