© 1997 Oxford University Press

Human Molecular Genetics, 1997, Vol. 6, No. 11 **1791–1801**

ARTICLE

Epigenetic variation illustrated by DNA methylation patterns of the fragile-X gene *FMR1*

Reinhard Stöger[1,*], Tina M. Kajimura[1], W. Ted Brown[2] and Charles D. Laird[1,3,*]

[1]Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, C3-168, Seattle, WA 98109, USA, [2]Institute for Basic Research in Developmental Disabilities, Staten Island, NY, USA and [3]Department of Zoology, University of Washington, Seattle, WA, USA

Received May 5, 1997; Revised and Accepted July 2, 1997

Genomic methylation patterns of mammals can vary among individuals and are subject to dynamic changes during development. In order to gain a better understanding of this variation, we have analyzed patterns of cytosine methylation within a 200 bp region at the CpG island of the human *FMR1* gene from leukocyte DNA. *FMR1* is normally methylated during inactivation of the X chromosome in females and it is also methylated and inactivated upon expansion of CGG repeats in fragile-X syndrome. Patterns of methylation (epigenotypes) were determined by the sequencing of bisulfite-treated alleles from normal males and females and alleles from a family of five brothers who are methylation mosaics and are affected to various degrees by the fragile-X syndrome. Our data indicate that: (i) methylation of individual CpG cytosines is strikingly variable in hypermethylated epigenotypes obtained from a single individual, suggesting that maintenance of cytosine methylation is a dynamic process; (ii) methylation of non-CpG cytosines in the region studied may occur but is rare; (iii) mosaicism of methylation in the analyzed fragile-X males is remarkably similar to that found for the active X and inactive X alleles in normal females, suggesting that the methylation mosaicism of some fragile-X males reflects similar on and off states of *FMR1* expression that exist in normal females; (iv) hypermethylation is slightly more pronounced on fragile-X alleles than on normal inactive X alleles of females; (v) the general dichotomy of hypo- and hypermethylated alleles persisted over the 5 year period that separated samplings of the fragile-X males; (vi) methylation variability was most pronounced at a consensus binding sequence for the α-PAL transcription factor, a sequence that may play a role in regulating expression of *FMR1*.

INTRODUCTION

The DNA sequence of an individual human is generally found to be invariant throughout development and life, with the exception of somatic mutations and rearrangements at the immunoglobulin loci in a minority of cells. With most current sequencing methods, however, only the four major bases in DNA are detected. A fifth base, 5-methylcytosine (5-MeC), is present in the DNA sequence of many organisms, including humans. It is usually found in the symmetrical sequence 5-MeCpG. Like the primary nucleotide sequence, the overall methylation blueprint can differ among individuals and is inherited in Mendelian fashion (1). Stability of methylation is apparent in the propagation of clonal patterns that maintain the inactive X chromosome in its silenced state in somatic cells of female mammals (reviewed in 2–4).

Other aspects of DNA methylation are more dynamic. For example, fluctuations in cytosine methylation occur during early development of the mouse, where genome-wide demethylation is followed by an extensive wave of *de novo* methylation prior to gastrulation (5,6). Reduced levels of methyltransferase activity in mice result in an embryonic lethal phenotype and this result indicates that an intact DNA modification system plays a vital role in mammals (7). Aberrant methylation patterns and their clonal propagation in cell lineages are correlated with certain conditions in humans, such as the Prader–Willi, Angelman and Fragile-X syndromes (8–14).

Methylation of DNA is believed to be involved in establishing and maintaining a particular state of gene expression during development, a model originally proposed in 1975 (15–17). A role for cytosine methylation has been convincingly established for promoter elements, in which it mediates transcriptional silencing and gene repression (reviewed in 18). The promoters of ~60% of genes in the human genome coincide with dense clusters of potential methylation sites and these CpG islands remain unmethylated throughout development (19,20). Methylation of CpG islands, however, does occur *in vivo* at genes where haploid

*To whom correspondence should be addressed. Tel: +1 206 667 4032/4033; Fax: +1 206 667 6524; Email: rstoeger@fred.fhcrc.org or claird@fred.fhcrc.org

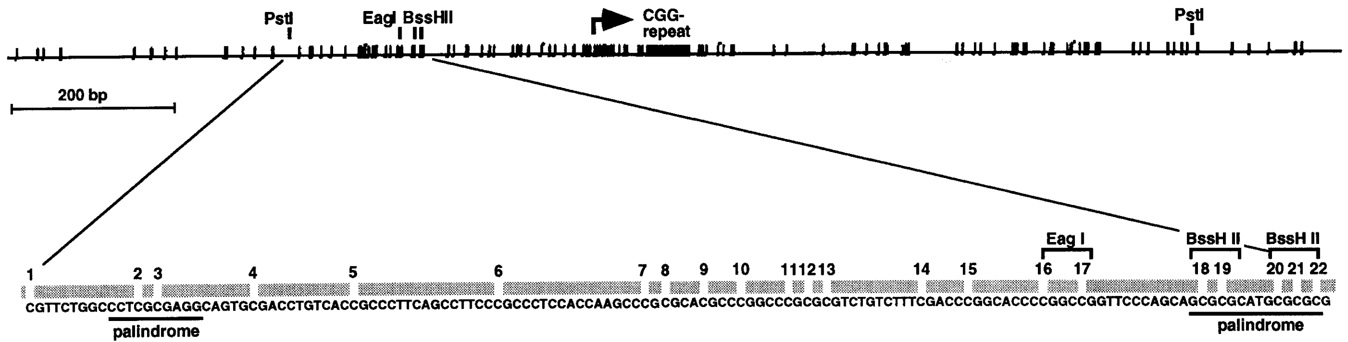


Figure 1. CpG island of the human *FMR1* gene. CpG sites and their distribution within 1.5 kb of the 5′-end of the gene are indicated as small vertical lines on the upper plot. The recognition sites for restriction enzymes *Pst*I, *Eag*I and *Bss*HII, which are frequently used to assess the size of the repeat and the methylation status of the promoter, are shown, as is the presumed RNA transcription start site (arrow) inferred from the published *FMR1* cDNA (accession no. X69962), and the CGG repeat. The 1030 bp region between the two *Pst*I sites is 66% GC-rich and contains equal amounts of CpG and GpC dinucleotides (distribution of GpC dinucleotides not shown). The gray bar in the lower part of the figure is a magnification of a 200 bp region that was analyzed by sequencing of bisulfite-treated, PCR-amplified genomic DNA. The DNA sequence is the upper strand of the *FMR1* promoter and corresponds to bases 2327–2466 of the published genomic DNA sequence (accession no. X61378). The 22 CpG sites are numbered and their relative positions are indicated as white boxes on the gray bar. Two palindromic elements and the recognition sites for methylation-sensitive restriction enzymes *Eag*I and *Bss*HII are also indicated

dosage and long-term repression of an allele is critical for normal development of an organism (for reviews see 21,22).

FMR1 is an example of an X-linked gene that is normally regulated by methylation of the promoter during X inactivation in somatic cells of females (11,23–25). Methylation at the *FMR1* promoter is also found if a CGG triplet repeat, located in the 5′-untranslated region, expands to more than ~200 repeats (13,24–26). This leads to gene silencing and insufficient synthesis of *FMR1* protein, which is the predominant cause of fragile-X syndrome, a common form of inherited mental retardation (27). Males and females are affected with cognitive function ranging from severe mental retardation to only mild learning disabilities (28–32). Many mildly affected individuals show mosaic methylation at the *FMR1* promoter and thus provide an example of epigenetic mosaicism in mammals.

The *FMR1* promoter and first exon are embedded within a typical CpG island that spans ~1 kb of DNA (33; Fig. 1). We analyzed a 200 bp sequence from this region that includes the recognition sites for the restriction enzymes *Eag*I and *Bss*HII, which are commonly used to clinically assess promoter methylation status. Patterns of cytosine methylation in peripheral leukocyte DNAs were analyzed for normal and fragile-X individuals, including five fragile-X males from a family originally described by Brown *et al.* (34). The family is remarkable in that the five brothers inheriting the fragile-X chromosome are affected to varying degrees ranging from only mild cognitive deficits to complete expression of the clinical phenotype (31). In a previous study the degree of *FMR1* promoter methylation was assessed in this fragile-X family by the ability of a methylation-sensitive restriction enzyme, *Eag*I, to digest leukocyte DNA. The percentage of DNA unmethylated at this *Eag*I site correlated well with cognitive ability (31). The use of methylation-sensitive restriction enzymes, however, is limited by the small number of potential methylation sites that can be analyzed and the relatively large amounts of DNA that are required. Development of a sensitive bisulfite-based technique now permits more complete analysis of methylation patterns of single DNA molecules (35,36).

We used bisulfite-treated, PCR-amplified genomic DNA to analyze the methylation status within the promoter region of *FMR1*. Examination of DNA from normal females and from fragile-X individuals enabled us to address the following questions about the nature of *in vivo* cytosine methylation in humans: (i) how variable are methylation patterns in leukocytes of an individual?; (ii) is methylation of non-CpG cytosines a prominent feature in this CpG island?; (iii) how is mosaicism of methylation within a population of cells manifested on single DNA molecules?; (iv) does cytosine methylation on a normal, inactive X chromosome differ from the pattern on a fragile-X chromosome?; (v) are methylation patterns maintained over time?; (vi) does the methylation status of a particular CpG site or sites within this promoter region indicate a salient role in gene regulation?

RESULTS

Region analyzed and sequencing approach

The 200 bp sequence whose methylation profile is described here contains 22 CpG dinucleotides and 45 non-CpG cytosines (Fig. 1). Sodium bisulfite catalyzes the conversion of cytosine to uracil residues in single-stranded DNA, whereas methylated cytosines remain unreactive under these conditions. The specificity of this reaction is remarkably high, as will be apparent in the presentation of data from females, where both hyper- and hypomethylated alleles are normally present in the same cell. Upon PCR amplification of the genomic region of interest the converted uracil residues are replicated as thymines instead of cytosines (35,36). A remaining cytosine in the sequence of the PCR product therefore indicates that this site was methylated on the template DNA. We cloned and sequenced individual PCR products from the coding strand; because methylation is usually symmetrical at CpG dinucleotides (37,38), we interpret our data as representing methylation patterns of individual alleles. Each methylation pattern thus represents the epigenetic profile of an allele in one cell, or an ‘epigenotype’.

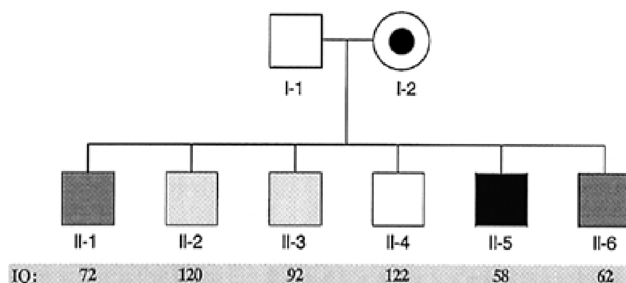


Figure 2. Pedigree of the fragile-X family. DNAs from a fragile-X carrier mother (I-2) and from her six sons (II-1–II-6) were analyzed. Five of the sons inherited the fragile-X syndrome chromosome (II-1, II-2, II-3, II-5 and II-6) and one of them inherited a normal X chromosome (II-4). The degree of methylation at the *Eag*I site, the sizes of the *FMR1* CGG repeat and the cognitive function of the six brothers have been analyzed (31); a summary of these results is presented here. The full scale IQ determined by the Wechsler Adult Intelligence Scale–Revised is listed below each of the six sons. The approximate degree of methylation at the *Eag*I site was determined by densitometric analysis on genomic DNA and is indicated here by gray shading in the boxes of the pedigree tree. Individual II-1, affected, variable repeat size, ~45% of total DNA methylated at *Eag*I; individual II-2, non-penetrant, variable repeat size, ~3% of total DNA methylated at *Eag*I; individual II-3, non-penetrant, variable repeat size, ~3% of total DNA methylated at *Eag*I; individual II-4, normal *FMR1* allele, 0% total DNA methylated at *Eag*I; individual II-5, affected, 100% of total DNA methylated at *Eag*I; individual II-6, affected, variable repeat size, ~45% of total DNA methylated at *Eag*I.

Normal and aberrant methylation patterns in males

DNA from six brothers of a fragile-X family was used to analyze methylation patterns of normal and fragile-X alleles in males (Fig. 2). We determined 180 epigenotypes from these six brothers, representing for each brother 15 epigenotypes from each of two samples collected 5 years apart. Male II-4 inherited a normal X chromosome and was shown to have a completely unmethylated *Eag*I site at *FMR1*, as expected for a CpG island of an X chromosomal locus in males (31). DNA from individual II-4 can thus be used to assess the methylation status for normal, transcriptionally active *FMR1* alleles in males (39). Our results confirm that this region of the *FMR1* CpG island is markedly hypomethylated in normal male DNA. Twenty seven of 30 alleles were unmethylated at all sites, two alleles were methylated at one site each (sites 19 and 22) and one allele was methylated at three sites (sites 19, 21 and 22; Fig. 3). Similar hypomethylation was observed for epigenotypes of a normal male, unrelated to the fragile-X family (data not shown).

Individual II-5, who has a full scale IQ of 58, is the most affected of the brothers, manifesting numerous clinical symptoms of fragile-X syndrome. The *Eag*I site from his DNA appears to be completely methylated by restriction digestion analysis (31). We find that methylation patterns of individual alleles from this male indeed indicate that the *FMR1* promoter is predominately hypermethylated in most of his leukocytes (Fig. 3). However, a completely unmethylated allele was also detected from each of his blood samples (clones 1986-9 and 1991-8) and these are likely to be transcriptionally active alleles.

The other four brothers are less affected by the fragile-X syndrome. DNAs of individuals II-1 and II-6 have been shown previously to be ~45% methylated at the *Eag*I site (31). Our finding of a mixture of hypo- (defined as ≤6 sites modified) and hypermethylated (defined as ≥16 sites modified) alleles is consistent with the partial methylation detected by restriction digestion; methylation mosaicism, therefore, represents markedly bimodal epigenetic patterns arising from different cells.

Individuals II-3 and II-2 were also identified as methylation mosaics, but only small proportions of their DNAs were found to be methylated at the *Eag*I site; both brothers were originally considered to have a non-penetrant phenotype, although they have substantially different IQ scores (31). II-2 has a full scale IQ score of 120, compared with 92 for II-3 (Fig. 2). Only hypomethylated alleles were isolated from DNA samples of individual II-2, whereas one hypermethylated allele was detected in each blood sample of individual II-3 (Fig. 3).

Methylation patterns of female DNA

To determine methylation patterns of normal *FMR1* alleles from active and inactive X chromosomes we examined DNA from two females (L2 and L7) who are not related to the fragile-X family (Fig. 4). In addition, DNA was analyzed from the mother (I-2) of the six brothers (Fig. 4). I-2 has one normal X chromosome and one X with an expanded CGG repeat in the pre-mutation range (~65 repeats) and is referred to as a 'carrier'; she is not affected by the fragile-X syndrome and does not exhibit aberrant methylation of the *Eag*I site in the *FMR1* promoter (31).

The general patterns of methylation among these three females are similar to those expected for a mixture of hypermethylated inactive and hypomethylated active alleles (24). No significant differences were apparent for the alleles from the pre-mutation female I-2 compared with the normal females, although more epigenotypes of this and neighboring regions are needed to test and extend this result. Surprisingly, all four samples of female DNA, including the two samples from I-2, yielded hypomethylated alleles more frequently than expected by chance (Fig. 4). Overall, 18 hypermethylated and presumably inactive alleles, 60 hypomethylated and presumably active alleles and four alleles of intermediate methylation status were observed (Fig. 4). Due to X chromosome inactivation, ~50% of normal and pre-mutation female DNA is digested at the *FMR1* promoter by methylation-sensitive restriction enzymes (24,33,40). The bias towards recovering unmethylated alleles likely occurred during the amplification or cloning procedures.

Comparison of normal X and fragile-X alleles

Most hypermethylated alleles were not methylated at all 22 CpG sites (Fig. 6). This was significantly more pronounced ($P < 0.001$, *t*-test) on normal, inactive X chromosomes from females, where hypermethylated alleles have an average of 2.9 ± 0.9 (±2 SE) unmodified CpG sites. In contrast, hypermethylated alleles from fragile-X chromosomes in males are found to have a mean of only 1.1 ± 0.4 unmethylated CpG sites (Figs 5 and 6). Complete methylation of all 22 CpGs occurred in 32% of hypermethylated alleles from the five fragile-X brothers, but only in 17% of the normal, inactive X alleles. Fewer than 1% (1/150) of the fragile-X alleles from the five males had intermediate levels of methylation (defined as ≥7 and ≤15 sites modified), whereas 5% (4/82) of alleles derived from females had intermediate methylation levels. Thus both the completeness of methylation of hypermethylated alleles and the frequency of alleles with intermediate methylation

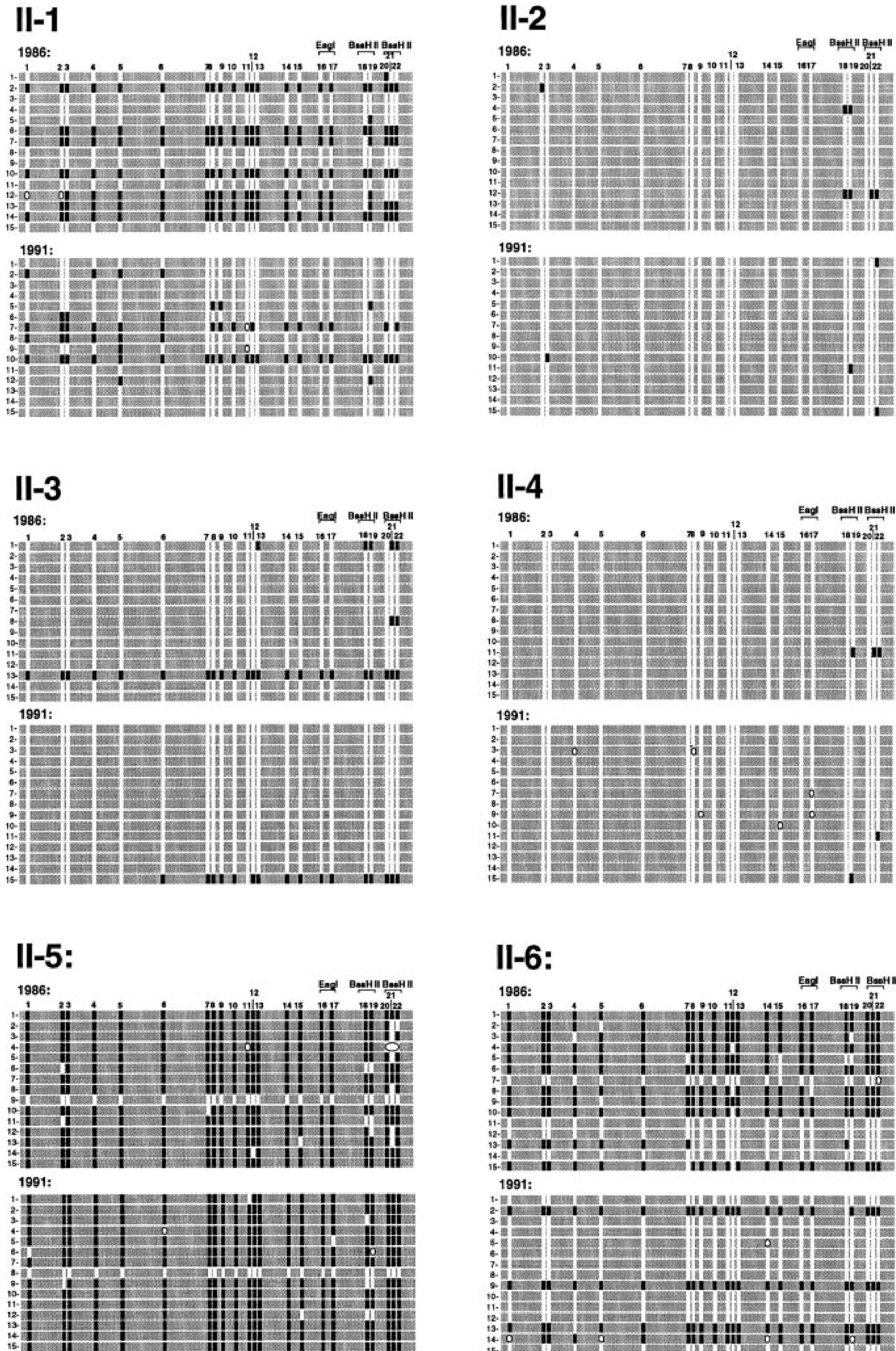
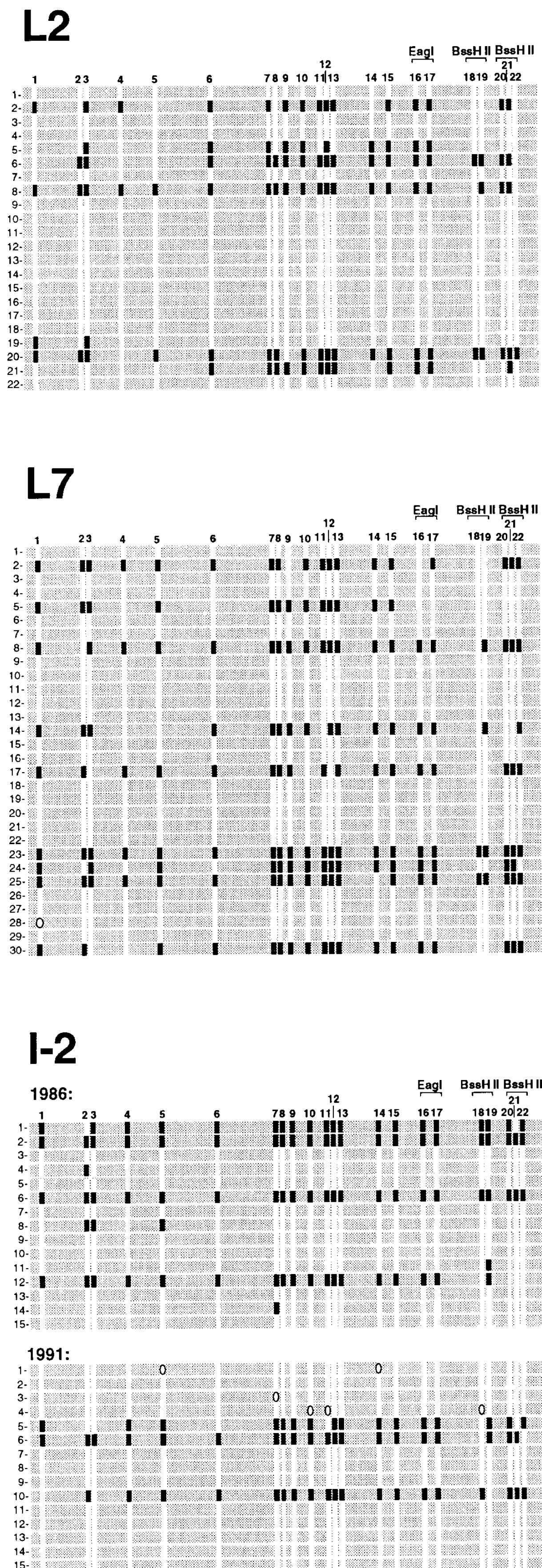


Figure 3. Epigenotypes of the six brothers (II-1–II-6) were determined at the *FMR1* promoter. For each individual 15 alleles were sequenced from a 1986 DNA sample and 15 alleles from a 1991 DNA sample. The relative position of a CpG site and its methylation status is indicated by either a white (unmethylated) or a black (methylated) box on the DNA strand (gray bar); ellipses indicate incomplete sequence information for a particular site. Individual clones are numbered on the left side of the gray bar.



status differed between fragile-X and normal alleles, with the fragile-X alleles exhibiting higher levels of methylation.

The hypomethylated alleles from normal female DNA and from DNA of the six males of the fragile-X family did not differ noticeably in their methylation states. Most hypomethylated alleles were completely devoid of 5-MeC; the frequency of single-site methylation was ~10-fold lower than for hypomethylated alleles with no methylation (Fig. 5). A small subset of hypomethylated alleles had two to six unconverted C residues at particular CpG sites that may represent rare cytosine methylation in active alleles (for example sites 19 and 22; Fig. 6).

Specificity of conversion of cytosine and resistance of methylated cytosine to conversion

Sequencing provided complete information for all 22 CpG sites on 246 of the 262 epigenotypes, representing 5412 CpG cytosines. Complete information for all 45 non-CpG cytosines was obtained for 188 epigenotypes, representing 8460 non-CpG cytosines. This data set permits accurate estimation of the specificity of the bisulfite reaction in terms of the efficiency of conversion of non-methylated cytosines to uracil and the degree of protection from conversion by methylation.

Non-CpG cytosines are expected to be rarely methylated and thus provide a good estimate of the efficiency of conversion: 99.4% of the non-CpG cytosines were converted in female DNA samples, irrespective of whether the alleles were hyper- or hypomethylated at CpG sites; conversion of non-CpG cytosines in DNA samples from the males of the fragile-X family was 99.6% efficient. Does the remaining 0.6 and 0.4% protection represent *bona fide* methylation of non-CpG cytosines? Inheritance of cytosine methylation at non-CpG sites, especially CpNpG trinucleotides in mammalian cells, has been reported previously (41). We therefore examined the location of the 43 events of unconverted non-CpG cytosines. These were found at highest frequency (16/43) to be at trinucleotides C̲pNpT, in particular at the C̲pApT (9/43) positioned immediately 3′ of CpG site 19 in a palindromic sequence. Resistance to conversion of this cytosine was most pronounced on hypomethylated alleles of males (5/9) in which one or more of the flanking CpG sites (sites 18–22) was also methylated. In addition, two hypermethylated male alleles and two intermediately methylated alleles (one female and one male) had this cytosine protected. Non-CpG cytosine protection was also detected on 14 CpNpG trinucleotides. Eleven of these cytosines were positioned at CCG triplets, with four being at the CCG identified by CpG site 10 [here referred to as C(CpG)-10], three at C(CpG)-7, two at C(CpG)-11, one at C(CpG)-5 and one at C(CpG)-6. The rest of the unconverted non-CpG cytosines occurred at CpNpC (12/43) and CpNpA (1/43) trinucleotides sites.

Figure 4. Epigenotypes of three females were determined at the *FMR1* promoter. Individuals L2 and L7 have normal CGG repeats and are not related to the fragile-X family. Individual I-2 is a fragile-X carrier and the mother of the six brothers (see also legend to Fig. 2). Individual alleles are represented as gray bars and are numbered on the left side. The relative position of a CpG site and its methylation status are indicated by either a white (unmethylated) or a black (methylated) box on the DNA strand (gray bar). Ellipses indicate incomplete sequence information for a particular site. Twenty two epigenotypes were analyzed for DNA of female L2, 30 for L7 DNA and 15 each for 1986 and 1991 DNA samples of female I-2.

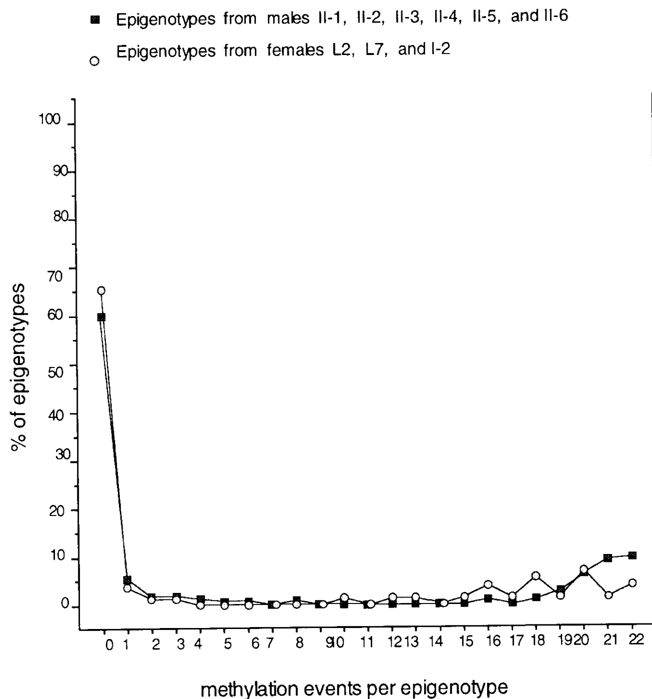


Figure 5. Frequency of CpG methylation on individual alleles. The occurrence of methylation events on individual alleles is plotted against the percentage of alleles where complete information for all 22 CpG sites was obtained. CpG cytosine methylation events (1021) detected on 167 alleles of the six brothers (filled squares) are compared with methylation events (401) detected on 78 informative alleles from females L2, L7 and I-2 (gray circles).

DNA from normal females and fragile-X mosaic males provides an especially useful control to monitor the specificity of the chemical conversion of cytosine at CpG sites because the reactivity can be examined for both hypo- and hypermethylated alleles in the same DNA sample. Examination of the hypermethylated alleles permits an estimate of the minimum degree of protection from bisulfite conversion conferred by cytosine methylation. As described above, we inferred an average methylation of 95% (20.9/22) of the CpGs in hypermethylated alleles of fragile-X males. Sites 7–17 exhibited an even greater degree of protection from bisulfite in both male (97%) and female (96%) alleles (Fig. 6). For male DNA eight sites were always protected in the 47 hypermethylated alleles analyzed. With an average methylation probability of 0.95, finding even two sites always protected in 47 epigenotypes would be significant ($P <$ 0.01); the likelihood of finding eight sites uniformly protected in 47 clones is even smaller ($P < 4 \times 10^{-9}$). Thus the degree of protection conferred by methylation in our assays is at least 95% and may be >99%; the latter figure is the more accurate one if, as seems likely, some sites in the region studied are methylated more frequently than others.

The hypomethylated alleles from females L2, L7 and I-2 provide an estimate of the efficiency of conversion of unmethylated CpG cytosines. Only eight of 1232 CpG cytosines remained unconverted, indicating that the efficiency of conversion was >99.3%, an estimate that is similar to that reached above for non-CpG cytosines.

Methylation status of individual CpG sites

The apparent specificity of the bisulfite reaction under appropriate conditions, as described above, together with the large number of alleles analyzed, permits analysis of the frequencies of methylation of individual sites. The incomplete methylation of most hypermethylated alleles allows us to determine which particular CpG sites have unusual frequencies of methylation in comparison with the general methylation status. This frequency varies from site to site for both fragile-X and inactive X alleles (Fig. 6). Among the hypermethylated fragile-X and inactive X alleles, representing 65 epigenotypes, the highest degree of modification was observed at sites 8 and 13, which were always methylated (Fig. 6). In contrast, site 18 had the lowest frequency of methylation on both inactive X and fragile-X hypermethylated alleles (0.45 and 0.8 respectively, compared with averages of 0.87 and 0.95 for these two allele classes). This cytosine marks the 5′-end of a 14 bp palindromic DNA sequence (Fig. 1) containing CpG sites that, as mentioned above, are occasionally methylated on hypomethylated alleles from males (Fig. 6).

Sites 2–4 in female hypermethylated alleles are also unusually low in their frequency of methylation, averaging 0.72 (Fig. 6). A palindrome spans the first two of these sites, each of which was found to be methylated in three hypomethylated male and two hypomethylated female alleles.

Stability of bimodal methylation patterns over time

The inactive state of X chromosomes is stable and inherited in a clonal fashion to daughter cells after DNA replication (42); methylation patterns are thought to be similarly inherited because of the properties of the maintenance methyltransferase (37,38). We thus looked for evidence of stable methylation patterns for CpG sites within the hypermethylated alleles derived from an individual. DNAs from the six brothers were isolated in both 1986 and 1991, thus providing the opportunity to determine if cytosine methylation patterns change in lymphocytes of affected individuals over a 5 year period. The marked bimodal methylation status of alleles was maintained over these years in all of the DNA samples. The same distributions of hypo- and hypermethylated alleles were observed for both sets of samples from individuals II-2, II-3 and II-5 (Fig. 3). The 1991 DNA samples from II-1 and II-6 have higher proportions of hypomethylated alleles compared with the 1986 samples (Fig. 3). The decrease in the fraction of hypermethylated clones is marginally significant for II-6 and for II-6 and II-1 taken together, using a one-tailed χ^2 test ($P < 0.05$). More epigenotypes would need to be characterized to assess the significance of this finding.

Infrequent clonality of detailed methylation patterns

As mentioned above, the broad methylation patterns of hypo- or hypermethylation were maintained over a 5 year period. Most hypermethylated alleles, however, were characterized by unique modification patterns that did not indicate strict clonal propagation of detailed methylation patterns. In only one case did we observe an identical, detailed pattern in which the same three sites lacked methylation (1986-6 and 1991-9 from individual II-5; Fig. 3). Several examples of completely methylated alleles from an individual were observed (for example 1986-2, 1986-6 and 1991-10 from individual II-1; Fig. 3), but complete methylation patterns provide less sensitivity to address questions of clonality

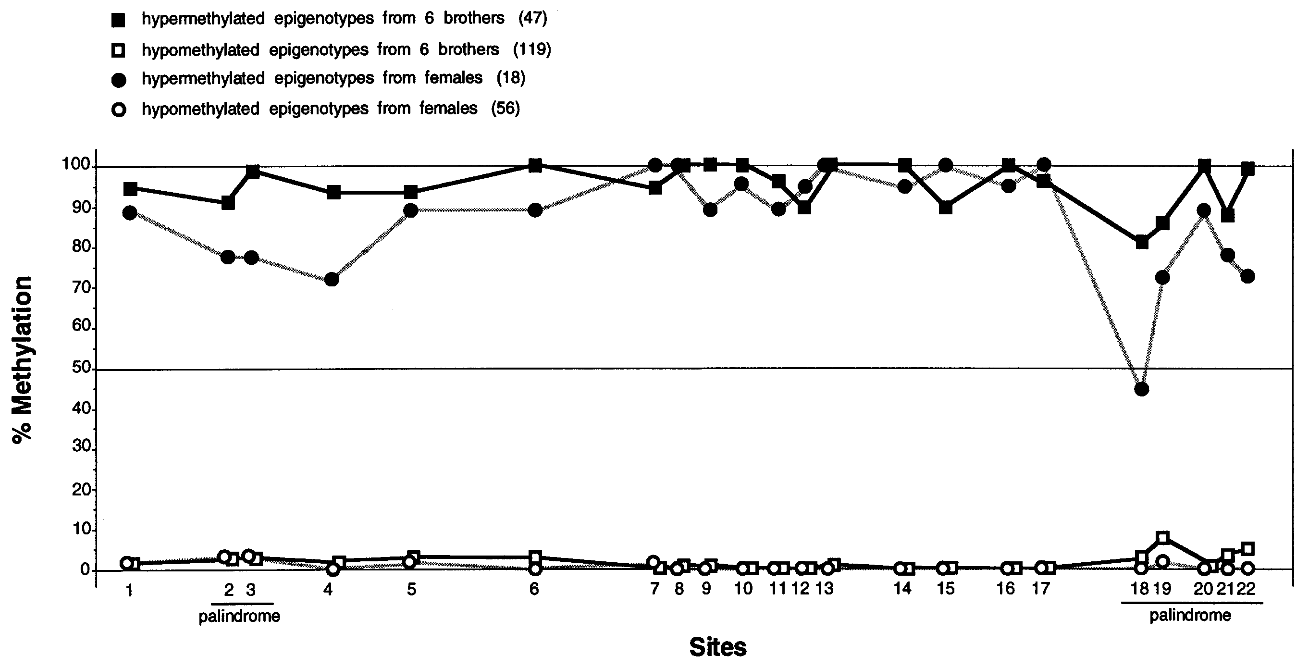


Figure 6. The degree of CpG cytosine methylation is site specific at the *FMR1* promoter. The percentage of methylation for individual CpG sites is shown for hypermethylated alleles (≥16 sites methylated) of females and fragile-X males (filled circles and squares) and for hypomethylated alleles (≤6 sites methylated) from females and the six brothers (empty circles and squares).

than do partially methylated alleles. Thus only a small subset of the hypermethylated epigenotypes characterized are compatible with a model of strict clonal propagation of methylation patterns.

DISCUSSION

We analyzed *in vivo* methylation patterns at the resolution of individual DNA molecules, using the CpG island of *FMR1* as a model to study epigenetic heterogeneity in humans. Our findings underline the complexity of epigenetic variations within a mammalian species (1,43). Some of these variations may affect the phenotype of an individual, as exemplified by the mosaicism of methylation in the fragile-X family described here. The description of cytosine methylation and its variations is thus of interest for a better understanding of epigenetic control of human phenotype and the rules that govern the propagation of methylation patterns.

Bimodal distribution of methylation density

The methylation patterns of the *FMR1* promoter in cells of normal females and of mosaic fragile-X individuals are markedly bimodal in being either hypo- or hypermethylated. From previous studies on methylation and transcript levels in cells from normal and fragile-X males we may correlate these hypo- and hypermethylated patterns with active and inactive alleles (39). For fragile-X individuals the proportion of hypermethylated alleles is correlated with the degree of mental retardation. Within this 200 bp region of the *FMR1* promoter the qualitative aspects of methylation mosaicism in fragile-X males are remarkably similar to that in normal females, except that normal females maintain an active *FMR1* allele in each cell, leading to normal levels of *FMR1* protein, compared with a variable fraction of cells in a mosaic male in which *FMR1* alleles are active.

Of the 262 epigenotypes that we characterized, only five (2%) showed intermediate levels of methylation, which we defined as methylation of between seven (32%) and 15 (68%) of the 22 CpG sites. With random methylation of these 22 sites 41% of epigenotypes would be expected to fall within this range. The methylation states at *FMR1* thus appear to be binary, reflecting an epigenetic control that leads to largely stable active or stable inactive loci.

Stability of bimodality over time

The proportion of hypermethylated *FMR1* alleles in fragile-X individuals did not change dramatically over a 5 year period. This result is consistent with long-term silencing of *FMR1* and, therefore, may be similar to the stable inactivation of genes on the normal, inactive X chromosome. The reduced frequency of hypermethylated alleles that may have occurred in the leukocyte population during the 5 year period between DNA sampling of males II-1 and II-6 may have arisen from mild selection against methylated *FMR1* alleles. The limited number of clones examined and the observed bias towards recovering unmethylated clones, however, preclude strong quantitative arguments on selection for these males. Selection has been reported against cells with inactivated expanded *FMR1* alleles on the active X chromosome during female adulthood (44) and against expanded *FMR1* CGG repeats during male germ cell development (45).

Microvariation within the broad patterns of bimodal methylation: considerations of clonality and methyltransferase processivity

Complete methylation of all CpGs, which was observed on only a subset of hypermethylated alleles, is consistent with a processive enzymatic activity of the maintenance methyltransferase, whereby every potential CpG site becomes modified. The majority of hypermethylated alleles, however, have complex, seemingly random methylation patterns that are best described as concerted. Among the clones from each patient the patterns were markedly diverse (Figs 3 and 4). This variability was only seen at CpG cytosine sites and was therefore not induced by PCR errors during the amplification and cloning procedure of bisulfite-treated DNA. We found little evidence for faithful clonality of specific methylation patterns and conclude that detailed *in vivo* methylation patterns are not inherited in an entirely faithful manner during the process of DNA replication, suggesting that the maintenance methyltransferase has low fidelity. Heterogeneous methylation patterns have also been observed in tissue culture (46–48), in tumor tissue (49) and in filamentous fungi (50,51), all of which are experimental systems where maintenance of particular epigenotypes was expected to be stable. Our analysis at the resolution of individual DNA molecules from human leukocytes supports the finding that broad variability of methylation patterns is typical for hypermethylated alleles of a given cell population (47).

Clinical significance of occasional lack of methylation at sites on hypermethylated alleles

The degree to which individual CpGs are methylated in hypermethylated alleles depends on the site and the nature of the allele. Most CpG sites have a lower degree of methylation on normal, inactive X alleles when compared with fragile-X alleles (Fig. 6). For normal inactive X alleles both ends of the region analyzed showed reduced methylation compared with the central portion between sites 7 and 17. Fragile-X hypermethylated alleles mirrored this pattern for the central and right end regions. Site 18 is the clearest example of a site-dependent difference in methylation frequency, as it has the lowest degree of methylation on both sets of hypermethylated alleles. In contrast, sites 8 and 13 were always methylated on both inactive X and fragile-X alleles. The overall higher degree of methylation on fragile-X alleles may indicate a more pronounced stimulation of *de novo* methylation activity, triggered by an expanded CGG repeat and a regional change in the DNA structure. The difference in the degree of hypermethylation confirms an earlier study where densitometric analysis of DNA blots indicated that certain restriction sites are methylated at a somewhat higher frequency in DNA from alleles with large CGG repeat expansions compared with normal, inactive X (24).

These site-specific differences in methylation frequencies may influence some aspects of the diagnosis of individuals tested for fragile-X syndrome. The restriction enzyme *Eag*I, widely used to assess methylation status of *FMR1* (40), includes sites 16 and 17 in its recognition sequence. We find that these sites accurately reflect the methylation status of an allele, thus validating the use of this enzyme. In contrast, restriction enzyme *Bss*HII, which is occasionally used to assess the methylation status (45,52), includes sites 18 and 19 in its recognition sequence. We find that in hypermethylated alleles these sites are generally methylated at a lower frequency than other sites in leukocyte DNA. The use of *Bss*HII could therefore overestimate the proportion of unmethylated alleles in a fragile-X patient.

Potential transcription factor binding site that influences methylation

Of the 22 CpGs analyzed on the *FMR1* promoter, sites 18–22 exhibited the most intriguing variations, with sites 18 and 19 being methylated less frequently than average on hypermethylated normal female and fragile-X alleles (Fig. 6). Site-specific variations in methylation have also been observed in hypermethylated CpG islands of the phosphoglycerate kinase (*PGK1*) and the hypoxanthine phosphorybosyltransferase (*HPRT*) genes (53,54). The CpGs of Sp1 binding sites are almost completely devoid of methylation on an otherwise hypermethylated *HPRT* promoter (54). Binding elements for the Sp1 transcription factor protect CpG islands from methylation and may attract factors for removal of methylated cytosines (48,55,56). Although the sequence around sites 18–22 of the *FMR1* promoter does not correspond to an Sp1 binding element, these sites are part of the 14 bp palindromic DNA sequence 5′-GCGCGCATGCGCGC-3′. A database search revealed that this sequence has high similarity with the α-PAL binding element 5′-TGCGCATGCGCA-3′ (57). α-PAL is a palindrome binding transcription factor with strong sequence similarity to ewg, a protein that directs the expression of neuronal and flight muscle genes in *Drosophila* embryos (58,59). We speculate that α-PAL or a related transcription factor binds to the DNA sequence containing sites 18–22 and influences expression of *FMR1*. In general, detailed methylation patterns may identify *cis*-acting DNA promoter elements that are likely to function *in vivo*.

Cytosine methylation at non-CpG sites and rare CpG methylation events on hypomethylated alleles: do they reflect *in vivo* events?

Interpretation of rare non-conversion events is a general concern when the bisulfite genomic sequencing technique is applied because not all cytosine residues are deaminated with the same efficiency during the chemical conversion (36). Even with our large data set we cannot easily distinguish inefficient bisulfite conversion at a particular site from occasional *in vivo* cytosine methylation. CpG sites 19 and 21, for example, had conversion rates of 92.5% and 96.2% on hypomethylated male alleles, compared with an average conversion rate of 99.5% (Fig. 6). These two sites were frequently converted on hypermethylated alleles however, indicating that they are not particularly resistant to the chemical reaction. As mentioned, CpG sites 18–22 are part of a palindromic DNA element that has the potential to form a cruciform structure. The potential of this element to form a hairpin structure may lower the conversion efficiency, as the sodium bisulfite conversion procedure relies on single-stranded DNA. Alternatively, the occasional CpG non-conversion events on otherwise hypomethylated alleles may reflect *in vivo* methylation. Cytosines in cruciform DNA structures such as those at sites 19 and 21 can be methylated *de novo* by the human methyltransferase at efficiencies close to that of methylation of hemimethylated DNA (60); this process presumably occurs *in vivo* as well as *in vitro*.

Similar ambiguity exists for the occasional failure of chemical conversion of non-CpG cytosines, estimated to occur at an overall frequency of 0.5%. We observed an especially high non-conversion rate of 5% for the non-CpG cytosine embedded within the 14 nt palindromic sequence spanning sites 18–22. The other two protected non-CpG cytosines were found at CpCpG sites 7 and 10, which are not within a palindrome.

These sites are good candidates for non-CpG cytosines that are occasionally methylated *in vivo* in mammalian cells (41). The very low frequency of these non-CpG methylation events, however, indicates that this kind of methylation may not have a strong biological function at the *FMR1* promoter in leukocytes.

Mechanisms of methylation propagation

The epigenetic variations reported here provide insight into the rules that govern the propagation of methylation patterns. During certain phases of development a genomic locus may acquire an 'archetype' status for methylation; this status represents one of two pronounced bimodalities, hypo- or hypermethylated, and is inherited by daughter cells in a clonal and stable fashion. However, methylation of individual CpG sites within such an archetype is not strictly clonal and can differ among alleles. The methylation status of an individual CpG site is thus influenced by the methylation fate of the genomic region and also by specific DNA sequences, such as the α-PAL consensus region in the *FMR1* promoter.

The mechanisms that give rise to these properties of DNA methylation are not well understood. Silva and White proposed that the frequency of modification of a certain CpG site depends on *cis*- and *trans*-acting factors that modulate the activity of the maintenance methyltransferase in a concerted and tissue-specific fashion (49). Steric hindrance of transcription factors bound to DNA may lower the activity of the maintenance methyltransferase (19,47), as could be the case for CpG site 18 of the *FMR1* promoter. It has been proposed that enzymatic or ribozymatic demethylation of CpG islands occurs *in vivo* (47,60,61), which may involve promoter binding proteins that attract an otherwise non-specific demethylase, leading to reduced levels of modification at individual sites (47,48,55,56).

In these two models of factors that either inhibit the methyltransferase or attract a demethylase to a hypermethylated promoter, methylation would be lost over time if the maintenance methyltransferase were strictly copying patterns of hemi-methylated DNA. Low levels of *de novo* methylation are consequently required to maintain the hypermethylated state in dynamic balance with loss of methylation, as suggested by Pfeifer *et al*. (47). The mammalian methyltransferase has indeed been demonstrated to have a *de novo* methylation activity (63).

Concluding remarks

Our observations on methylation of the *FMR1* locus in leukocytes indicate that:

(i) the variability in epigenotypes from cells of each individual is striking, suggesting that methylation patterns in both fragile-X and normal individuals are dynamic rather than completely stable;

(ii) methylation of non-CpG cytosines in the region studied may occur but is rare, making *in vivo* methylation difficult to distinguish from inefficient bisulfite conversion at these sites;

(iii) the bimodal nature of methylation of the *FMR1* promoter in mosaic fragile-X males is remarkably similar to that found for the active X and inactive X alleles in normal females, suggesting that the methylation mosaicism of some fragile-X males reflects the same on and off states of *FMR1* expression that exist in normal females;

(iv) subtle differences distinguish the populations of methylated normal alleles from the populations of methylated expanded alleles, thus verifying previous observations that the CGG expansion in fragile-X individuals can augment the degree of methylation above that observed for normal X inactivation;

(v) the general pattern of bimodality in methylation of individual alleles persisted over a 5 year period that separated samplings of five fragile-X males; strict clonality, which would be indicated by the persistence of specific epigenotypes over this period, was rare;

(vi) methylation variability was most pronounced at a consensus binding sequence for transcription factor α-PAL, a sequence that we propose has a role in regulating expression of *FMR1*.

MATERIALS AND METHODS

DNA

Genomic DNAs from the fragile-X family were isolated from peripheral blood samples as described by McConkie-Rosell *et al*. (31). The 1986 DNA samples were used to determine the percentage of cytosine methylation at the *Eag*I site in the McConkie-Rosell *et al*. study after double digestion with *Eco*RI and *Eag*I and subsequent hybridization with probe StB12.3 (31). The following data on *FMR1* CGG repeats from males of the fragile-X family are taken from table 2 of the McConkie-Rosell *et al*. study (31): II-1 is mosaic with unmethylated alleles of ~66 and 300 repeats and methylated alleles of ~300 repeats; II-2 is mosaic with unmethylated alleles ranging between ~130 and 200 repeats and methylated alleles of ~170 repeats; II-3 is mosaic with unmethylated alleles ranging between ~100 and 270 repeats and methylated alleles of ~170 repeats; II-4 has a normal allele with ~29 repeats; II-5 is mosaic with methylated alleles ranging between ~130 and 670 repeats; II-6 is mosaic with unmethylated alleles ranging between 170 and 200 repeats and methylated alleles of ~200 and 530 repeats; I-2 has a normal allele with ~29 CGG repeats and an allele with an expanded repeat in the pre-mutation range (~65 repeats). Leukocyte DNAs from females L2 and L7 were isolated following a standard phenol/chloroform extraction method. L2 has two alleles with normal CGG repeats (~29 repeats), as determined by double digestion of DNA with *Eco*RI and *Eag*I and subsequent hybridization with probe StB12.3 (data not shown). L7 has two alleles with normal CGG repeats (~29 repeats), as determined by amplification of the *FMR1* CGG repeat by PCR (data not shown).

Sodium bisulfite conversion

With a few minor modifications, the bisulfite conversion of genomic DNA was carried out following the protocol developed and described by Clark *et al*. (36). Approximately 10 ng unsheared, genomic DNA and 2 µg yeast tRNA (carrier nucleic acid) were denatured by adding freshly prepared NaOH to a final concentration of 0.3 M in a 10 µl reaction volume and incubated at 42°C for 30 min. Increasing the stringency of the denaturing

conditions (42°C/30 min instead of 37°C/15 min) improved the conversion rate of non-CpG cytosines. The more stringent denaturing conditions are perhaps necessary because the *FMR1* promoter sequence has a high GC content (66% GC-rich) and therefore an overall higher T_m. Fresh solutions of 3.8 M sodium bisulfite (S-8890; Sigma), adjusted to pH 5 with NaOH, and 20 mM hydroquinone (H-9003; Sigma) were prepared by gentle mixing at 37°C. Final concentrations of 3.4 M sodium bisulfite and 1 mM hydroquinone were added to the denatured DNA to a final volume of 100 μl. The DNA was gently mixed in this sodium bisulfite/hydroquinone solution, overlaid with mineral oil and incubated in the dark at 55°C for 6 h. After recovering the aqueous phase from under the oil, unbound bisulfite was removed from the DNA by use of MicroSpin S-200 HR columns (Pharmacia Biotech). The purified DNA sample was subsequently mixed and incubated with freshly prepared NaOH (0.3 M final concentration) at 37°C for 20 min. NaOH was removed by use of MicroSpin S-200 HR columns and the flow-through (~100 μl), containing the converted DNA, was frozen and stored until aliquots were used as PCR templates.

PCR

The primers were designed to amplify bisulfite-converted DNA (top strand) of the *FMR1* promoter and have the following sequences: 5′-GGAATTTTAGAGAGGT$^{C}/_{T}$GAATTGGG-3′ (1F), positions 2246–2270; 5′-GTTATTGAGTGTATTTTTGTAGAAATGGG-3′ (2F), positions 2296–2325; 5′-CCCTCTCTCTTCAAATAACCTAAAAAC-3′ (3R), positions 2492–2466. (The numbering of the base residues corresponds to the unconverted, genomic human *FMR1* DNA sequence, accession no. X61378.) The fragment of interest was amplified by semi-nested PCR in 25 μl reaction mixtures. The first PCR contained 10 μl bisulfite-treated genomic DNA, 200 μM dNTPs, 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 0.5 mM primers (1F and 3R) and 1.25 U Taq polymerase, overlaid with 25 μl mineral oil. The amplification was performed in a Hybaid Omn-E thermal cycler under the following conditions: 94°C for 2 min for one cycle; 94°C for 15 s, 57°C for 15 s, 72°C for 1 min for 29 cycles; 72°C for 2 min for one cycle. The second PCR contained 1 μl reaction mixture from the first PCR, 200 μM dNTPs, 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 3 mM $MgCl_2$, 0.01% gelatin, 0.5 mM primers (2F and 3R) and 1.25 U Taq polymerase and was overlaid with 25 μl mineral oil. The amplification was performed in a Hybaid Omn-E thermal cycler under the following conditions: 94°C for 2 min for one cycle; 94°C for 15 s, 65°C for 15 s, 73°C for 30 s for 29 cycles; 72°C for 2 min for one cycle.

This semi-nested PCR is extremely sensitive; minute amounts of template DNA can be amplified. Special care was taken during the sodium bisulfite conversion and the subsequent PCR procedures in order to avoid contamination. At least two DNA-negative PCR controls were performed for every DNA sample that was amplified.

Cloning, sequencing and analysis of PCR products

The PCR products generated with primer pair 2F/3R were run on a 2% agarose gel and isolated using the MERmaid kit (Bio 101). Isolated fragments were ligated into plasmid pCR2.1 and the ligation product transformed into competent *Escherichia coli* (INVaF′) according to the protocols of the manufacturer (TA cloning kit; Invitrogen). DNA of randomly picked clones was isolated (QIAprep; Qiagen) and sequenced using either the Sequenase version 2.0 DNA sequencing kit (US Biochemials) or the ABI PRISM dye terminator cycle sequencing ready reaction kit (Perkin Elmer). Sequence analysis was performed with computer programs of the Wisconsin Package, V.8.1-Unix (Genetics Computer Group, Madison, WI) and SIGNAL SCAN (64). DNA samples from the fragile-X family were coded and the epigenotypes matched with the individuals only after the bisulfite conversion and cloning had been performed.

ACKNOWLEDGEMENTS

We thank Scott Hansen, Stan Gartler, Douglas Kim and members of the Laird laboratory for critical comments on the manuscript, Susan Clark for helpful advice on the bisulfite sequencing method and Tracy Stapp for high quality sequencing. R.S. was supported by an Erwin-Schrödinger fellowship from the Austrian Science Foundation and a fellowship from the FRAXA Research Foundation. This work was supported by National Institutes of Health grant GM53805.

REFERENCES

1. Silva,A.J. and White,R. (1988) Inheritance of allelic blueprints for methylation patterns. *Cell*, **54**, 145–152.
2. Gartler,S.M., Dyer,K.A., Graves,J.A.M. and Rocchi,M. (1985) A two step model for mammalian X-chromosome inactivation. In Cantoni,G.L. and Razin,A. (eds), *Biochemistry and Biology of DNA Methylation*. Alan R.Liss, New York, NY, pp. 223–235.
3. Lyon,M.F. (1988) X-chromosome inactivation and the location and expression of X-linked genes. *Am. J. Hum. Genet*., **42**, 8–16.
4. Riggs,A.D. and Pfeifer,G.P. (1992) X-chromosome inactivation and cell memory. *Trends Genet*., **8**, 169–174.
5. Monk,M., Boubelik,M. and Lehnert,S. (1987) Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development*, **99**, 371–382.
6. Kafri,T., Ariel,M., Brandeis,M., Shemer,R., Urven,L., McCarrey,J., Cedar,H. and Razin,A. (1992) Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev*., **6**, 705–714.
7. Li,E., Bestor,T.H. and Jaenisch,R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915–926.
8. Dittrich,B., Robinson,W.P., Knoblauch,H., Buiting,K., Schmidt,K., Gillessen-Kaesbach,G. and Horsthemke,B. (1992) Molecular diagnosis of the Prader–Willi and Angelman syndromes by detection of parent-of-origin specific DNA methylation in 15q11–13. *Hum. Genet*., **90**, 313–315.
9. Driscoll,D.J., Waters,M.F., Williams,C.A., Zori,R.T., Glenn,C.C., Avidano,K.M. and Nicholls,R.D. (1992) A DNA methylation imprint, determined by the sex of the parent, distinguishes the Angelman and Prader–Willi syndromes. *Genomics*, **13**, 917–924.
10. Reis,A., Dittrich,B., Greger,V., Buiting,K., Lalande,M., Gillessen-Kaesbach,G., Anvret,M. and Horsthemke,B. (1994) Imprinting mutations suggested by abnormal DNA methylation patterns in familial Angelman and Prader–Willi syndromes. *Am. J. Hum. Genet*., **54**, 741–747.
11. Bell,M.V., Hirst,M.C., Nakahori,Y., MacKinnon,R.N., Roche,A., Flint,T.J., Jacobs,P.A., Tommerup,N., Tranebjaerg,L., Froster-Iskenius,U., Kerr,B., Turner,G., Lindenbaum,R.H., Winter,R., Pembrey,M., Thibodeau,S. and Davies,K.E. (1991) Physical mapping across the fragile X: hypermethylation and clinical expression of the fragile X syndrome. *Cell*, **64**, 861–866.
12. Vincent,A., Heitz,D., Petit,C., Kretz,C., Oberlé,I. and Mandel,J. (1991) Abnormal pattern detected in fragile X patients by pulsed field gel electrophoresis. *Nature*, **349**, 624–626.
13. Oberlé,I., Rousseau,F., Heitz,D., Kretz,C., Devys,D., Hanauer,A., Boué,J., Bertheas,M.F. and Mandel,J.L. (1991) Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science*, **252**, 1097–1102.
14. Zeschnigk,M., Schmitz,B., Dittrich,B., Buiting,K., Horsthemke,B. and Doerfler,W. (1997) Imprinted segments in the human genome—different DNA methylation patterns in the Prader–Willi/Angelman syndrome region as determined by the genomic sequencing method. *Hum. Mol. Genet*., **6**, 387–395.

15. Holliday,R. and Pugh,J.E. (1975) DNA modification mechanisms and gene activity during development. *Science*, **187**, 226–232.
16. Riggs,A.D. (1975) X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet*., **14**, 9–25.
17. Sager,R. and Kitchin,R. (1975) Selective silencing of eukaryotic DNA. *Science*, **189**, 426–433.
18. Meehan,R., Lewis,J., Cross,S., Nan,X., Jeppesen,P. and Bird,A. (1992) Transcriptional repression by methylation of CpG. *J. Cell. Sci*., **16** (suppl.), 9–14.
19. Bird,A. (1986) CpG-rich islands and the function of DNA methylation. *Nature*, **321**, 209–213.
20. Antequera,F. and Bird,A. (1993) Number of CpG islands and genes in human and mouse. *Proc. Natl. Acad. Sci. USA*, **90**, 11995–11999.
21. Razin,A. and Cedar,H. (1994) DNA methylation and genomic imprinting. *Cell*, **77**, 473–476.
22. Eden,S. and Cedar,H. (1994) Role of DNA methylation in the regulation of transcription. *Curr. Opin. Genet. Dev*., **4**, 255–259.
23. Wöhrle,D. and Steinbach,P. (1991) Fragile X expression and X inactivation. II. The fragile site at Xq27.3 has a basic function in the pathogenesis of fragile X-linked mental retardation. *Hum. Genet*., **87**, 421–424.
24. Hansen,R.S., Gartler,S.M., Scott,C.R., Chen,S.-H. and Laird,C.D. (1992) Methylation analysis of CGG sites in the CpG island of the human *FMR1* gene. *Hum. Mol. Genet*., **1**, 571–578.
25. Hornstra,I.K., Nelson,D.L., Warren,S.T. and Yang,T.P. (1993) High resolution methylation analysis of the *FMR-1* gene trinucleotide repeat region in Fragile X syndrome. *Hum. Mol. Genet*., **2**, 1659–1665.
26. Dietrich,A., Kioschis,P., Monaco,A.P., Gross,B., Korn,B., Williams,S.V., Sheer,D., Heitz,D., Oberlé,I., Toniolo,D., Warren,S.T., Lehrach,H. and Poustka,A. (1991) Molecular cloning and analysis of the fragile X region in man. *Nucleic Acids Res*., **19**, 2567–2572.
27. Pieretti,M., Zhang,F.P., Fu,Y.H., Warren,S.T., Oostra,B.A., Caskey,C.T. and Nelson,D.L. (1991) Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell*, **66**, 817–22.
28. Hagerman,R., Kemper,M. and Hudson,M. (1985) Learning disabilities and attentional problems in boys with the fragile X syndrome. *Am. J. Dis. Child*., **139**, 674–678.
29. Veenema,H., Veenema,T. and Geraedts,J.P.M. (1987) The fragile X syndrome in a large family. II Psychological investigations. *J. Med. Genet*., **24**, 32–38.
30. Loesch,D.Z., Huggins,R., Hay,D.A., Gedeon,A.K., Mulley,J.C. and Sutherland,G.R. (1993) Genotype–phenotype relationships in fragile X syndrome: a family study. *Am. J. Hum. Genet*., **53**, 1064–1073.
31. McConkie-Rosell,A., Lachiewicz,A.M., Spiridigliozzi,G.A., Tarleton,J., Schoenwald,S., Phelan,M.C., Goonewardena,P., Ding,X. and Brown,W.T. (1993) Evidence that methylation of FMR-1 locus is responsible for variable phenotypic expression of the fragile x syndrome. *Am. J. Hum. Genet*., **53**, 800–809.
32. de Vries,B.B., Jansen,C.C., Duits,A.A., Verheij,C., Willemsen,R., van Hemel,J.O., van den Ouweland,A.M., Niermeijer,M.F., Oostra,B.A. and Halley,D.J. (1996) Variable FMR1 gene methylation of large expansions leads to variable phenotype in three males from one fragile X family. *J. Med. Genet*., **33**, 1007–1010.
33. Heitz,D., Rousseau,R., Devys,D., Saccone,S., Abderrahim,H., LePaslier,D., Cohen,D., Vincent,A., Toniolo,D., Della Valle,G., Johnson,S., Schlessinger,D., Oberlé,I. and Mandel,J.L. (1991) Isolation of sequences that span the Fragile X and identification of a Fragile X-related CpG island. *Science*, **251**, 1236–1239.
34. Brown,W.T., Gross,A., Goonewardena,P., Ferrando,C., Dobkin,C. and Jenkins,E.C. (1991) Detection of fragile X non-penetrant males by DNA marker analysis. *Am. J. Med. Genet*., **38**, 292–297.
35. Frommer,M., McDonald,L.E., Millar,D.S., Collis,C.M., Watt,F., Grigg,G.W., Molloy,P.L. and Paul,C.L. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA*, **89**, 1827–1831.
36. Clark,S.J., Harrison,J., Paul,C.L. and Frommer,M. (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Res*., **22**, 2990–2997.
37. Gruenbaum,Y., Cedar,H. and Razin,A. (1982) Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature*, **295**, 620–622.
38. Bestor,T.H. and Ingram,V.M. (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc. Natl. Acad. Sci. USA*, **80**, 5559–5563.
39. Feng,Y., Lakkis,L., Devys,D. and Warren,S.T. (1995) Quantitative comparison of FMR1 gene expression in normal and premutation alleles. *Am. J. Hum. Genet*., **56**, 106–113.
40. Rousseau,F. *et al*. (1991) Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. *New Engl. J. Med*., **325**, 1673–1681.
41. Clark,S.J., Harrison,J. and Frommer,M. (1995) CpNpG methylation in mammalian cells. *Nature Genet*., **10**, 20–27.
42. Lyon,M. (1961) Gene action in the X chromosome of the mouse (*Mus musculus* L.). *Nature*, **190**, 372–373.
43. McGowan,R., Campbell,R., Peterson,A. and Sapienza,C. (1989) Cellular mosaicism in the methylation and expression of hemizygous loci in the mouse. *Genes Dev*., **3**, 1669–1676.
44. Rousseau,F., Heitz,D., Oberlé,I. and Mandel,J.-L. (1991) Selection in blood cells from female carriers of the fragile X syndrome: inverse correlation between age and proportion of active X chromosomes carrying the full mutation. *J. Med. Genet*., **28**, 830–836.
45. Malter,H.E., Iber,J.C., Willemsen,R., de Graaff,E., Tarleton,J.C., Leisti,J., Warren,S.T. and Oostra,B.A. (1997) Characterization of the full fragile X syndrome mutation in fetal gametes. *Nature Genet*., **15**, 165–169.
46. Reis,R.J. and Goldstein,S. (1982) Interclonal variation in methylation patterns for expressed and non-expressed genes. *Nucleic Acids Res*., **10**, 4293–4304.
47. Pfeifer,G.P., Steigerwald,S.D., Hansen,R.S., Gartler,S.M. and Riggs,A.D. (1990) Polymerase chain reaction-aided genomic sequencing of an X chromosome-linked CpG island: methylation patterns suggest clonal inheritance, CpG site autonomy, and an explanation of activity state stability. *Proc. Natl. Acad. Sci. USA*, **87**, 8252–8256.
48. Macleod,D., Charlton,J., Mullins,J. and Bird,A.P. (1994) Sp1 sites in the mouse *aprt* gene promoter are required to prevent methylation of the CpG island. *Genes Dev*., **8**, 2282–2292.
49. Silva,A.J., Ward,K. and White,R. (1993) Mosaic methylation in clonal tissue. *Dev. Biol*., **156**, 391–398.
50. Selker,E.U., Fritz,D.Y. and Singer,M.J. (1993) Dense nonsymmetrical DNA methylation resulting from repeat-induced point mutation in *Neurospora*. *Science*, **262**, 1724–1728.
51. Goyon,C., Nogueira,T.I. and Faugeron,G. (1994) Perpetuation of cytosine methylation in *Ascobolus immersus* implies a novel type of maintenance methylase. *J. Mol. Biol*., **240**, 42–51.
52. Meijer,H., de Graaff,E., Merckx,D.M., Jongbloed,R.J., de Die-Smulders,C.E., Engelen,J.J., Fryns,J.P., Curfs,P.M. and Oostra,B.A. (1994) A deletion of 1.6 kb proximal to the CGG repeat of the *FMR1* gene causes the clinical phenotype of the fragile X syndrome. *Hum. Mol. Genet*., **3**, 615–620.
53. Pfeifer,G.P., Tanguay,R.L., Steigerwald,S.D. and Riggs,A.D. (1990) *In vivo* footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes Dev*., **4**, 1277–1287.
54. Hornstra,I.K. and Yang,T.P. (1994) High-resolution methylation analysis of the human hypoxanthine phosphoribosyltransferase gene 5′ region on the active and inactive X chromosomes: correlation with binding sites for transcription factors. *Mol. Cell. Biol*., **14**, 1419–1430.
55. Brandeis,M., Frank,D., Keshet,I., Siegfried,Z., Mendelsohn,M., Nemes,A., Temper,V., Razin,A. and Cedar,H. (1994) Sp1 elements protect a CpG island from *de novo* methylation. *Nature*, **371**, 435–438.
56. Weiss,A., Keshet,I., Razin,A. and Cedar,H. (1996) DNA demethylation *in vitro*: involvement of RNA. *Cell*, **86**, 709–718.
57. Jacob,W.F., Silverman,T.A., Cohen,R.B. and Safer,B. (1989) Identification and characterization of a novel transcription factor participating in the expression of eukaryotic initiation factor 2 alpha. *J. Biol. Chem*., **264**, 20372–20384.
58. DeSimone,S.M. and White,K. (1993) The *Drosophila erect wing* gene, which is important for both neuronal and muscle development, encodes a protein which is similar to the sea urchin P3A2 DNA binding protein. *Mol. Cell .Biol*., **13**, 3641–3649.
59. Efiok,B.J., Chiorini,J.A. and Safer,B. (1994) A key transcription factor for eukaryotic initiation factor-2 alpha is strongly homologous to developmental transcription factors and may link metabolic genes to cellular growth and development. *J. Biol. Chem*., **269**, 18921–18930.
60. Laayoun,A. and Smith,S.S. (1995) Methylation of slipped duplexes, snapbacks and cruciforms by human DNA(cytosine-5)methyltransferase. *Nucleic Acids Res*., **23**, 1584–1589.
61. Antequera,F., Boyes,J. and Bird,A. (1990) High levels of *de novo* methylation and altered chromatin structure at CpG islands in cell lines. *Cell*, **62**, 503–514.
62. Frank,D., Keshet,I., Shani,M., Levine,A., Razin,A. and Cedar,H. (1991) Demethylation of CpG islands in embryonic cells. *Nature*, **351**, 239–241.
63. Bestor,T.H. (1992) Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J*., **11**, 2611–2617.
64. Prestridge,D.S. (1991) SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *CABIOS*, **7**, 203–206.