The EMBO Journal vol.5 no.9 pp.2223-2229, 1986

Comparison of Factor IX methylation on human active and inactive X chromosomes: implications for X inactivation and transcription of tissue-specific genes

Catherine Ruta Cullen, Pamela Hubberman, David C.Kaslow and Barbara R.Migeon

Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

Communicated by G.G.Brownlee

Maintenance of dosage compensation for housekeeping genes on the human X chromosome is mediated through differential methylation of clustered CpG nucleotides associated with these genes. To determine if methylation has a role in maintaining inactivity of X-linked genes which show tissue-specific expression, we examined the locus for blood clotting Factor IX. The analysis encompassed 91% of the *Hpa*II and *Hha*I sites in the 41-kb region that includes the presumed promoter region, 5 kb of 5'- and 4 kb of 3'-flanking sequences. Although there are sex differences in methylation of the locus in leukocytes, the methylation pattern in liver, where the gene is expressed, is essentially the same for loci on the active and inactive X chromosome. The lack of differences in methylation of active and inactive genes makes it unlikely that methylation within the locus has a role in expression of the Factor IX gene. These findings, along with the absence of clustered CpG dinucleotides within the Factor IX locus, suggest that functional differences in DNA methylation related to X chromosome dosage compensation may be limited to CpG clusters. In any event, dosage compensation seems to be maintained regionally, rather than locus by locus.

Key words: dosage compensation/CpG clusters/DNA polymorphism/hemophilia B

Introduction

In organisms with an XX/XY sex chromosome complex, compensatory mechanisms have evolved to eliminate sex differences in dosage of genes on the X chromosome. Dosage compensation in mammals is achieved by silencing one of the X chromosomes in female somatic cells, resulting in a single active X in both sexes (Lyon, 1972). Although the molecular basis for mammalian X dosage compensation is unknown, it is clear that initiation of the process, and maintenance of the single active X from one cell to its progeny, are separate events. Dosage compensation is not initiated simultaneously in all tissues (Monk and Harper, 1979) but is programmed along with other tissue-specific functions. After tissue differentiation, only one X chromosome has the potential for transcriptional activity; the others are inactive. Except for oocytes and extraembryonic membranes (Migeon *et al.*, 1985, 1986) the maintenance of transcriptional inactivity is highly stable: rare reactivation events are localized, and do not affect the chromosome as a whole.

There have been numerous suggestions that DNA methylation has a role in silencing the inactive X chromosome (Riggs, 1975; Sager and Kitchin, 1975; Holliday and Pugh, 1975). Assumptions that the inactive chromosome is relatively hypermethylated have been based on observations that promoter and enhancer regions of transcribed autosomal genes tend to be hypomethylated (reviewed in Doerfler, 1983), that methylation of CpG dinucleotides *in vitro* can inhibit transcription (Busslinger *et al.*, 1983) and that methylation inhibitors like 5-azacytidine can derepress genes on the silent X (Jones *et al.*, 1982). The hypothesis is attractive, as modifying the inactive X chromosome by DNA methylation would provide a way to maintain differential chromosome activity.

In fact, comparisons of active and inactive X DNA show that maintenance of X dosage compensation is not associated with global differences in methylation of the chromosome. In regions probed by random DNA sequences, both X chromosomes are methylated similarly (Wolf and Migeon, 1982; Lindsay *et al.*, 1985); consistent sex differences in methylation have been reported only for X-linked housekeeping genes such as glucose-6-phosphate dehydrogenase, hypoxanthine phosphoribosyl transferase and phosphoglycerate kinase. Clusters of CpG dinucleotides associated with these genes are hypomethylated in active genes and methylated in inactive ones (Wolf *et al.*, 1984a,b; Riggs *et al.*, 1985). Paradoxically, except for these clusters, the active gene is extensively methylated whereas the inactive gene is relatively less methylated (Wolf *et al.*, 1984a; Yen *et al.*, 1984; Toniolo *et al.*, 1984).

On the other hand, methylation patterns observed for housekeeping genes on the active X chromosome are not unique features of X chromosomal genes (reviewed by Bird, 1986). Autosomal genes, such as adenine phosphoribosyl transferase (Stein *et al.*, 1983) and dihydrofolate reductase (Shimada and Nienhuis, 1985) have extensive 5' clusters of CpG dinucleotides that are characteristically hypomethylated. That CpG clusters in genes on the inactive X chromosome are methylated and nuclease insensitive, but are specifically demethylated and acquire nuclease sensitivity when the silent allele is reactivated (Wolf *et al.*, 1984a, b; Wolf and Migeon, 1985) indicates a relationship between the state of methylation of these CpG clusters and transcription of the relevant gene. In fact, the comparison of methylation of active and inactive housekeeping loci (only possible for X-linked genes) has revealed which aspects of the differential methylation are functional and which are not (Wolf *et al.*, 1984a). Such observations clearly indicate that DNA methylation within CpG clusters helps to maintain the silence of housekeeping genes on the inactive X.

To determine if DNA methylation also has a role in maintaining silence of X-linked genes that are not ubiquitously expressed, we examined the Factor IX locus which codes for a plasma glycoprotein in the intrinsic clotting pathway. The gene is present in single copy with no evidence of pseudogenes, overlapping or internal genes (Anson *et al.*, 1984; Yoshitake *et al.*, 1985); the locus is on the distal long arm of the X chromosome at Xq27.3 (Boyd *et al.*, 1984; Buckle, 1986). Although direct evidence that the locus is subject to dosage compensation is not available (i.e. demonstration of cells with mutant and others with normal activity in liver from carriers of hemophilia B), the existing evidence strongly suggests that the gene on the inactive X is silent: no

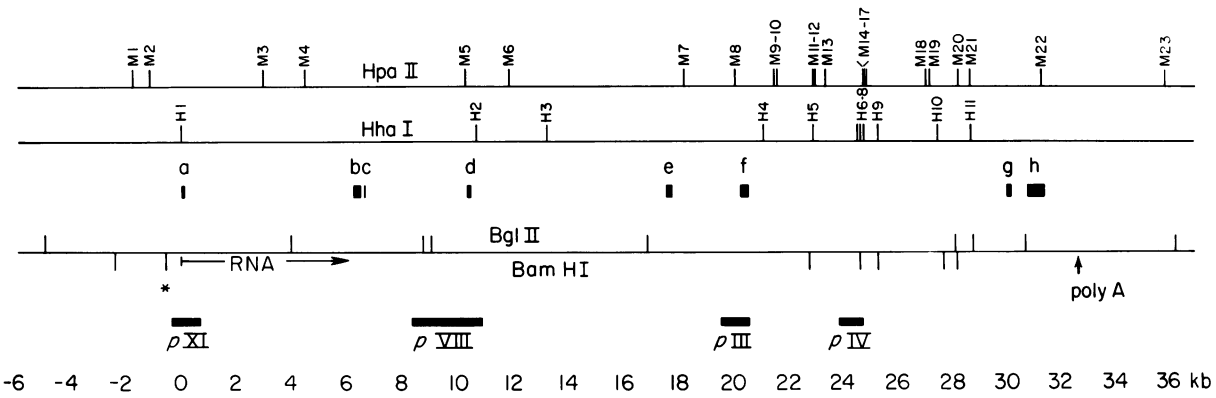

Fig. 1. Map of the Factor IX locus showing restriction enzyme sites, introns, exons (a−h) and location of genomic probe (pXI, pVIII, pIII, pIV).

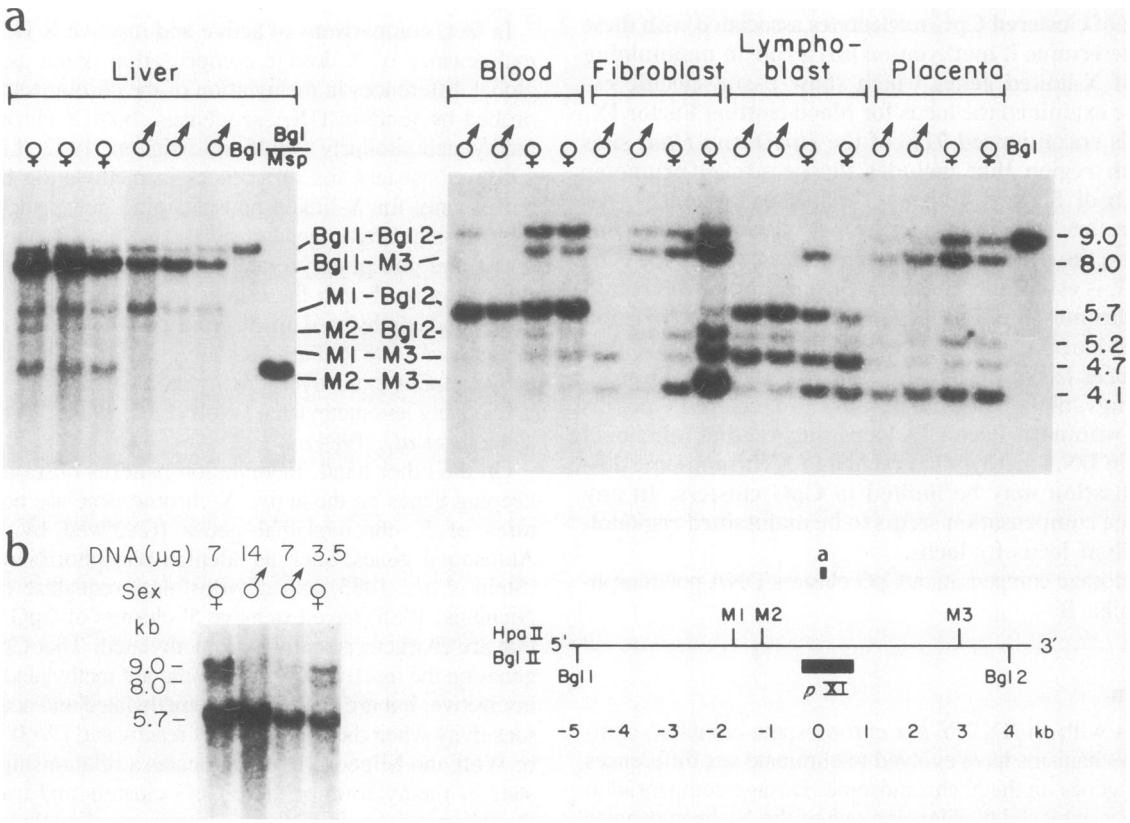

Fig. 2. Southern blot showing methylation of 5′ *Hpa*II sites (M1−M3) in males (♀) and females (♂). Size markers are given in kb. **(a)** Restriction fragments from liver, leukocyte, fibroblast, lymphoblast and placenta DNA digested with *Bgl*II/*Hpa*II, probed with pXI. **(b)** Effect of varying the amount of *Bgl*II/*Hpa*II digested leukocyte DNA, probed with pXI.

sex differences in Factor IX activity have been noted, and females known to be heterozygous for the deficiency (having normal sons as well as sons with hemophilia B) frequently have <50% Factor IX activity (Kasper *et al.*, 1977; Thompson, 1977) and may manifest abnormal clotting (Thompson, 1977) — unlikely if both alleles were active. The cDNA probes for the locus were isolated from liver libraries and antibodies to Factor IX localize the production of antigen to hepatocytes (Wion *et al.*, 1985). Unlike Factor VIII, which is synthesized at non-hepatic sites, Factor IX is vitamin K-dependent and, therefore, generally considered to be synthesized exclusively in liver (discussed by Wion *et al.*, 1985). Furthermore, there are probes for the entire locus and promoter region (Anson *et al.*, 1984), and the nucleotide sequence of the complete locus is known (Anson *et al.*, 1984; Yoshitake *et al.*, 1985). Our strategy was to compare the methylation of CpGs at the Factor IX locus on active and inactive X chromosomes in liver and other tissues. The analysis reveals inconsistent sex differences in methylation of the locus, and significant tissue variation. The differential methylation does not correlate with expression of the locus.

Results

Map of locus and location of probes

Figure 1 shows a map of the locus including relevant restriction enzyme sites. The locations of these sites were derived from published restriction maps (Anson *et al.*, 1984), sequence data (Yoshitake *et al.*, 1985) and results of our restriction enzyme

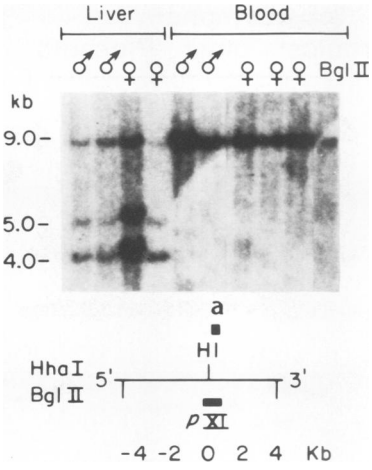

Fig. 3. Southern blot showing tissue differences in methylation of *Hha*I site H1. *Bgl*II/*Hha*I digest of leukocyte and liver DNA probed with pXI. Note absence of sex difference.

digests. The genomic probes permitted us to examine the 5′ non-transcribed region, the first six exons and five intervening sequences, and the cDNA was used for the seventh and largest exon as well as 4 kb of the 3′-flanking region.

The presumptive promoter region of Factor IX differs from those of X-linked housekeeping genes, as it has several potential TATA signals and lacks a CpG-rich region (Anson *et al.*, 1984; Yoshitaki *et al.*, 1985). The 41 kb of DNA analysed contained 281 CpGs, including 23 *Hpa*II sites (M1−23) and 11 *Hha*I sites (H1−11). However, even in the area richest in these sites (in the 2300 bp extending from M11 to H9), the ratio of CpGs to GpCs (40:100) was inconsistent with criteria suggested by Bird (1986) for CpG clusters. We were able to examine 91% of the *Hpa*II sites and 91% of the *Hha*I sites; the remaining were inaccessible usually because neighboring sites were unmethylated.

Overview of Factor IX methylation in males

If the locus on all active X chromosomes within a tissue were methylated similarly, and the genomic probes had no internal *Hpa*II (or *Hha*I) sites, then DNA from males, who have only a single locus, should yield only a single *Hpa*II (or *Hha*I) fragment. However, the presence of several fragments in most tissues (see Figures 2−5) indicates that all cells within the tissue are not methylated the same. Overall, the locus is heavily methylated (Figure 5) except for a few sites in some cells; these specific unmethylated sites vary among tissues. Liver differs in methylation of 5′ sites probed by pXI (Figures 2, 3); *Hpa*II site M3 and *Hha*I site H1, which are methylated in most tissues, are unmethylated in approximately half of the genes from liver cells.

Overview of Factor IX methylation in females

The pattern in females was expected to be complex as there are two loci per cell, each with a potentially unique methylation pattern. However, most *Hpa*II and *Hha*I fragments from female cells are the same as those from male cells — with each showing a dosage effect; that is, more intense than the corresponding fragment in male DNA (Figures 2a, 3, 4b and 5) indicating that the gene is methylated similarly on most active and inactive X chromosomes. The sex differences in methylation of a few *Hpa*II sites are prominent in blood cells, not liver — so are not related to transcript activity, and the differential sites are inconsistently methylated in most tissues (see Figures 2, 4a and 5, and below).

Specific details of the Factor IX methylation

The 5′ gene region. Methylation in this region varies considerably among tissues. *Bgl*II/*Hpa*II digests probed with pXI (Figure 2a) show that site M3 in liver and placenta, sites M1 and M3 in lymphoblasts and site M1 in blood are predominantly unmethylated. Furthermore, the multiple bands within individual specimens indicate that the population of cells is heterogeneous. For example, in liver, the major band is at 8.0 kb, so that the site at M3 is unmethylated, but the weaker bands at 4.1, 4.7, 5.2 and 5.7 kb indicate that a variety of sites may be unmethylated in small populations of cells. In blood, the major band in males is the 5.7-kb fragment (resulting from unmethylated M1) so that the majority of cells have this pattern; however, unlike most males, females have additional 8.0- and 9.0-kb bands. Dosage studies (Figure 2b) show that absence of these bands in males was not because they have less X DNA; increasing the amount of male DNA, or decreasing the amount of female DNA did not alter the results. On the other hand, these bands must originate from a small number of cells, as the major band in females (the 5.7-kb band) shows the dosage expected if both genes in most cells were methylated the same as the single gene in the male.

Examination of H1, the only 5′ *Hha*I site, reveals a striking difference between liver and blood cells (Figure 3). The 9.0-kb fragment in leukocytes indicates that H1 is consistently methylated in males and females. However, the strong 5.0-kb and 4.0-kb fragments in liver cells show that H1 is unmethylated in at least half the genes from both males and females.

The body of the gene. Figure 4a shows fragments obtained when *Bgl*II/*Hpa*II digests from various tissues were hybridized with pIII. Methylation in this 11.3-kb region varies within and between tissues. Liver DNA has a predominant 11.3-kb band, more abundant than in other tissues, resulting from methylation of all *Hpa*II sites in this region. The prominent fragments in other tissues (11.3, 6.5, 6.1 and 4.7 kb) reflect variation in methylation of sites M11−13; some of this variation represents a sex difference (see below). The same digests probed with pVIII or the cDNA show that M6 is another variable site, being unmethylated in liver, but methylated in blood (band 2.7 in Figure 5). Other *Hpa*II sites in this region examined with genomic probe pIV (data not shown) are predominantly methylated (see below). *Hha*I sites are variably methylated, but no sex differences are apparent (Figure 4b).

The 3′ region. Lacking suitable genomic probes, we used the 1.9-kb cDNA to examine the 3′ region of the gene. The most intense band in *Bgl*II/*Hpa*II digests (the 5.7-kb and 1.9-kb bands in Figure 5), originate from the vicinity of exons g and h (see Figure 1). The only *Hpa*II sites in this region (M22 and M23) are usually methylated in liver and blood from males and females.

Blots probed with the cDNA also provide the best overview of methylation of this locus. The similarity between the *Hpa*II/*Bgl*II digests in lanes 1−7 and the complete *Bgl*II digest shown in lane 8 of Figure 5 indicate that the predominant pattern for liver in both males and females is one in which all *Hpa*II sites in the region from M5 to M23 are methylated. This blot also shows the significant exceptions: the 6.5-kb fragment present in blood cells (and less intense in liver) of females, but not males, reflects the sexual dimorphism in the body of the gene, and the 2.7-kb fragment, present in liver, but not in blood, reflects the differential methylation of M6 (seen best in *Bgl*II/*Hpa*II digests probed with pVIII, not shown).

Sex differences in methylation. Some sites within the 5′ region and body of the gene are differentially methylated in males and

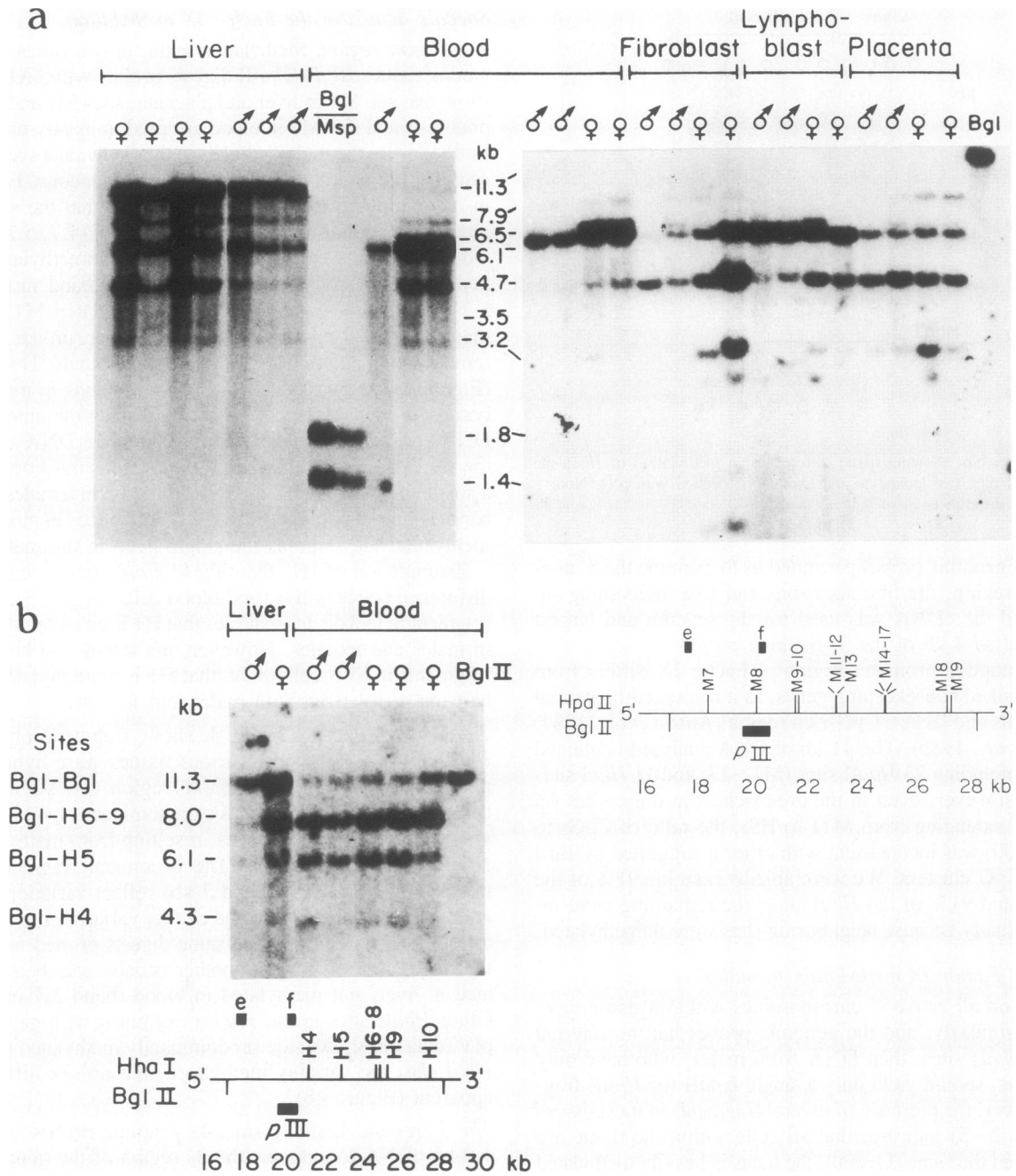

Fig. 4. Southern blot showing *Hpa*II sites M7−M19, and *Hha*I sites H4−H10 from males and females, and sex difference in methylation of M11−12. **(a)** *Bgl*II/*Hpa*II digests of liver, leukocyte, fibroblast, lymphoblast and placenta DNA probed with pIII. **(b)** *Bgl*II/*Hha*I digests of liver and leukocyte DNA probed with pIII.

females. Site M1, unmethylated in all male blood cells, is methylated in a small number of female blood cells (Figure 2b). The more striking sex difference, which involves sites M11−12, is also most prominent in DNA of leukocyte origin. Figure 4a shows that M11−12 are consistently cleaved in males, producing a 6.1-kb band in the *Bgl*II/*Hpa*II digests. Yet, these sites are methylated in at least 50% of the chromosomes in female samples, based on the presence of 6.5-, 7.9- and 11.3-kb bands (consult Figure 3 for origin of fragments). In liver, the sex difference is much less prominent: the 6.5-kb band, absent in males, is weak in females. In other tissues, the sex difference is inconsistent; in DNA from fibroblasts, lymphoblasts and placenta, the 6.5-kb band is present in some males, and absent in some females (Figure 4a).

To determine the extent of sex differences, we examined additional *Hpa*II sites and the six *Hha*I sites in this region: *Bgl*II/*Hha*I digests from liver and blood, probed with pIII (Figure 4b), show that sites H4−10 are methylated similarly in both sexes. *Hpa*II sites M14−19, inaccessible when M11−12 are unmethylated, can be examined in *Bgl*II/*Hpa*II digests probed with pIV (see Figure 1); these sites are methylated with no detectable sex differences (data not shown).

Methylation in expressing tissue. The unique feature is that, in contrast to blood and cultured cells, *Hpa*II sites M3 and M6, and *Hha*I site H1 are unmethylated in a significant population of liver cells (Figures 2, 3 and 5). The intensity of the relevant *Hpa*II/*Bgl*II fragments is greater in DNA from females than

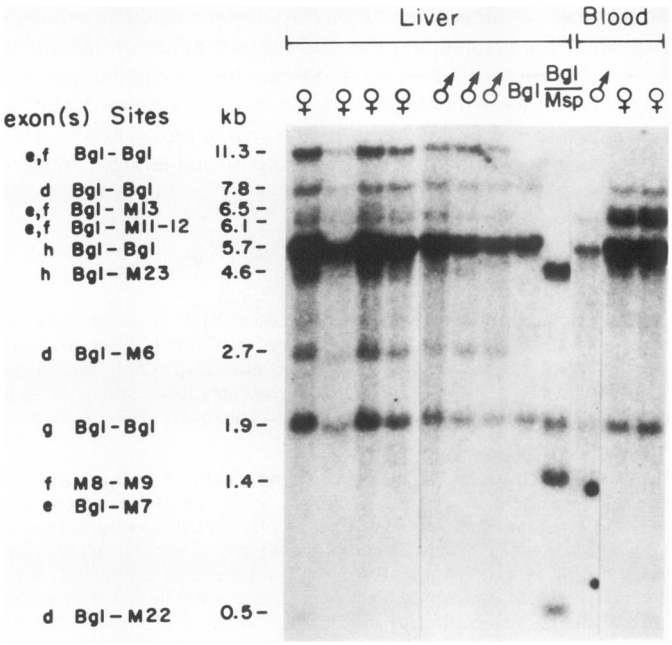

Fig. 5. Southern blot showing that Factor IX is extensively methylated *Bgl*II/*Hpa*II digests from liver and leukocytes, probed with cDNA. See Figure 1 for location of methyl-sensitive sites. The origin of the various fragments indicated on the left is derived from studies with genomic probes. For technical reasons the 11.3-kb *Bgl*II fragment in lane 8 is not detectable in this blot.

males, so that these sites are unmethylated on both active and inactive X chromosomes.

Restriction fragment length polymorphisms at the Factor IX locus

A by-product of this analysis is the discovery of two common restriction fragment length polymorphisms in Blacks. Some individuals have an additional *Bam*HI site (marked with an asterisk in Figure 1). The polymorphic site results from a C to T transition at nucleotide −587 in the sequence reported by Yoshitake *et al.* (1985), creating a novel *Bam*HI site. In *Bam*HI/*Msp*I digests probed with pXI, the additional *Bam*HI site produces a 3.5-kb band that replaces the 4.1-kb band in hemizygous males, and is present in addition to the 4.1-kb band in heterozygous females (Figure 6). We observed the polymorphism in nine of the 18 chromosomes of Black origin, giving frequencies of 0.5 for either allele. The 3.5-kb variant was seen in only one of the 16 X chromosomes of Caucasian origin. This 5′ polymorphism can be detected in *Bam*HI/*Msp*I digests probed with cDNA (Figure 6). In these digests we noted an *Msp*I polymorphism that also occurs frequently in Blacks: seven of 15 chromosomes yield a 6.3-kb variant band, reflecting the absence of an *Msp*I site. In contrast, none of the 16 chromosomes from Caucasians had the polymorphism. Based on the size of *Msp*I fragments (6.3, 2.3 kb) and *Eco*RI/*Msp*I fragments (4.8, 2.2 kb), the polymorphic site should be at nucleotide 15 948, in the sequence of Yoshitake *et al.* (1985). This polymorphism may be the same as that reported by Camerino *et al.* (1985), as a 5.8-kb variant with a frequency of 0.2 in Europeans. In any event, the presence of the 6.3-kb variant was independent of the 5′ 3.5-kb variant (Figure 6). Along with those reported previously (Gianelli *et al.*, 1984; Camerino *et al.*, 1984; Winship *et al.*, 1984) the two polymorphisms that occur so frequently in Blacks should provide useful markers for pre-natal diagnosis of hemophilia B in this population.

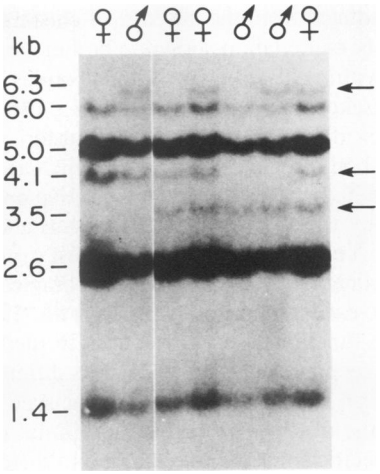

Fig. 6. Autoradiograph of *Bam*HI/*Msp*I digests probed with the cDNA, showing 4.1- and 3.5-kb (*Bam*HI polymorphism) and 6.3-kb (*Msp*I polymorphism) fragments. The 2.3-kb *Msp*I fragment is obscured in this blot by the intense 2.6-kb non-polymorphic *Msp*I band.

Discussion

Methylation is tissue specific, but unrelated to expression of gene

Based on evidence from *in vitro* assays showing that methylation of certain CpGs inhibits transcription, and observations that some sites are relatively less methylated in transcribing tissues, DNA methylation is considered to have a role in the transcription of tissue-specific genes. However, the abnormal cellular environment in which *in vitro* assays are carried out, and the considerable background methylation and tissue variation *in vivo*, make it difficult to distinguish methylation patterns that are imprints of more generalized tissue differentiation from those directly involved in expression of the relevant locus. To determine if tissue-specific patterns are in fact functional, it helps to compare active with inactive gene within the same cell; demethylation events occurring on the inactive as well as the active X chromosome cannot influence function directly, and may not even be necessary for expression of the locus.

If CpG methylation had the same role in the transcription of Factor IX as for X-linked housekeeping genes, we expect to find sex differences in methylation of the gene in liver cells where the locus is expressed. In fact, sex differences occurring in this tissue originate from only a minor population of cells, and the sex difference is considerably more prominent in blood. Moreover, methylation of active and inactive genes does not differ at the sites where the striking tissue difference occurs. Therefore, it seems that hypomethylation of sites M3, M6 and H1 in liver relative to other tissues, if related to transcriptional activity, is not sufficient to induce activity of the locus on the inactive X chromosome. The similar methylation of active and inactive Factor IX genes within the same cells in expressing tissue supports the arguments put forth by Bird (1986), that methylation within genes that are expressed in a tissue-specific manner has little relevance to expression. It is unlikely that our assay has missed 'the critical' CpGs, that might be inaccessible to restriction enzyme analysis as genomic sequencing has not revealed such critical sites in the maize alcohol dehydrogenase gene to explain tissue-specific expression (Nick *et al.*, 1986). Furthermore, the restriction enzyme assay has been able to detect functional differences in methylation of housekeeping genes (see Introduction).

Functional DNA methylation is localized to X-linked CpG clusters

The lack of functional differences in Factor IX methylation might

have been predicted from the absence of clustered CpGs at this locus. The only methylation shown to be functional in X-linked genes is that within clustered CpG dinucleotides that are associated with housekeeping genes (Wolf *et al.*, 1984a, b; Riggs *et al.*, 1985). Sex differences outside the cluster occur most frequently in the body of the gene (at sites which are unmethylated on the inactive X, but methylated on the active one); such differences are highly variable between individuals and tissues (Wolf *et al.*, 1984a; Yen *et al.*, 1984) and persist even when the inactive late replicating X chromosome is completely reactivated and switches to early replication (Migeon *et al.*, 1986). It is likely that these non-functional sex differences in methylation of the body of housekeeping genes, as well as sex differences in Factor IX that were seen in some tissues, reflect the state of the chromosome at the time of *de novo* methylation in the relevant tissue; such tissue-specific imprints may be due to differential accessibility of methylatable sites in the gene on active and inactive chromosomes, resulting from global differences in configuration and/or behavior of the homologous X chromosome when *de novo* methylation occurred.

Conceivably, the only methylation capable of affecting expression of a locus *in vivo* is that occurring in CpG clusters. Estimates for the frequency of such clusters suggest that they occur throughout the genome, perhaps as frequently as every 100 kb (Bird *et al.*, 1985). That the most common location reported for clustered CpGs is in promoter regions of housekeeping genes and growth factors may reflect biased ascertainment. Because the analysis has been limited, we do not know if CpG clusters are infrequent in tissue-specific genes. Moreover, as only the gene and immediate flanking sequences were examined, our analysis of Factor IX does not preclude CpG clusters in the vicinity that might affect the function of the locus from a distance.

Dosage compensation without consistent differences in methylation within the Factor IX locus

As we observed no significant sex differences in hepatocytes, methylation within the locus cannot be responsible for maintaining dosage compensation of Factor IX. Nor have our studies identified other features of the locus (i.e. sequence rearrangements, insertions) to explain the differential expression of the homologous genes in liver cells. Therefore, X-linked tissue-specific genes seem to differ from X-linked housekeeping genes with respect to the role of methylation in maintaining differential transcriptional activity. Perhaps only housekeeping genes use methylation of 5′ CpG clusters to maintain silence of the locus on the inactive X, whereas this function in other genes may be mediated by tissue-specific transcription factors; regulatory signals in the promoter region clearly differ for these classes of genes. On the other hand, perhaps differences are more apparent than real; methylation might also affect the expression of tissue-specific genes, but the effect could be mediated more indirectly, through the agency of sequences outside the locus that can influence transcriptional activity of the locus. One candidate regulatory sequence is the ubiquitous CpG cluster. These clusters are found in vertebrates (Cooper *et al.*, 1983) and have been conserved in mammals even though these genomes are relatively deficient in CG dinucleotides (Nusinov, 1981). That clustered CpGs, unlike isolated ones, are generally hypomethylated and therefore not subject to deamination may explain this paradox. Their abundance suggests that these clusters serve an important and perhaps complex function in eukaryotic cells. Pre-dating the evolution of mammalian X dosage compensation, clustered CpGs may have taken on additional functions during evolution of eutherian mammals, possibly serving to regulate domains of chromatin. Methylation of these clusters along the X chromosome could be a unique adaptation for maintaining the inactivity of genes on the silent X. Ascertainment of the location of these clusters and the relationship between their methylation status and expression of nearby tissue-specific genes should clarify the role of these clusters (and DNA methylation) in transcriptional activity and dosage compensation.

Materials and methods

DNA sources and extraction

DNA was purified from blood leukocytes, placenta, cultured fibroblasts, lymphoblasts (Wolf *et al.*, 1980) and liver (Kunkle *et al.*, 1978) from Black and Caucasian males and females. The liver samples were obtained up to 24 h post-mortem from subjects without liver disease (based on gross appearance of the specimen and clinical history) and were frozen and ground to a fine powder before extraction.

Probes

The Factor IX cDNA and Factor IX genomic probes III, IV, VIII and XI have been described elsewhere (Anson *et al.*, 1984). Briefly, the cDNA and genomic probes have been inserted into the plasmid pAT153/PVUII/8, except for probe XI, which is in pATX, a modification of this plasmid. The cDNA probe, containing all eight exons (a–h) and ~500 bp of 3′ non-translated sequence, is 1.9 kb in length. Probe III contains the sixth exon (f) and ~900 bp of intron, probe VIII contains the fourth exon (d) and 2.5 kb of intron. The 5′ probe, XI, includes 200 bp of presumed promoter sequence, the first exon (a) and ~600 bp of the first intron. Probe IV contains 1.4 kb of non-coding sequence located in intron seven. A 780-bp segment of probe IV, excised with *Pst*I/*Hha*I was used to examine *Hpa*II sites in intron 6 (lying between f and g).

Probes were labelled with ^{32}P by nick translation (Rigby *et al.*, 1977), or oligoradiolabelling (Feinberg and Vogelstein, 1984).

DNA methylation analysis

To identify the methylation status of CpG sites in the Factor IX gene, DNA was cleaved with *Bgl*II, *Pst*I, *Hin*dIII, *Bam*HI or *Eco*RI to yield large fragments, and then cut with *Hpa*II or *Hha*I, methyl-sensitive restriction enzymes that contain CpG in their recognition sequences. *Msp*I digests were used to control for polymorphic restriction sites.

DNA digestion and Southern analysis

DNA fragments, obtained by digestion with various restriction enzymes, were separated by gel electrophoresis in 0.9% agarose and transferred to nitrocellulose filters (Southern, 1975) or Gene Screen Plus (DuPont – NEN, Reed and Mann, 1985) and hybridized with labelled probes in the presence of heparin sulfate (Singh and Jones, 1984).

Acknowledgements

We thank George G.Brownlee and his colleagues for generously providing the Factor IX probes. This work was supported by NIH grant HD05465, and an NIH postdoctoral fellowship to D.C.K.

References

Anson,D.S., Choo,K.H., Rees,D.J.G., Gianelli,F., Gould,K., Huddleston,J.A. and Brownlee,G.G. (1984) *EMBO J.*, **3**, 1053–1060.

Bird,A.P. (1986) *Nature*, in press.

Bird,A.P., Taggart,M., Frommer,M., Miller,O.J. and Macleod,D. (1985) *Cell*, **40**, 91–99.

Boyd,Y., Buckle,V.J., Munro,E.A., Choo,K.H., Migeon,B.R. and Craig,I.W. (1984) *Ann. Hum. Genet.*, **48**, 145–152.

Buckle,V. (1986) D.Phil. Thesis, Oxford.

Bussliner,M., Hurst,J. and Flavell,R.A. (1983) *Cell*, **34**, 197–206.

Camerino,G., Grzeschik,K.H., Jaye,M., De La Salle,H., Tolstoshev,P., Lecocq, J.P., Heilig,R. and Mandel,J.L. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 498–502.

Camerino,G., Oberle,I., Drayna,D. and Mandel,J.L. (1985) *J. Hum. Genet.*, **71**, 79–81.

Cooper,D.N., Taggart,M.H. and Bird,A.P. (1983) *Nucleic Acids Res.*, **11**, 647–658.

Doerfler,W. (1983) *Annu. Rev. Biochem.*, **52**, 93–124.

Feinberg,A.P. and Vogelstein,B. (1984) *Anal. Biochem.*, **137**, 266–267.

Gianelli,F., Choo,K.H., Winship,P.R., Rizza,C.R., Anson,D.S., Rees,D.J.G., Ferrari,N. and Brownlee,G.G. (1984) *Lancet*, **1**, 239–241.

Holliday,R. and Pugh,J.E. (1975) *Science*, **187**, 226–232.

Jones,P.A., Taylor,S.M., Mohandas,T. and Shapiro,L.J. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 1215–1219.

Kasper,C.K., Osterud,B., Minami,J.Y., Shonick,W. and Rapaport,S.I. (1977) *Blood*, **50**, 351–366.
Kunkle,L.M., Miller,O.J., Breg,W.R., Jones,H.W. and Rory,J.M. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 1245–1249.
Lindsay,S., Monk,M., Holliday,R., Huschtscha,L., Davies,K.E., Riggs,A.D. and Flavell,R.A. (1985) *Ann. Hum. Genet.*, **49**, 115–127.
Lyon,M.F. (1972) *Biol. Rev.*, **42**, 1–35.
Migeon,B.R., Wolf,S.F., Axelman,J., Kaslow,D.C. and Schmidt,M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3390–3394.
Migeon,B.R., Schmidt,M., Axelman,J. and Ruta Cullen,C. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2182–2186.
Monk,M. and Harper,M.I. (1979) *Nature*, **281**, 311–313.
Nick,H., Bowen,B., Ferl,R.J. and Gilbert,W. (1986) *Nature*, **319**, 243–246.
Nusinov,R. (1981) *J. Mol. Biol.*, **149**, 125–131.
Reed,K.C. and Mann,D.A. (1985) *Nucleic Acids Res.*, **13**, 7207–7221.
Rigby,P.W.J., Dieckmann,M., Rhodes,C. and Berg,P. (1977) *J. Mol. Biol.*, **113**, 237–251.
Riggs,A.D. (1975) *Cytogenet. Cell Genet.*, **14**, 9–25.
Riggs,A.D., Singer-Sam,J. and Keith,D.H. (1985) In Razin,A. and Cantoni,G.L. (eds), *Biology and Biochemistry of DNA Methylation*. Alan R.Liss, Inc., NY, pp. 211–222.
Sager,R. and Kitchin,R. (1975) *Science*, **189**, 426–433.
Shimada,T. and Nienhuis,A.W. (1985) *J. Biol. Chem.*, **260**, 2468–2474.
Singh,L. and Jones,K.W. (1984) *Nucleic Acids Res.*, **12**, 5627–5637.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503–517.
Stein,R., Sciaky-Gallili,N., Razin,A. and Cedar,H. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2422–2425.
Thompson,A.R. (1977) *Thromb. Res.*, **11**, 193–203.
Toniolo,D., D'Urso,M., Martini,G., Persico,M., Tufanov,V., Battistuzzi,G. and Luzzatto,L. (1984) *EMBO J.*, **3**, 1987–1995.
Winship,P.R., Anson,D.S., Rizza,C.R. and Brownlee,G.G. (1984) *Nucleic Acids Res.*, **12**, 8861–8872.
Wion,K.L., Kelly,D., Summerfield,J.A., Tuddenham,E.G.D. and Lawn,R.M. (1985) *Nature*, **317**, 726–729.
Wolf,S.F. and Migeon,B.R. (1982) *Nature*, **295**, 667–671.
Wolf,S.F. and Migeon,B.R. (1985) *Nature*, **314**, 467–469.
Wolf,S.F., Mareni,C. and Migeon,B.R. (1980) *Cell*, **21**, 95–102.
Wolf,S.F., Jolly,D.J., Lunnen,K.D., Friedmann,T. and Migeon,B.R. (1984a) *Proc. Natl. Acad. Sci. USA*, **81**, 2806–2810.
Wolf,S.F., Dintzis,S., Toniolo,D., Persico,G., Lunnen,K., Axelman,J. and Migeon,B.R. (1984b) *Nucleic Acids Res.*, **12**, 9333–9348.
Yen,P.H., Patel,P., Chinault,A.C., Mohandas,T. and Shapiro,L.J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1759–1763.
Yoshitake,S., Schach,B.G., Foster,D.C., Davie,E.W. and Kurachi,K. (1985) *Biochemistry*, **24**, 3736–3750.

Received on 13 May 1986; revised on 23 June 1986