MOLECULAR AND CELLULAR BIOLOGY, Aug. 1993, p. 4523–4530
0270-7306/93/084523-08$02.00/0

Vol. 13, No. 8

Developmental Differences in Methylation of Human *Alu* Repeats

UTHA HELLMANN-BLUMBERG,[1] MARY F. McCARTHY HINTZ,[2] JOE M. GATEWOOD,[3] AND CARL W. SCHMID[1,4]*

Department of Chemistry,[1] Department of Biochemistry,[2] and Department of Genetics,[4] University of California, Davis, Davis, California 95616, and Life Sciences Division-2, Los Alamos National Laboratory, Los Alamos, New Mexico 87545[3]

Received 19 December 1992/Returned for modification 19 January 1993/Accepted 12 April 1993

***Alu* repeats are especially rich in CpG dinucleotides, the principal target sites for DNA methylation in eukaryotes. The methylation state of *Alu*s in different human tissues is investigated by simple, direct genomic blot analysis exploiting recent theoretical and practical advances concerning *Alu* sequence evolution. Whereas *Alu*s are almost completely methylated in somatic tissues such as spleen, they are hypomethylated in the male germ line and tissues which depend on the differential expression of the paternal genome complement for development. In particular, we have identified a subset enriched in young *Alu*s whose CpGs appear to be almost completely unmethylated in sperm DNA. The existence of this subset potentially explains the conservation of CpG dinucleotides in active *Alu* source genes. These profound, sequence-specific developmental changes in the methylation state of *Alu* repeats suggest a function for *Alu* sequences at the DNA level, such as a role in genomic imprinting.**

A significant fraction of mammalian DNA consists of short interspersed repeats (SINEs) such as the roughly 1 million *Alu* sequences which are broadly distributed throughout the human genome. Several *Alu* subfamilies which were retrotransposed at different evolutionary times can be recognized on the basis of their sequence divergence and diagnostic base changes (references 1, 2, 9, 21, 22, and 31 and references therein). Members of the youngest subfamily, termed PV here, closely match their consensus and share five diagnostic differences compared to the next older subfamily, Precise. These two subfamilies comprise approximately 0.1 and 10% of all *Alu*s, respectively. Most *Alu*s belong to the Major subfamily.

The average CpG dinucleotide content for human DNA is less than 1% (5), whereas *Alu*s contain up to 9% CpGs (16, 30). In vertebrates, CpG dinucleotides are frequently methylated at position 5 of cytosine, which results in a high rate of transition to either TpG or CpA (5). Consequently, older *Alu*s have lost many of their CpGs as a result of mutation. Source genes encoding young, CpG-rich *Alu* repeats presumably either are protected from germ line methylation and rapid transitions of CpG to TpG or are actively selected as source genes because they have intact CpGs (9, 31). CpGs of young *Alu*s examined by restriction enzyme cleavage are almost completely methylated in normal spleen DNA (30). Because of their large number and interspersed nature, the methylation status of *Alu* CpGs may reflect a biological function. In particular, developmental processes are often accompanied by changes in methylation, and since *Alu* repeats contribute a major fraction of CpGs, we examined their methylation status in tissues representing different developmental stages.

To investigate differences in the degree of *Alu* methylation in various human tissues, *Alu* fractions were released from total human DNA with a methyl-sensitive restriction enzyme, blotted, and hybridized to an oligonucleotide probe. Many *Alu* consensus restriction sites contain CpGs (Fig. 1), and most of the corresponding enzymes are inhibited by cytosine methylation. However, since these sites are frequently inactivated by mutations rather than methylation in older, diverged *Alu*s, restriction enzyme digestion has previously not been useful for studying *Alu* methylation. Therefore, this study focuses on young *Alu*s which have mostly intact consensus restriction sites. Also, young *Alu*s can be selectively detected against the vast *Alu* background by hybridization with specific oligonucleotide probes (1, 2, 21, 22). Oligonucleotide probe 51 (Fig. 1) incorporates two diagnostic mutations that distinguish the PV subfamily from the Precise subfamily. At high stringency (i.e., 65°C, the elution temperature for its exact complement), oligonucleotide 51 hybridizes only to PV *Alu*s, whereas at low stringency (i.e., 50°C, the elution temperature for sequences with two mutations), it partially cross-hybridizes to *Alu*s belonging to the Precise subfamily (13a). Like similar oligonucleotide probes, it does not hybridize to *Alu*s from the more abundant older subfamilies which have additional mutations (1, 2, 13a, 21, 22). The presence of older *Alu*s can be demonstrated by hybridization with nonselective, full-length *Alu* probes. Thus, a fraction of Major subfamily *Alu*s which retain the targeted restriction sites can also be sampled.

The PV, Precise, and Major *Alu* subfamily consensus sequences share a *Bst*UI site on the 5′ end and an *Asp*I site on the 3′ end (Fig. 1). While *Asp*I is not affected by methylation, *Bst*UI having the recognition sequence CpGpCpG is inhibited, so that a *Bst*UI-*Asp*I double digest releases a 265-bp fragment only if the *Bst*UI site is unmethylated. Also, because the *Bst*UI site consists of two CpGs, younger *Alu* repeats should be enriched in the *Bst*UI-*Asp*I fraction (16). Another consensus restriction enzyme, *Taq*I, which recognizes the sequence TpCpGpA (Fig. 1), is not inhibited by cytosine methylation, so that the intensity of the 200-bp *Taq*I-*Asp*I *Alu* band provides a relative measure of the total amount of *Alu* DNA.

MATERIALS AND METHODS

DNA preparations from tissue samples. DNA was prepared by standard methods (27) from sperm, HeLa cells, and

* Corresponding author.

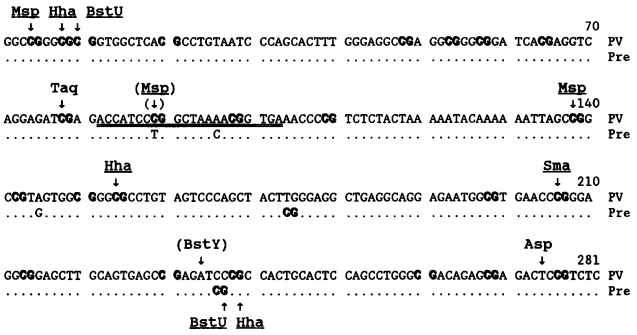

FIG. 1. Sequence of the PV *Alu* consensus, with Precise subfamily-specific mutations (1, 2, 21, 22) indicated below the sequence and CpGs indicated in boldface. Consensus restriction sites for *Msp*I-*Hpa*II, *Hha*I, *Bst*UI, *Taq*I, *Sma*I, *Bst*YI, and *Asp*I are marked by arrows. Methyl-sensitive sites are underlined, PV subfamily-specific sites are in parentheses, and Precise subfamily-specific sites are indicated below the dotted line. The position of oligonucleotide probe 51 (ACCATCCCGGCTAAAACGGTGA) is indicated by double underlining of the corresponding PV *Alu* sequence.

normal adult spleen, liver, brain, and placenta. A total of six different sperm samples was obtained from four unrelated donors. We also examined a variety of germinal cancer tissues and abnormal gestations (two testicular seminomas, a malignant testicular teratoma, a male retroperitoneal tumor [a mature teratoma], a mature cystic ovarian teratoma, and a hydatidiform mole).

Genomic blot hybridization analysis. DNA was first digested with *Asp*I at 37°C; two equal aliquots were then used for double digestion with either *Bst*UI or *Taq*I at 62°C. Each aliquot contained at least 3 μg of DNA. Control experiments demonstrate that the bands examined here result from double digestion with *Bst*UI-*Asp*I or *Taq*I-*Asp*I (data not shown). The digests were fractionated by conventional agarose gel electrophoresis using 1.4 or 1.8% agarose and blotted onto 0.2-μm-pore-size nitrocellulose (27).

Oligonucleotide hybridization and washing were performed as previously described in 5× SSPE buffer (22, 30). Oligonucleotide hybridization probe 51, having the sequence ACCATCCCGGCTAAAACGGTGA, which exactly matches the PV *Alu* subfamily consensus sequence and differs at two sites from the Precise subfamily consensus (Fig. 1), was ^{32}P end labeled with polynucleotide kinase. As discussed previously, 65°C corresponds to the elution temperature from exactly paired PV *Alu* complements and 50°C corresponds to the elution temperature from the Precise *Alu* complements (13a). These values are similar to those previously reported for a slightly different oligonucleotide probe to this region (22, 30). A nonselective, full-length hybridization probe was synthesized by polymerase chain reaction labeling a gel-purified restriction fragment containing the *Alu* PV92 (22) during amplification with 3′ and 5′ *Alu* consensus primers in the presence of [α-^{32}P]dATP. The resulting 281-bp *Alu* probe was hybridized in 3× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and washed in 0.5× SSC at 65°C.

End label detection of restriction digest. Total DNA and isolated 265-bp *Bst*UI-*Asp*I or 200-bp *Taq*I-*Asp*I restriction fragments were digested with *Msp*I or *Hpa*II isoschizomers and end labeled with [α-^{32}P]dCTP by using the Klenow fragment of *Escherichia coli* DNA polymerase and unlabeled dGTP, dATP, and TTP (23). The labeled products were fractionated by electrophoresis on 10% polyacrylamide gels and directly visualized by radioautography of the dried gels.

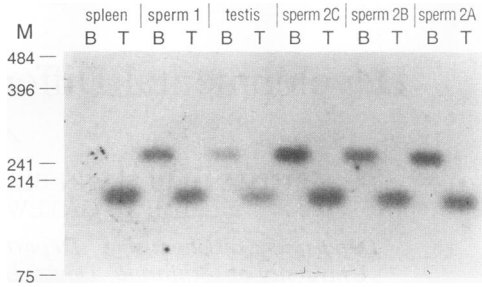

FIG. 2. Genomic blot hybridization of DNAs from spleen, sperm 1, testis, and sperm 2A, 2B, and 2C (samples from a second donor obtained at successive 2-day intervals). Lanes: T, *Taq*I-*Asp*I digests; B, *Bst*UI-*Asp*I digests of equal aliquots for each DNA sample. Washing was performed at stringent (65°C) conditions. Marker lengths (M) are indicated in base pairs on the left.

RESULTS

The *Bst*UI site of PV *Alu*s is unmethylated in sperm. The methylation status of the *Bst*UI restriction site on the 5′ end of PV *Alu*s was investigated by comparing double digests of human DNA from various sources on genomic blots which were probed with oligonucleotide 51 and washed under stringent (PV-specific) conditions (Fig. 2). The *Bst*UI-*Asp*I digest does not release a discernible 265-bp *Alu* band from spleen DNA, whereas the control *Taq*I-*Asp*I digest releases a prominent 200-bp band corresponding to cleavage at the consensus sites (Fig. 2, sperm 2A lanes). This finding confirms previous results showing that *Alu* CpGs in spleen DNA are extensively methylated (30). In contrast, the intensities of bands released by *Bst*UI-*Asp*I (lanes B) and *Taq*I-*Asp*I (lanes T) digestion of sperm DNA are approximately equal, suggesting that the *Bst*UI sites of PV *Alu*s are completely unmethylated in sperm. Virtually identical results are obtained for sperm from different donors (Fig. 2, sperm 1 and sperm 2A, 2B, and 2C; see also Fig. 4 and 5) and samples collected at different times from the same donor (Fig. 2, sperm 2A, 2B, and 2C).

For a more quantitative estimate of demethylation of the *Bst*UI site, the extent of hybridization to the 265- and 200-bp bands was determined by PhosphorImager analysis (Table 1). The average ratio for the sperm samples is 107%, showing that in sperm DNA, the *Bst*UI site of the vast majority of PV *Alu*s is unmethylated. For spleen DNA, the ratio approximates zero because the intensity of the 265-bp band approaches the background hybridization. A ratio of approximately zero is consistent with previous results which show that PV *Alu*s are almost completely methylated in spleen

TABLE 1. Hypomethylation of the 5′ *Bst*UI site in PV *Alu*s[a]

Tissue	% Unmethylated
Spleen	8.4
Sperm (1)	109
Testis	63
Sperm (2A)	117
Sperm (2B)	94
Sperm (2C)	108

[a] The intensities of the bands apparent in Fig. 2 were determined by PhosphorImager analysis. A sample-specific background which varied from 1/10 to 1/5 of the signal was subtracted in each case. The ratios of the corrected intensities of the *Bst*UI-*Asp*I band to the *Taq*I-*Asp*I band indicate the percentages of the *Bst*UI sites which are unmethylated.

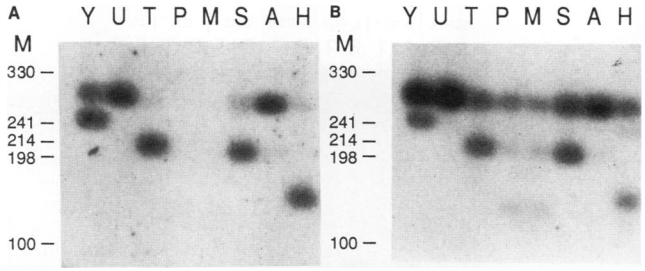

FIG. 3. (A) The 265-bp *Bst*UI-*Asp*I band from sperm DNA was digested with diagnostic and methyl-sensitive enzymes. *Bst*YI (Y) is diagnostic for the PV subfamily; uncut DNA (U) and DNA cut with *Taq*I (T) provide controls; *Hpa*II (P) does not cleave if the CpG is methylated; *Msp*I (M) is an *Hpa*II isoschizomer which cleaves regardless of CpG methylation; *Sma*I (S) is CpG methyl sensitive; *Psp*AI (A) is a failed digest; *Hha*I (H) is CpG methyl sensitive. The blot was probed with oligonucleotide 51 and washed at 50°C. (B) The blot shown in panel A was stripped and hybridized with a nonspecific full-length *Alu* probe.

DNA (30). In testis DNA, the relative intensity of the 265-bp *Bst*UI-*Asp*I band compared with the 200-bp *Taq*I-*Asp*I band suggests partial methylation (Fig. 2; Table 1). In addition to germ cells in various stages of development, testes contain somatic cells. In monkey testis, at least 40% of the cells do not belong to the germ line lineage (36).

Several CpGs in different *Alu* subfamilies are unmethylated in sperm. The experiment described above (Fig. 2) shows that a single restriction site containing 2 of 24 possible CpGs in PV *Alu*s is unmethylated in sperm DNA. In the following analysis, this result is extended to other *Alu* CpGs and other subfamilies. The 265-bp *Bst*UI-*Asp*I fraction from sperm was isolated and examined with additional restriction enzymes to determine its subfamily composition and the methylation status of other sites. The PV consensus has a diagnostic *Bst*YI site (Fig. 1). Digestion with *Bst*YI (Fig. 3A, lane Y) followed by low-stringency probing with oligonucleotide 51 (which allows for hybridization of Precise subfamily *Alu*s) reveals that a large fraction of the 265-bp band remains uncut and therefore consists of Precise subfamily *Alu*s. As expected, washing at high stringency eliminates most of the hybridization to the uncut fraction (data not shown). The vast majority of *Alu*s in the 265-bp band are cut by *Taq*I (Fig. 3A, lane T), indicating that in the case of PV and Precise *Alu*s, the *Taq*I-*Asp*I fraction used as a control in Fig. 2, 4, and 5 includes practically the entire *Bst*UI-*Asp*I fraction. Most of the 265-bp band is cleaved with the methyl-sensitive enzymes *Hpa*II (lane P), *Sma*I (lane S), and *Hha*I (lane H), showing that other *Alu* CpGs are unmethylated in sperm DNA. In particular, the digests with *Hpa*II and its methyl-insensitive isoschizomer *Msp*I appear identical, indicating that all of the corresponding CpGs are unmethylated. *Alu*s not cut with *Sma*I are presumably Precise subfamily *Alu*s which have lost the hexanucleotide recognition sequence through mutations. High-stringency washing eliminates the signal from *Alu*s not cut by *Sma*I, confirming this assignment (data not shown).

While the experimental logic used in the foregoing analysis exploits properties of young *Alu* subfamilies, the ensuing results are not restricted to a specific subfamily. Evidence presented above (Fig. 3A) and below (Fig. 5 and 6) shows that demethylation affects both the PV and Precise subfamilies. The following data indicate demethylation of CpGs in even older *Alu* subfamilies. All blots were reexamined by hybridization to a nonselective, full-length *Alu* probe which, in addition to increasing the relative intensity of the Precise *Alu* signal, detects the presence of older, more diverged *Alu* repeats. In particular, nonselective reprobing of secondary digests of the 265-bp band from sperm DNA (Fig. 3B) reveals additional *Alu*s that have lost consensus restriction sites (compare Fig. 3A and B). Even though additional bands in the *Hpa*II and *Msp*I digests become visible (Fig. 3B), the two digests are indistinguishable, showing that mutation, rather than methylation, is primarily responsible for the inactivation of consensus restriction sites in older *Alu*s. Consistent with this interpretation, nonselective reprobing (Fig. 3B) reveals that relatively large fractions of *Alu*s remain uncut for each enzyme tested (*Sma*I, *Hha*I, *Bst*YI, and *Taq*I), regardless of methyl sensitivity. These findings, in contrast to the results from the oligonucleotide hybridization, demonstrate that most *Alu*s in the 265-bp *Bst*UI-*Asp*I band belong to older subfamilies which make up the majority of *Alu* repeats. Although we have sampled only a small portion of the diverged *Alu*s, we conclude that all *Alu*s in the *Bst*UI-*Asp*I fraction are, within the accuracy of this method, completely unmethylated in sperm, as was observed for PV subfamily *Alu*s. Additional secondary digests show that in placenta and HeLa DNAs, *Alu*s in the *Bst*UI-*Asp*I fraction are not unmethylated but hypomethylated compared with DNAs from somatic tissues such as spleen, which are almost completely methylated (see below).

***Alu*s released by *Bst*UI-*Asp*I represent a distinct subset.** Secondary restriction digests, in analogy to those in Fig. 3, reveal that the 200-bp *Taq*I-*Asp*I band from sperm DNA is partially methylated, in contrast to the nearly complete demethylation of *Alu*s in the *Bst*UI-*Asp*I fraction (data not shown). A possible explanation of this difference in methylation is that the *Bst*UI-*Asp*I fraction may be significantly less methylated than the average for *Alu*s in sperm DNA. This interpretation is confirmed by comparing the intensities of *Alu*-specific bands in *Hpa*II and *Msp*I digests of both the *Taq*I-*Asp*I and *Bst*UI-*Asp*I fractions (Fig. 4). In this experiment, *Alu* consensus bands are detected by direct end labeling (23), which avoids complications due to blot hybridization efficiencies of *Alu* subfamilies.

The principal *Hpa*II and *Msp*I fragments can be identified in these digests according to their lengths predicted from the *Alu* consensus sequence (Fig. 1), but for simplicity, we focus the analysis on the 70-nucleotide band which is common to both fractions (Fig. 4, arrow). For sperm DNA, the *Taq*I-*Asp*I band is significantly less sensitive to *Hpa*II cleavage than is the *Bst*UI-*Asp*I fraction. The relative intensity of the *Hpa*II and *Msp*I bands is 0.3 for the *Taq*I-*Asp*I band but is 0.9 in the case of the *Bst*UI-*Asp*I fraction from sperm DNA (Fig. 4, legend). Since the appearance of the *Hpa*II band requires cleavage of two *Hpa*II sites, we estimate that CpGs are 55% unmethylated in the *Taq*I-*Asp*I fraction and at least 90% unmethylated in the *Bst*UI-*Asp*I fraction from sperm DNA. *Hpa*II and *Msp*I digestions of total sperm DNA gave results similar to those for the *Taq*I-*Asp*I band (data not shown), and both are consistent with blot hybridization results of Kochanek et al., who estimate that sperm *Alu* CpGs are 30% unmethylated (19). These direct labeling results confirm the conclusion that *Alu* CpGs are almost completely (at least 90%) unmethylated in the sperm DNA *Bst*UI-*Asp*I fraction and also show that *Bst*UI-*Asp*I digestion selects a distinct *Alu* subset having significantly lower methylation than the majority of sperm DNA *Alu*s.

Comparing the *Hpa*II and *Msp*I digests of the *Taq*I-*Asp*I bands from sperm, placenta, and spleen DNAs, we found

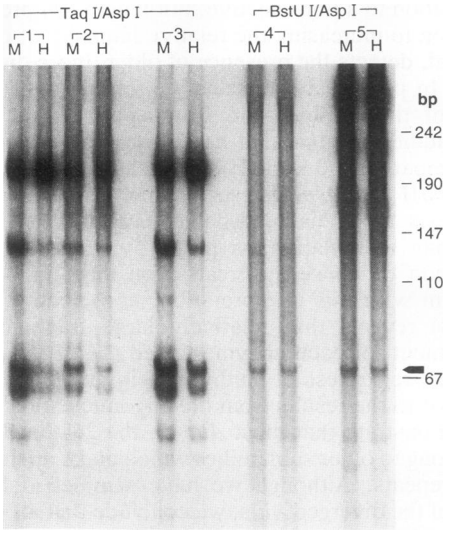

FIG. 4. Autoradiograph of end-labeled *Hpa*II (H) and *Msp*I (M) digests of *Taq*I-*Asp*I fractions from spleen (lanes 1), placental (lanes 2), and sperm (lanes 3) DNAs and of *Bst*UI-*Asp*I fractions from sperm (lanes 4) and placental (lanes 5) DNAs. Major bands can be assigned according to the consensus *Alu* restriction map (Fig. 1). Numerical values of the intensities of the 70-nucleotide bands (marked by an arrow) common to all fractions were obtained from duplicate measurements of two gels by PhosphorImager analysis. After lane-specific background subtraction, the ratios of intensities of the *Hpa*II and *Msp*I bands for the *Taq*I-*Asp*I fractions are 0.1 for spleen, 0.3 for placental, and 0.3 for sperm DNAs. For the *Bst*UI-*Asp*I fractions, these ratios are 0.6 for placental and 0.9 for sperm DNAs. The lengths of marker bands are indicated on the right.

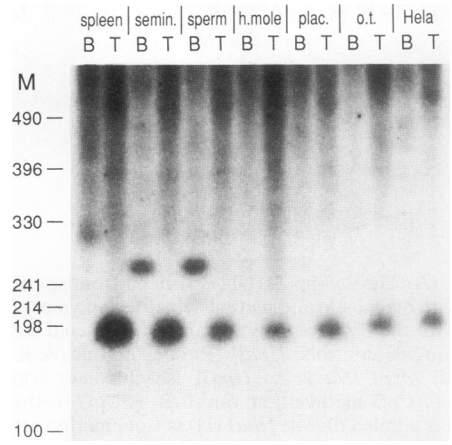

FIG. 5. Genomic blot of *Bst*UI-*Asp*I (lanes B) and *Taq*I-*Asp*I (lanes T) digests of different genomic DNAs probed with oligonucleotide 51. Pairwise, from left to right: adult male spleen, seminoma (semin.), mature sperm, hydatidiform mole (h.mole), placenta (plac.), ovarian teratocarcinoma (o.t.), and HeLa cells. The blot was hybridized to oligonucleotide 51 at 42°C and washed at 50°C. Marker lengths (M) in base pairs are indicated on the left.

that placental and sperm *Alu*s are methylated to very similar extents and significantly hypomethylated compared with spleen DNA *Alu*s (Fig. 4). Whereas young *Alu* repeats in spleen DNA are almost completely methylated (30), a small but significant fraction (approximately 10%) of the *Taq*I-*Asp*I band is cleaved by *Hpa*II, showing that some spleen *Alu*s are incompletely methylated (Fig. 4). Although the overall extents of methylation of *Alu*s in placental and sperm DNAs are very similar, the *Bst*UI-*Asp*I fraction is not noticeably enriched for a subset of unmethylated *Alu*s from placental DNA (Fig. 4). The relative intensity of the *Hpa*II and *Msp*I products released from the placental DNA *Bst*UI-*Asp*I fraction is about the same as the relative intensity of these products from the *Taq*I-*Asp*I band (approximately 30% [legend to Fig. 4 and data not shown]).

Methylation status of *Alu* CpGs in various tissues and germinal tumors. To test the possibility that the *Bst*UI-*Asp*I *Alu* subset is generally unmethylated in (male) germ line tissues and methylated in somatic tissues, the methylation state of the *Bst*UI site in additional DNA samples was determined. Blots (Fig. 5 and 6) were probed with oligonucleotide 51 at low stringency, thus detecting both PV and Precise subfamilies. Sperm and spleen DNAs examined under these conditions gave results similar to those reported above (Fig. 2), with three important differences.

(i) In the case of low-stringency probing, the 265-bp *Bst*UI-*Asp*I band from sperm is noticeably less intense than the corresponding 200-bp *Taq*I-*Asp*I band (Fig. 5; Table 2). At high stringency, the ratio of hybridization to the 265-bp band and the 200-bp band is approximately 100% for sperm, whereas at low stringency, the ratio is consistently and significantly lower (Tables 1 and 2 and data not shown). The high-stringency data are a fairly accurate measure of the fraction of PV subfamily *Alu*s demethylated at the *Bst*UI site, since PV *Alu*s closely match their consensus, i.e., are likely to retain equal numbers of *Bst*UI and *Taq*I sites. The observed difference between the low-stringency (50°C) and high-stringency (65°C) values (Table 2) is consistent with a higher degree of mutational loss of *Bst*UI sites compared with *Taq*I sites in older *Alu* subfamilies, as the *Bst*UI site consists of two highly mutable CpGs, whereas the *Taq*I site contains a single CpG. Data presented above show that in older *Alu*s, some of the other CpG-containing restriction

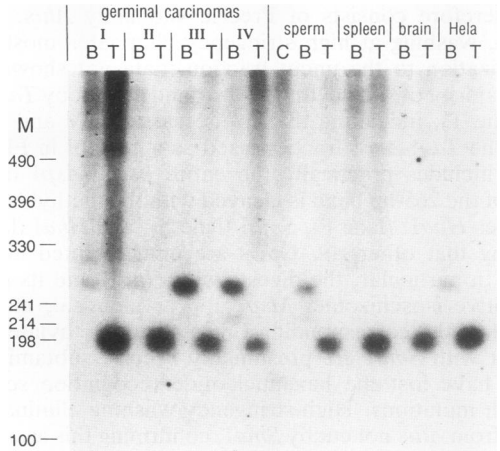

FIG. 6. Genomic blot prepared as for Fig. 4. Lanes B and T represent *Bst*UI-*Asp*I and *Taq*I-*Asp*I double digests, respectively. In the case of sperm DNA, an additional *Bst*UI-*Taq*I digest was performed (lane C); it shows the predicted 160-bp Precise *Alu* band. The germinal carcinomas are a male retroperitoneal teratoma (I), a malignant testicular teratoma (II), and testicular seminomas (III and IV). In addition, the blot shows DNAs from normal spleen, brain, and cultured HeLa cells.

TABLE 2. Hypomethylation of *Alu*s in various tissues

Tissue or cell type	Intensity (%)[a]		
	50°C (a)	50°C (b)	65°C (b)
Spleen	4.6	3.7	3.9
Seminoma	36		
Sperm	87	63	114
Hydatidiform mole	30		
Placenta	15	24	31
Ovarian teratocarcinoma	6.6		
HeLa	15	8.8	19

[a] Values represent relative intensities of the unmethylated *Alu* bands from different tissues after washing at low (50°C) or high (65°C) stringency. The blot depicted in Fig. 4 (a) and another blot which contained three times more DNA (b) were counted directly on an AMBIS radioanalytic imager. After subtraction of a lane-specific background, the ratio of the 265-bp *Bst*UI-*Asp*I band over the 200-bp *Taq*I-*Asp*I band was calculated.

sites also are inactivated by mutation rather than methylation (Fig. 3B).

(ii) An additional 230-bp *Bst*UI-*Bst*UI band corresponding to cleavage at a second *Bst*UI site present in the Precise subfamily consensus (Fig. 1) is observed in the *Bst*UI-*Asp*I digests of sperm and seminoma DNA after low-stringency washing (Fig. 5 and 6). As expected for Precise subfamily *Alu*s, this band disappears after washing at high stringency (e.g., Fig. 2). Restriction mapping of the 230-bp band also confirms this assignment (data not shown). The presence of the 230-bp Precise subfamily band provide additional direct evidence that the demethylation pattern reported here is not restricted to the PV subfamily.

(iii) Low-stringency washing reveals a more intense smear of *Alu* hybridization in addition to the bands of interest (Fig. 5 and 6) in comparison with high-stringency washing (Fig. 2). The intensity of the smear hybridization is comparable in the corresponding B and T lanes, showing that it results at least in part from mutational loss of consensus restriction sites.

Comparison of the 265-bp *Bst*UI-*Asp*I bands obtained after digestion of DNAs from various human tissues shows marked tissue specific differences in the methylation of the 5′ *Bst*UI site (Fig. 5, lanes B). Again, this band is scarcely detectable in spleen DNA (lane B). DNAs from liver (data not shown) and brain (Fig. 6) give results identical to those for spleen, suggesting that *Alu* repeats are extensively methylated in all adult somatic tissues. Genomic digests of other DNAs show varying intensities of the 265-bp band. This band is most intense in the B lanes of mature sperm and seminoma, a tumor derived from male germ cells (26). Other tissues showing a 265-bp *Alu* band are a hydatidiform mole, from an abnormal gestation with predominantly paternal DNA, and normal placenta, a tissue in which paternal genes are differentially expressed (33). Thus, the tissues in which *Alu*s are hypomethylated (seminoma, sperm, hydatidiform mole, and placenta) are associated with the male germ line or differential expression of paternal DNA.

A faint 265-bp band is also visible in DNA from cultured HeLa cells and, to a lesser extent, in several male and female germinal tumors (Fig. 5, ovarian teratocarcinoma, an embryonal-type tumor derived from female gametes; Fig. 6, male germ line tumors I and II, which are also teratomas). The level of demethylation of the 5′ *Bst*UI site in these tissues was estimated by direct counting of genomic blots (Table 2). Because the 50°C values include hybridization signals from both PV and the somewhat less well conserved Precise *Alu*s (see above), the *Bst*UI-*Asp*I/*Taq*I-*Asp*I ratios could underestimate demethylation of non-PV *Alu*s. However, differences in the blotting efficiencies of the 265- and 200-bp bands could lead to an overestimate of *Alu* demethylation. The resulting values show that the levels of *Alu* demethylation in seminoma, hydatidiform mole, placenta, and HeLa cells are intermediate between the extremes of sperm and spleen DNA (Table 2).

The methylation status of *Alu*s in the *Bst*UI-*Asp*I fraction from HeLa and placental DNAs differ from the extremes of spleen and sperm *Alu*s, as they are partially methylated. The isolated 265-bp *Bst*UI-*Asp*I fragments from placental and HeLa DNAs were analyzed by blot hybridization with additional restriction enzymes as described above for sperm (Fig. 3). In the cases of placental and HeLa DNAs, *Hpa*II cut with lower efficiency than did *Msp*I, indicating partial methylation of conserved *Alu* CpGs (data not shown). This finding agrees with the results of the *Hpa*II end labeling experiment, which show that *Alu* CpGs in the *Bst*UI-*Asp*I fraction from placental DNA are partially methylated (Fig. 4). The intermediate methylation state of *Alu* CpGs in HeLa DNA may be an artifact of particular cell culture conditions. However, *Alu* hypomethylation in placental DNA may be related to the methylation status of *Alu*s in male germ line and hydatidiform mole DNA (Fig. 5; Table 2).

Demethylation of human *Alu* repeats, including the *Bst*UI subset in particular, evidently occurs prior to spermatogenesis, as judged by the band released from seminoma DNA. Virtually identical results are obtained for testicular seminomas (Fig. 6, III and IV) and a metastasized seminoma (data not shown). Surprisingly, male teratomas from either retroperitoneum or testis (Fig. 6, I and II), which are also germinal carcinomas, do not show a high degree of *Alu* demethylation, in contrast to the seminomas, which are almost completely demethylated at the 5′ *Bst*UI site. Other tissues in this experiment (sperm, spleen, brain, and HeLa cells) provide calibrations for the degree of methylation. *Alu* demethylation in DNAs from teratomas (I and II) is somewhere between that of spleen DNA and HeLa DNA, whereas in seminoma DNAs (III and IV), demethylation is comparable to that of sperm or testis DNA (Fig. 2). The methylation pattern of the teratomas from male germ cells is also similar to that of ovarian teratocarcinoma (Fig. 5). In contrast to seminomas, teratomas contain morphologically distinct, differentiated cells indicative of embryo-like development (26).

DISCUSSION

Why *Bst*UI might identify a distinct *Alu* subset. Results of this investigation agree with the finding of Kochanek et al. that a significant fraction of *Alu* CpGs are undermethylated in sperm DNA (19). In particular, we find that an *Alu* subset released by *Bst*UI-*Asp*I digestion is almost entirely unmethylated at all CpGs tested and that this subset is enriched in young *Alu* repeats, as shown by its inclusion of almost the entire PV *Alu* subfamily.

The finding that *Bst*UI digestion reveals a unique subset of *Alu*s in sperm DNA seems surprising. Possible explanations for the identity of the *Bst*UI subset include functionality of sequences in and around the *Bst*UI site, a statistical bias toward extremely hypomethylated *Alu*s imposed by the *Bst*UI recognition sequence, and the effects of chromosomal locations; the *Bst*UI site is coincidentally positioned within the polymerase III (Pol III) promoter A box of *Alu* repeats, and mutation of these CpGs drastically reduces *Alu* template

activity (20). Also, methylation, including specific methylation of the *Bst*UI site in the promoter A box, represses *Alu* template activity (20). Transcription factors might protect actively transcribed *Alu*s from methylation, and the most active templates having intact *Bst*UI sites would be the least methylated. As a different possibility, the probability of cleaving the *Bst*UI site (CpGpCpG) depends on the fourth power of the probability that each individual CpG is unmethylated. Assuming that individual *Alu*s are hypomethylated to different degrees, cleavage by *Bst*UI would significantly enrich for those *Alu*s that happen to be unmethylated at other sites. This statistical explanation does not address why nearly all PV *Alu*s are contained in the *Bst*UI subset. DNA methylation patterns are also determined by the context of surrounding sequences, and the *Bst*UI subset might correspond to those *Alu*s which happened to insert into very hypomethylated regions of sperm DNA. Given that *Bst*UI might enrich for *Alu*s that are unmethylated at other sites, the combined effects of position and *Bst*UI enrichment could generate a *Bst*UI subset. However, even this combination of effects is insufficient to explain why almost all PV *Alu*s are contained in this subset unless we further postulate that new *Alu* insertions are targeted toward these regions which are unmethylated in the male germ line. The three possibilities discussed above are not exclusive and might concertedly define the *Bst*UI subset. *Bst*UI digestion enriches for the most hypomethylated *Alu*s as primarily determined by their sequence context, but because these *Alu*s have intact *Bst*UI sites and are hypomethylated, they include the most active templates, thereby reinforcing their hypomethylation.

Implications of the *Bst*UI subset for *Alu* activity. *Alu* repeats have long been recognized to transpose via an RNA intermediate, and they are readily transcribed by RNA Pol III in vitro (15). The mechanisms regulating transcription and transposition have not yet been determined. However, methylation represses RNA Pol III transcription of adenovirus VAI (17) and tRNA (3) genes. Similarly, methylation represses in vitro transcription of *Alu* repeats by RNA Pol III (19, 20). Major differences in the methylation state of *Alu* repeats as observed here could alter their expression in vivo and thus the probability of their transposition. The transposition of a PV *Alu* which is thought to have occurred in the paternal germ line is consistent with the demethylation of *Alu*s in the male germ line observed here (35).

Since young *Alu*s have mostly intact CpGs, whereas in older *Alu*s these residues have decayed into TpGs because of methylation, source genes encoding young *Alu*s must also have intact CpG residues. CpGs in source genes might be protected from germ line methylation, or intact CpGs might provide a selective transcriptional and transpositional advantage favoring certain potential source genes (9, 20, 31). The *Bst*UI subset exemplifies the possibility that potential source genes are protected from methylation during at least part of the male germ line development, and it is especially notable that this subset includes most members of the very young, transcriptionally and transpositionally active PV *Alu* subfamily. As discussed above, this subset might be selectively transcribed in comparison with both methylated *Alu*s and *Alu*s lacking the *Bst*UI site, thereby further insuring its unmethylated status in the germ line.

Possible role for *Alu* DNA in developmental regulation. The biological function of human *Alu* and other mammalian SINE families such as rodent BI repeats is unknown. If these and other SINEs serve a common function, then this function is accomplished by very different sequences. DNAs of nonhomologous mammalian SINEs do share two features: they are CpG rich and ubiquitously distributed throughout their respective genomes.

Alu repeats contribute a significant fraction of CpGs in human DNA (30). The demethylation of a select subset these sequences is especially remarkable when contrasted to the increased level of methylation of many genes and other sequences in human and mouse sperm DNA (10, 14, 18, 23, 32; reviewed in reference 24). For example, interspersed mouse L1 repeats are slightly less methylated in oocytes than in sperm (14). The absence of CpG methylation in the *Bst*UI subset sperm DNA *Alu*s and the overall reduction of *Alu* methylation in the male germ line are therefore highly sequence specific. Tandemly organized satellite sequences are also hypomethylated in sperm DNA (12, 25, 28), but the ubiquity of unmethylated *Alu* repeats potentially imprints the entire paternal genome.

As judged from the result obtained with normal testis, which contains mostly immature sperm in addition to various other cell types, the *Bst*UI *Alu* subset is already unmethylated at the beginning of spermatogenesis. This interpretation is supported by the observation on seminoma, which is also derived from primary spermatocytes. In extraembryonic tissues, placental tissue and hydatidiform mole tissue, whose formation is largely directed by the expression of the paternal genome (33), the *Alu* repeats are hypomethylated in comparison with adult somatic tissue (spleen, liver, and brain). In contrast, the methylation state of *Alu*s in teratomas originating from male (or female) germ cells is very similar to that in HeLa cells or adult somatic tissue, showing that *Alu* methylation is either present in very early germ cells (before sex-specific gametogenesis) or reestablished during embryonic development. *Alu*s are methylated in ovarian teratocarcinoma, but we do not know their methylation state in the female germ line. A major difference in the methylation state of *Alu* repeats in male and female germ lines would be a uniquely suitable determinant of genomic imprinting.

The high transition frequency of *Alu* CpGs to TpGs (16) indicates that these dinucleotides are methylated in the germ line. Since the *Bst*UI subset is either unmethylated or methylated for only a relatively brief period in male germ line development, we suspect that they might be methylated in the female germ line.

Magnitude of developmental changes in *Alu* methylation. The methylation state of *Alu* repeats affects the total abundance of 5-methylcytosine in human DNA. Depending on the tissue type, about 0.7 to 1.0 mol% of human DNA consists of 5-methylcytosine residues. As specific examples, this value is 0.84% in sperm DNA and is 0.98% in brain DNA (11). Assuming that the human genome consists of 2.5×10^9 bp, 1% corresponds to 5×10^7 5-methylcytosines. There are 25 CpGs in the Precise *Alu* subfamily consensus, so that the estimated 100,000 members of this subfamily would account for as many as 5×10^6 potential methylation sites. In older subfamilies, many of the CpGs have been lost and overall the number of CpGs is highly variable (16). Typical values range from 3 to 10 CpGs per *Alu*. Assuming a nominal average of six CpGs for the remaining 900,000 members of older subfamilies results in another 10^7 potential methylation sites. While these estimates are not very accurate, *Alu* repeats could contribute as many as 1.5×10^7 methylation sites to human DNA, which typically contains 5×10^7 methylated cytosines. Kochanek et al. (19) estimate that 30% of the *Alu* CpGs in total sperm DNA are unmethylated. Results from the *Hpa*II end labeling experiment reported here are consistent with 55% of the CpGs being unmethylated in the

*Taq*I-*Asp*I band from sperm DNA. Given these values, *Alu* repeats are entirely sufficient to account for the observed differences in the 5-methylcytosine content of DNAs from sperm and more highly methylated tissues such as brain (11). The absence of these 5-methylcytosines in sperm DNA constitutes a major structural alteration from somatic DNA.

Changes in methylation have been linked to transcriptional regulation and often involve CpG-rich or HTF islands (6–8). Because they are CpG rich, *Alu*s might be considered mini-CpG islands. Like CpG islands, *Alu*s are enriched in Giemsa-negative chromosome bands, which contain the majority of genes (4). However, at present there is no evident relationship between *Alu* demethylation and transcriptional activation of neighboring genes. *Alu* CpGs are largely methylated in somatic tissues, whereas CpG islands are often unmethylated (7). The methylation levels of both *Alu*s and CpG islands may each independently influence chromatin structure and activity and, consequently, gene expression.

Different methylation patterns of paternal and maternal genomes have been implicated in genomic imprinting (13, 29, 34). Because of their ubiquity, the quantitative and sequence-specific differences between the methylation pattern of *Alu* repeats in sperm DNA and those in DNAs from other tissues makes them excellent candidates for this function.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM21346 and the Agricultural Experiment Station at UCD.

We thank Raymond Teplitz for insightful guidance concerning the significance of the tissue samples investigated. We also thank Wendy Erickson for providing valuable DNA samples and Raymond Teplitz, Judy Lund, and the Cooperative Human Tissue Network for tissue samples. We thank Carol Rubin for introducing the end labeling assay for *Alu* fragments, and we especially thank S. Kochanek for graciously providing a preprint of his important findings on *Alu* methylation.

REFERENCES

1. **Batzer, M. A., and P. L. Deininger.** 1991. A human-specific subfamily of Alu sequences. Genomics **9**:481–487.
2. **Batzer, M. A., G. E. Kilroy, P. E. Richard, T. H. Shaikh, T. D. Desselle, C. L. Hoppens, and P. L. Deininger.** 1990. Structure and variability of recently inserted Alu family members. Nucleic Acids Res. **18**:6793–6798.
3. **Besser, D., F. Gotz, K. Schulze-Forster, H. Wagner, H. Kroger, and D. Simon.** 1990. DNA methylation inhibits transcription by RNA polymerase III of a tRNA gene, but not of a 5S rRNA gene. FEBS Lett. **269**:358–362.
4. **Bickmore, W. A., and A. T. Sumner.** 1989. Mammalian chromosome banding—an expression of genome organization. Trends Genet. **5**:144–148.
5. **Bird, A. P.** 1980. DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res. **8**:1499–1504.
6. **Bird, A. P.** 1986. CpG-rich islands and the function of DNA methylation. Nature (London) **321**:209–213.
7. **Bird, A. P.** 1992. The essentials of DNA methylation. Cell **70**:5–8.
8. **Cedar, H.** 1988. DNA methylation and gene activity. Cell **53**:3–4.
9. **Deininger, P. L., M. A. Batzer, C. A. Hutchison, and M. H. Edgell.** 1992. Master genes in mammalian repetitive DNA amplification. Trends Genet. **8**:307–311.
10. **Driscoll, D. J., and B. R. Migeon.** 1990. Sex difference in methylation of single-copy genes in human meiotic germ cells: implications for X chromosome inactivation, parental imprinting, and origin of CpG mutations. Somatic Cell Mol. Genet. **16**:267–282.
11. **Ehrlich, M., M. A. Gama-Sosa, L.-H. Huang, R. M. Midgett, K. C. Kuo, R. A. McCune, and C. Gehrke.** 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissue or cells. Nucleic Acids Res. **10**:2709–2721.
12. **Gama-Sosa, M. A., R. Y.-H. Wang, K. C. Kuo, C. W. Gehrke, and M. Ehrlich.** 1983. The 5-methylcytosine content of highly repeated sequences in human DNA. Nucleic Acids Res. **11**: 3087–3095.
13. **Groudine, M., and K. F. Conkin.** 1985. Chromatin structure and de novo methylation of sperm DNA: implications for activation of the paternal genome. Science **228**:1061–1068.

13a. **Hellmann-Blumberg, U.** Unpublished data.

14. **Howlett, S. K., and W. Reik.** 1991. Methylation levels of maternal and paternal genomes during preimplantation development. Development **113**:119–127.
15. **Jelinek, W. R., and C. W. Schmid.** 1982. Repetitive sequences in eukaryotic DNA and their expression. Adv. Biochem. **51**:813–844.
16. **Jurka, J., and A. Milosavljevic.** 1991. Reconstruction and analysis of human Alu genes. J. Mol. Evol. **32**:105–121.
17. **Jüttermann, R., K. Hosokawa, S. Kochanek, and W. Doerfler.** 1991. Adenovirus type 2 VAI RNA transcription by polymerase III is blocked by sequence-specific methylation. J. Virol. **65**: 1735–1742.
18. **Kafri, T., M. Ariel, M. Brandeis, R. Shemer, L. Urven, J. McCarrey, H. Cedar, and A. Razin.** 1992. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. Genes Dev. **6**:705–714.
19. **Kochanek, S., D. Renz, and W. Doerfler.** 1993. DNA methylation in the Alu sequences of diploid and haploid primary human cells. EMBO J. **12**:1141–1151.
20. **Liu, W. M., and C. W. Schmid.** 1993. Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. Nucleic Acids Res. **21**:1351–1359.
21. **Matera, A. G., U. Hellmann, M. F. Hintz, and C. W. Schmid.** 1990. Recently transposed Alu repeats result from multiple source genes. Nucleic Acids Res. **18**:6019–6023.
22. **Matera, A. G., U. Hellmann, and C. W. Schmid.** 1990. A transpositionally and transcriptionally competent *Alu* subfamily. Mol. Cell. Biol. **10**:5424–5432.
23. **Monk, M., M. Boubelik, and S. Lehnert.** 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. Development **99**:371–382.
24. **Monk, M., and M. Grant.** 1990. Preferential X-chromosome inactivation, DNA methylation and imprinting. Development **1990**(Suppl.):55–62.
25. **Ponzetto-Zimmerman, C., and D. J. Wolgemuth.** 1984. Methylation of satellite sequences in mouse spermatogenic and somatic DNAs. Nucleic Acid Res. **12**:2807–2822.
26. **Roth, L. M., and J. J. Gillespie.** 1980. Pathology and ultrastructure of germinal neoplasms of the testis, p. 1–28. *In* L. H. Einhorn (ed.), Testicular tumors. Mason Publishing, New York.
27. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
28. **Sanford, J., L. Forrester, and V. Chapman.** 1984. Methylation patterns of repetitive DNA sequences in germ cells of Mus musculus. Nucleic Acids Res. **12**:2823–2835.
29. **Sapienza, C., A. C. Peterson, J. Rossant, and R. Balling.** 1987. Degree of methylation of transgenes is dependent on gamete or origin. Nature (London) **328**:251–254.
30. **Schmid, C. W.** 1991. Human Alu subfamilies and their methylation revealed by blot hybridization. Nucleic Acids Res. **19**: 5613–5617.
31. **Schmid, C. W., and R. J. Maraia.** 1992. Transcriptional regulation and transpositional selection of active SINE sequences. Curr. Opin. Genet. Dev. **2**:874–882.
32. **Shemer, R., T. Kafri, A. O'Connell, S. Eisenberg, J. L. Breslow, and A. Razin.** 1991. Methylation changes in the apolipoprotein AI gene during embryonic development of the mouse. Proc. Natl. Acad. Sci. USA **88**:11300–11344.

33. **Surani, M. A. H., S. C. Barton, and M. L. Norris.** 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature (London) **308:**546–550.
34. **Swain, J. L., T. A. Stewart, and P. Leder.** 1987. Parental legacy determines methylation and expression of an autosomal transgene: a molecular mechanism for parental imprinting. Cell **50:**719–727.
35. **Wallace, M. R., L. B. Andersen, A. M. Saulino, P. E. Gregory, T. W. Glover, and F. S. Collins.** 1991. A de novo Alu insertion results in neurofibromatosis type 1. Nature (London) **353:**864–866.
36. **Weinbauer, G. F., U. Fingscheidt, and E. Nieschlag.** 1991. Human follicle-stimulating hormone exerts a stimulatory effect on spermatogenesis, testicular size and serum inhibin levels in the gonadotropin-releasing hormone antagonist-treated nonhuman primate (*Macaca fascicularis*). Endocrinology **129:**1831–1839.