Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 10763–10768, September 1998
Genetics

A cis-acting element that directs the activity of the murine methylation modifier locus *Ssm1*

PETER ENGLER*, LYNN T. DOGLIO†, GRAZYNA BOZEK, AND URSULA STORB

Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637

Communicated by Hewson Swift, University of Chicago, Chicago, IL, June 26, 1998 (received for review December 15, 1997)

ABSTRACT **Silencing of chromosomal domains has been described in diverse systems such as position effect variegation in insects, silencing near yeast telomeres, and mammalian X chromosome inactivation. In mammals, silencing is associated with methylation at CpG dinucleotides, but little is known about how methylation patterns are established or altered during development. We previously described a strain-specific modifier locus, *Ssm1*, that controls the methylation of a complex transgene. In this study we address the questions of the nature of *Ssm1*'s targets and whether its effect extends into adjacent sequences. By examining the inheritance of methylation patterns in a series of mice harboring deletion derivatives of the original transgene, we have identified a discrete segment, derived from the *gpt* gene of *Escherichia coli*, that is a major determinant for *Ssm1*-mediated methylation. Methylation analysis of sequences adjacent to a transgenic target indicates that the influence of this modifier extends into the surrounding chromosome in a strain-dependent fashion. Implications for the mechanism of *Ssm1* action are discussed.**

Silencing of chromosomal domains was first described in *Drosophila* almost 70 years ago (1). An analogous process, telomere position effect, has been described in yeast (2). By using genetic approaches, a number of modifier loci have been identified in *Drosophila* (3) and yeast (4) that affect silencing and fundamental aspects of chromatin structure. In mammals, chromosomal silencing is involved in processes such as X chromosome inactivation, parental imprinting, and oncogenesis. Little is known about the genes and proteins involved in silencing chromosomal regions in higher organisms.

The vertebrate genome is divided into two compartments: one that is heavily methylated at CpG dinucleotides and transcriptionally silent and another that is undermethylated and transcribed (5). These patterns of methylation are established during embryogenesis, erased during gametogenesis and early development, and reestablished anew each generation (6). How certain domains are targeted for, or protected from, methylation remains largely unanswered. It is clear, however, that CpG methylation plays an essential role during mammalian development as shown by the inability of mice deficient in methyltransferase to complete development (7).

We have previously identified a locus, *Ssm1* (strain-specific modifier), on distal mouse chromosome 4 that has a dramatic effect on the methylation of a particular transgene (8). When this transgene is carried in certain genetic backgrounds, such as C57BL/6, it is highly methylated, but it becomes unmethylated after several crosses with other strains such as DBA/2. When an unmethylated transgene is crossed back into C57BL/6 (or other methylating strains) it becomes fully modified within one generation. A detailed analysis of transgenic embryonic stem cells and embryos has shown that methylation occurs around the time of implantation, coincident with major global methylation changes of endogenous loci (9). Although no endogenous targets of *Ssm1* have yet been identified, a possible human homologue of *Ssm1*, *MEMO-1*, has been mapped to 1p35–36.1, a region syntenic with mouse distal chromosome 4 (10). Deletions of 1p35–36.1 are associated with HLA class I methylation changes seen in many neuroblastomas.

Determination of the properties of the *Ssm1* modifier/target model system should shed light on basic mechanisms involved in silencing chromosomal domains in higher organisms. In this study we focus on two related questions: do discrete elements within a transgene serve to target *Ssm1*-mediated methylation, and does this modification spread beyond the target?

MATERIALS AND METHODS

Transgenic Mice. The construction of pHRD and its promoterless and enhancerless derivatives has been described (11, 12) (Fig. 1). The gpt transgene contains 0.9 kb from the *gpt* gene of *E. coli*. This same fragment was deleted from pHRD to yield the HRDΔG transgene. Nucleotide sequences and details of construction of the transgenes are available on request. The transgenic lines were made by microinjection into the male pronucleus of (C57BL/6J × SJL/J)F_1 × C57BL/6J zygotes and were maintained by breeding with normal DBA/2J or C57BL/6J mice; all mice were hemizygous for the transgene locus. For the mapping experiments, hemizygous transgenic males with hypomethylated transgenes were mated with females from the BXD/Ty recombinant inbred set, and the transgenic progeny were assayed for transgene methylation. The transgene copy number in the 7 lines studied in this report was estimated from hybridization intensity to be between 2 and 10.

Methylation Analysis. Tail or kidney DNA was digested with *Hpa*II and Southern blots were probed with transgene-specific probes (*gpt* for the ΔP, ΔE, and gpt lines or SV40 splice and poly(A) segments for the ΔG lines). The level of transgene CpG methylation has previously been shown to be very similar or identical in all tissues of adult mice (8).

Genotype Determination. The *Ssm1* genotype of selected mice was determined by using PCR assays to type at two simple sequence repeat loci (13) flanking *Ssm1*. Ongoing mapping work has shown that *D4Mit259* is 1.0 centimorgan (cM) proximal to *Ssm1* and *D4Mit342* is 0.6 cM distal (unpublished results). The C57BL/6 (B), DBA/2 (D), and SJL (S) alleles yield different sized PCR fragments in both assays.

Cloning the Integration Site from the ΔP5303 Line. DNA from the ΔP5303 line was digested with *Bam*HI and electrophoresed on a preparative agarose gel. Fragments of ≈1.6 kb

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: SV40, simian virus 40.

*To whom reprint requests should be addressed at: Department of Molecular Genetics and Cell Biology, University of Chicago, 920 East 58th St., Chicago, IL 60637. e-mail: jpengler@midway.uchicago.edu.
†Present address: Department of Pediatrics, Northwestern University, Chicago, IL 60614.

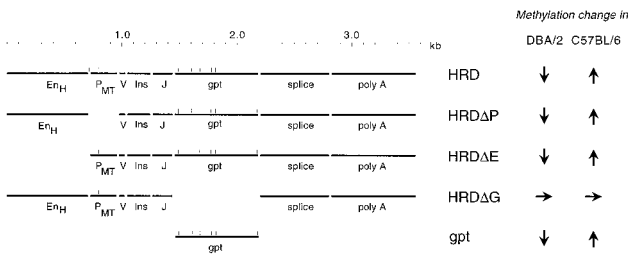

FIG. 1. Transgene maps and summary of methylation patterns in transgenic mice. En_H, mouse Ig heavy chain enhancer; P_{MT}, mouse metallothionein-1 promoter; V, recombination recognition sequences from a mouse variable region κ gene; Ins, portion of the rat preproinsulin gene containing the initiation codon; J, recombination recognition sequences from a mouse κ joining region segment; gpt, xanthine–guanine phosphoribosyltransferase gene from *E. coli*; splice, splicing signals from simian virus 40 (SV40); polyA, polyadenylation signals from SV40. *Hpa*II/*Msp*I sites used for methylation analysis are indicated by vertical ticks above each map. Arrows indicate the trend of methylation (data summarized from Figs. 2 and 3).

were isolated, ligated into the *Bam*HI site of pZErO-1 (from Invitrogen), and electroporated into *E. coli* MC1061. Colonies reacting with a probe from the 3′ end of the transgene contained a 1.6-kb insert. A clone lacking transgene sequences was made by PCR with *Pfu* polymerase (Stratagene) and primers from the flank and from the vector.

RESULTS

Inheritance of Methylation Patterns in Mice with Transgene Deletion Derivatives. The HRD transgene has been shown to be methylated in a strain-specific fashion; it is highly methylated in C57BL/6 (B) founder mice, becomes progressively less methylated [correlated with increased transcription (14) and site-specific recombination (15)] when crossed into a DBA/2 (D) or SJL (S) strain background, and becomes fully methylated when bred back into B, the modifying strain. This effect is because of a strain-specific modifier locus, *Ssm1*, on distal chromosome 4 (8). We now report a similar genetic analysis of transgenic mice harboring derivatives of the original HRD transgene. By systematically eliminating pieces of this complex transgene, we have determined that a discrete region, from the *gpt* gene of *E. coli*, is critical for directing methylation.

Seven independent transgenic lines, of roughly similar copy number (between 2 and 10), produced from four different HRD derivatives, were analyzed by Southern blotting of *Hpa*II-cleaved DNA. Two independent lines derived from a promoterless version of HRD (HRDΔP; Fig. 1) showed similar behavior (Figs. 2*A* and 3): the transgene was highly methylated in the founders but became hypomethylated after several crosses with D. Upon crossing with B, the transgene became completely methylated. Thus, this derivative behaved just as the complete HRD transgene. Genotype determination, by typing selected mice for two loci flanking *Ssm1*, gave the results expected for methylation being determined by a single dominant gene. The transgene was highly methylated in mice with one or two $Ssm1^b$ alleles but became hypomethylated after being in a $Ssm1^{d/d}$ or $Ssm1^{d/s}$ background for one or two generations. Transgene methylation was restored after introduction of a single copy of $Ssm1^b$. Similar results were found in two transgenic lines lacking the enhancer (HRDΔE; Fig. 1). In the 5319 line (Figs. 2*B* and 3) the results were virtually identical to the HRDΔP (and HRD) analysis. The 5326 line (Fig. 2) was somewhat different in that when hypomethylated transgenes (in $Ssm1^{d/s}$ mice) were crossed back into strain B complete methylation was not always observed in the first-generation offspring (in every case, however, there was a

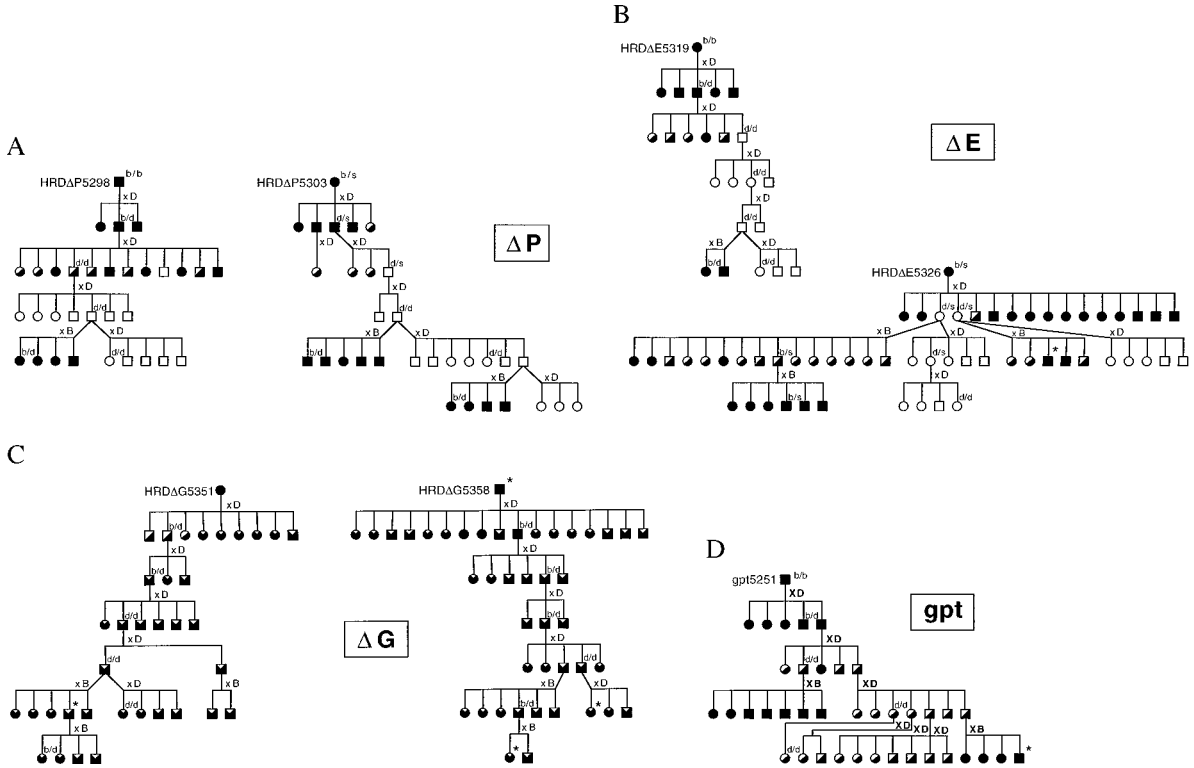

FIG. 2. Inheritance of methylation patterns in pedigrees of transgenic mouse lines containing deletion derivatives of the HRD transgene. Filled symbols indicate methylated transgene arrays, open symbols indicate unmethylated arrays, and partially filled symbols indicate partially methylated transgenes. The partially methylated transgenes seen in the ΔG pedigrees were in general substantially more methylated than those in other lines (see Fig. 3). The *Ssm1* genotype is indicated by lowercase letters (b, d, or s). Those mice used for the analysis presented in Fig. 3 are marked with genotype designations or with asterisks. (*A*) Two HRDΔP lines. (*B*) Two HRDΔE lines. (*C*) Two HRDΔG lines. (*D*) A gpt-only line.

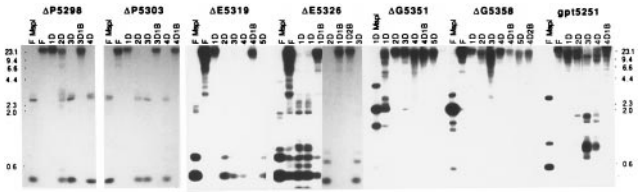

FIG. 3. Analysis of transgene methylation in seven mouse lines. F designates founder animals. DNA from spleen or kidney was cleaved with *Hpa*II (or *Msp*I where indicated). The number of generations of crossing to DBA/2J (D) or to C57BL/6J (B) is noted. Sizes, in kb, of DNA marker fragments are noted.

dramatic increase in methylation). Whether this is a common property of $Ssm1^{d/s}$ heterozygotes or is due to other factors has not been determined. The general pattern was the same, however; transgenes were highly methylated in a B, but hypomethylated in a D background. These results rule out the possibility that promoter or enhancer elements, by themselves, are responsible for the strain-specific methylation of the HRD transgene.

In contrast to the results obtained for the HRD, ΔP, and ΔE transgenes, a different pattern was seen during analysis of transgenes lacking *gpt* sequences (HRDΔG; Figs. 1, 2*C*, and 3). As in the other lines, transgene methylation was high in the founders, but upon crossing with D, only modest (as in the first D generation of the ΔG5351 line), if any, reduction in transgene methylation was observed. Even maintenance of the transgene in a $Ssm1^{d/d}$ background for several generations did not result in complete hypomethylation. Crossing the transgene into a B strain background also did not significantly change the degree of methylation. The mouse-to-mouse and generation-to-generation variations are minor and appear to be independent of genetic background. The marked difference in strain-specific transgene methylation between lines containing and lines lacking *gpt* sequences suggests that this region is critical for establishing methylation patterns.

To determine whether *gpt* sequences alone can direct strain-specific methylation, a transgenic mouse line harboring only this critical region was examined. The same strain-specific methylation differences were indeed observed in a *gpt*-only line (Figs. 2*D* and 3). In this case crossing the transgene into the D background was accompanied with a very significant reduction, rather than elimination, of methylation. Genotyping the mice for loci flanking *Ssm1* confirmed that the transgene had been in an $Ssm1^{d/d}$ background for at least two generations (Fig. 2*D*). When the transgene was crossed back into the B background, virtually complete methylation resulted. Thus the gpt transgene is subject to a similar strain effect as the intact HRD transgene. We conclude that *gpt* by itself is essential for targeting strain-specific methylation.

Mapping the Locus Responsible for Strain-Specific Methylation. The preceding analysis indicated that methylation of transgenes containing the *gpt* target sequence is affected in a strain-specific manner. To determine whether *Ssm1*, previously shown to be located on distal chromosome 4 and to control the methylation of the complete HRD transgene, is responsible for strain-specific methylation of the derivatives, genetic mapping using recombinant inbred (RI) mice was undertaken. These mice are useful for this purpose because each locus has a characteristic strain distribution pattern (SDP) among members of a RI series (16). Loci that are linked have similar SDPs, whereas identical SDPs are consistent with identical (or very closely linked) loci. When mice carrying hypomethylated transgenes were crossed to strains from the BXD recombinant inbred series, transgene methylation in the offspring followed one of two general patterns: methylation either remained unchanged relative to the parent or increased dramatically (Table 1). The SDPs for all the deletion derivatives were concordant with those for the HRD342 transgene, consistent only with methylation being determined by a locus on distal chromosome 4 near *Fv1*. To confirm and refine the mapping based on RI strains, and to position *Ssm1* with respect to loci useful for positional cloning, a backcross segregating *Ssm1* is being examined. Over 500 offspring from an HRD342 cross and 38 HRDΔP5303 progeny have been typed at numerous loci on distal chromosome 4 and assayed for transgene methylation. The results confirm that a single locus on distal chromosome 4, most probably *Ssm1*, controls the methylation of numerous transgenes that contain *gpt* sequences.

Influence of the Transgenic Target on Methylation of Flanking Chromosomal Regions. In some other instances of chromosomal silencing, the silenced region has been shown to "spread" along the chromosome. To determine whether the influence of *Ssm1* extends beyond its transgenic target, we cloned a transgene insertion site and used these chromosomal sequences to assess the methylation of this region in transgenic mice of both methylating and nonmethylating strain backgrounds.

Using the SV40 sequences at the 3′ end of the transgene as a probe, we cloned a 1.6-kb *Bam*HI "junction fragment" from the HRDΔP5303 line (Fig. 4). The DNA sequence of this fragment showed ≈0.5 kb of SV40 (lacking only 30 nt of the microinjected DNA) contiguous with ≈1 kb of new sequence (with no significant matches in the GenBank database), comprising the flanking chromosomal DNA. PCR, using primers near either end of the flank sequence with DNA from several common inbred strains as well as from HRDΔP5303 transgenic DNA, gave indistinguishable 0.7-kb fragments (data not shown), suggesting that no major structural changes occurred during integration of the transgene.

We focused on several methylation-sensitive restriction sites downstream of the transgene: an *Ava*I site ≈0.6 kb away (≈1.9 kb from the end of *gpt*) and a group of two *Hpa*II sites ≈3.4 kb away (≈4.7 kb from *gpt*).

The transgenic and wild-type alleles can be distinguished by Southern blotting after cleavage with *Bam*HI + *Bsm*I (Figs. 4

Table 1. Strain distribution patterns of transgene methylation in the BXD recombinant inbred series

Transgenic line	Methylation in BXD strain																								
	1	2	5	6	8	9	11	12	13	14	15	16	18	19	21	22	23	24	25	27	28	29	30	31	32
HRD342	−	+	−	+	−	+	+	−	−	−	+	−	−	+	−	+	+	+	+	−	−	+	−	+	−
ΔP5298		+	−		−	+	+		−			−	−	+			+	+		−	−	+			
ΔP5303		+	−		−	+	+		−			−	−	+			+	+		−	−	+			−
ΔE5326		+	−			+	+			−		−			−		+			−				+	−
ΔE5319			−	+															+			+	−		
gpt5251	−				−		+	−			+				−	+	+		+		−	+	−	+	

Males with hypomethylated transgene arrays were mated with BXD/Ty females, and transgene methylation in the offspring was assayed by restriction endonuclease digestion of DNA, prepared from kidneys of young animals, with *Hpa*II and probing Southern blots with transgene-specific probes. The data on the HRD342 line have been previously reported (6). −, No increase in transgene methylation compared with parent; +, transgene methylation shifted to complete, or nearly complete, pattern.

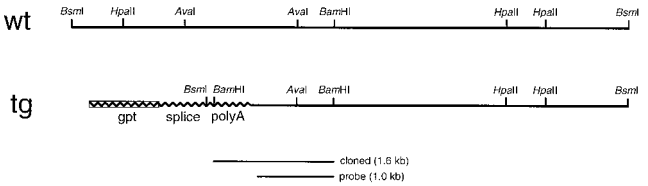

FIG. 4. Restriction site maps of the HRDΔP5303 transgene locus. Both the nontransgenic (wild type; wt) and transgenic (tg) alleles are shown. Only the 3′ end of the HRDΔP transgene (*gpt* and splice/polyA as in Fig. 1) is indicated.

and 5*A*); the transgenic allele yields a 1.6-kb fragment, whereas the wild-type allele is associated with a 3.6-kb fragment. The 1.6-kb fragment contains a single *Ava*I restriction site and the 3.6-kb fragment contains two sites. *Ava*I cleaves only when the CpG dinucleotide in its recognition site is unmethylated. These *Ava*I sites are substantially, though not completely, methylated on the nontransgenic (wt) chromosome in both kidney and spleen cells. Furthermore, the degree of methylation at these *Ava*I sites is not greatly influenced by *Ssm1* genotype (compare Fig. 5 *Upper* and *Lower*). In contrast, *Ssm1* genotype has a pronounced effect on the methylation of the analogous *Ava*I site on the transgenic chromosome. This site is partially, though substantially, methylated in mice carrying at least one copy of $Ssm1^{b}$ (Fig. 5*A Upper*). Such mice have previously been shown to carry fully methylated transgene arrays. Mice that are $Ssm1^{d/d}$, however, show a clear reduction in methylation at the *Ava*I site on the transgenic (tg) chromosome (Fig. 5*A Lower*). Thus the transgene does indeed affect the methylation of flanking chromosomal DNA in a strain-dependent manner.

Methylation at the more distant *Hpa*II sites was assayed in a similar manner, using *Bsm*I fragment length differences to distinguish the transgenic allele from its wild-type counterpart (Fig. 5*B*). In this case the transgene-associated *Bsm*I fragment has two *Hpa*II sites, whereas there are three sites in the corresponding wild-type fragment. Again, these *Hpa*II sites are substantially, though not completely, methylated on the wild-type chromosome. A similar degree of methylation is seen at the corresponding sites on the transgenic chromosome when in an $Ssm1^{b}$ genetic background (these sites are significantly less methylated in mouse C094 for reasons that are not understood). A clear reduction of methylation at the *Hpa*II sites near the transgene is seen in $Ssm1^{d/d}$ mice, although the magnitude of the effect is not as great as at the closer *Ava*I site. Again it appears that these sites downstream of the transgene are less

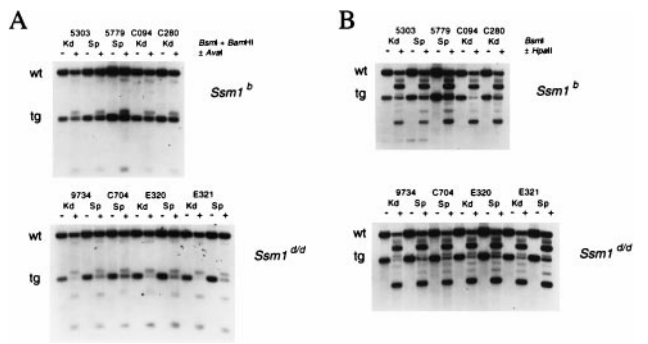

FIG. 5. Methylation analysis of chromosomal DNA flanking the HRDΔP5303 transgene insertion. Numbers at top are mouse identification codes. DNA from kidney (Kd) or spleen (Sp) was cut with *Bsm*I + *Bam*HI (*A*) or *Bsm*I alone (*B*) without (−) or with (+) the addition of *Ava*I (*A*) or *Hpa*II (*B*). Bands corresponding to the wild-type (wt; 3.6 kb in *A*; 7.4 kb in *B*) and transgenic (tg; 1.6 kb in *A*; 5.6 kb in *B*) alleles are indicated. In *Upper* are mice that have at least one copy of $Ssm1^{b}$ (and therefore have methylated transgene arrays); homozygous $Ssm1^{d/d}$ mice (with unmethylated transgenes) are in *Lower.*

methylated when in a D strain background. Thus, it appears that methylation of sites some distance away from a transgenic target is affected in a transgene-dependent and strain-dependent manner.

DISCUSSION

Nature of the Target. The methylation status of the HRD transgene has been shown to be determined by the *Ssm1* locus (8). Mice that are $Ssm1^{b/b}$ or $Ssm1^{b/d}$ have methylated transgenes, whereas the same transgene is hypomethylated in homozygous $Ssm1^{d/d}$ or $Ssm1^{s/s}$ mice. Partially methylated transgenes often occur during breeding the transgene from a methylating strain to a nonmethylating strain. When the transgene is transferred in the other direction, however, immediate and complete methylation is observed. Results presented here show that the three different transgene derivatives containing sequences from the *E. coli gpt* gene (HRDΔP, HRDΔE, and gpt) all exhibit strain-dependent behavior similar to what had been reported for the HRD transgene. In contrast, the derivative lacking *gpt* sequences (HRDΔG) is not subject to these strain effects. This marked difference is unlikely to be due to position effects (integration being presumably random) or to differences in copy number (all mice in this study have roughly equivalent numbers of transgenes) but suggests that *gpt* sequences themselves are responsible for the observed strain effects. We conclude that *gpt* is a major determinant for strain-specific methylation.

When the HRD, HRDΔP, and HRDΔE transgenes are bred into a D strain background, virtually complete hypomethylation results. The *gpt*-only transgene, in contrast, shows very significant, but incomplete, hypomethylation. The presence of high molecular weight DNA (uncut by *Hpa*II) suggests that the *gpt* transgene remains methylated in a subset of cells. It is not possible to conclude whether this occurs in a particular lineage or is a random event. Nonetheless, a clear strain effect, likely due to *Ssm1* (see below), is also seen with the *gpt* transgene.

It was conceivable that modifier loci unlinked to *Ssm1* might be responsible for determining the strain-specific behavior of the different transgene deletion derivatives. The finding that the BXD strain distribution patterns for all transgene derivatives are concordant with those for the intact HRD transgene makes it very unlikely that this is the case. *Ssm1* had been previously mapped to distal chromosome 4 by crossing HRD transgenic mice with 25 strains from the BXD recombinant inbred set (8). For the gpt5251 line, where 13 of 13 concordance with the HRD342 results was found, the probability of linkage or identity is 0.962 (17). For the ΔP transgenic lines, with complete concordance among 14 (for ΔP5298) or 15 (for ΔP5303) strains, the probability is even higher (0.979 or 0.988). Fewer BXD strains were bred with the ΔE derivatives, but again complete concordance with the HRD342 results was found, again suggesting that distal chromosome 4 harbors the modifier(s). If indeed modifiers other than *Ssm1* affect methylation of the derivatives, they would have to be quite closely linked to *Ssm1*. For example, in the BXD27 strain the DBA/2 genomic interval on distal chromosome 4 is less than 5 centimorgans (P.E., unpublished results and http://www.informatics.jax.org/). Because all transgenic lines (HRD342, ΔP5298, ΔP5303, and ΔE5326) crossed with BXD27 yielded hypomethylated progeny, any additional modifier(s) would have to be in this small interval. We conclude that the strain effect seen with *gpt*-containing transgenes is likely to be due to *Ssm1* alone.

What features of *gpt* make it a target for *Ssm1*-mediated modification? This segment, derived from *E. coli*, is clearly foreign to the mouse genome, but it seems unlikely that *Ssm1* recognizes all prokaryotic (or other foreign) DNA sequences. Although systematic examination of methylation levels has not been reported, a number of prokaryotic genes have been

successfully expressed, and therefore presumably are hypomethylated in strains known to methylate the HRD transgene. For example, *lacZ* from *E. coli* can be successfully expressed after introduction into the mouse genome of various strains, including C57BL/6 and 129 (18), another methylating strain (unpublished results). Similarly, the *neo* gene from Tn5, or a *lacZ-neo* fusion, is able to function in embryos of the 129 genetic background (19). These embryonic stages are well after *Ssm1*-mediated methylation of the HRD transgene is complete (9). In the cases just cited it is possible that certain elements, such as strong promoters or enhancers, may be able to overcome methylation's repressive effect, so these conclusions about the specificity of *Ssm1* must be regarded as tentative.

Examination of the target sequence reveals that *gpt* is moderately rich in C+G nucleotides (53% versus 43% for the ΔG transgene), with CpG dinucleotides being particularly abundant (CpG/GpC ratio of 1.1 versus 0.3 for ΔG). This composition is dissimilar to the overall composition of mammalian DNA, which is lower in C+G content and is deficient in CpG dinucleotides. The CpG/GpC ratio is similar to, although the C+G content is lower than, that seen in vertebrate CpG islands (20), which are important for maintaining methylation-free stretches around genes and which have been implicated in differential gene expression in parental imprinting (21). No Sp1 binding sites are present in the *gpt* segment (see below). It is tempting to speculate that *Ssm1* acts upon a set of endogenous sequences with an unusual base composition.

Sequences important for determining the methylation status of another transgene have been described (22). Expression and methylation of the RSVIgmyc transgene are subject to parental imprinting. Deletion analysis of this transgene implicated a small portion, from an Ig switch region, as being critical for the imprinted behavior. The switch region is composed of repetitive elements, a particularly interesting finding in light of the frequent association of repetitive DNA with imprinted loci (21). There is no apparent similarity of these sequences to *gpt* and the possible connection to *Ssm1* is not clear.

Influence of the Transgene on Flanking Chromosomal Sequences. When sites downstream of a *gpt*-containing transgene are examined, methylation is clearly affected in a strain-specific manner. These sites are partially methylated in mice carrying one copy of $Ssm1^b$ (in which the transgene is methylated) but are much less methylated in homozygous $Ssm1^d$ mice (which harbor hypomethylated transgenes). There is a steady decline of the *Ssm1* effect with distance from *gpt*: the sequences flanking *gpt* in the transgene show an all-or-none effect (8), the near *Ava*I site is less completely, but still highly affected, and the distant *Hpa*II sites show the least effect. This is clearly a cis effect because there is no dramatic strain-specific difference in methylation of the analogous sites on the non-transgenic chromosome. It seems likely that *Ssm1* is responsible for the strain-specific differences at these chromosomal sites as well as at sites in the transgene. Most likely, *gpt* serves as a focus for *Ssm1* action which then affects the flanking sequences in a manner reminiscent of the "spreading" seen in other systems.

Implications for the Mechanism of *Ssm1* Action. It has been previously shown that the HRD transgene is methylated in strain B, unmethylated in D, and methylated in $(B\times D)F_1$ hybrids. These results suggest that the B allele of *Ssm1* is dominant and are compatible with ideas that *Ssm1* might direct methylation to sequences in the transgene.

Two lines of evidence presented in this paper suggest that a different mechanism, dosage-dependent protection, should be considered. First, when *gpt* sequences are eliminated from the HRD transgene, methylation results. Thus, methylation could be viewed as a "default" state and the product of the D allele of *Ssm1* could be acting to protect *gpt*-containing transgenes against methylation. Other similar mechanisms, involving demethylation by Ssm1-d, could also be proposed. Second, the methylation of chromosomal sequences is affected in a strain- and transgene-dependent manner but, when the transgene is present, these sequences become less methylated in a D background rather than more methylated in a B background. In other words, there is apparent "spreading" of hypomethylation. This mechanism for *Ssm1* action may be analogous to the methylation protection afforded by Sp1 binding to its site in the promoter of housekeeping genes (23, 24). Its presence prevents methylation of the promoter during the *de novo* methylation process that targets most tissue-specific genes around the time of embryo implantation. It is perhaps significant that *Ssm1*-mediated methylation of transgenic targets also occurs at this time (9). In a different system, evidence has been presented for the existence of modifier genes that block transgene methylation (22). In the case of the RSVIgmyc transgene, imprinted behavior (expression and hypomethylation when inherited from the father) is seen on certain genetic backgrounds, but other backgrounds cause hypermethylation regardless of parental legacy. Because imprinting is restored in F_1 hybrids between the permissive and repressive strains it seems that the modifier loci responsible for blocking methylation of the paternally imprinted transgene must act in a dominant fashion. This mechanism of dominant blocking of methylation is in contrast to what is observed in the *Ssm1*/*gpt* system (see below).

While these two observations may seem to suggest mechanisms involving protection from methylation (or mechanisms involving demethylation), some other observations are more compatible with a targeting mechanism for *Ssm1* action. First is the observation that the HRD transgene is methylated in $(B\times D)F_1$ hybrids. The simplest interpretation is that the B allele dominantly directs methylation to transgenic sequences. For protection/demethylation mechanisms it is necessary to invoke dosage dependence: a single copy of $Ssm1^d$ has no effect but two copies are required for protection. Another observation difficult to reconcile with protection models is that transgene methylation is slowly lost on sequential crosses to D (nonmethylating strain) but methylation is immediately acquired after a single cross with strain B. Even after two sequential crosses with D mice, when 25% of the mice are homozygous $Ssm1^{d/d}$, only a small proportion of the mice have completely unmethylated transgenes (8) (Fig. 2). Because *Ssm1* acts around the time of implantation, if it encoded a protective factor, newly synthesized DNA containing the *gpt* sequence would be associated with the protective factor and rapid demethylation should occur in $Ssm1^{d/d}$ mice. Also, the observation that transgene methylation is complete in the first generation after crossing from D to B mice is more easily reconciled with an inducing effect of $Ssm1^b$.

The apparent spreading of undermethylation to sequences flanking the transgene may be compatible with a methylation-inducing effect of $Ssm1^b$ according to the following scheme. The transgene, perhaps because of the CpG island qualities of *gpt*, may induce a local demethylation around its integration site. Thus, in an $Ssm1^d$ background the transgene would be unmethylated and its flanks would be less methylated than in the absence of the transgene. In the presence of $Ssm1^b$, however, *gpt* would be targeted for methylation resulting in complete transgene methylation and an increase in the methylation of the flanking DNA.

In addition to the two extreme models of "protection" and "targeting," a number of other mechanisms are possible. For example, both the Ssm1-b and Ssm1-d products could bind to their target sequences but recruit methyltransferase with greatly different efficiencies. Definitive proof of the mechanism of action of this modifier awaits the cloning of the *Ssm1* gene.

We thank A. Weng, Z. Wang, and N. Michael for comments on the manuscript. This work was supported by grants from the National Institutes of Health (AI 24780; U.S.), the National Science Foundation (MCB-9506784; P.E.), the National Multiple Sclerosis Society (L.T.D.), and the University of Chicago Cancer Research Center (P.E.).

1. Spradling, A. & Karpen, G. (1990) *Genetics* **126,** 779–784.
2. Gottschling, D., Aparicio, O., Billington, B. & Zakian, V. (1990) *Cell* **63,** 751–762.
3. Spofford, J. (1976) in *The Genetics and Biology of* Drosophila, eds. Ashburner, M. & Novitsky, E. (Academic, New York), Vol. 1c, pp. 955–1009.
4. Aparicio, O., Billington, B. & Gottschling, D. (1991) *Cell* **66,** 1279–1288.
5. Bird, A. P. (1986) *Nature (London)* **321,** 209–213.
6. Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H. & Razin, A. (1992) *Genes Dev.* **6,** 705–714.
7. Li, E., Bestor, T. H. & Jaenisch, R. (1992) *Cell* **69,** 915–926.
8. Engler, P., Haasch, D., Pinkert, C. A., Glymour, M., Brinster, R. & Storb, U. (1991) *Cell* **65,** 939–947.
9. Weng, A., Magnuson, T. & Storb, U. (1995) *Development (Cambridge, U.K.)* **121,** 2853–2859.
10. Cheng, N., Chan, A., Beitsma, M., Speleman, F., Westerveld, A. & Versteeg, R. (1996) *Hum. Mol. Genet.* **5,** 309–317.
11. Engler, P. & Storb, U. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4949–4953.
12. Engler, P., Roth, P., Kim, J. & Storb, U. (1991) *J. Immunol.* **146,** 2826–2835.
13. Dietrich, W., Katz, H., Lincoln, S. E., Shin, H.-S., Friedman, J., Dracopoli, N. C. & Lander, E. S. (1992) *Genetics* **131,** 423–447.
14. Weng, A., Engler, P. & Storb, U. (1995) *Mol. Cell. Biol.* **15,** 572–579.
15. Engler, P., Weng, A. & Storb, U. (1993) *Mol. Cell. Biol.* **13,** 571–577.
16. Bailey, D. W. (1981) in *The Mouse in Biomedical Research*, eds. Foster, H., Small, J. & Fox, J. (Academic, San Diego), Vol. 1, pp. 224–236.
17. Neumann, P. E. (1990) *Genetics* **126,** 277–284.
18. Gossler, A., Joyner, A., Rossant, J. & Skarnes, W. (1989) *Science* **244,** 463–465.
19. Friedrich, G. & Soriano, P. (1991) *Genes Dev.* **5,** 1513–1523.
20. Antequera, F. & Bird, A. (1993) in *DNA Methylation: Molecular Biology and Biological Significance*, eds. Jost, J. & Saluz, H. (Birkhauser, Basel), pp. 169–185.
21. Neumann, B., Kubicka, P. & Barlow, D. (1995) *Nat. Genet.* **9,** 12–13.
22. Chaillet, J. R., Bader, D. S. & Leder, P. (1995) *Genes Dev.* **9,** 1177–1187.
23. Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A. & Cedar, H. (1994) *Nature (London)* **371,** 435–438.
24. Macleod, D., Charlton, J., Mullins, J. & Bird, A. (1994) *Genes Dev.* **8,** 2282–2292.