

## Epigenetic and Genotype-specific Effects on the Stability of *de Novo* Imposed Methylation Patterns in Transgenic Mice\*

Received for publication, June 5, 2000, and in revised form, August 17, 2000  
Published, JBC Papers in Press, August 22, 2000, DOI 10.1074/jbc.M004839200

Axel Schumacher, Paul A. Koetsier‡, Jennifer Hertz§, and Walter Doerfler¶

From the Institute of Genetics, University of Cologne, Weyertal 121, D-50931 Koeln, Germany

**The chloramphenicol acetyltransferase gene under the control of the late E2A promoter of adenovirus type 2 (Ad2) was introduced as transgene into the B6D2F1 mouse strain with mixed genetic background and became extensively *de novo* methylated. The methylation of this *pAd2E2AL-CAT* (7-1A) transgene was regulated in a strain-specific manner apparently depending on the site of integration. Transmission of the 7-1A transgene into an inbred DBA/2, 129/sv, or FVB/N genetic background led to a significant loss of methylation in the transgene, whereas C57BL/6, CB20, and Balb/c backgrounds favored the *de novo* methylation in very specific patterns. The newly established patterns of *de novo* methylation were transmitted to the offspring and remained stable for many generations, regardless of the heterozygosity of strain-specific DNA sequences present in these mouse strains. Segregation analyses showed a non-mendelian transmission of methylation phenotypes and suggested the involvement of dominant modifiers of methylation. The genotype-specific modifications of the transgene were followed for 11 backcross generations. These observations reflect an evolutionarily conserved mechanism directed against foreign, e.g. viral or bacterial, DNA at least in the chromosomal location of the 7-1A transgene. In seven additional mouse lines carrying the same transgene in different chromosomal locations, strain-specific alterations of methylation patterns were not observed.**

Studies on the fate of foreign DNA introduced into an established mammalian genome are of considerable general interest. Upon integration, foreign (viral) DNA becomes frequently *de novo* methylated in very specific patterns (1–3) that are characterized by the gradual spreading of DNA methylation (4, 5). Following fertilization, similar epigenetic changes are known to affect a number of transgenes in mice as well as in transgenic fish and plants (6–12). The changes take place postzygotically in the early mouse embryo even before a stable somatic pattern is established (13). Already during early gastrulation, shortly after the genome-wide loss of methylation

before the blastocyst stage, many DNA sequences become heavily remethylated (14–17).

In our laboratory, the adenovirus system has been used as a model to gain insight into consequences of foreign DNA integration. We have shown that the insertion of foreign DNA can lead to changes in cellular DNA methylation patterns even remote from the site of transgene integration (18, 19).

The regulation of *de novo* methylation of foreign DNA presumably involves a series of genes, sometimes referred to as modifier genes. Their putative function has frequently been ascribed to the demethylation of transgenes in DBA/2, FVB/N, and 129/sv backgrounds, whereas the Balb/c and C57BL/6 genotypes are often associated with extensive *de novo* methylation (20–25). Some transgenes show parent-of-origin effects that affect the genomic imprint at individual loci (for reviews see Refs. 26–28). The phenomenon of genomic imprinting at specific loci has first been demonstrated at randomly integrated transgene loci (8, 9, 29–32).

In this study, changes in the methylation of the transgene 7-1A (*pAd2E2AL-CAT*) have been investigated. This transgene has become methylated in a strain-specific manner in mice (12). We have studied the dynamics of 7-1A methylation for up to 12 generations in different mouse strains. We have also asked additional questions. Are the observed changes in transgene methylation restricted to specific 5'-CG-3' sites? How many factors influence the observed alterations in methylation? Is there a correlation between transgene methylation and the copy number of heterologous genome segments present?

### EXPERIMENTAL PROCEDURES

**Mouse Strains**—The origins and relationships of the inbred mouse strains DBA/2, 129, FVB/N, C57BL/6, Balb/c, and CBA are reviewed in Beck *et al.* (33). A chart of inbred strain genealogies is available at the Mouse Genome Database.

**Transgenic Mice**—The plasmid *pAd2E2AL-CAT*, which was used to generate the transgenic animals, was described previously (34). A complete map of the construct is shown in Fig. 2*b*. The male founder animal 7-1 was obtained after microinjection of unmethylated DNA. The 7-1 mouse line bearing two transgene arrays was subsequently separated into two substrains, 7-1A and 7-1B. Embryos used for microinjection resulted from matings between B6D2F<sub>1</sub> mice. 129/sv and CB20 mice were obtained from the mouse facility of the Institute of Genetics, University of Cologne, and all other strains (DBA/2, C57BL/6, Balb/c, FVB/N, and B6D2F<sub>1</sub>) were received from Charles River, Sulzfeld, Germany.

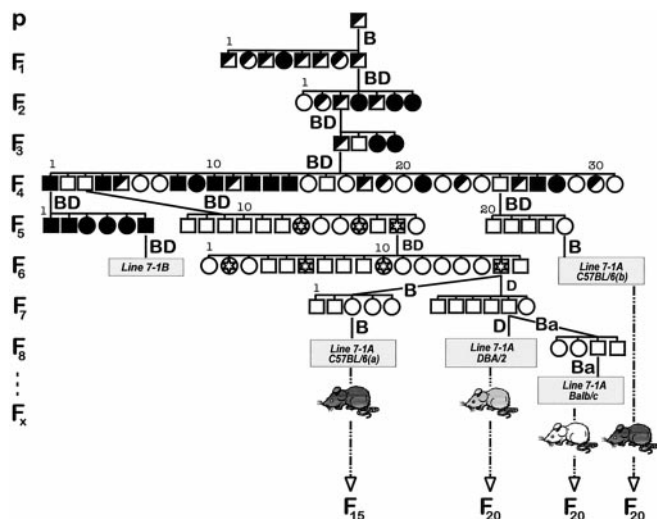
**DNA Extraction and Southern Blotting**—Animals were screened for the presence of the transgene by cutting tail tips of 3–4-week-old mice. Tail and organ tissues were frozen in liquid N<sub>2</sub>, mortar-pulverized, and then incubated at 55 °C for a minimum of 12 h in lysis buffer (100 mM Tris-HCl, pH 8.5, 5 mM EDTA, 200 mM NaCl, 0.2% SDS, and 500 μg of proteinase K per ml). Proteinase K was then inactivated for 30 min at 65 °C. Subsequently, RNase A (20 μg/ml) was added, and incubation was continued for 1 h at 37 °C. The DNA was then purified by one phenol:chloroform:isoamyl alcohol (25:24:1) extraction followed by one phenol:chloroform (1:1) treatment. The DNA samples were ethanol-precipitated and resuspended in TE (10 mM Tris, pH 7.5, 1 mM EDTA, pH 8.0), and 20 μg of each sample was used to identify transgenic

\* This work was supported in part by Deutsche Forschungsgemeinschaft Grant SFB274-A1, by the Center for Molecular Medicine Koeln, TV13, and EU Contract BMH4-CT96-0050. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: Qiagen GmbH, Max-Volmer Strasse 4, D-40724 Hilden, Germany.

§ Present address: 1135 Stanyan St., San Francisco, CA 94117. Supported by a fellowship of the Alexander von Humboldt Stiftung.

¶ To whom correspondence should be addressed: Institute of Genetics, University of Cologne, Weyertal 121, D-50931 Koeln, Germany. Tel: 49-221-470-2386; Fax: 49-221-470-5163; E-mail: doerfler@scan.genetik.uni-koeln.de.



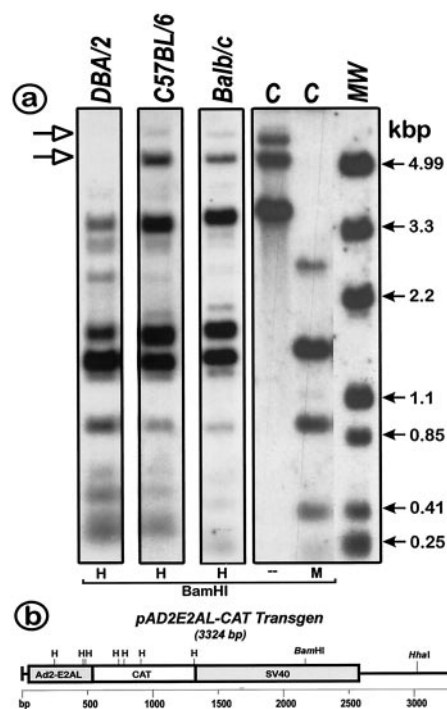
**FIG. 1. Pedigree of 7-1 transgenic mice.** All lines originated from a B6D2F<sub>1</sub> hybrid genetic background. Transgenes were transmitted by heterozygous animals. After establishing a transgenic line, all further crosses were performed with heterozygous males. Non-transgenic animals were not included in this figure. *BD* indicated mating to non-transgenic B6D2F<sub>1</sub> animals, *B* mating to C57BL/6, *Ba* mating to Balb/c, and *D* mating to DBA/2 animals. *Open squares* designate male and *open circles* designate female animals with 7-1A transgenes; *asterisk* indicates animals with demethylated transgenes (<30% methylated); *filled squares* indicate male, and *filled circles* indicate female animals with 7-1B transgenes; *half-filled squares* indicate male, and *half-filled circles* indicate female animals with 7-A/B transgenes. For comparisons of the methylation dynamics of the 7-1A transgenes, all mouse lines were derived from the same animal 16F<sub>6</sub> with the exception of mouse line C57BL/6(b) that originated from animal 24F<sub>5</sub>.

animals. The methylation levels were determined by cleaving the DNA with *Bam*HI (10 units/ $\mu$ g DNA) for 5 h followed by either *Hpa*II (10 units/ $\mu$ g DNA) incubation overnight. The fragments were separated by electrophoresis on 0.8% agarose gels in 1 $\times$  TAE (40 mM Tris acetate, 1 mM EDTA, pH 8.0) for about 14 h at 30–40 V. Gels were stained with ethidium bromide (1  $\mu$ g/ml) for 5–10 min, destained in water, and then prepared for downward Southern transfer (35) to Nylon plus membranes (Qiagen). The *pAd2E2AL-CAT* construct with flanking pBR322 vector sequences was used as <sup>32</sup>P-labeled hybridization probe in all hybridization experiments. The membranes were then exposed to x-ray films for 1–3 days. Methylation indices were determined by comparing the intensities of the *Hpa*II (methylated) with the *Msp*I (unmethylated) bands on the autoradiogram. The autoradiograms were scanned and analyzed by two different computer programs, Cybertech CS-1 Image Processing Software and TINA version 2.08. Some of the blots were also evaluated with the help of a PhosphorImager (Bio-Imaging Analyzer BAS 1000).

## RESULTS

**Establishing the 7-1A Mouse Line**—In earlier work in our laboratory, six mouse lines transgenic for the *pAd2E2AL-CAT* were established (34). The founder animal of line 7-1 originated from a microinjection experiment with an unmethylated transgene and had a mixed C57BL/6  $\times$  DBA/2 genetic background. Mouse line 7-1 was characterized by the stable integration of an estimated number of 10 transgene copies in a *de novo* methylated state. These animals had initially been mated to C57BL/6 animals for one generation. Out of this offspring, a transgenic male was selected to initiate systematic crossing experiments (Fig. 1, animal F<sub>1</sub>, number 8).

The methylation patterns of all animals from the ensuing crosses were routinely screened by *Bam*HI/*Hpa*II restriction and subsequent Southern blot hybridization. Transgenic male animals were crossed further with non-transgenic hybrid strain B6D2F<sub>1</sub> females (C57BL/6  $\times$  DBA/2) for at least 4 generations. Although these transgenic animals were heterozygous for the *pAd2E2AL-CAT* transgene, they transmitted



**FIG. 2. Methylation patterns of the 7-1A mouse lines.** *a*, tail tip DNA (20  $\mu$ g/lane) was cut with *Bam*HI and the isoschizomeric endonuclease *Hpa*II (*H*) or *Msp*I (*M*) as indicated in the figure. Cleavage at the seven *Hpa*II sites (see map in *b*) revealed a loss of methylation in the DBA/2 line, as evidenced by the almost complete absence of the 4.8- and 5.3-kbp off-size bands (arrows on the left) which represented the junctions linking the transgene to genomic DNA. The 3.3-kbp transgene repeat fragment in the C57BL/6 and Balb/c lines exhibited about double intensity when compared with DBA/2 DNA. This fragment was created by the cleavage at the unique *Bam*HI restriction site. Densitometric evaluation (see “Experimental Procedures”) of the cleavage patterns revealed that the transgene copies in DBA/2 mice were approximately 20% methylated, whereas the C57BL/6 lines (about 55% methylation) and Balb/c (about 62% methylation) favored *de novo* methylation of their transgenes. *C*, control animal DBA/2. *b*, the *pAd2E2AL-CAT* construct contained seven *Hpa*II (*H*) restriction sites, one *Bam*HI and a single *Hha*I site. The transgene consisted of two short pBR322 vector sequences at both ends, as well as a simian virus 40 (SV40) poly(A) sequence with the CAT gene as reporter gene, driven by the adenoviral E2A late (*Ad2-E2AL*) promoter. This construct was used as hybridization probe in all hybridization experiments.

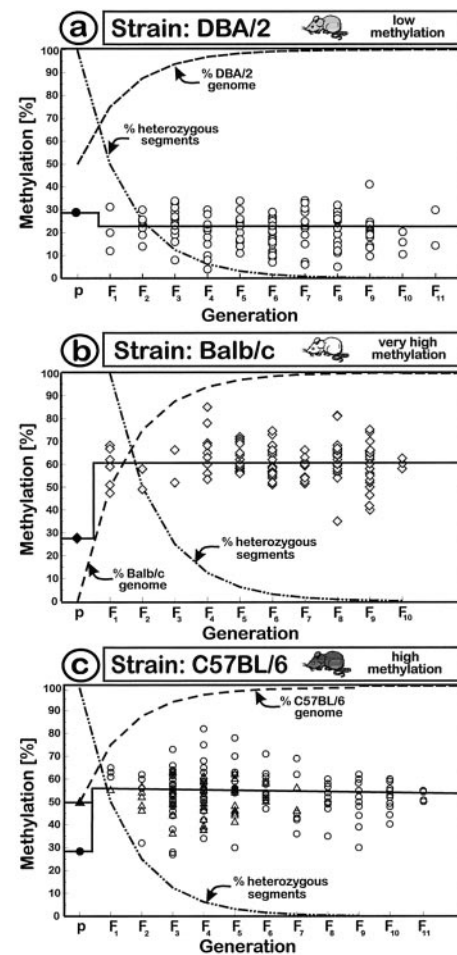
transgenes usually to about 70% of the offspring. Thus, the transgenes were located on separate chromosomes. As previously shown, the 7-1 mouse line carried two short arrays of four (A) and six (B) transgenes, respectively, in two different chromosomal locations (12). Starting from the F<sub>4</sub> cross, the separate transgene arrays were isolated on individual loci, resulting in the mouse lines 7-1A and 7-1B. Most of the six transgenes in the 7-1B animals were completely methylated at the 5'-CCGG-3' (*Hpa*II) sites, whereas transmission of the 7-1A transgenes in a mixed B6D2F<sub>1</sub> background sometimes led to a significant decrease in transgene methylation. Males from both 7-1 mouse lines were subsequently mated to non-transgenic C57BL/6, DBA/2, or Balb/c females. These crosses revealed that the methylation status of the 7-1A locus was influenced by the genetic background of the mouse strain. The DBA/2 background favored the loss of methylation from the 7-1A transgenes, whereas the C57BL/6 and Balb/c backgrounds not only maintained the methylation pattern but led to an increase in transgene methylation. Strain-specific effects were, however, not found for the 7-1B line. This finding indicated position effects on the methylation of the *pAd2E2AL-CAT* transgene. Starting with animals F<sub>5</sub> number 18 and F<sub>5</sub> number 24 (Fig. 1), consecutive matings with animals from different mouse strains

(C57BL/6, Balb/c, DBA/2, CB20, FVB/N, or 129/sv) were performed to assess the effect of the inbred genetic background on the methylation patterns of the transgene in the 7-1A progeny.

**DNA Methylation of the 7-1A Transgenes**—The breeding experiments generated four main 7-1A lines that were screened for over 11 generations as follows: two C57BL/6 lines with moderate to high *de novo* methylation of the transgene, very high methylation levels in the Balb/c line, and the decrease of methylation in the DBA/2 line. Typical transgene methylation profiles of the 7-1A transgene in animals from different strains are presented in Fig. 2a. Among different animals, there is some variability in the extent of transgene methylation. In the experiment shown here, densitometric evaluation of transgene methylation revealed 22% for the DBA/2 animals, 55% for C57BL/6, and 62% for Balb/c mice.

The *Hpa*II-*Msp*I cleavage patterns were evaluated by densitometry. The ratio of *Hpa*II bands representing methylated transgene segments, most of which were larger than 2.0 kbp,<sup>1</sup> to *Msp*I fragments reflecting unmethylated transgene segments facilitated an assessment of the levels of 5'-CCGG-3' methylation in the transgenes of these mice. The cleavage patterns of the seven *Hpa*II sites (Fig. 2b) revealed that the transgene in the DBA/2 line lost methylation, as evidenced by the almost complete absence of the 4.8- and 5.3-kbp off-size bands (Fig. 2a). These bands carried the junctions that linked the transgene to genomic DNA. Cleavage at the unique *Bam*HI restriction site in the transgene produced a strong signal for the 3.3-kbp transgene repeat band, which exhibited about twice the intensity in the DNA from the highly methylating mouse strains C57BL/6 and Balb/c (Fig. 2a). The transgene copies with low levels of methylation were approximately 20% methylated, whereas the hypermethylated copies in the C57BL/6 strain reached levels of methylation between 40 and 60%. In the Balb/c genetic background, the methylation of the *pAd2E2AL-CAT* transgene was in excess of 80%. The high degree of 5'-CCGG-3' methylation correlated with the loss of CAT activity in the transgene. These methylation and expression patterns were identical in all organs, except when the transgenic construct had become demethylated and regained CAT activity (34). The increasing loss of methylation from the transgene in the germ line might signal an epigenetic reprogramming event at early stages of gametogenesis.

**Maintenance and Variability of the Strain-specific Methylation in 7-1A Transgenes**—Obviously, in the chromosomal location of the 7-1A transgene, the genetic background of the animals played a major role in determining the maintenance or alterations of the methylation status in a foreign DNA transgene that had been introduced into the founder generation of hybrid C57BL/6-DBA/2 mice. The kind and number of genetic strain-specific factors affecting *de novo* methylation and its alterations in transgenes were unknown. Therefore, we performed a large number of crosses to transfer the transgene from the hybrid strain into a more homogeneous background and to keep it there for several generations. For example, starting with the F<sub>7</sub> animal 10 (Fig. 1) nine generations of backcrosses to DBA/2 mice resulted in a 99.8% DBA/2 genetic background. Since a DBA/2 background favored the loss of methylation of the 7-1A transgenes, matings were set up to investigate to what extent, after the prolonged crossing to DBA/2 animals, methylation in the transgene was decreased and whether the expression of the CAT gene could be reactivated. With successive backcrossings in a genetic DBA/2 background, methylation of the 7-1A transgene was reduced to and



**FIG. 3. Methylation status of foreign DNA in different inbred mouse strains.** The *pAd2E2AL-CAT* transgene was crossed from a B6D2 hybrid strain into homozygously inbred strains and was transmitted by heterozygous males. The resulting mouse lines were screened for up to 11 backcross generations. None of the analyzed mouse strains exhibited a correlation of transgene methylation and heterozygous DNA segments or strain-specific DNA sequences present (dash-double dot line). a, after crossing the 7-1A transgene into a DBA/2 background, transgene methylation remained at a low level. Although some variability in transgene methylation in different mice was found, the transgene methylation remained stable at a level of 21% (regression line, solid horizontal line) and was transmitted to the following generations. Open circles represent the methylation status of transgenic animals. Filled circle, founder animal. b, crossing the 7-1A transgene from a mixed genetic background ([C57BL/6 × DBA/2] × DBA/2)F<sub>1</sub> into a Balb/c background led to the *de novo* methylation in the transgene. The methylation level then remained constant at around 61.5%. Each diamond represents the methylation status of one transgenic animal. Filled diamond, founder animal. c, each open circle represents a transgenic animal derived from founder animal 24F<sub>5</sub> (mouse line C57BL/6-b, see Fig. 1), whereas a triangle represents an animal derived from founder animal 16F<sub>6</sub> (mouse line C57BL/6-a). Filled triangle, founder animal. Both transgenic lines harbored hypermethylated 7-1A transgenes (fluctuating around 55%) after initial *de novo* methylation. The dashed line (dashes only) indicates the theoretical values for percentages of original DBA/2 (a), Balb/c (b), and C57BL/6 (c) genomes remaining after increasing numbers of crosses.

maintained at the 10–30% level (Fig. 3a). In these breedings, the transgene was always transmitted by the males to rule out a putative cytoplasmic reprogramming mechanism of the transgene during oogenesis. However, no significant variations in methylation were observed when transgenic 7-1A females were crossed with males of the same strain (data not shown). Regression analyses of methylation levels (Fig. 3a) demonstrated that in the DBA/2 animals transgene methylation remained stable at a level of 21%. Among siblings of the same

<sup>1</sup> The abbreviations used are: kbp, kilobase pairs; CAT, chloramphenicol acetyltransferase; Ad2, adenovirus type 2.

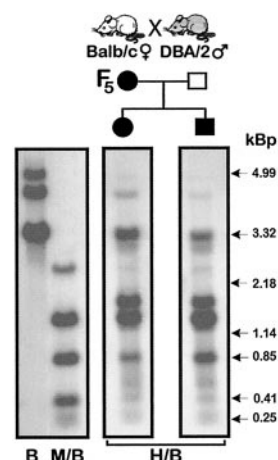
litter, a 10% variation in transgene methylation was found. However, these differences were not transmitted to the offspring. Even in a pure genetic background of >99%, the transgene methylation remained mosaic and was associated with the silencing of the *pAd2E2AL-CAT* transgene.

All transgenes in animals resulting from Balb/c  $\times$  7-1A-Balb/c or C57BL/6  $\times$  7-1A-C57BL/6 crosses remained hypermethylated after the initial *de novo* methylation in the F<sub>1</sub> generation (Fig. 3, *b* and *c*). The transgenes in the C57BL/6 animals showed comparable methylation levels at around 55%; transgenic animals from backcrosses into Balb/c had a methylation level of 61.5% in the transgenes. In earlier work, we had shown in crosses with mixed genetic background BDF<sub>1</sub>  $\times$  BDF<sub>1</sub> (Fig. 1) that in only 10% of the offspring transgene methylation was lost, and in the following generations the grandparental methylation patterns were usually reestablished (12). The ensemble of results (Fig. 3) suggests that both mouse strains, C57BL/6 and Balb/c, supply dominant factors favoring *de novo* methylation. The methylation dynamic of the 7-1A transgene resembles that of the *TKZ751* transgene in mice (22, 23) in which the Balb/c genetic background led to stronger *de novo* methylation than the C57BL/6 background.

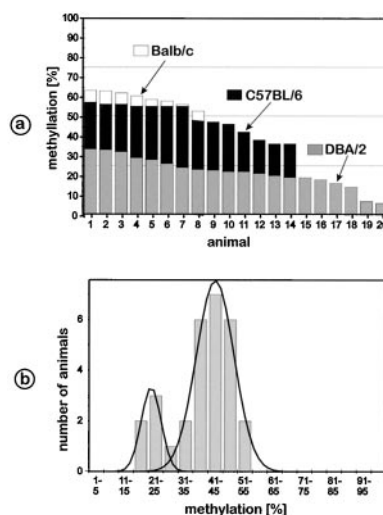
In addition to these B6 mouse lines, an additional C57BL/6-(b) line was established. In contrast to C57BL/6-(a), this line originated from animal number 24F<sub>5</sub> which was derived from animals showing no loss of methylation in their transgenes. After the first backcross, however, these animals exhibited a methylation pattern similar to that in the C57BL/6-(a) line. In both lines, the variation of transgene methylation was somewhat higher than in the other strains analyzed. For a more detailed analysis of these reproducibly observed strain-specific modifications, we backcrossed the 7-1A transgene into mouse strains FVB/N, 129/sv, or CB20 (data not shown). Starting from a DBA/2 animal with a hypomethylated transgene, its transfer into the FVB/N strain did not further affect the low level of transgene methylation. This observation paralleled the behavior of the *RSV1gmyc* transgenes in the same mouse strain (36). Similar results were obtained with 129/sv mice. Again like for other transgenes (22), the 129/sv background favored a moderate loss of methylation of the transgene. In testis, the transgene was fully demethylated, an effect typical for many strain-specifically methylated transgenes (8, 21, 34). In contrast to the FVB/N and 129/sv strains, the CB20 background led to *de novo* methylation. The CB20 mouse strain originated from Balb/c and C57BL/6 animals and might also harbor factors promoting methylation.

**Stability of the 7-1A Methylation Patterns**—The observed strain-specific variations of transgene methylation are presumably subject to the control of modifying genes that are different in individual mouse strains. To test this hypothesis, a series of crosses was analyzed. At first, we investigated to what extent the highly methylated transgene phenotype in Balb/c mice could be reversed. Hence, we crossed the highly methylated transgene in the Balb/c animals into a DBA/2 background (Fig. 4). Conversely, the hypomethylated transgene from DBA/2 animals was backcrossed into a Balb/c background (Fig. 5*a*). In both experiments, the transgenes in the offspring showed the hypermethylated phenotype of the Balb/c mice. Therefore, the Balb/c genetic background seemed to play a dominant role in this process. At least after the first backcross, the methylation phenotypes were not immediately reversed.

We also tested whether the genetic factors would work in an independent fashion or, when inherited as heterozygous loci, by complementing each other. The hypermethylated transgene from Balb/c males was, therefore, crossed into a C57BL/6 genetic background. Intriguingly, all offspring showed the typical



**FIG. 4. Genetic dominance of the Balb/c background.** Crosses of 7-1A Balb/c animals with non-transgenic DBA/2 mice always resulted in the maintenance of the hypermethylated state of the transgene (>50% methylated). Off-size band and 3.3-kbp transgene repeat bands were not cut. The same results were seen when transgenic DBA/2 mice were crossed with non-transgenic Balb/c animals (see Fig. 5*a*). The Balb/c genetic background seems to play a dominant role in the generation of the 7-1A methylation pattern. Tail DNA was cut with *Bam*HI (lane 1) or with *Bam*HI and *Hpa*II (lanes 3 and 4). Double cleavage with *Bam*HI and *Msp*I (lane 2) served as reference for unmethylated fragments. Symbols are as follows: *H*, *Hpa*II; *M*, *Msp*I; *B*, *Bam*HI. Open squares, male, non-transgenic animals; filled squares, male; filled circles, female mice carrying the 7-1A transgene.



**FIG. 5. Epigenetic stability of the hypomethylated transgene copies.** *a*, methylation status of transgenic progeny from crosses with 7-1A DBA/2 males with hypomethylated transgene. The following crosses were analyzed: C57BL/6  $\times$  7-1A-DBA/2 (black bars), Balb/c  $\times$  7-1A-DBA/2 (white bars), and DBA/2  $\times$  7-1A-DBA/2 (dark gray bars). All male transgenic animals were randomly chosen for analyses and stemmed from the 9th DBA/2 backcross generation with an estimated 99.8% pure genetic background. All transgenes that were maintained in a DBA/2 background remained hypomethylated, whereas transmissions into Balb/c or C57BL/6 genetic backgrounds always led to significant *de novo* methylation of the transgenes. Therefore, both genetic backgrounds, Balb/c and C57BL/6, may contribute at least one dominant factor directing the *de novo* methylation of the 7-1A transgene. *b*, non-mendelian segregation of methylation phenotypes in progeny of crosses (C57BL/6  $\times$  DBA/2)F<sub>1</sub>  $\times$  7-1A-DBA/2. One-third of the F<sub>1</sub> animals exhibited the typical DBA/2 pattern (0–30% methylation), and two-thirds showed the C57BL/6 phenotype in their 7-1A transgenes (31–60% methylation). Methylation indices were determined as described under “Experimental Procedures.”

B6 methylation pattern of the 7-1A transgene (Fig. 6). These data indicated that the B6 allele was not only dominant over the DBA/2 allele (Fig. 5*a*) but also over the Balb/c allele. This

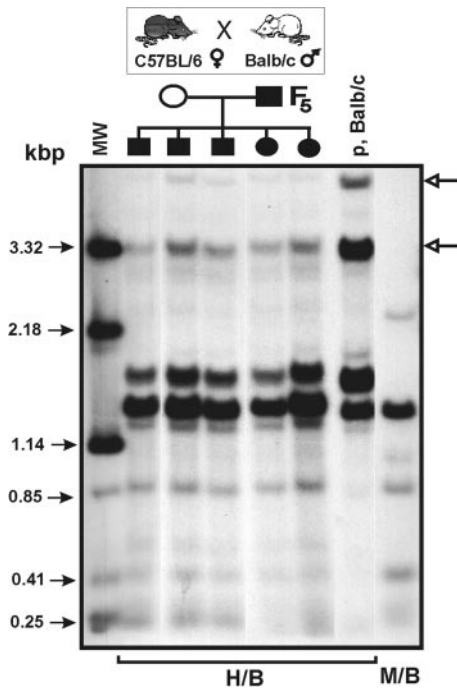


FIG. 6. Factors affecting *de novo* methylation in C57BL/6 and Balb/c genetic backgrounds cannot complement each other. To test whether factors can work in an independent fashion or, when transmitted as heterozygous loci, by complementing each other, the hypermethylated 7-1A transgene was crossed from the Balb/c genetic background into the C57BL/6 background. Tail DNA (20  $\mu$ g) of mice derived from the cross C57BL/6  $\times$  7-1A-Balb/c was cut with *Bam*HI and the methylation-sensitive endonuclease *Hpa*II. The parent animal (p.Balb/c) exhibited a typical Balb/c methylation status of 73%. After the cross, all transgenes underwent a loss of methylation which led to the typical C57BL/6 methylation pattern (45–58% methylation). No increase in 5'-CCGG-3' methylation was found. Symbols are as follows: p, parental animal; H, *Hpa*II; M, *Msp*I; B, *Bam*HI; open circles, female, non-transgenic animal; filled squares, male; filled circles, female mice carrying the 7-1A transgene. The arrows designate the positions of the 4.8- and 3.4 kbp fragments that decreased in intensities in the DNA from the offspring of this cross.

result suggests the existence of a dominant factor affecting levels of DNA methylation in the C57BL/6 genetic background.

*Segregation of the High and Low Methylation Phenotypes in the Progeny of 7-1A males with (B6  $\times$  D2)F<sub>1</sub> Females*—The initial strategy was to correlate the methylation in the offspring of crosses between DBA/2 males and (B6  $\times$  D2)F<sub>1</sub> females with previously characterized polymorphic loci. But instead of either the B6 or the D2 methylation phenotypes, additional methylation phenotypes were observed (data not shown). This result indicated that a multitude of factors could contribute to the establishment of distinct methylation phenotypes. Similarly complex patterns of methylation were encountered in segregation analyses with the transgenes *E36* (25, 37, 38) and *TKZ751* (22).

Further evidence for the complex role of genetic backgrounds in transgene methylation was adduced by crossing (C57BL/6  $\times$  DBA/2) F<sub>1</sub> females with 7-1A males of the DBA/2 background. Consistent with the notion of multiple methylation modifier loci, a non-mendelian segregation of methylation levels was apparent in the progeny (Fig. 5b). One-third of the offspring carried hypermethylated transgenes, whereas two-thirds of the progeny showed a *de novo* methylation that represented the B6 phenotype. These results indicate that the effect of several factors combine to modulate the transgene methylation in the 7-1A lines.

Interestingly, the observed *de novo* methylation might be directed to specific 5'-CG-3' dinucleotides (Fig. 7). *Hpa*II and

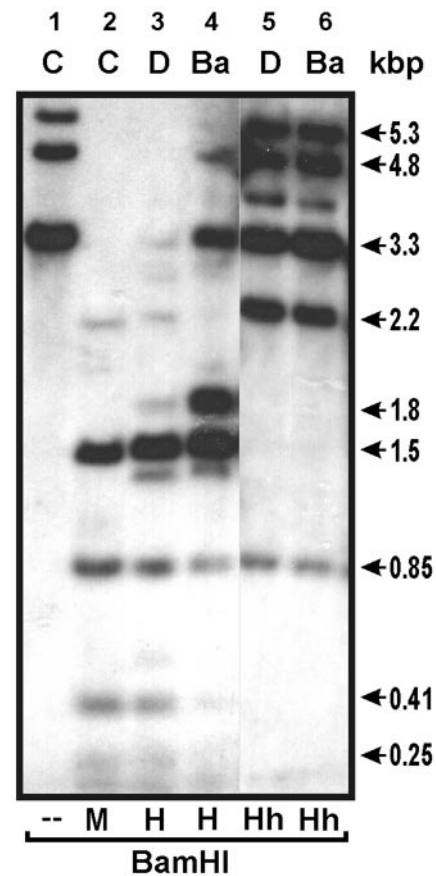


FIG. 7. The unique *Hha*I (GCGC) site escapes strain-specific loss of methylation. After *Hpa*II treatment, the tail DNA was cleaved differentially, depending on the genetic background. The 3.3-kbp transgene repeat band was cut to almost 100% in transgenes in the DBA/2 (D) background (lane 3), whereas a large portion remained uncleaved in the Balb/c (Ba) genetic background (lane 4). When the same DNA was treated with *Hha*I, however, the cleavage patterns were not influenced by the genetic background (lanes 5 and 6). Symbols are as follows: H, *Hpa*II; M, *Msp*I; Hh, *Hha*I; D, DBA/2; C, control animal (DBA/2).

*Hha*I cleavage patterns of transgenic DNA from DBA/2, Balb/c, and C57BL/6 (not shown) were compared among animals of both sexes, which were randomly chosen from the sixth and seventh backcross generations. Although striking differences in *Hpa*II-(CCGG) methylation were observed, the same animals showed no difference in their *Hha*I-(GCGC) cleavage patterns. This *Hha*I sequence is the only such site located within the pBR322 sequences of the construct. We conclude that the process of strain-specific transgene methylation seems to be a very specific modification that is dependent on more than one host-specific factor as well as on position effects.

#### DISCUSSION

*Factors Affecting Transgene Methylation*—In recent years, a large number of transgenic mice have been generated to study various issues in genetics, like gene function or gene therapy. It is of considerable interest to understand the events accompanying or following the insertion of foreign DNA into established mammalian genomes (39). Not only the transgene affects the biology of the host, the transgene itself is often the target for several modifications. One of the major changes is the *de novo* methylation of the transgene upon insertion into the recipient genome. This heritable genetic modification modulates the overall genomic patterns of chromatin organization and gene expression without altering the nucleotide sequence. The extent of *de novo* methylation of the transgene depends on the mouse strain bearing the transgene. The 7-1A transgene used

TABLE I  
Strain-specific methylation of transgenes

Transgene methylation, which often correlates with silencing of the transgene construct, is usually observed in specific mouse strains (C57BL/6 and Balb/c). Other mouse strains like DBA/2, FVB/N, and 129 seem to favor the loss of or decrease in methylation of foreign DNA and therefore facilitate transgene expression.

Transgene	Mouse strains								
	C57BL/6	Balb/c	DBA/2	FVB/N	129	CB20	CBA	C3H	SJL
7-1A <sup>1</sup>	↑	↑	↓	↓	↓	↑			
TKZ751 <sup>2</sup>	↑↓ <sup>4</sup>	↑	↓		↓		↑↓		
pHRD <sup>3</sup>	↑		↓						↓
RSVlgmyc <sup>4</sup>	↑			↓					
HBsAg <sup>5</sup>	↓	↑	↓					↓	
CAT17 <sup>6</sup>	↑	↑↓ <sup>4</sup>	↓				↓		
Troponin I 379 <sup>7</sup>	↑		↓						

\* A segregation of methylation phenotype may be due to a CFLP (TKZ751) or CBA (CAT17) background in the founder animals.

<sup>1</sup> This study.

<sup>2</sup> From Ref. 22.

<sup>3</sup> From Ref. 24.

<sup>4</sup> From Ref. 42.

<sup>5</sup> From Ref. 25.

<sup>6</sup> From Ref. 23.

<sup>7</sup> From Ref. 21. Up arrows indicate high methylation, and down arrows indicate low methylation.

in this report consists of the CAT reporter gene driven by the adenovirus type E2A late promoter which is silenced after integration into the mouse genome. In early work from this laboratory, an inverse correlation between the methylation of the E2A late promoter and other adenovirus promoters and patterns of gene expression has been documented (2, 3, 34, 40, 41). The *de novo* methylation of transgenes is generally associated with altered chromatin organization that is inhibitory for transcription (6, 31). Nevertheless, the mechanisms behind this reprogramming remain an intriguing problem. Various factors are relevant in determining the extent and pattern of DNA methylation of the *pAd2-E2AL-CAT* transgene and of foreign DNA in general as follows: (a) position effects; (b) the mode of recombination; (c) specific bacterial or viral motifs in the foreign DNA; (d) the nucleotide sequence of the inserted construct; (e) the timing of insertion; and (f) the influence of other genes in the recipient cell or organism.

In this study, we have shown that the *Ad2-E2AL-CAT* transgene methylation is influenced in a strain-specific manner by a very coordinated regulation. In the mouse strains Balb/c, CB20, and C57BL/6, the *7-1A* transgenes become highly *de novo* methylated, whereas a decrease in methylation is observed when the transgene is transmitted by DBA/2, 129/sv, or FVB/N animals. Obviously, the genetic background of the mouse strain can be a crucial factor in determining the expression levels of integrated foreign DNA. The expression of the transgene loci is inversely correlated with the extent of their methylation. Strain-specific effects are, of course, not restricted to our transgene model; other transgenes, regardless of their nature, can show similar characteristics when introduced into these mouse strains (Table I). The *TKZ751* transgene (22) becomes repressed and methylated when *TKZ751* males fertilize Balb/c eggs but not DBA/2 eggs. The brief transient exposure of C57BL/6 male pronuclei to the DBA/2 egg cytoplasm in the mouse compromises the ability of these nuclei in preimplantation development even when these nuclei are transplanted back into C57BL/6 recipient eggs (43). However, some transgenes are expressed in a strain-specific manner and are not *de novo* methylated upon passage into these backgrounds.

*The Dynamics of 7-1A Methylation and the Role of Modifying Genes*—The dynamics of the *de novo* methylation of the *7-1A* transgene might shed light on the amount and nature of strain-specific factors. Immediately after the first cross of the *7-1A* transgene in the Balb/c or C57BL/6 backgrounds, *de novo*

methylation was initiated. The newly established methylation pattern was subsequently transmitted to the offspring and remained stable in the following generations. This dominant methylation was a unique event with no further modifications, and indicated that both mouse strains supplied at least one, perhaps several, dominant allele that affected the stability of the pattern of methylation.

In 1991, Engler and co-workers (24) reported that a modifying locus regulated the methylation phenotype of a *pHRD* transgene that encoded a target sequence for the VDJ recombinase. Another interesting observation in the *7-1A* lines was the decrease in methylation in the transgenes in the DBA/2, 129/sv, and FVB/N strains. Since these mice show no maintenance of the transgene methylation as compared with the methylation status in the founder animals, the loss of methylation might be an active process or the consequence of a failure in the maintenance mechanism. Thus, the methylation of the transgene could be considered the “default” state of the surrounding genomic sequences. Similar events were proposed for the *RSVlgmyc* transgene (36) which became hypomethylated and expressed in an inbred FVB/N genetic background but hypermethylated and silenced in inbred C57BL/6 mice. In contrast to the *7-1A* construct, the *RSVlgmyc* transgene was also regulated by genomic imprinting, with the transgene being undermethylated when transmitted by the father and hypermethylated and silent when transmitted by the mother. Furthermore, the imprint was lost in the C57BL/6 genetic background but was restored by breeding the transgene into the FVB/N strain. Variation in the imprinting behavior in the C57BL/6 background suggested that elements in addition to the strain-specific factors had influenced *de novo* methylation during embryogenesis.

In mice carrying the *7-1A* transgene, no significant variations in transgene methylation could be observed, when the transgene was transmitted by females instead of *7-1A* males. Furthermore, the methylation pattern remained stable, when the transgene was transmitted as homozygous *7-1A* locus (data not shown). These data suggest the following two alternative mechanisms affecting the methylation of the *7-1A* transgene: an active decrease in methylation or methylation protection and a dominant mechanism supplied by the C57BL/6, Balb/c, and CB20 backgrounds resulting in the *de novo* methylation of the *7-1A* transgene. The observation that the *7-1A* transgene

became dominantly *de novo* methylated in (B6 × D2)F<sub>1</sub> and (Balb/c × D2)F<sub>1</sub> hybrids was more compatible with the latter mechanism. However, the protection from *de novo* methylation would be incomplete and dosage-dependent. Successive backcrossings of the transgene into the DBA/2 background did not lead to the complete loss of methylation in the 7-IA transgene, although the decrease in transgene methylation could have been the result of changes in the proportion of cells carrying differentially methylated transgenes. In addition, the offspring of the highly inbred DBA/2 animals crossed with Balb/c animals always showed direct *de novo* methylation of the transgene.

The mechanism of *de novo* methylation might be related to an ancient defense system targeted against foreign DNA (1, 44–46). The 7-IA transgene consists of viral sequences (Ad2 and SV40) as well as bacterial DNA (pBR322; Fig. 2b). The importance of specific DNA sequences for the methylation status of transgenes has been previously discussed (36, 47). A short repetitive IgA sequence in the *RSVIgmyc* transgenes conveyed parental imprinting in a strain-specific manner. In another *pHRD* transgene, an *E. coli* sequence (*gpr*) might be recognized by the host and could therefore have served as a focus for the strain-specific modifier action with spreading of DNA methylation into the flanking sequences (47). It is unknown whether a specific sequence or repetitive structure within the 7-IA transgene can help trigger strain-specific methylation.

We have previously reported that the *de novo* methylation of the *pAd2E2AL-CAT* transgene is influenced by the position of integration (12). It is important to emphasize that only one of the six transgenic mouse lines then studied (34) showed strain-specific effects. These results are reminiscent of the phenomenon of position effect variegation in *Drosophila* where proximity to heterochromatic regions predisposes to variegation. Models have been proposed to explain the transgene inactivation by factors similar to the *Drosophila Su(Var)* genes that function in the epigenetic control of gene expression, presumably by modifying higher order chromatin structures (48). Recently, two vertebrate homologues, the human SUV39H1 and the mouse Suv39hl proteins, were isolated (49). The mammalian homologues contained highly conserved sequence motifs, the chromo and SET domains, a combination that was also preserved in the *Schizosaccharomyces pombe* silencing factor *clr4* (50). Therefore, based on the notion of chromatin modeling and the fact that mammalian heterochromatin is often enriched in 5'-CG-3' methylation, the methylation status observed in the 7-IA transgene in different mouse strains might reflect secondary events following the formation of heterochromatin.

Our data also underscore the importance to choose the right mouse strain for gene targeting or embryogenic stem cell generation. The inbred genetic background may be decisive in determining the stability and continued expression of foreign DNA. Since the Balb/c and C57BL/6 backgrounds favor frequent inactivation of transgene constructs, mice from the DBA/2, FVB/N, or 129 strains would be preferable for targeting experiments.

Much further work will be necessary to understand the role of genetic background in the maintenance or alteration of the methylation status that affects the phenotypes associated with mutations in mice and in human disease. DNA methylation in mammalian genomes has frequently been interpreted as an epigenetic effect (51). Since patterns of DNA methylation in specific segments of a mammalian genome with a specific over-

all genetic background are stable and inherited, it might be more realistic to view these patterns as true genetic, rather than epigenetic, signals.

## REFERENCES

- Sutter, D., Westphal, M., and Doerfler, W. (1978) *Cell* **14**, 569–585
- Sutter, D., and Doerfler, W. (1980) *Cold Spring Harbor Symp. Quant. Biol.* **44**, 565–568
- Sutter, D., and Doerfler, W. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 253–256
- Toth, M., Lichtenberg, U., and Doerfler, W. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 3728–3732
- Toth, M., Müller, U., and Doerfler, W. (1990) *J. Mol. Biol.* **214**, 673–683
- McGowan, R., Campbell, R., Peterson, A., and Sapienza, C. (1989) *Genes Dev.* **3**, 1669–1676
- Surani, M. A., Kothary, R., Allen, N. D., Singer, P. B., Fundele, R., Ferguson-Smith, A. C., and Barton, S. C. (1990) *Development* **3**, (suppl.) 89–98
- Chaillet, J. R., Vogt, T. F., Beier, D. R., and Leder, P. (1991) *Cell* **66**, 77–83
- Sasaki, H., Hamada, T., Ueda, T., Seki, R., Higashinakagawa, T., and Sasaki, Y. (1991) *Development* **111**, 573–581
- Kakutani, T., Jeddeloh, J. A., and Richards, E. J. (1995) *Nucleic Acids Res.* **11**, 130–137
- Martin, C. C., and McGowan, R. (1995) *Genet. Res.* **65**, 21–28
- Koetsier, P. A., Mangel, L., Schmitz, B., and Doerfler, W. (1996) *Transgenic Res.* **5**, 235–244
- Monk, M. (1990) *Trends Genet.* **6**, 110–114
- Monk, M., Boubelik, M., and Lehnert, S. (1987) *Development* **99**, 371–382
- Sanford, J. P., Clark, H. J., Chapman, V. M., and Rossant, J. (1987) *Genes Dev.* **1**, 1039–1046
- Howlett, S. K., and Reik, W. (1991) *Development* **113**, 119–127
- Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H., and Razin, A. (1992) *Genes Dev.* **6**, 705–714
- Heller, H., Kämmer, C., Wilgenbus, P., and Doerfler, W. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 5515–5519
- Remus, R., Kämmer, C., Heller, H., Schmitz, B., Schell, G., and Doerfler, W. (1999) *J. Virol.* **73**, 1010–1022
- Lacy, E., Roberts, S., Evans, E. P., Burtenshaw, M. D., and Constantini, F. D. (1983) *Cell* **34**, 343–358
- Sapienza, C., Paquette, J., Tran, T. H., and Peterson, A. (1989) *Development* **107**, 165–168
- Allen, N. D., Norris, M. L., and Surani, M. A. (1990) *Cell* **61**, 853–861
- Reik, W., Howlett, S. K., and Surani, M. A. (1990) *Development* **3**, (suppl.) 99–106
- Engler, P., Haasch, D., Pinkert, C. A., Doglio, L., Glymour, M., Brinster, R., and Storb, U. (1991) *Cell* **65**, 939–947
- Schweizer, J., Valenza-Schaerly, P., Goret, F., and Pourcel, C. (1998) *DNA Cell Biol.* **17**, 427–435
- Latham, K. E. (1995) *Differentiation* **59**, 269–282
- Reik, W., and Walter, J. (1998) *Curr. Opin. Genet. & Dev.* **8**, 154–164
- Reik, W. (1998) *Int. J. Dev. Biol.* **42**, 951–954
- Reik, W., Collick, A., Norris, M. L., Barton, S. C., and Surani, M. A. (1987) *Nature* **328**, 248–251
- Sapienza, C., Peterson, A. C., Rossant, J., and Balling, R. (1987) *Nature* **328**, 251–254
- Swain, J. L., Stewart, T. A., and Leder, P. (1987) *Cell* **50**, 719–727
- DeLoia, J. A., and Solter, D. (1990) *Development* **3**, (suppl.) 73–79
- Beck, J. A., Lloyd, S., Hafezparast, M., Lennon-Pierce, M., Eppig, J. T., Festing, M. F. W., and Fisher, E. M. C. (2000) *Nat. Genet.* **24**, 23–25
- Lettmann, C., Schmitz, B., and Doerfler, W. (1991) *Nucleic Acids Res.* **19**, 7131–7137
- Koetsier, P. A., Schorr, J., and Doerfler, W. (1993) *BioTechniques* **15**, 260–262
- Chaillet, J. R., Bader, D. S., and Leder, P. (1995) *Genes Dev.* **9**, 1177–1187
- Hadchouel, M., Farza, H., Simon, D., Tiollais, P., and Pourcel, C. (1987) *Nature* **329**, 454–456
- Wu, X., Hadchouel, M., Farza, H., Amar, L., and Pourcel, C. (1992) *Mechanisms of Eukaryotic DNA Recombination*, pp. 69–73, Academic Press, New York
- Doerfler, W. (1999) *Foreign DNA in Mammalian Systems*, Wiley-VCH Publishers, New York
- Vardimon, L., Neumann, R., Kuhlmann, I., Sutter, D., and Doerfler, W. (1980) *Nucleic Acids Res.* **8**, 2461–2473
- Müller, U., and Doerfler, W. (1987) *J. Virol.* **61**, 3710–3720
- Weichman, K., and Chaillet, J. R. (1997) *Mol. Cell. Biol.* **17**, 5269–5274
- Latham, K. E., and Solter, D. (1991) *Development* **113**, 561–568
- Jähner, D., Stuhlmann, H., Stewart, C. L., Harbers, K., Lohler, J., Simon, I., and Jaenisch, R. (1982) *Nature* **298**, 623–628
- Doerfler, W. (1991) *Biol. Chem. Hoppe-Seyler* **372**, 557–564
- Yoder, J. A., Walsh, C. P., and Bestor, T. H. (1997) *Trends Genet.* **13**, 335–340
- Engler, P., Doglio, L. T., Bozek, G., and Storb, U. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10763–10768
- Sapienza, C. (1990) *Development* **3**, (suppl.) 107–113
- Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P. B., Reuter, G., and Jenuwein, T. (1999) *EMBO J.* **18**, 1923–1938
- Ivanova, A. V., Bonaduce, M. J., Ivanov, S. V., and Klar, A. J. S. (1998) *Nat. Genet.* **19**, 192–195
- Holliday, R. (1987) *Science* **238**, 163–170