DNA methylation represses transcription *in vivo*

Zahava Siegfried[1], Sharon Eden[1], Monica Mendelsohn[2], Xu Feng[1], Ben-Zion Tsuberi[1] & Howard Cedar[1]

DNA in somatic tissue is characterized by a bimodal pattern of methylation, which is established in the animal through a series of developmental events[1]. In the mouse blastula, most DNA is unmethylated, but after implantation a wave of *de novo* methylation modifies most of the genome, excluding the majority of CpG islands, which are mainly associated with housekeeping genes. This genomic methylation pattern is broadly maintained during the life of the organism by maintenance methylation[2], and generally correlates with gene expression. Experiments both *in vitro*[3–5] and *in vivo*[6–9] indicate that methylation inhibits transcription. It has not yet been possible, however, to determine the role of DNA methylation on specific sequences during normal development. *Cis*-acting regulatory elements and *trans*-acting factors appear to be involved in both stage- and tissue-specific demethylation processes[10,11]. Sp1-like elements have a key role in protecting the CpG island of *Aprt* (encoding adenine phosphoribosyltransferase) from *de novo* methylation, and when these elements are specifically mutated, the *Aprt* CpG island becomes methylated in transgenic mice[12,13]. We have now characterized an embryo-specific element from the CpG island sequence upstream of *Aprt* that can protect itself from *de novo* methylation in transgenic mice as well as reduce methylation of flanking sequences. We placed this element on a removable cassette adjacent to a human *HBB* (encoding β-globin) reporter and generated a transgene whose methylation pattern can be switched *in vivo*. Analysis of globin transcription in this system showed that methylation in *cis* inhibits gene expression in a variety of tissues, indicating that DNA modification may serve as a global genomic repressor.

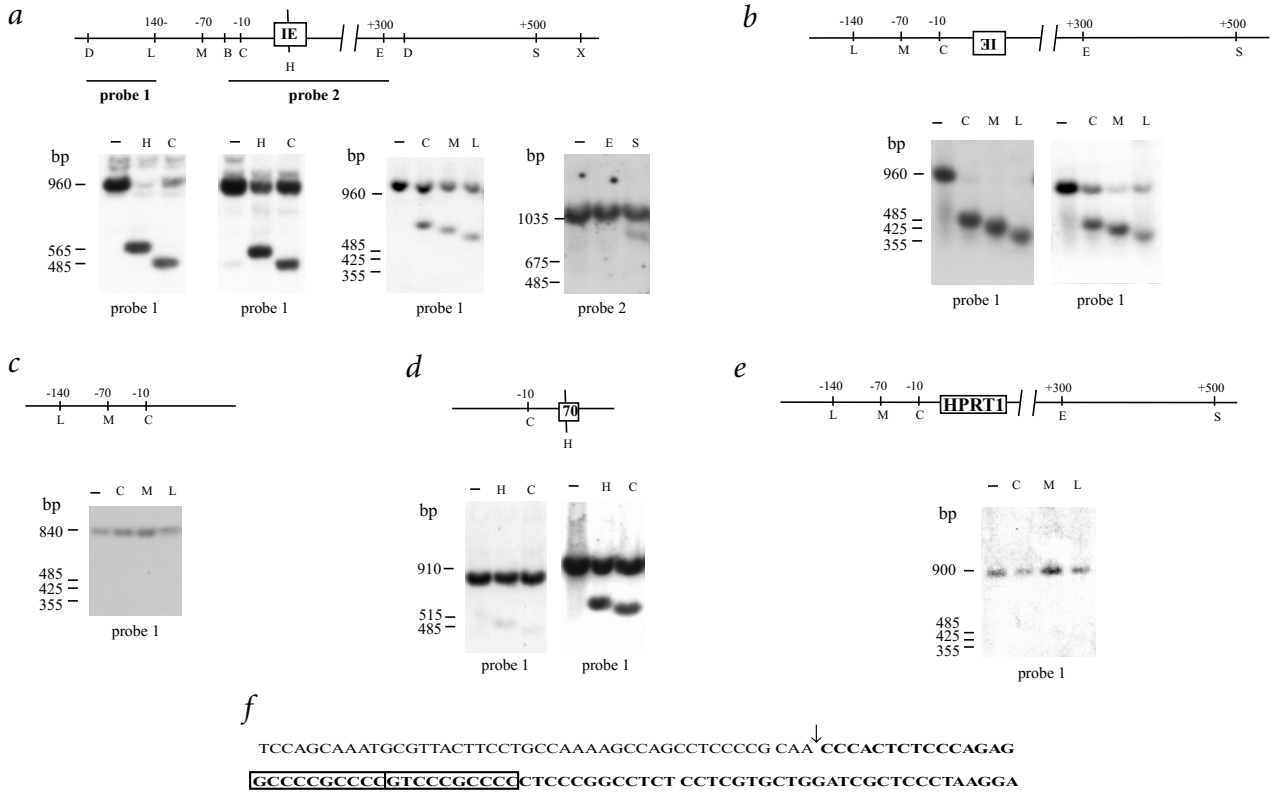

Fig. 1 The IE prevents methylation. We made transgenic embryos using a test plasmid containing various inserts and tested DNA for methylation at specific restriction sites. ***a***, The IE (12 founders, average 60% unmethylated). ***b***, The IE in reverse orientation (2 founders, average 90% unmethylated). ***c***, No insert (3 founders, all <5% unmethylated). ***d***, A 70-bp fragment from the IE (6 founders, average 20% unmethylated). ***e***, A 60-bp fragment from human *HPRT1* (***e***; 2 founders, all <5% unmethylated). We digested total founder embryo DNA with either *Dra*I alone (–) or together with *Cfo*I, *Sma*I, *Sal*I or *Hpa*II, or with *Ban*II and *Xba*I alone (–) or together with *Eag*I or *Sac*II. We performed Southern-blot analysis using probes 1 or 2 as labelled. The sizes of expected bands are indicated on the side of each autoradiogram. Significant undermethylation is seen exclusively in constructs with the IE in either orientation (*a* and *b*) or the smaller 70-bp fragment of the IE (*d*). The farther sites, *Eag*I and *Sac*II, remained methylated in all constructs tested (*a*, and data not shown). It should be noted that both probes contain human globin sequences, which may cross-react slightly with endogenous mouse globin, possibly explaining the additional minor bands seen in some blots. Data shown are representative of results obtained with several transgenic founder mice for each construct. Two separate founder mice are shown in (*b*,*d*) and three separate founder mice in (*a*,*f*).***f***, Sequence of the 120-bp IE, with the 2 Sp1-like elements in the boxed region. The 70-bp fragment includes the sequences 3′ to the arrow. B, *Ban*II; C, *Cfo*I; D, *Dra*I; E, *Eag*I; H, *Hpa*II; L, *Sal*I; M, *Sma*I; S, *Sac*II; X, *Xba*I.

[1]*Department of Cellular Biochemistry, Hebrew University Medical School, Jerusalem 91120, Israel.* [2]*Center for Neurobiology and Behavior, Columbia University Medical School, New York, New York 10032, USA. Correspondence should be addressed to H.C. (e-mail: cedar@md2.huji.ac.il).*

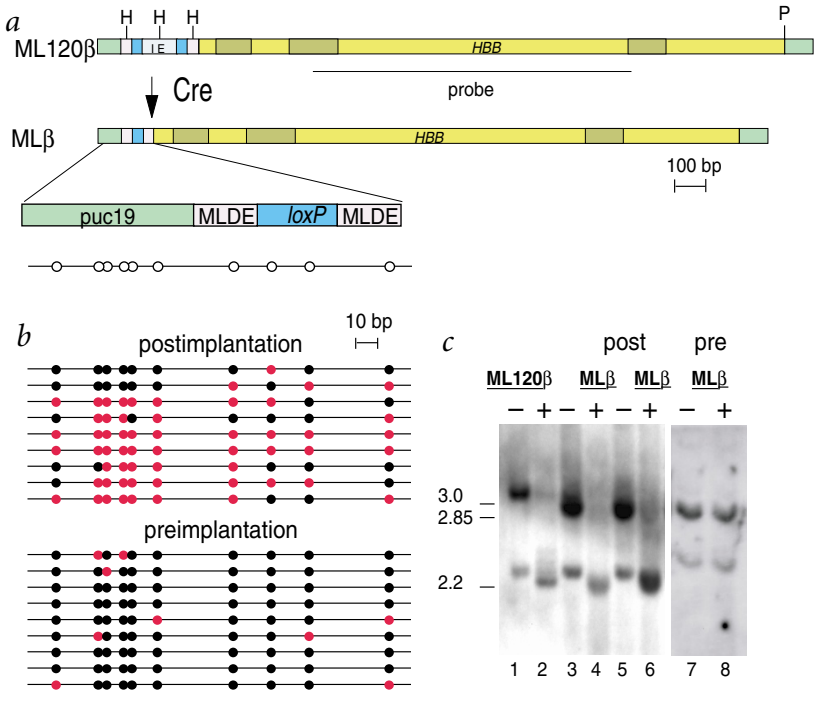

Fig. 2 IE removal switches methylation patterns *in vivo*. **a**, Diagram of the construct ML120β showing human *HBB* sequences (yellow with exons marked), including a 48-bp minimal promoter, two 30-bp MLDE sequences (pink), two 34-bp *loxP* elements (dark blue), the 120-bp IE (light blue) and flanking plasmid sequences (green). After Cre expression, the IE is removed, leaving one *loxP* element (MLβ). Circles indicate the locations of CpG residues. One *Pst*I site (P) is located 3′ to the *HBB* coding sequences, with another *Pst*I site in the mouse sequence 5′ to the integration site of the transgene. This yields a 3-kb band, which is reduced to 2.85 kb after removal of the IE. The probe is a 920-bp *Eco*RI–*Bam*HI fragment from *HBB*. H, *Hpa*II; P, *Pst*I. **b**, We analysed the region upstream of *HBB* for methylation at CpG residues in liver or spleen cells using the bisulfite technique[20]. The top set of molecules was derived from mice in which the IE was removed after implantation, whereas the bottom set was derived from mice in which the IE was deleted before implantation. In these animals, similar results were also obtained using tail DNA. Black dots indicate a methylated CpG, whereas red dots indicate a lack of methylation. The next 2 CpG residues in the 3′ direction are located approximately 200 bp from MLDE, and were found by bisulfite analysis to be unaffected by the IE (data not shown). **c**, We digested tail or spleen DNA from transgenic mice with either *Pst*I alone (–) or *Pst*I and *Hpa*II (+). Lanes 1,2, DNA from the founder ML120β; lanes 3,4, DNA from mice containing the transgene and a gene encoding Cre under inducible control, following induction with interferon at age 2 months; lanes 5,6, the same as 3 and 4 except the mouse was sacrificed three months after induction of Cre; lanes 7,8, DNA from mice containing the transgene and a gene encoding Cre under constitutive control, which is expressed before implantation. Southern-blot analysis using the probe labelled in (*a*) detects methylation in both MLDE elements together. The smaller band seen in the lanes digested with *Pst*I alone is due to cross-hybridization of mouse globin sequences to the human probe. The methylated or undermethylated state was retained throughout the life of the animal, and was re-established after mating (data not shown).

To characterize a minimal element capable of protecting CpGs from *de novo* methylation, we inserted a 120-bp sequence containing 2 Sp1 sites from the hamster *Aprt* CpG island into a vector[12] consisting of non-CpG island DNA (Fig. 1). We then engineered recognition sites for methylation-sensitive restriction endonucleases at various distances from this element. We generated transgenic embryos (13.5–15.5 days post-coitum) containing this construct and assayed embryonic genomic DNA to examine the methylation status of vector sequences. Both the *Hpa*II site in the island element and adjacent test sites were undermethylated (Fig. 1*a*), indicating that this island element (IE) was able to protect itself from *de novo* methylation as well as bring about undermethylation on sequences up to 150 bp away. In contrast, sites 300 and 500 bp from the insertion site were almost completely methylated in every founder (Fig. 1*a*), as were all sites analysed in the control construct lacking the IE (Fig. 1*c*). Although levels of undermethylation varied between mice (60% on average), all sites underwent demethylation to approximately the same degree in given founders, suggesting that the IE works in a regional manner.

We then reversed the IE and showed that it prevented *de novo* methylation of non-island sequences independent of orientation (Fig. 1*b*). A shorter fragment of 70 bp (Fig. 1*f*) contained within the 120-bp sequence also induced undermethylation, but was less effective (20% on average) than the larger fragment (Fig. 1*d*). Moreover, neither Sp1-like binding sites alone (data not shown) nor a sequence from the gene encoding human X-linked hypoxanthine phosphoribosyltransferase (*HPRT1*, containing a cluster of four Sp1-binding sites; ref. 14) prevented *de novo* methylation (Fig. 1*e*). This suggests that such elements are necessary[12], but not sufficient, for demethylation *in vivo*.

To test whether DNA methylation inhibits gene transcription *in vivo*, we used the IE to artificially undermethylate a tissue-specific promoter region and examined expression of a reporter. Our reporter construct consisted of two copies of the downstream enhancer[15,16] (MLDE) of *Lzm-s1* (encoding mouse M lysozyme) separated by an IE and linked to *HBB* driven by its own minimal promoter (ML120β; Fig. 2*a*). The IE was flanked by *loxP* elements to allow Cre-mediated excision[17]. Although factors that bind MLDE are expressed ubiquitously, MLDE is methylated in most cell types and appears to interact with these factors only in macrophage cells, where it is specifically unmethylated[15,16].

We generated several founders using this construct and carried out methylation and expression analysis in one single-copy line (ML120β). The CpGs in the two MLDE sequences of this transgene were unmethylated (Fig. 2*c*). To generate mice carrying transgenes identical in sequence but either methylated or unmethylated, the *loxP*-flanked IE was deleted by mating carriers with two different Cre-expressing lines. In the first line, Cre is expressed constitutively before implantation[18]. In mice carrying this construct and the *loxP*-flanked IE, the IE was deleted before the wave of *de novo* methylation; MLDE sites became methylated and remained that way in all examined tissues (Fig. 2*c*, lanes 7,8). The second Cre-expressing line carries interferon-inducible Cre[19]. In mice with this gene and the *loxP*-flanked IE, the IE remained present during implantation in the absence of interferon application and protected adjacent regions from methylation. By treating adult animals with interferon, we induced Cre expression and removed the IE, generating an unmethylated version of the transgene. Even one year after the IE was deleted, the transgene remained unmethylated at the MLDE (Fig. 2*c*, and data not shown), showing that CpG island undermethylation is maintained in the absence of elements required for its establishment.

To determine more precisely whether the IE influences DNA methylation in a regional manner, we subjected liver or spleen

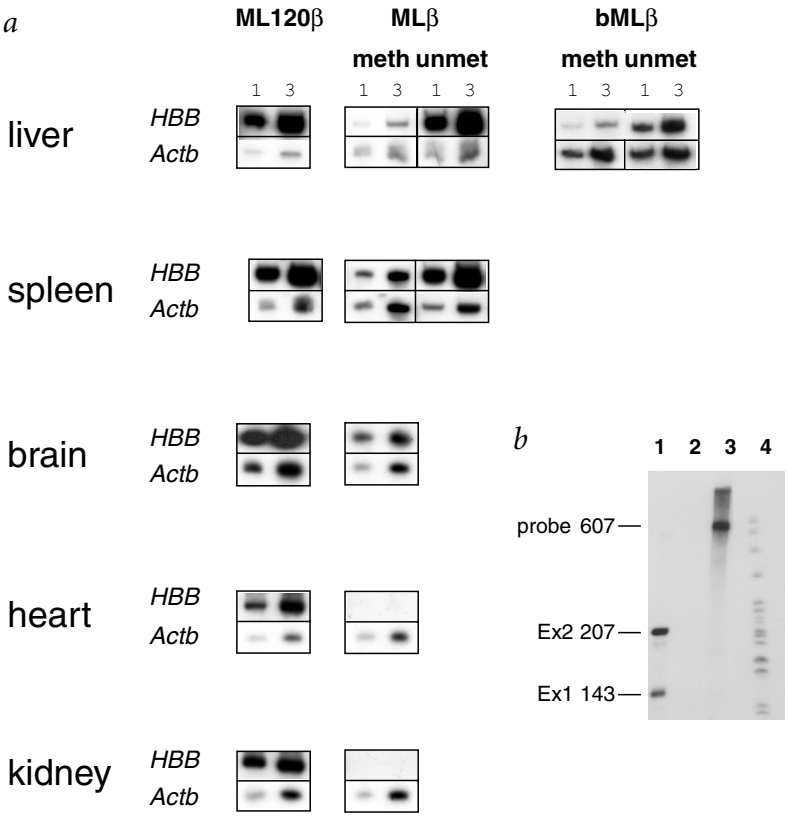

Fig. 3 DNA methylation represses background transcription. ***a***, We examined expression of the transgene and endogenous *Actb* by RT-PCR in liver, spleen, brain, kidney and heart. We measured *Actb* expression to control for the amount of cDNA in the reaction. DNA methylation reduced transcription in liver (30-fold), spleen (8-fold), heart and kidney (>50-fold), and brain (5-fold). We crossed another founder (bML120β) with Cre-containing mice to induce deletion (bMLβ), then demonstrated that these mice carried the appropriate methylated or unmethylated pattern similar to MLβ (data not shown). In this founder, bMLβ, DNA methylation inhibited basal transcription (sixfold) as well. ***b***, We transfected the ML120β construct into tk⁻ mouse L-cells and subjected total RNA to an RNase protection assay[28]. The 607-nt antisense probe (*Xba*I/*Bam*HI fragment) extended from the middle of exon 2 to a point upstream of the HBB promoter. Hybridization of this probe (42 °C overnight) to total RNA from transfected cells yielded 2 protected products which represent the start of transcription including exon 1 (143 bp) and the beginning of exon 2 (207 bp). Lane 1, L-cell RNA; lane 2, control yeast RNA; lane 3, yeast RNA assayed without RNase treatment; lane 4, marker DNA (pBR322 cut with *Msp*I).

DNA from both transgenic mouse lines to bisulfite sequencing, which allowed us to evaluate the methylation state of every cytosine[20]. The region upstream of MLDE sites was only 10% unmethylated in mice lacking the IE at implantation, whereas the same cytosine residues were 65% unmethylated on average in animals that retained the IE during early development (Fig. 2*b*). We did not observe methylation of cytosines not contained within CpG dinucleotides in any instance (data not shown).

Using mice containing the methylated or unmethylated transgenes, we then asked whether methylation status of the upstream region had an effect on expression of the linked *HBB* gene. We first ascertained that this construct directed transcription from the *HBB* promoter. RNase protection assays showed that correctly initiated RNA was synthesized in cells transfected with this plasmid (Fig. 3*b*). Although the same bands were observed weakly in tissue samples with methylated IE-deleted transgene (MLβ), levels were low and not amenable to quantitation (data not shown). For this reason, we used RT-PCR to analyse transcription. Transgene expression was considerably lower than that of a β-actin control. To accomodate this difference, we analysed RT-PCR products of *HBB* after 32 cycles and β-actin RT-PCR products after 18 cycles. We found that the level of expression was approximately 30-fold higher in the livers of mice containing the unmethylated transgene (Fig. 3*a*); a similar but more modest effect was observed in the spleen. We also saw inhibition (sixfold) of background transcription (Fig. 3*a*) in a second transgenic founder line (bML120β) that was genetically manipulated in the same manner.

Induction of Cre expression by interferon is most efficient (>90% deletion) in liver and spleen[19]; we were only able to compare expression of identical transgenes with different methylation states in those tissues. Because the level of unmethylated transgene expression was influenced little by the presence of the IE (Fig. 3*a*), it was possible to examine transcription levels in other tissues by comparing expression of ML120β with that of MLβ. DNA methylation also also inhibited expression in these cell types (Fig. 3*a*).

Previous studies showed that expression of specific genes is activated in DNA methyltransferase-deficient mouse embryos in which the entire genome is undermethylated[6–9]. These findings may be due to pleiotropic effects, but our results demonstrate that DNA methylation can operate in *cis* to significantly reduce background transcription in a range of different tissues *in vivo*. This repression may be partially mediated through the effects of methylation on MLDE binding itself[16], but modification of nearby cytosine residues (Fig. 2*a*) may also effect repression by altering histone acetylation and chromatin structure[21–23]. Although our test vector was constructed from heterologous gene elements, we propose that the results reflect the effects of DNA methylation on endogenous gene regions throughout the genome. An alternative interpretation of these experiments is that the IE functions to set up a heritable chromatin state early in development, with undermethylation being a consequence of this structure. We consider this possibility unlikely because the IE seems to lack enhancer activity (Fig. 3*a*) and these same island sequences have already been shown to be capable of directing specific demethylation on non-chromatin substrates *in vitro*[24].

The concept that methylation has a role in controlling gene expression during normal development has been challenged by the suggestion that this modification may function exclusively to inhibit the expression of foreign sequence elements or transposons within the genome[25,26]. We believe that our results may help resolve this issue. The wave of *de novo* methylation which takes place at the time of implantation serves as an essential developmental step in gene regulation. Given that the mouse has a pre-programmed mechanism to ensure that CpG islands at the 5′ ends of all housekeeping genes remain unmethylated, this may represent an inherent regulatory system aimed at modifying

genomic DNA non-specifically, whereas allowing gene expression where needed. Our studies show that this modification can repress the level of RNA synthesis associated with tissue-specific genes in the absense of transcriptional enhancers. It is likely that the methylation pattern covering a large fraction of DNA in every cell type serves as a global mechanism for reducing transcriptional activity. Within this context, foreign sequences that lack the elements necessary to become unmethylated during development will adopt a methylated pattern *in vivo* and be subjected to the same methyl-mediated gene repression mechanism[9].

Methods

Construction of plasmids. We used plasmid M234 (ref. 12) to generate the series of plasmids used here. This plasmid consists of 2 kb of non-island sequences with a single *Pst*1 site into which the inserts were cloned. We used site-directed mutagenesis to create *Cfo*I, *Sma*I, *Sal*I, *Eag*I and *Sac*II sites using the Transformer Site-Directed Mutagenesis kit (Clontech). The inserts included in each construct are as follows: IE, nt 168–286 relative to the hamster *Aprt* CpG island (Gb_Ro:X03603); 70, nt 212–286 relative to the hamster *Aprt* CpG island; Hprt, nt −218 to −157 relative to the translation initiation codon[14] from human *HPRT*. We created plasmid ML120β by cloning human *HBB* with its minimal promoter, including nt −48 to 2,200 relative to the transcription initiation site (Gb_Pr:J00179), between the *Xba*I and *Pst*I sites of pUC19. We synthesized two tandem MLDE sites with *Kpn*I and *Bam*HI ends and a *Pac*I site in the center (top strand, 5′–GATCCCGGAAGTCCCACCTTCTACTTCCTGTTAATTAACCGGAAGTCCCACCTTCTACTTCCTGGTAC–3′; bottom strand, 5′–CAGGAAGTAGAAGGTGGGACTTCCGGTTAATTAACAGGAAGTAGAAGGTGGGACTTCCGG–3′). We cloned the IE, flanked by *loxP* sequences, into the *Pac*I site and placed this entire sequence between the *Kpn*I and *Bam*HI sites of pUC19.

Transgenic mice and methylation analysis. We generated transgenic mice by injection of an *Eco*RI-*Hin*dIII DNA fragment (plasmids in Fig. 1) or a *Pvu*II fragment (ML120β) into the pronuclei of fertilized eggs according to standard procedures[27]. We extracted DNA from total founder embryos (14 dpc), liver, spleen or tail and used Southern-blot analysis to detect transgenes and analyse their methylation patterns. Bisulfite methylation analysis was carried out as described[20]. Following treatment, the DNA was amplified using forward (5′–ACTACAAAACGATTAAATTAAA–3′ and 5′–ACTACAAAACAATTAAATTAAA–3′) and reverse (5′–TTTGTTTAG GGTTTTATTATT–3′) primers. We cloned the resulting products and then sequenced individual colonies. Cre-expressing mice[17–19] were bred into (C57Black×BALB/c)F1 mice for at least five generations.

RT-PCR. We extracted total RNA from liver, spleen, brain, kidney and heart with TriPure Isolation reagent (Boehringer), treated total RNA (250 ng) with DNaseI (Promega) and made cDNA using M-MLV reverse transcriptase (Promega) and random hexanucleotide p$(dN)_6$ (Pharmacia) in a reaction volume (20 µl) under conditions recommended by the manufacturer. To normalize for the relative amount of cDNA synthesized, we analysed mouse *Actb* (encoding β-actin) expression. cDNA was diluted (1:5) and used for PCR reactions (1, 3 or 9 µl). We first calibrated PCR amplification to use the least number of cycles and still attain concentration dependence in the presence of ^{32}PαdCTP (Amersham). The sequence of the *HBB* upstream primer was 5′–GCTTCTGACACAACTGTGTTC–3′ and the downstream primer was 5′–CTGAAGTTCTCAGGATCCACG–3′, which spans the first intron and allows one to distinguish between cDNA and genomic DNA. The upstream and downstream primers used for assaying *Actb* cDNA were 5′–CAGCTTCTTTGCAGCTCCTT–3′ and 5′–TCACCCACATAGGAGTCCTT–3′, respectively. Conditions for PCR were as follows: 95 °C for 1 min, 55 °C for 1 min, 72 °C for 2 min (18 cycles for *Actb* and 32 cycles for *HBB*). RT-PCR fragments were separated on 5% polyacrylamide gels and exposed for autoradiography. The level of transcription was judged by quantitative scanning of autoradiograms and correction for the amounts of *Actb*. Sizes of RT-PCR products are 218 bp for *Actb* and 356 bp for *HBB*.

Acknowledgements

We thank E. Rand and T. Jakubowicz for help in preparing the manuscript and figures, and F.A. Asimakopoulos for help in the bisulfite analysis. This work was supported by grants from the NIH (H.C.), Council for Tobacco Research (H.C.), Israel Cancer Research fund (H.C.,Z.S.) and the Israel Ministry of Science (H.C.,Z.S.).

Received 2 December 1998; accepted 10 May 1999.

1. Razin, A. & Shemer, R. DNA methylation in early development. *Hum. Mol. Genet.* **4**, 1751–1755 (1995).
2. Gruenbaum, Y., Cedar, H. & Razin, A. Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* **292**, 620–622 (1982).
3. Yisraeli, J. *et al.* Muscle-specific activation of a methylated chimeric actin gene. *Cell* **46**, 409–416 (1986).
4. Busslinger, M., Hurst, J. & Flavell, R.A. DNA methylation and the regulation of the globin gene expression. *Cell* **34**, 197–206 (1983).
5. Yisraeli, J., Frank, D., Razin, A. & Cedar, H. Effect of in vitro DNA methylation on β globin gene expression. *Proc. Natl Acad. Sci. USA* **85**, 4638–4642 (1988).
6. Panning, B. & Jaenisch, R. DNA methylation can activate Xist expression and silence X-linked genes. *Cell* **90**, 907–916 (1997).
7. Beard, C., Li, E. & Jaenisch, R. Loss of methylation activates Xists in somatic but not in embryonic cells. *Genes Dev.* **9**, 2325–2334 (1995).
8. Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**, 362–365 (1993).
9. Walsh, C.P., Chaillet, J.R. & Bestor, T.H. Transcription of IAP endogenous retrovirus is constrained by cytosine methylation. *Nature Genet.* **20**, 116–117 (1998).
10. Kirillov, A. *et al.* A role for nuclear NF-κB in B-cell-specific demethylation of the Igκ locus. *Nature Genet.* **13**, 435–441 (1996).
11. Lichtenstein, M., Keini, G., Cedar, H. & Bergman, Y. B-cell specific demethylation: a new role for the intronic κ-chain enhancer sequence. *Cell* **76**, 913–923 (1994).
12. Brandeis, M. *et al.* Sp1 elements protect a CpG island from *de novo* methylation. *Nature* **371**, 435–438 (1994).
13. Macleod, D., Charlton, J., Mullins, J. & Bird, A.P. Sp1 sites in the mouse Aprt gene promoter are required to prevent methylation of the CpG island. *Genes Dev.* **8**, 2282–2292 (1994).
14. Hornstra, I. & Yang, T.P. High resolution analysis fo 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 (1994).
15. Klages, S., Möllers, B. & Renkawitz, R. The involvement of demethylation in the myeloid-specific function of the mouse M lysozyme gene downstream enhancer. *Nucleic Acids Res.* **20**, 1925–1932 (1992).
16. Nickel, J., Short, M.L., Schmitz, A., Eggert, M. & Renkawitz, R. Methylation of the mouse M-lysozyme downstream enhancer inhibits heterotetrameric GABP binding. *Nucleic Acids Res.* **23**, 4785–4792 (1995).
17. Gu, H., Marth, J.D., Orban, P.C., Mossman, H. & Rajewsky, K. Deletion of a DNA polymerase β gene segment in T cell using cell type specific gene targeting. *Science* **265**, 103–106 (1994).
18. Lallemand, Y., Luria, V., Haffner-Krausz, R. & Lonai, P. Maternally expressed PGK-Cre transgene as a tool for early and uniform activation of the Cre site specific recombinase. *Transgenic Res.* **7**, 105–112 (1998).
19. Kuhn, R., Schwenk, F., Aguet, M. & Rajewsky, K. Inducible gene targeting in mice. *Science* **269**, 1427–1429 (1995).
20. Clark, S.J., Harrison, J., Paul, C.L. & Frommer, M. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* **22**, 2990–2997 (1994).
21. Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386–389 (1998).
22. Jones, P.L. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genet.* **19**, 187–191 (1998).
23. Eden, S., Hashimshony, T., Keshet, I. & Cedar, H. DNA methylation models histone acetylation. *Nature* **394**, 842 (1998).
24. Weiss, A., Keshet, I., Razin, A. & Cedar, H. DNA demethylation in vitro: involvement of RNA. *Cell* **86**, 709–718 (1996).
25. Yoder, J.A., Walsh, C.P. & Bestor, T.H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335–340 (1997).
26. Walsh, C.P. & Bestor, T.H. Cytosine methylation and mammalian development. *Genes Dev.* **13**, 26–34 (1999).
27. Hogan, B., Constantini, F. & Lacey, E. *Manipulating the Mouse Embryo* 92–94 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1986).
28. Calzone, F.J., Britten, R.J. & Davidson, E.H. Mapping of gene transcripts by nuclease protection assays and cDNA primer extension. *Methods Enzymol.* **152**, 611–632 (1987).