Cellular mosaicism in the methylation and expression of hemizygous loci in the mouse

Ross McGowan, Rob Campbell, Alan Peterson, and Carmen Sapienza
Ludwig Institute for Cancer Research, West Montreal, Quebec H3A 1A1 Canada

Proposed models for the inheritance of locus-specific methylation phenotypes in somatic cells include those in which there is stable inheritance of a methylation pattern such that all cells contain a similarly methylated locus, as well as models in which the inheritance of methylation can be variable. We investigated these possibilities by examining the methylation and expression of hemizygous loci in the mouse. Our results demonstrate that differences in both methylation and expression can exist between apparently identical cells and that such mosaicism is genetically controlled.

[Key Words: Methylation; mosaicism; transgene]
Received June 29, 1989; revised version accepted August 22, 1989.

The modification of mammalian DNA most commonly involves the addition of a methyl group to position 5 of cytosine, usually when that base precedes a guanosine (Vanyushin et al. 1970). Cytosine methylation within or adjacent to many loci is thought to affect, or be affected by, several biochemical processes involved in the control of gene expression, as well as mutation and genomic imprinting: Actively transcribed sequences tend to be undermethylated relative to those that are not being transcribed (Doerfler 1983); DNA sequences that are highly methylated, such as satellite DNA (Selig et al. 1988) or those on the inactive X chromosome (German 1962), tend to be late-replicating, the methylation of cytosine increases the rate of C-to-T transitions, resulting in point mutations at those sites at an increased frequency (Wang et al. 1982); and the methylation of some transgenes has been shown to be affected by genotype of origin and therefore to show a genomic imprint (Hatchouel et al. 1987; Reik et al. 1987; Sapienza et al. 1987; Swain et al. 1987).

Because the demonstration that the methylation of CpG sites is heritable through DNA replication and cell division (Wigler et al. 1981), it has been generally assumed that once the methylation pattern of a locus has been established in an embryonic cell, all of its descendants should faithfully reproduce the pattern such that they all remain identical with respect to the methylation of any particular CpG site. We demonstrated previously that the methylation phenotype of the 379 quail troponin I transgene locus could be affected by genetic and epigenetic factors to produce three different phenotypes in different individuals (Sapienza et al. 1989a). These different phenotypes could reflect methylation differences either between or within the transgene array.

To distinguish between these possibilities, we analyzed the methylation phenotype of hemizygous ecotropic virus loci and a lacZ-containing transgenic line. The expression characteristics of the lacZ-containing transgene were also analyzed using an in situ histochemical assay. The evidence we obtained suggests that cells sharing the same developmental and phenotypic lineage can express different DNA methylation phenotypes within an individual. Furthermore, such otherwise identical cells can differ in the expression of a transgene locus, resulting in a mosaic phenotype.

Results

In an attempt to identify whether methylation differences could exist between cells within a tissue, we assayed the methylation state of a Smal site adjacent to an endogenous ecotropic virus locus. Because different inbred mouse strains contain such loci at different chromosomal sites (Jenkins et al. 1982), F1 hybrids between strains that carry different ecotropic virus loci are hemizygous at each locus. This situation is analogous to that of hemizygous transgene loci but can be interpreted more easily because the ecotropic virus locus consists of only a single copy of the provirus.

The DDK inbred strain contains a single ecotropic virus locus (designated Env-DDK in Fig. 1), whereas the RF/J strain contains three such loci (designated Env-1, Env-16, and Env-17 in Fig. 1). Genetic analyses indicate that Env-16 and Env-17 are tightly linked (Jenkins and Copeland 1985), whereas Env-1 (Jenkins and Copeland 1985) and Env-DDK (C. Sapienza, unpubl.) segregate independently. F1 hybrids between DDK and RF/J are thus hemizygous at each locus. The single copy of the Env-
DDK provirus contained within each cell was contributed by the DDK female through an ovum.

Any particular methylation-sensitive restriction endonuclease cleavage site in or around those hemizygous loci must have been either methylated or unmethylated in the single gamete of origin. Because we are examining only a single restriction site within each cell, a somatically stable inheritance model predicts that the site should either be methylated in all cells and, therefore, completely insensitive to cleavage by Smal or unmethylated in all cells, and, therefore, completely sensitive to cleavage by Smal. However, if this site is methylated differently in different cells, such a tissue mosaic should give rise to a DNA sample that is partially sensitive to cleavage by Smal.

The PvuII and Smal cleavage patterns of the Emv-DDK, Emv-1, Emv-16, and Emv-17 loci in DDK-RF/J hybrids are shown in Figure 1. Whereas Emv-1, Emv-16, and Emv-17 appear to be completely insensitive to cleavage by Smal, the PvuII fragment corresponding to the Emv-DDK locus becomes reduced in intensity and a new band of ~8 kb appears.

Because each cell carries only one copy of each provirus, the partial Smal sensitivity of DNA isolated from whole tissue must reflect differences at the level of individual cells. Furthermore, the same cleavage pattern is found in a variety of somatic tissues whose last common developmental ancestor was primitive ectoderm (data not shown). Apparently identical proportions of methylated and unmethylated cells therefore exist in all tissues, implying either that methylation differences between cells are established early in development and are then faithfully propagated or that these differences are generated continuously.

Because such evidence for methylation differences at the level of individual cells is indirect and could conceivably be due to different cell types within a tissue, we sought to define an assay system capable of resolution at the level of the single cell. Such a system was available in the Tg4 transgenic line of mice. The Tg4 transgene locus consists of 15–20 copies of a mouse hsp68–Escherichia coli lacZ construct inserted at the dystonia musculorum locus on chromosome 1 (Kothary et al. 1988). In this particular line, the lacZ gene is expressed in the floor plate of the neural tube during days 10–14 of development. However, lacZ expression was absent in the neural tubes of some transgenic fetuses. Therefore, we investigated whether the methylation of Mspl sites at this transgene locus correlated with the expression of lacZ in the neural tube. Figure 2 shows that day-12 fetuses containing a hypomethylated transgene express the lacZ gene in the neural tube, whereas fetuses containing a hypermethylated transgene do not. As has been observed with other transgene loci, the methylation phenotype of this transgene was identical in all nonexpressing tissues, with the exception of testes and extraembryonic membranes (Reik et al. 1987; Sapienza et al. 1987, Swain et al. 1987).

Furthermore, within this transgenic line, as in the troponin I line described previously (Sapienza et al. 1989a), individuals can not only carry the transgene locus in high or low methylation states but also in intermediate states. The observed correlation between methylation and expression of this transgene (Fig. 2) allows us to distinguish between the two possible ways in which such an intermediate methylation phenotype could arise. If the intermediate transgene methylation phenotype observed in these mice is shared by all the cells in the neural floor plate, lacZ expression should also be identical in all cells. However, if the intermediate methylation phenotype were a composite of both hyper- and hypomethylated cells, the staining pattern should reveal this mosaicism. The data presented in Figure 3 support the latter possibility. In these individuals the intermediate methylation phenotype was associated with a reduced number of expressing cells in this region, which we interpret to represent a mosaic population of expressing and nonexpressing cells. Moreover, the proportion of cells that expressed lacZ was inversely correlated with the degree of methylation of the transgene. The intensity of staining appeared to be equivalent in all fetuses stained on the same occasion (whether mosaic or not), although the staining intensity did vary between trials.

In previous experiments with the 379 troponin I transgene, we observed the methylation phenotype of the transgene in an individual to be affected by the genotype of the nontransgenic female parent (Sapienza et al. 1989a). To determine whether the Tg4 transgene locus was similarly affected, we analyzed fetuses derived from crosses of transgenic males to females of different genotypes.
Strains. Figure 4 shows that offspring of the same male may exhibit different transgene methylation phenotypes. Matings to a C57BL/6 female produced a hypomethylated phenotype and 'complete' staining, whereas offspring of the BALB/c female cross were mosaic for lacZ expression and had an intermediate transgene methylation phenotype. Similar results have been obtained from matings involving five different Tg4 transgene-carrying males which produced a total of 52 transgenic progeny (22 from C57BL/6 females and 30 from BALB/c females). Only a single exception was found in which one fetus in a litter of completely staining siblings from a cross to a C57BL/6 female showed incomplete staining in its neural tube. These findings demonstrate that the methylation and expression of the transgene are under genetic control and that differences exist between common inbred strains of mice. Furthermore, this level of regulation must involve a trans-acting factor because the maternal genetic effect is realized after fertilization.

Discussion

We have previously described a transgene locus, the 379 troponin I gene [Sapienza et al. 1989a], that could display three different somatic methylation phenotypes. A similar situation has been identified for another transgenic locus, the Tg4 gene. These three phenotypes could reflect one or both of the situations illustrated in Figure 5.
Either (1) differences exist in the methylation of \textit{MspI} sites among different copies within a transgene array, but within an individual all the arrays are alike, such that each cell within a tissue carries a transgene array which is identically methylated [Fig. 5A], or (2) differences in the methylation of the transgene array exist between cells, such that some cells carry an array with one methylation pattern and others carry an array with a different methylation pattern [Fig. 5B].

The model in Figure 5A assumes that the methylation pattern of the array is established only once and that, once established, the pattern is stably propagated in all of the cells that give rise to somatic tissues. The model in Figure 5B assumes that cells with at least two different methylation phenotypes are established after fertilization and that both types of cells may contribute to somatic tissue [Sapienza et al. 1989b].

Analyses of a single methylatable site, the \textit{CpG} dinucleotide, contained in a \textit{SmaI} recognition sequence adjacent to a hemizygous ecotropic virus locus identified both a methylated and nonmethylated \textit{SmaI} site in the same tissue, consistent with the idea that different cells within the same tissue can be methylated differently. If the data in Figure 1 were an isolated case, one might explain these results as being due to a random failure to methylate the affected \textit{SmaI} site in a single descendant of a progenitor cell that contained a methylated \textit{SmaI} site. Such epigenetic 'mutations' are well known and relatively common [Holliday 1987]. However, the relative amounts of the two bands, methylated and non-methylated.
Figure 4. [Top] A pedigree is presented and shows the mating of a single transgenic male to different inbred females. The ventral portion of the neural tube stained for lacZ expression and HpaII-cleaved DNAs of 12-day fetuses are shown beneath the pedigree and demonstrate that the same male can give rise to progeny with different methylation and expression phenotypes, depending on the genotype of the nontransgenic female. Matings to C57BL/6 females produce progeny showing extensive lacZ expression throughout the neural tube floor plate, whereas matings to BALB/c females produce progeny that show mosaic expression.

Histochemical staining for the presence of lacZ in the Tg4 transgenic line allowed for an analysis of expression at the level of single cells and demonstrated a correlation between the expression and methylation phenotype. Because the intermediate expression phenotype in some transgenic animals reflects changes in the number of cells expressing lacZ rather than the level of lacZ expression in all cells, it seems likely that the intermediate methylation phenotype reflects changes in the number of hypermethylated versus hypomethylated
Figure 5. Two models demonstrating the possible means of producing the three methylation phenotypes observed for the two transgene loci (line 379-TNI and line Tg4). In both models the transgene array is represented along the top, the cell configuration in the center ([open circles] cells that contain a hypomethylated allele; [hatched circles] cells that contain a partially methylated allele; [cross-hatched circles] cells that contain a hypermethylated allele) and the methylation-sensitive restriction endonuclease hybridization pattern along the bottom. (A) One possibility in which all cells within an individual have the array methylated similarly but different individuals have a different degree of transgene array methylation; (B) the other possibility in which different cells within an individual are methylated differently.

cells [Fig. 5B]. The methylation of the 379 troponin I transgene also displays different methylation phenotypes (Sapienza et al. 1989a) similar to that of the Tg4 transgene and is therefore also consistent with the model presented in Figure 5B.

This model requires the establishment of at least two cell-specific methylation patterns. The mechanism by which such differences are established between apparently identical cells is unknown, but data presented here [Fig. 4], as well as previously published data [Sapienza et al. 1989a], indicate that this process is under genetic control. Because tissues of ectodermal, endodermal, and mesodermal origin from each individual have the same overall transgene methylation phenotype, we presume that the same relative numbers of cells with each methylation phenotype are present in each tissue. These observations suggest the possibility that the methylation mosaicism is already established in the last common ancestor of these tissues, i.e., by primitive ectoderm formation at day 4.5 of development [Hogan et al. 1986]. However, we cannot say whether the methylation pattern is established only once and then propagated faithfully or is generated continuously. Either possibility will produce the observed pattern.

The genetic influence of one inbred mouse strain (C57BL/6) on the methylation phenotype of the two transgenic lines we studied appears to be different. In the 379-TNI line (Sapienza et al. 1989a), the transgene becomes more methylated under the influence of this genotype, whereas the converse is true in the Tg4 line. This is presumably a result of either the site of insertion or inherent differences between the two sequences and demonstrates that different domains or gene structures can be treated differently in the same genetic background.

It is interesting to note that several other transgenes
have been reported to exhibit cellular mosaicism in expression when maintained in a hemizygous state, even though the tissue-specific site of expression is correct. Among these are a myelin basic protein antisense cDNA (Katsuki et al. 1988), a liver fatty acid-binding protein—human growth hormone construct (Sweetser et al. 1988a) and an intestinal fatty acid-binding protein—human growth hormone construct (Sweetser et al. 1988b). In addition, another lacZ-containing transgene shows expression differences in preimplantation embryos, apparently as a function of genetic background (M.H.A. Surani, pers. comm.). The expression phenotype we see with the Tg4 transgene, therefore, is not a unique consequence of sequence or site of insertion. In this regard it is noteworthy that the Tg4 insertion disrupts the dt locus and, therefore, lies within or adjacent to an endogenous locus (Kothary et al. 1988).

Allele-specific methylation differences have been demonstrated in cultured cells of human origin (Chandler et al. 1987) and heritability of locus-specific methylation patterns have been demonstrated in human pedigrees (Silva and White 1988) but neither of these observations shows an effect of genotype on either the methylation or expression pattern.

Somatic mosaicism has been implicated in the establishment or progression of a variety of clinical disorders (for review, see Hall 1988). Such mosaicism has been interpreted as arising from random events that occur in developing somatic tissue such that daughter cells of the affected somatic cell carry a mutation. Our analysis of the methylation phenotypes associated with transgenes, as well as those of an ecotropic retroviral locus, suggest that methylation mosaicism may arise normally during development. Furthermore, the different cells in the methylation mosaics are apparently not identical in their ability to express all of the genetic information they contain. Such expression differences may reflect true mutations rather than epigenetic changes, but if this were the case somatic mutation rates must be very high, and either the creation or allocation of cells containing such mutations must be under genetic control.

If the allele inactivation process we observe at hemizygous transgene loci also operates at endogenous loci containing two alleles, one predicts that some individuals will be functionally hemizygous at affected loci in some fraction of cells (Scrable et al. 1989). The genetic consequences of such mosaicism (Sapienza 1989) would be that mutant alleles at loci subject to methylation and inactivation would behave in a dominant fashion but would exhibit incomplete penetrance and variable expressivity. These possibilities are being investigated.

Materials and methods

Mice

Inbred strains of mice were obtained from The Jackson Laboratory (Bar Harbor, Maine). The two transgenic lines of mice used have been described previously. 379-TNI is a quail tropolin I fast-fiber muscle-specific gene (Hallauer et al. 1988); Tg4 is a hsp68–lacZ construct (Kothary et al. 1988). Timed matings were accomplished by placing estrus females with transgenic males and identifying copulatory plugs on the following morning (day 0 of pregnancy). Females were killed at 12 days gestation, and fetuses were removed. For the Tg4 line, fetal heads were used for DNA extraction, and the bodies were fixed and stained for β-galactosidase activity.

Southern hybridization

DNA was prepared from tail biopsies or fetal tissues, as described previously (Sapienza et al. 1987), and cleaved with the indicated restriction endonucleases, as suggested by the manufacturer (Pharmacia). The extent of digestion was tested by the addition of control plasmid DNA to a sample of the experimental digest. Digestion was considered to be complete when the internal control gave a digestion pattern identical to the plasmid alone when digested with the same enzyme. Gel electrophoresis, transfer of DNA to nylon membranes (Nytran, Amersham Corp.), preparation of 32P-labeled probes, hybridizations, and autoradiography were also described previously (Sapienza et al. 1987).

E. coli β-galactosidase assays

β-Galactosidase activity was determined using a modification of the method described by Kothary et al. (1988). Fetal tissues were fixed in 0.5% paraformaldehyde, 0.5% glutaraldehyde in PBS (0.07 M Na2HPO4, 0.03 M NaH2PO4, 0.5 M NaCl) for 1 hr at 4°C. The tissues were rinsed in PBS and incubated in a staining mixture consisting of 3.1 mM K+ ferricyanide, 3.1 mM K+ ferrocyanide, 10 mM NaPO4 [pH 7.2], 0.5 M NaCl, 1.0 mM MgCl2, and 0.4 mg/ml hologenated indolyl-β-D-galactoside [Bluo-gal, BRL] for 12–18 hr at 37°C.

Acknowledgments

We are grateful to Janet Rossant for the Tg4 transgenic mice, to Thu-Hang Tran, Jean Paquette, Julie Trembly, and Susan Albrechtson for technical assistance, to Terri Genio for typing the manuscript, and to Linda Sapienza and Robert Derval for artwork.

References

Chandler, L.A., H. Ghazi, P.A. Jones, P. Boukamp, N.E. Fusenig. 1987. Allele-specific methylation of the human c-Ha-raf-1 gene. Cell 50: 711–717.

Doerfler, W. 1983. DNA methylation and gene activity. Annu. Rev. Biochem. 52: 93–124.

German, J.L. 1962. DNA synthesis in human chromosomes. Trans. N.Y. Acad. Sci. 24: 395–407.

Hadchouel, M., H. Farza, D. Simon, P. Tiollais, and C. Pourcel. 1987. Maternal inhibition of hepatitis B surface antigen gene expression in transgenic mice correlates with de novo methylation. Nature 329: 454–456.

Hall, J. 1988. Review and hypothesis: Somatic mosaicism: Observations related to clinical genetics. Am. J. Hum. Genet. 43: 355–363.

Hallauer, P., K. Hastings, and A. Peterson. 1988. Fast skeletal muscle-specific expression of a quail tropolin I gene in transgenic mice. Mol. Cell. Biol. 8: 5072–5079.

Hogan, B., F. Constantini, and E. Lacy. [ed.] 1986. Summary of mouse development. In Manipulating the mouse embryo: A
McGowan et al.

laboratory manual, pp. 17–70. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Holliday, R. 1987. The inheritance of epigenetic defects. Science 238: 163–170.

Jenkins, N. and N.G. Copeland. 1985. High frequency germline acquisition of ecotropic MuLV proviruses in SWR/J-RF/J hybrid mice. Cell 43: 811–819.

Jenkins, N.A., N.G. Copeland, B.A. Taylor, and B.K. Lee. 1982. Organization, distribution and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of Mus musculus. J. Virol. 43: 26–36.

Katsuki, M., M. Sato, M. Kimura, and M. Yokoyama. 1988. Conversion of normal behaviour to shiverer by myelin basic protein antisense cDNA in transgenic mice. Science 241: 593–595.

Kothary, R., S. Clapoff, A. Brown, R. Campbell, A. Peterson, and J. Rossant. 1988. A transgene containing lacZ inserted into the dystonia locus is expressed in neural tube. Nature 335: 435–437.

Reik, W., A. Collick, M. Norris, S. Barton, and M. Surani. 1987. Genomic imprinting determines methylation of parental alleles in transgenic mice. Nature 328: 248–251.

Sapienza, C. 1989. Genome imprinting and dominance modification. Ann. N.Y. Acad. Sci. 564: 24–38.

Sapienza, C., J. Paquette, T.-H. Tran, and A. Peterson. 1989a. Epigenetic and genetic factors affect transgene methylation imprinting. Development 106 [in press].

Sapienza, C., A. Peterson, J. Rossant, and R. Balling. 1987. Degree of methylation of transgenes is dependent on gamete of origin. Nature 328: 251–254.

Sapienza, C., T.-H. Tran, J. Paquette, R. McGowan, and A. Peterson. 1989b. A methylation mosaic model for mammalian genome imprinting. Prog. Nucleic Acids Res. Mol. Biol. 36: 145–152.

Scrabble, H., W. Cavenee, F. Ghavimi, M. Lowell, K. Morgan, and C. Sapienza. 1989. A model for embryonal rhabdomyosarcoma tumorigenesis which involves genome imprinting. Proc. Natl. Acad. Sci. 86 [in press].

Silva, A.J. and R. White. 1988. Inheritance of allelic blueprints for methylation patterns. Cell 54: 145–152.

Selig, S., M. Ariel, R. Gotein, M. Marcus, and H. Cedar. 1988. Regulation of mouse satellite DNA replication time. EMBO J. 7(2): 419–426.

Swain, J.L., T.A. Stewart, and P. Leder. 1987. Parental legacy determines methylation and expression of an autosomal transgene: A molecular mechanism for parental imprinting. Cell 50: 719–727.

Sweetser, D., S. Haut, P. Hoppe, and E. Birkenmeier. 1988a. Transgenic mice containing intestinal fatty acid-binding protein-human growth hormone fusion genes exhibit correct regional and cell-specific expression of the reporter gene in their small intestines. Proc. Natl. Acad. Sci. 85: 9611–9615.

Sweetser, D., E. Birkenmeier, P. Hoppe, D. McKeel, and J. Gordon. 1988b. Mechanisms underlying generation of gradients in gene expression within the intestine, an analysis using transgenic mice containing fatty acid binding protein–human growth hormone fusion genes. Genes Dev. 2: 1318–1332.

Vanyushin, B.F., S.G. Tkacheva, and A.N. Belozersky. 1970. Rare bases in animal DNA. Nature 225: 948–949.

Wang, R.Y.-H., K.C. Kuo, C.W. Gehrke, L.-H. Huang, and M. Ehrlich. 1982. Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA. Biochim. Biophys. Acta 697: 371–377.

Wigler, M., D. Levy, and M. Perucho. 1981. The somatic replication of DNA methylation. Cell 24: 33–40.
Cellular mosaicism in the methylation and expression of hemizygous loci in the mouse.

R McGowan, R Campbell, A Peterson, et al.

*Genes Dev.* 1989, 3:
Access the most recent version at doi:10.1101/gad.3.11.1669