The EMBO Journal vol.3 no.9 pp.1987–1995, 1984

Specific methylation pattern at the 3′ end of the human housekeeping gene for glucose 6-phosphate dehydrogenase

Daniela Toniolo, Michele D'Urso, Giuseppe Martini, Maria Persico, Vincenzo Tufano, Giorgio Battistuzzi[1] and Lucio Luzzatto[1]

International Institute of Genetics and Biophysics, CNR, Via Marconi 10, 80125 Naples, Italy, and [1]Department of Haematology, Hammersmith Hospital, London W12 OHS, UK

Communicated by L.Luzzatto

During detailed restriction enzyme mapping of the human X-linked gene *Gd*, specifying the enzyme glucose 6-phosphate dehydrogenase (G6PD), we have observed the presence, over a 14-kb DNA region spanning across the 3′ end of the G6PD transcript, of a large number of methylatable sites. These include 60 *Hpa*II sites, 13 *Sma*I sites, 22 *Ava*I sites and 46 *Hha*I sites. In male leukocyte DNA the majority of *Hpa*II sites are resistant to digestion, indicating that they are in the C^{m5}CGG form. However, a few sites are found reproducibly unmethylated in 24 samples analyzed. By double and triple digestions we have mapped five unmethylated sites, four of which are within the gene transcript and one distal to the end of transcription. We have also identified a number of sites which are fully methylated, whereas for others the methylation status could not be positively assessed. Thus, in a housekeeping gene expressed in leukocytes, the 3′ end is extensively methylated, but some specific sites are unmethylated. In female leukocyte DNA, we found that all sites methylated in males were also methylated. However, of the five sites that are unmethylated in males two are partly methylated in females. This additional site-specific methylation involves ~50% of the female leukocyte DNA, and we show evidence that it is associated with the inactive X-chromosome.

Key words: glucose 6-phosphate dehydrogenase/X-chromosome/*Gd* gene/3′ end/methylation

Introduction

DNA methylation correlates with the expression of a variety of differentiation genes including those for globins (McGhee and Ginder, 1979; Van der Ploeg and Flavell, 1980), ovalbumin, conalbumin and ovomucoid (Mandel and Chambon, 1979; Kuo *et al.*, 1979). The significance of the correlation observed is not clear, since in some cases individual cytosine residues in a particular gene are less methylated in cells expressing that gene, but in other cases the reverse is true (review by Ehrlich and Wang, 1981; Cooper, 1983). DNA methylation has been implicated in the phenomenon of X-chromosome inactivation since treatment of cells with the cytidine analogue 5-azacytidine (azaC) that inhibits C-methylation can induce reactivation of some genes on the inactive X-chromosome (Venolia *et al.*, 1982). However, the action of azaC is not limited to X-linked genes and it may not be limited to affecting methylation.

The availability of specific probes for the human X-linked enzyme, glucose 6-phosphate dehydrogenase (G6PD) (Persico *et al.*, 1981), has enabled us to test whether individual cytosine residues are randomly or non-randomly methylated in the DNA that codes for a household gene, as well as to determine directly whether methylation of specific sites correlates with the presence of an inactive chromosome in female DNA. The 3′ end of the G6PD gene is very rich in CpG sequences, and the majority of these are regularly in the ^{m5}CpG form in DNA from normal human leukocytes. However, five specific cytosine residues (the same in all unrelated subjects thus far tested) are regularly and fully unmethylated in male DNA and two of these residues are 50% methylated in female DNA.

Results

The isolation of G6PD-specific cDNA clones has been previously reported (Persico *et al.*, 1981) and the use of clone pGD6405 to screen a genomic human DNA library will be published in detail elsewhere. A restriction map of the 3′ end of the *Gd* gene has been constructed (Figure 1). For the sake of clarity, we shall use as origin of the linear map the 3′ end of the G6PD mRNA, with positive numbers indicating distances 3′ from the origin and negative numbers indicating distances 5′ from the origin, according to the direction of transcription.

In eukaryotes most of the DNA methylation is found in the sequence ^{m5}CpG. We have mapped, on the G6PD genomic clones, those CpG sequences that can be recognized by restriction enzymes, such as *Hpa*II, *Hha*I, *Sma*I and *Ava*I. Digestion of the available genomic clones with the restriction enzyme *Hpa*II identified 60 sites over a 14-kb stretch of DNA, spanning from −7 to +7 kb (see Figure 1). Twenty of these are located outside the transcribed sequence. The remaining 40 sites are within the primary transcript and their positions have been confirmed and precisely located by sequencing. For ease of identification, these sites have been numbered sequentially from H 1 to H 60 in the 5′→3′ direction. *Sma*I sites (S 1 – S 13), *Ava*I sites (*Ava* 1 – *Ava* 22) and *Hha*I sites (*Hha* 1 – *Hha* 46) have been similarly mapped (Figure 1).

Methylation status of HpaII sites in male DNA

Because *Hpa*II cleaves a CCGG sequence only if the second C is not methylated (Bird and Southern, 1978), we have used it to test the methylation of genomic DNA from human male leukocytes. Many of these *Hpa*II sites are located within the DNA homologous to the two probes, pGD3 and pGD1.4 (see Figure 1). Nevertheless, on Southern blots the former yields only a doublet of bands at 3.2 and 3.0 kb (ratio ~9:1; see Figure 2, lane 1), and the latter only a single major band at 6.8 kb (see Figure 3, lane 3), indicating that most of the intervening sites are methylated, but some individual sites are largely unmethylated. To map the unmethylated sites, further analyses by Southern blots were carried out on DNA samples subjected to a number of double and triple digestions.

Starting from the results obtained with probe pGD3, the 3.2-kb band observed could arise from the cleavage of the

© IRL Press Limited, Oxford, England.

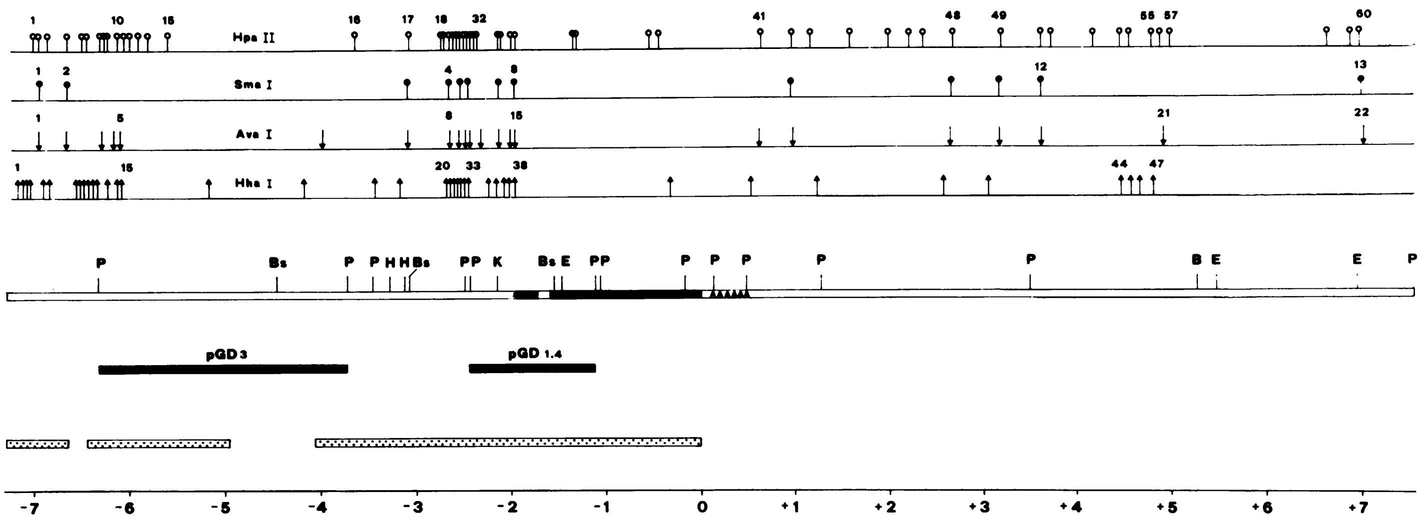

Fig. 1. Diagram of 3′ end of human *Gd* gene. From bottom to top. The scale on the abscissa is marked in kilobases, with the origin at the end of the transcript. The dotted segments indicate sequenced DNA. The dark segments indicate the two probes used in this work. The double line is genomic DNA, which is blacked in for known exons. The length of the intron on the 5′ side of −1.9 kb is not known. On this double line a partial restriction map is shown, with the following symbols: P: *Pst*I; Bs: *Bst*EII; H: *Hind*III; K: *Kpn*I; R: *Eco*RI; Bam: *Bam*HI. The four lines above give a detailed map of the four methylation-sensitive restriction enzymes, as shown. The *Hpa*II sites have been numbered from left to right, in order to ease reference to them in the text.

following pairs of *Hpa*II(H) sites: H 4 and H 16; H 10 and H 17; H 13−15 and H 18 (or one of its immediate neighbours). In a *Hind*III-*Hpa*II double digest the 3′ end of the 2.6-kb fragment obtained (Table I and Figure 2, lane 2) is identified by our restriction map with the *Hind*III site located at −3.3, thus the 5′ end of the fragment must be site H 14. This conclusion was cross-checked by using the enzyme *Bst*EII (see lane 17 in Figure 2). Here a 5.2-kb fragment arises from cleavages of a *Bst*EII site at −9.7 kb (not shown in Figure 1) and of the *Bst*EII site at −4.5 kb; and a 1.5-kb fragment arises from cleavages at the latter site and the site located at −3.0 kb on the map. When *Hpa*II is added (Figure 3, lane 18), the 1.5-kb band persists, confirming that *Hpa*II does not cleave H 16: indeed, the signal becomes stronger because of superimposition of another fragment of the same size spanning from the *Bst*EII site at −4.5 and H 14. The additional 1.3-kb fragment agrees with the distance between the same *Bst*EII site and H 15. Finally, a *Pst*I-*Hpa*II double digest (Figure 2, lane 6) yields a fragment of 2.15 kb which must arise from cleavage of the *Pst*I site at −3.8 and site H 14. Whenever a minor band is seen in these double digests (e.g., lanes 6 and 18 in Figure 2), it fits with cleavage at H 15. We conclude that the 3.2- and 3.0-kb fragments originally observed with *Hpa*II alone originate from cleavage at site H 18 (or one of its near 3′ neighbours) and at sites H 14−H 15. This means that H 18 and H 14 are unmethylated, H 16 and H 17 are fully methylated, and H 15 is partially unmethylated.

The boundaries of the 6.8-kb *Hpa*II fragment seen with probe pGD1.4 can also be similarly mapped. The 5′ end of the fragment was identified with H 36 by the finding of a 0.9-kb fragment in a *Pst*I-*Hpa*II double digest (Figure 3, lane 4: the 3′ end of this fragment is the *Pst*I site at −1.2), and this was confirmed by a *Pst*I-*Eco*RI-*Hpa*II triple digest (see Table I and Figure 3, lane 2). Therefore, the 3′ end of the 6.8-kb *Hpa*II fragment must be one of the sites H 55−H 57. All sites between H 36 and H 55 must be methylated. Results from additional double digestions are consistent with this conclusion (Table I and Figure 3).

Table I. Summary of results of single, double and triple digestions with *Hpa*II and other restriction enzymes

Restriction enzymes	Size of fragments, kb			
	Probed with pGD3		Probed with pGD1.4	
	Alone	+ *Hpa*II	Alone	+ *Hpa*II
None		3.2; 3.0		6.8
*Hind*III	4.2	2.6	12	
*Pst*I	2.75	2.15	1.4	0.9
*Bst*EII	5.7; 1.5	1.5; 1.3	6.5; 1.5	6.5
*Bam*HI	18	3.2	18	6.8
*Kpn*I	17	3.2; 3.0	11	6.8
*Pst*I + *Eco*RI			1.0; 0.4	0.45; 0.4

Analysis of the sites from H 19 to H 35 is difficult because they are so close together, and they are lined by the two unmethylated sites H 18 (or one of its closest neighbours) and H 36. The distance between these two sites is ~0.75 kb, but no fragment of such size is seen when *Hpa*II digests are probed with pGD1.4 (Figure 3, lanes 3 and 11), suggesting that smaller fragments are formed. Moreover, a 0.5-kb fragment is not found in an *Hpa*II-*Pst*I double digest (Figure 3, lane 4: absence of this band is not due to technical reasons, since even smaller fragments are seen on the same blot, lane 2). This confirms that not all sites H 19−H 32 are fully methylated. One possibility is that several of them are partially unmethylated, and the resulting fragments are both small and heterogeneous in size, and therefore below the threshold for detection by Southern blotting. A similar situation applies to sites H 5−H 13.

The detailed analysis we have reported above was carried out on DNA from three male subjects. Samples from four additional males were digested with *Hpa*II alone and in combination with *Kpn*I, *Pst*I, *Bst*EII and probed with pGD3; and samples from 21 additional males were digested with *Hpa*II alone and in combination with *Kpn*I and probed with pGD1.4. All these experiments display reproducibly the pattern we have described (see Figures 4 and 6).

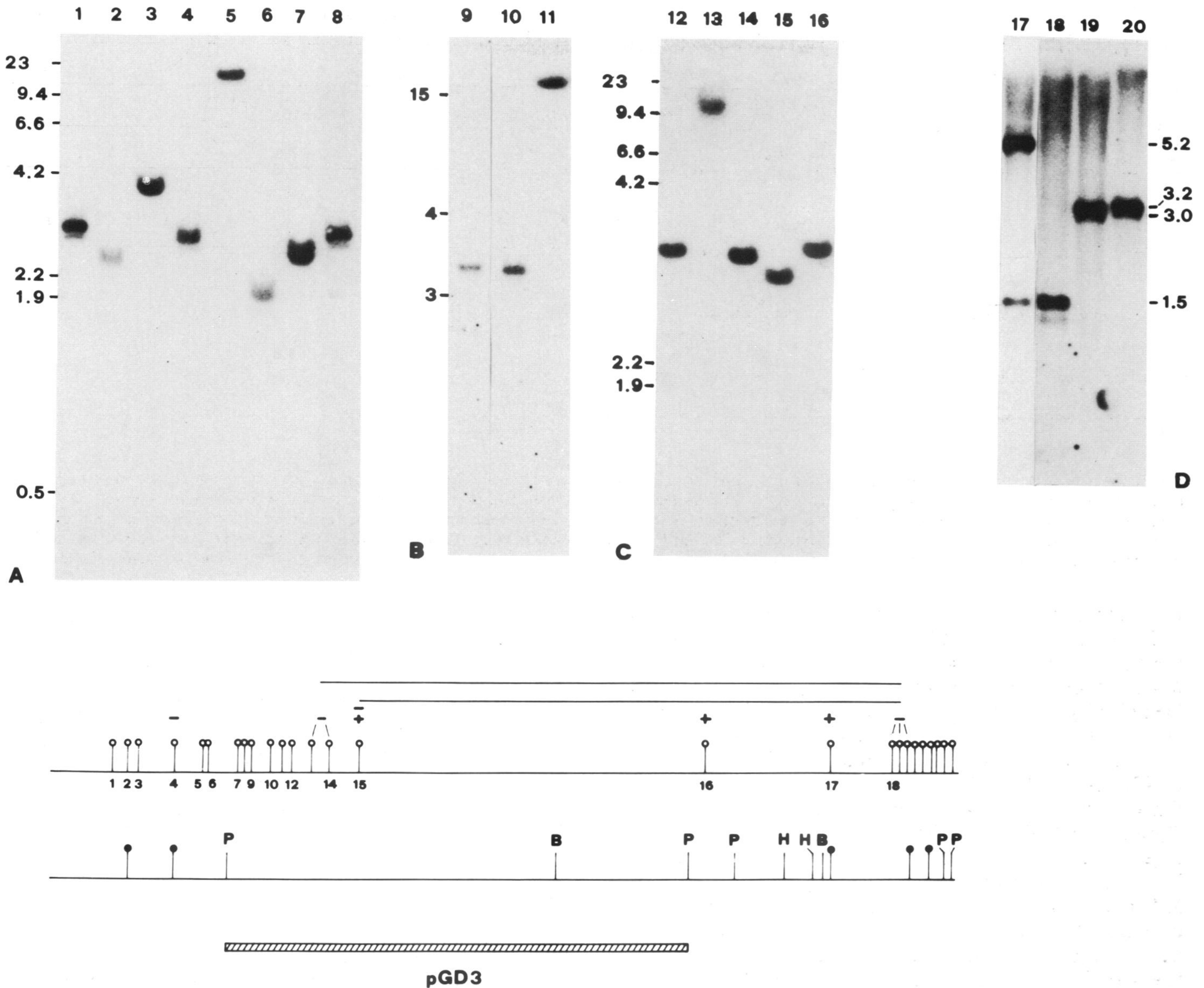

Fig. 2. Southern blot analysis of digests of human male leukocyte DNA probed with pGD3. Figures on the left indicate the position of appropriate size markers. Lanes are numbered at the top, and they contain DNA digested as follows: **1, 8, 9, 20:** *Hpa*II; **2:** *Hpa*II + *Hin*dIII; **3, 16:** *Hin*dIII; **4:** *Hpa*II + *Bam*HI; **5:** *Bam*HI; **6:** *Hpa*II + *Pst*I; **7:** *Pst*I; **10:** *Hpa*II + *Kpn*I; **11, 13:** *Kpn*I; **12:** *Sma*I; **14:** *Sma*I + *Kpn*I; **15:** *Sma*I + *Hin*dIII; **17:** *Bst*EII; **18:** *Hpa*II + *Bst*EII; **19:** *Hpa*II + *Kpn*I. The diagram at the bottom is an expanded version of the portion of the restriction map depicted in Figure 1, pertaining to the region that surrounds the probe pGD3. The vertical bars headed by full circles are *Sma*I sites, and those headed by empty circles are *Hpa*II sites. P indicates a *Pst*I site and B a *Bst*EII site. A + symbol indicates that the site is fully methylated, and a − symbol that the site is mostly unmethylated. For reasons explained in the text, we infer that site H 15 is partially methylated, H 14, more likely than H 13, is unmethylated, and the unmethylated site near H 18 is most likely H 20. The thin lines above the *Hpa*II sites represent the span of the 3.2- and 3.0-kb restriction fragments seen in *Hpa*II digests (e.g., in **lane 1**).

Other methylated sites in male DNA

In the *Gd* DNA region we have analyzed, there are numerous sites for other methylation-sensitive restriction enzymes; namely 13 *Sma*I sites, 22 *Ava*I sites and 46 *Hha*I sites. Here again some individual sites are uniformly methylated, as inferred from the fact that digestions with these enzymes yield restriction fragments which cannot arise from cleavage of immediately adjacent sites (see Table II; Figure 2, lanes 12−14; Figure 3, lanes 5−8). Of the bands seen with probe pGD3, that obtained with *Ava*I would be compatible with cleavage of sites *Ava* 5 and *Ava* 8 (corresponding approximately to H 10 and H 20); and that obtained with *Hha*I would be compatible with cleavage of sites *Hha* 15 and *Hha* 20. *Hha*I digests probed with pGD1.4 show a 6.8-kb fragment only, indicating that *Hha*I sites from 38 to 45 are all methylated. Since all *Sma*I sites (S) are also *Hpa*II sites, this enzyme was used to confirm some of our previous data. Specifically, in *Sma*I digests hybridized with pGD3, a 4-kb fragment is seen. This could be either from S 1 to S 3 or from S 2 to S 4. However, in a *Hin*dIII-*Sma*I double digest the size is reduced to 3.4 kb. Therefore, we conclude that sites S 2 and S 4 are unmethylated. These are identical to sites H 4 and H 20. Thus, H 20 is the near neighbour to H 18 which is cleaved by *Hpa*II (H 4 has not been tested by *Hpa*II because it lies outside the 3-kb fragment H 15−H 20). Unmethylation of site H 36, which is also a *Sma*I site, was also confirmed by the finding that *Sma*I-*Pst*I double digests both yield an identical 0.9-kb fragment (see lanes 4 and 8 in Figure 3). All *Sma*I sites distal to the transcript are methylated (see Figure 3, lane 7).

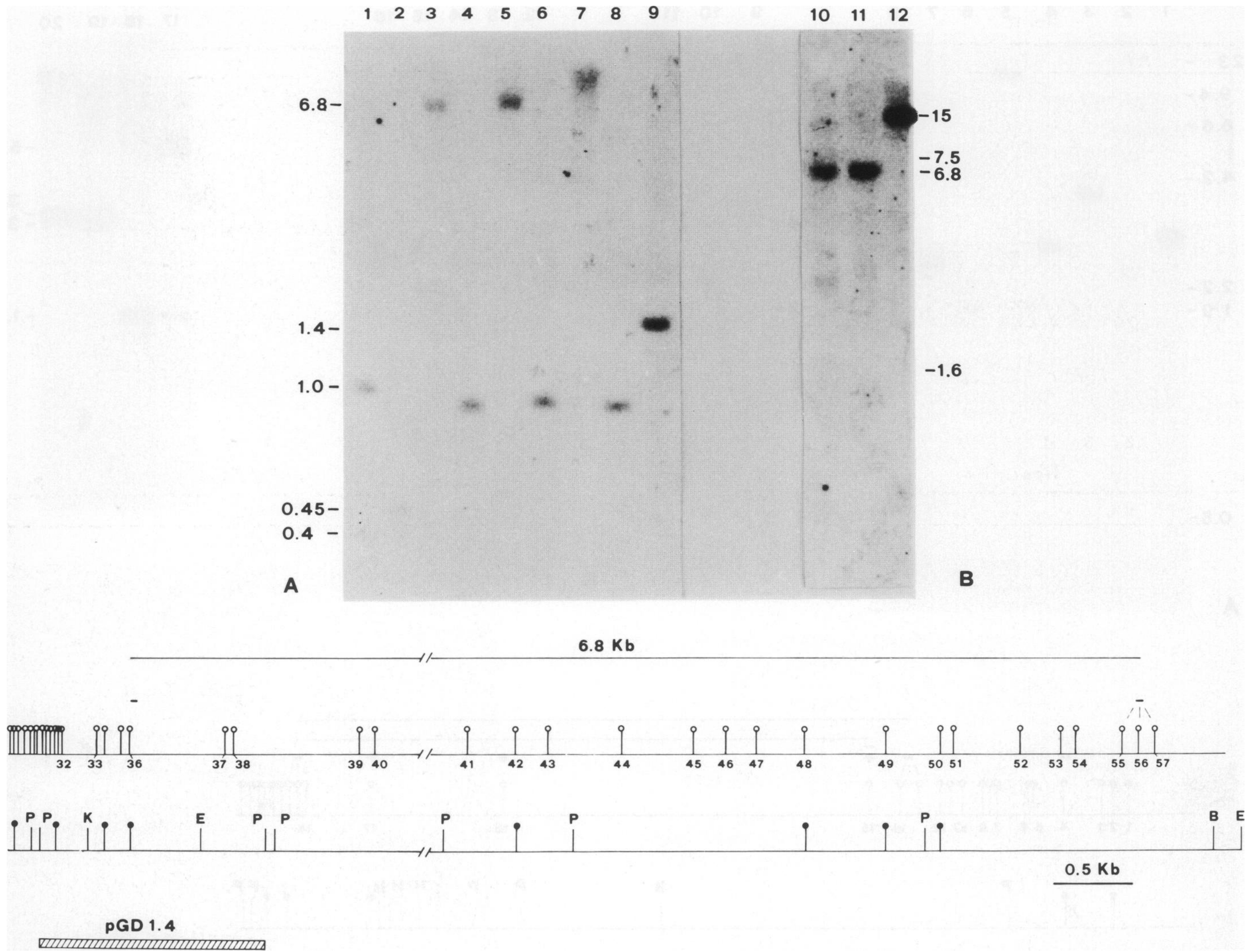

Fig. 3. Southern blot analysis of digests of human male leukocyte DNA probed with pGD1.4. Gel **A** was 1.5% agarose, while gel **B** was 0.8% agarose. The various lanes contain DNA digested as follows: **1:** *Pst*I + *Eco*RI; **2:** *Hpa*II + *Pst*I + *Eco*RI; **3, 11:** *Hpa*II; **4:** *Hpa*II + *Pst*I; **5:** *Hha*I; **6:** *Hha*I + *Pst*I; **7:** *Sma*I; **8:** *Sma*I + *Pst*I; **9:** *Pst*I; **10:** *Hpa*II + *Kpn*I; **12:** *Kpn*I. Symbols are as in Figure 2; in addition, K stands for *Kpn*I and E for *Eco*RI.

Table II. Summary of results of digestions with other methylation-sensitive restriction enzymes

Restriction enzyme	Size of fragments, kb							
	Probed with pGD3				Probed with pGD1.4			
	Alone	+ *Sma*I	+ *Hha*I	+ *Ava*I	Alone	+ *Sma*I	+ *Hha*I	+ *Ava*I
None: expected		3.6	0.8; 1.2	2.2; 0.9		0.5; 3.0	1.6	
observed		4.0	3.4	3.4		8.9	6.8	
*Pst*I	2.75	2.75	2.4; 2.5	2.4; 2.5	1.4	0.9	0.9	0.9
*Bam*HI	18				18		6.8	
*Hind*III	4.2	3.4			12			

Methylation status of HpaII sites in female leukocyte DNA

To find out if and how the complex methylation pattern we have observed is affected by X-inactivation, Southern blots were prepared from several male and female DNA samples run side by side (Figure 4). All restriction fragments seen in male DNA are also present in female DNA. The existence of a fragment larger than 3.2-kb (Figure 4B) revealed by pGD3 indicates that either one or both of sites H 14 and H 18 are methylated in the inactive X. From the finding that an extra fragment is still present after *Pst*I-*Hpa*II double digestion we infer that H 14 must be methylated, whereas H 18 may or may not be. These results have been confirmed in *Hind*III-*Hpa*II double digests. By the same reasoning we infer from

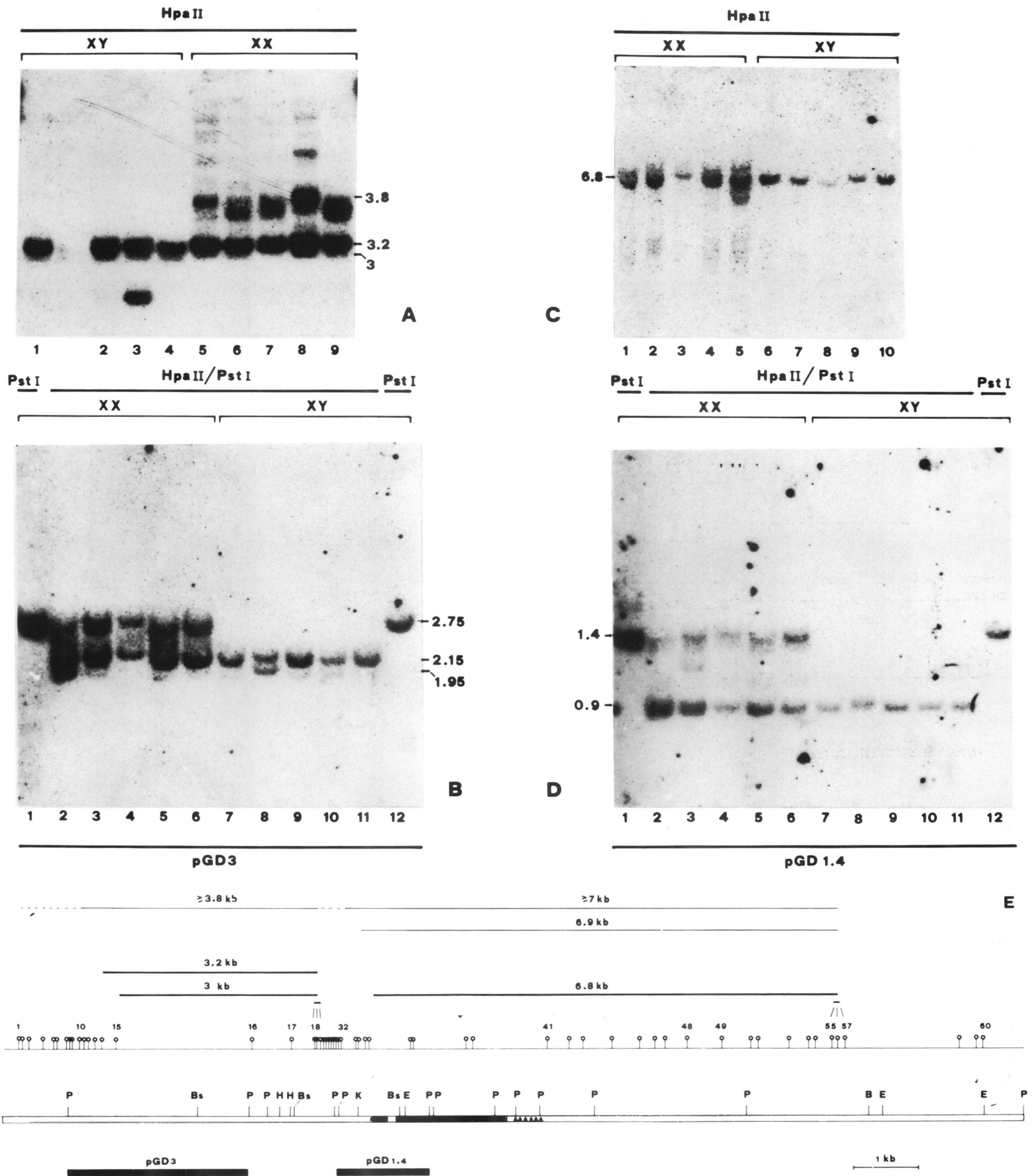

Fig. 4. Southern blot analysis of DNA from leukocytes from several male and female subjects. Restriction enzymes used are indicated on top of each panel. The probe used was pGD3 for **panels A** and **B**, and pGD1.4 for **panels C** and **D**. Agarose gel concentration was 0.8% for **panels A** and **C**, and 1.0% for **panels B** and **D**. Symbols in **panel E** as in Figure 2. Dark lines: bands common to male and female DNA; faint lines: bands found only in female DNA.

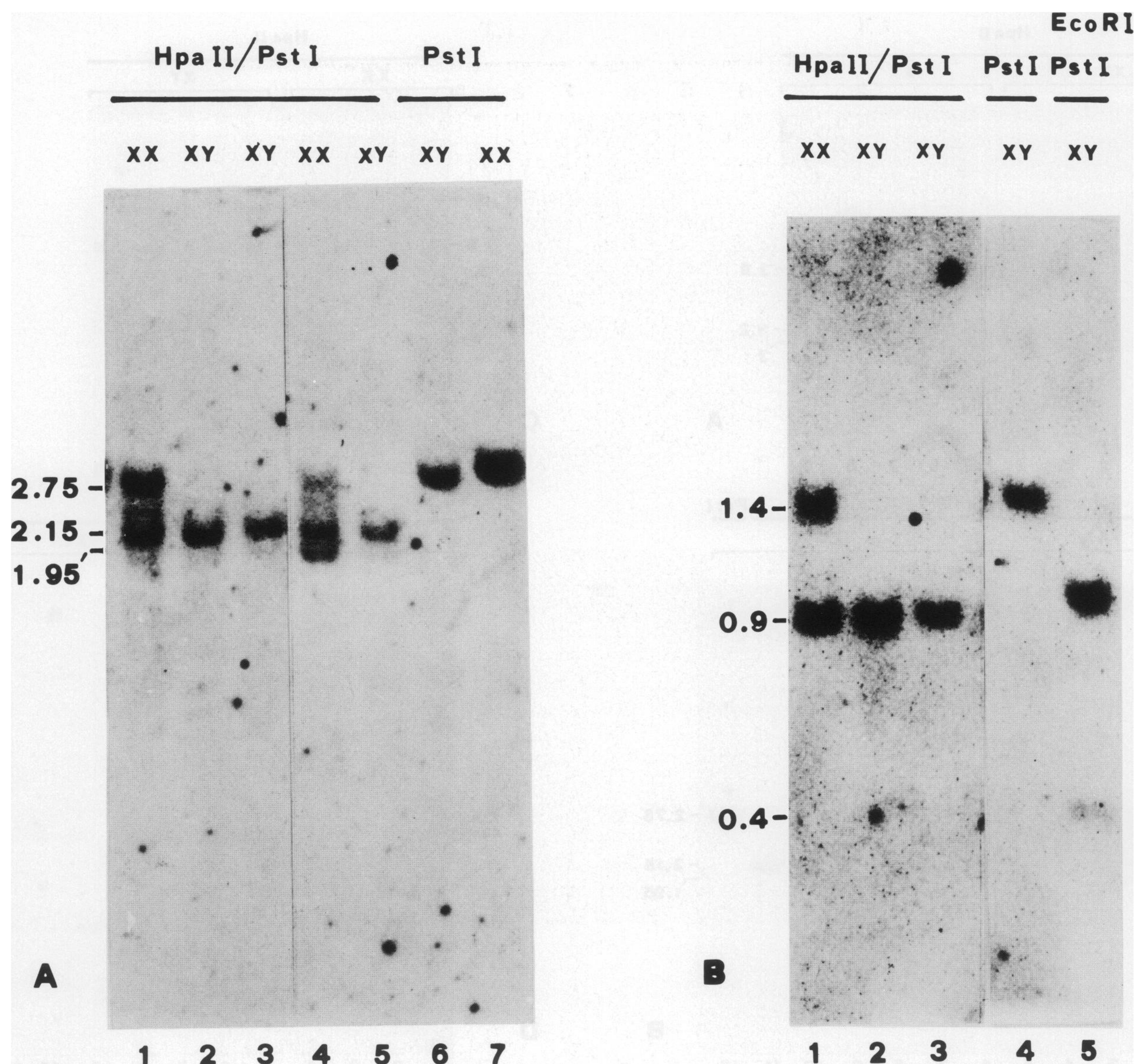

Fig. 5. Southern blot analysis of DNA from a female heterozygous for G6PD deficiency **(lanes A1, B1)**, and her two sons **(A2, 3; B2, 3)**. **Lanes A4, A7** are from another (control) female subject and **lanes A5, A6, B4, B5** are from another (control) male subject. **(A)** probe pGD3, **(B)** probe pGD1.4.

the 1.4-kb fragment seen in *Pst*I-*Hpa*II double digests probed with pGD1.4 that site H 36 is methylated in the inactive X.

In these double digests it is apparent that the intensity of the bands which are common to male and female DNA is similar. In addition, in each female DNA the intensity of these bands is similar to the intensity of the large-size bands (see Figure 4B and D). These quantitative relationships are consistent, of course, with the two bands corresponding to the active and the inactive X, respectively.

To analyse directly the same DNA when belonging to either the active or the inactive X-chromosome we have blotted samples from a family consisting of a mother heterozygous for 6GPD deficiency (genotype Gd^+/Gd^-) and her two sons, one of them Gd^+ and the other Gd^-. While the mother's DNA shows the female pattern we have just described, it is seen (Figure 5) that DNA from the two sons, carrying the two different *Gd* alleles, yields the same male pattern.

Methylation of G6PD gene DNA in other tissues

The male-female differences we have observed are not limited to leukocyte DNA. Rather, they are general to all tissues tested thus far (Figure 6). Tissue-specific methylation patterns of male DNA will be reported in full elsewhere (Battistuzzi *et al.*, unpublished). The analysis of the numerous sites for methylation-sensitive restriction enzymes other than *Hpa*II has been carried out on female DNA as well. The bands obtained after digestion with *Ava*I and *Sma*I are compatible with the pattern of methylation described above for *Hpa*II. However, DNA digested with *Hha*I shows only identical bands in male and female subjects (data not shown).

Discussion

The data reported here outline some remarkable features of the 3′ side of the human *Gd* gene. (1) This portion of the gene

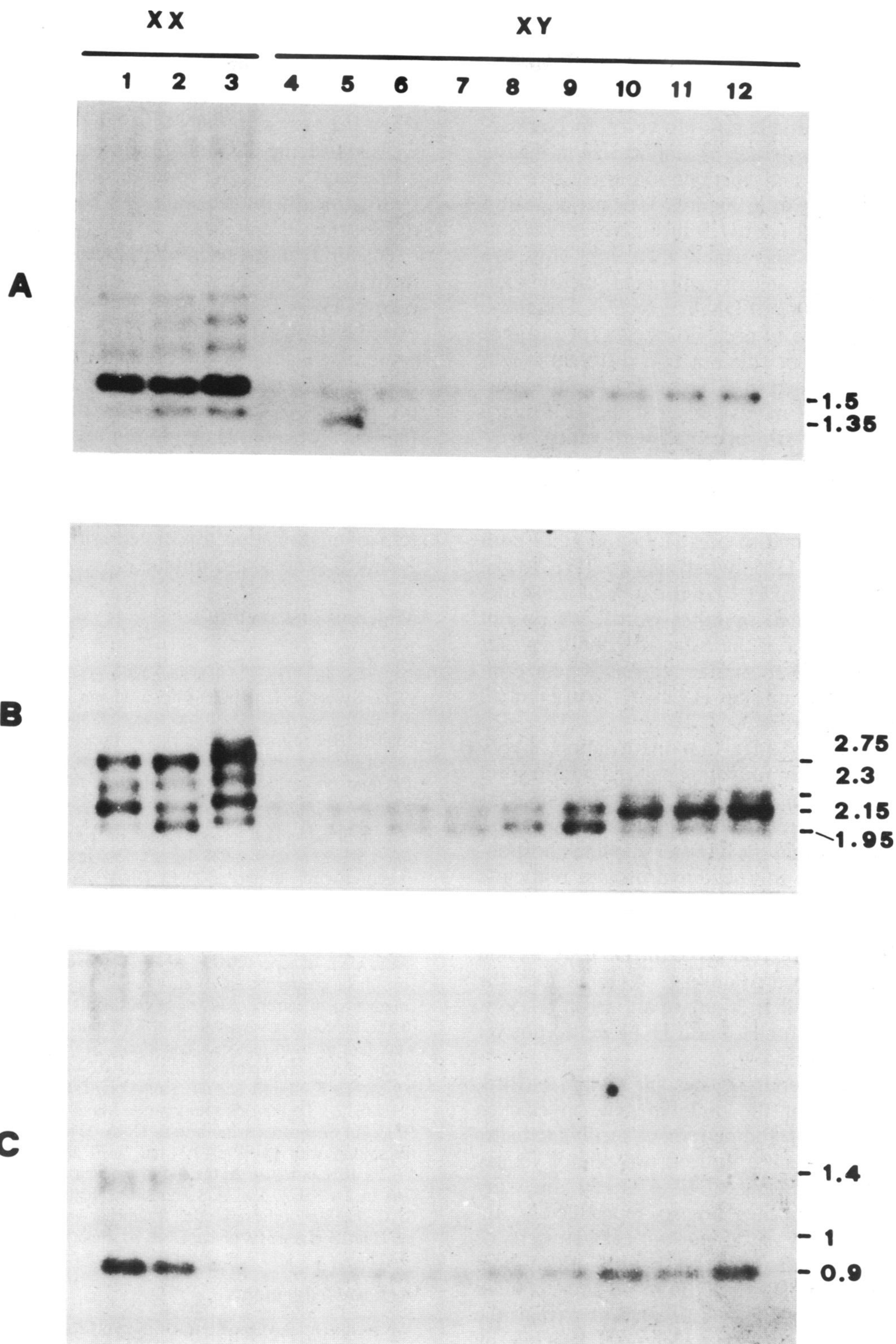

Fig. 6. Southern blot analysis of digests of male and female DNA from various tissues hybridized to probe pGD3 and pGD1.4. **1**: liver; **2**: skeletal muscle; **3**: cardiac muscle; **4–6**: adrenal; **7–9**: skeletal muscle; **10–12**: brain; the first, second and third DNA in each group of males corresponds to the same individual. Digestions were as follows: **A**, digestion with *Hpa*II + *Bst*EII probed with pGD3; **B**, digestion with *Hpa*II + *Pst*I probed with pGD3; **C**, digestion with *Hpa*II + *Pst*I probed with pGD1.4

is very rich in CpG sequences susceptible to methylation. Indeed, two clusters of such sequences can be identified: one between −7.3 and −5.8 kb, and one between −2.9 and −2.05 kb. Numerous sites are also present in the non-transcribed region from +1 to +4.7 kb, but these are less tightly packed. (2) While many of these sites are methylated, five specific *Hpa*II sites have been identified which are reproducibly unmethylated in human male leukocytes. With reference to Figure 1, these are H 4, H 14, H 20, H 36 and H 56. (3) Among the sites that are specifically unmethylated in male DNA, two (H 14 and H 36) are methylated in ~50% of female DNA and they are likely to correspond to the gene

on the inactive X-chromosome.

It is important to point out the limitations of the approach we have used. Firstly, if certain sites were partially unmethylated (i.e., in some DNA molecules but not in all), we might have missed the resulting bands. However, by comparing the intensity of signal in bands of similar size in digests with methylation-sensitive and methylation-insensitive enzymes (e.g., lanes 12 *versus* 16 in Figure 2), or in single and double digest (e.g., lanes 19 *versus* 20 in Figure 2), we find that in most cases it is not appreciably different. Thus, we estimate that for sites which we call methylated, methylation must be present in at least 80% of DNA molecules. Secondly, where H sites are very close to each other (see for instance sites H 18 – H 32), we cannot rule out that different neighbouring sites are unmethylated in different DNA molecules. Indeed, in some blots *Hpa*II bands appear slightly less sharp than bands of comparable size obtained with other digestions, suggesting that heterogeneity of methylation within a cluster may occur. This is especially obvious with respect to the doublet of bands of unequal intensity seen with probe pGD3, which can be attributed to sites H 15 and H 14 both being unmethylated, but H 15 only partially so. (We do not know the methylation status of H 14 in the 10% of molecules that have H 15 unmethylated: in other words, we do not know whether unmethylation of these two neighbouring sites is alternative or additive.) For similar reasons we have indicated as jointly unmethylated sites H 18 – H 20 in Figure 2 and sites H 55 – H 57 in Figure 3.

It is clear from these findings that methylation, hypomethylation or unmethylation must be analyzed with respect to individual sites rather than in terms of an entire gene. Our results are in general agreement with those of Stein *et al.* (1983), who have reported that the hamster adenine phosphoribosyl transferase and the mouse dihydrofolate reductase genes are methylated heavily throughout all but their respective 5′ ends. Methylation of C in CpG dinucleotides is widespread in both intron and exon regions of the *Gd* gene, as well as in the 3′-untranscribed flanking region. Of the unmethylated sites, H 20 and H 36 are in an intron, H 56 is distal to transcription, and for H 4 and H 14 we do not yet have the information.

Unlike the case of ribosomal genes in *Xenopus* (Bird, 1978), and like the case of the β-globin gene in chicken (McGhee and Ginder, 1979), the pattern of methylated sites in the *Gd* gene is highly specific. Indeed, we cannot yet say whether it is more significant that so many sites are methylated or that certain sites are so reproducibly unmethylated. It is especially striking that two clusters of methylatable regions have emerged, and that in both of them the sites at both ends of the cluster tend to be demethylated (H 4, H 14 – 15, H 20, H 36). It is remarkable that the specifically unmethylated residues which are detected by *Hpa*II, *Hha*I and *Ava*I are very close to each other (*Sma*I cannot be counted because all *Sma*I sites are automatically *Hpa*II sites). In addition, the methylation-demethylation pattern in *Gd* DNA from peripheral blood leukocytes is much more reproducible than it was found to be in cultured fibroblasts with another (anonymous) X-linked gene (Wolf and Migeon, 1982).

Our data show that methylation or unmethylation of specific sites can be related to X-chromosome activity. Sites H 14 – 15 and H 36 appear reproducibly methylated only in part of female DNA. The DNA methylated at sites H 14 – 15 and H 36 is likely to correspond to the G6PD gene on the inactive X-chromosome for the following reasons. (1) The sites are methylated in ~50% of all seven female DNA samples studied. (2) The methylation pattern is consistently found in all tissues analyzed. (3) This extra-methylation is reversed when the very same DNA is inherited by a male (Figure 6). Cells where the *Gd* gene on the inactive X-chromosome has been reactivated either spontaneously (Migeon *et al.*, 1982) or by treatment with azaC are now being analyzed to determine if the same sites are demethylated as well.

Also in females, as in males, we cannot assess precisely the methylation pattern of all sites, especially those in the two clusters H 1 – H 14 and H 18 – H 36. On the other hand, the extreme specificity of the changes associated with X-chromosome inactivation is strikingly illustrated by considering sites H36 and *Hha* 38. Here, within the sequence GCGCCCCGG, on which both CpG doublets are unmethylated in males, the first becomes methylated in females whereas the second does not. The type of gene regulation associated with X-chromosome inactivation differs from most others in that it affects only one allele at each locus. Our results indicate that the pattern of methylation associated with the inactive state of an X-linked gene is also unique.

Materials and methods

Probes

G6PD-specific cDNA clones have been previously described (Persico *et al.*, 1981). By using clone pGD 6405 a human genomic DNA library in phage λCh 28 (Hieter *et al.*, 1980) was screened, and two recombinant phages were isolated. *Pst*I restriction fragments of these were subcloned in pBR322, and two of these are pGD3 and pGD1.4. Details of isolation, subcloning, restriction maps and sequence data (see Figure 1) will be published in full elsewhere.

Southern blot analysis

Human blood leukocytes were separated from freshly drawn blood by the addition of 0.2 volumes of 3% dextran T500 in saline followed by sedimentation of red cells for 30 min at 37°C. DNA was then extracted and purified as previously described (Persico *et al.*, 1981). Digestions with restriction enzymes were carried out under the conditions recommended by the suppliers (Boehringer, Mannheim, FRG; Biolabs, Beverley, MA; Research Products, Amersham, UK). In general, 20 μg of DNA were digested with 80 units of enzyme for 4 h (a greater excess of enzyme did not alter the pattern of restriction fragments). Electrophoresis of DNA digests was carried out in agarose gels (0.8% unless otherwise stated) in TEB buffer (Maniatis *et al.*, 1982). Gels were blotted onto nitrocellulose filters (Schleier and Schuell: see Southern, 1975), and filters were hybridized for 20 – 40 h at 65°C in 5 x SSPE (i.e., 0.9 M NaCl, 40 mM NaOH, 5 mM NaEDTA, 50 mM sodium phosphate, pH 7.0), 2 x Denhardt's solution (see Maniatis *et al.*, 1982), 0.1% SDS and 50 μg/ml of sheared salmon testis DNA. DNA probes were labelled by nick translation with [α-^{32}P]dATP and [α-^{32}P]dCTP to a specific activity of 1 – 2 x 10^8 c.p.m./μg. At the end of the hybridization, the filters were washed first in 2 x SSC 0.1% SDS and subsequently in 0.2 x SSC, 0.1% SDS at 65°C. The size of the restriction fragments was determined by running in each gel *Hind*III digested λ-phage DNA as a marker.

Acknowledgements

We thank Dr. B.Migeon and Dr. S.Wolf for many helpful discussions, and Ms. Maria Terracciano, Ms. Carmela Salzano, Ms. Concetta Sole and Ms. Jane Faithful for excellent technical assistance. G.B. was supported by an EMBO long-term Research Fellowship. This work was carried out with the financial support of the two projects 'Controllo della crescita neoplastica' and 'Ingegneria genetica e basi molecolari delle malattie ereditarie', of CNR, Italy, and from a Programme Grant by MRC, Great Britain.

References

Bird,A.P. (1978) *J. Mol. Biol.*, **118**, 49-60.
Bird,A.P. and Southern,E.M. (1978) *J. Mol. Biol.*, **118**, 27-47.
Cooper,D.N. (1983) *Hum. Genet.*, **64**, 315-333.
Ehrlich,M. and Wang,R.Y.H. (1981) *Science (Wash.)*, **212**, 1350-1357.
Hieter,P.A., Max,E.E., Seidman,J.G., Maizel,J.V.,Jr. and Leder,P. (1980) *Cell*, **22**, 197-207.

Kuo,M.T., Mandel,J.L. and Chambon,P. (1979) *Nucleic Acids Res.*, **7**, 2105-2112.
Mandel,J.L. and Chambon,P. (1979) *Nucleic Acids Res.*, **7**, 2081-2103.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual,* published by Cold Spring Harbor Labortory Press, NY.
McGhee,J.D. and Ginder,G.D. (1979) *Nature,* **280**, 419-420.
Migeon,B.R., Wolf,S.F., Mareni,C. and Axelman,J. (1982) *Cell,* **29**, 595-600.
Persico,M.G., Toniolo,D., Nobile,C., D'Urso,M. and Luzzatto,L. (1981) *Nature,* **294**, 778-780.
Southern,E.M. (1975) *J. Mol. Biol.,* **98**, 503-517.
Stein,R., Sciaki-Gallili,N., Razin,A. and Cedar,H. (1983) *Proc. Natl. Acad. Sci. USA,* **80**, 2422-2426.
Van der Ploeg,L.H.T. and Flavell,R.A. (1980) *Cell,* **19**, 947-958.
Venolia,L., Gartler,S.M., Wassman,E.R., Yen,P., Mohandas,T. and Shapiro,L.J. (1982) *Proc. Natl. Acad. Sci. USA,* **79**, 2352-2354.
Wolf,S.F. and Migeon,B.R. (1982) *Nature,* **295**, 667-671.

Received on 6 February 1984; revised on 13 June 1984