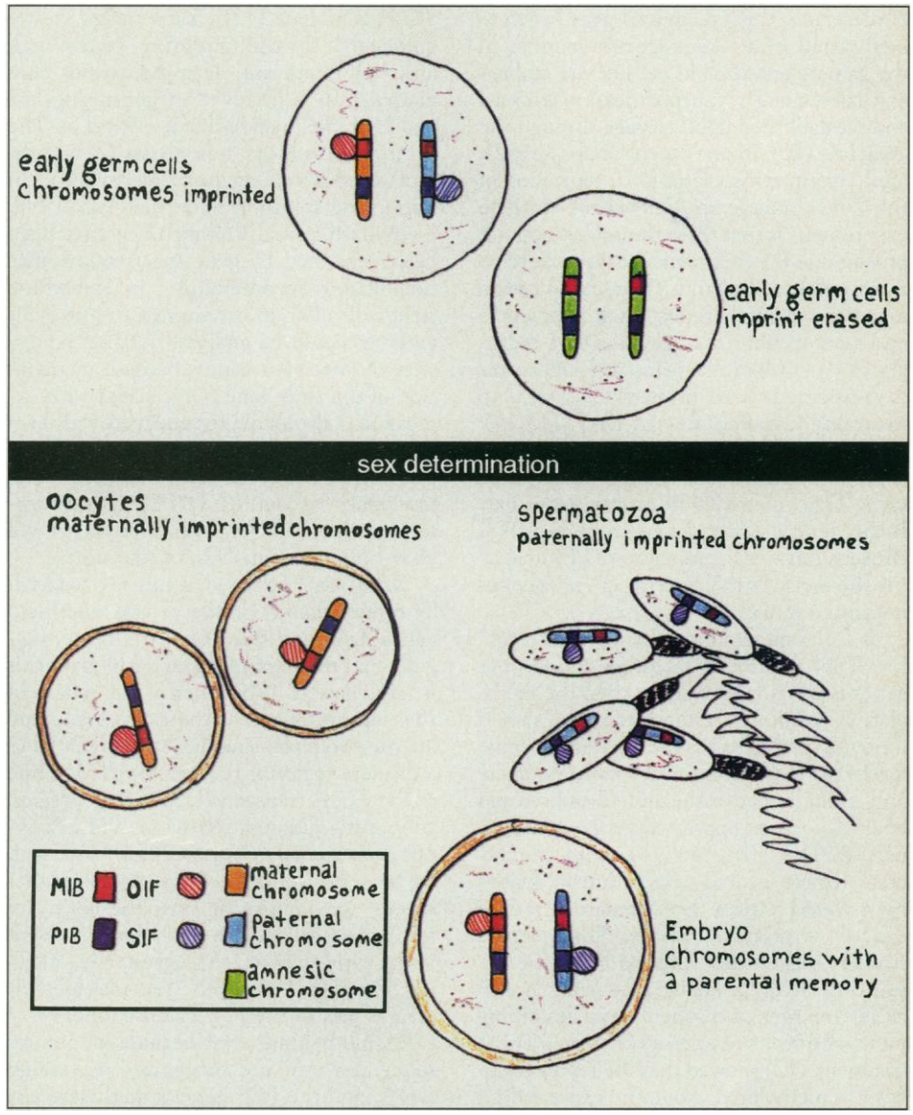
Methylation and Imprinting: From Host Defense to Gene Regulation?

Denise P. Barlow

Mammals inherit two complete sets of chromosomes from their parents and thus two copies of every autosomal gene. Normally both copies are expressed, but, in a minority of cases, a mechanism known as genomic imprinting causes the expression of a gene to vary according to its maternal or paternal origin (*1*). The *raison d'être* of imprinting is unclear. Nevertheless, working on the principle that once you know *how*, the *why* becomes self-evident, an increasing effort is being put into understanding the molecular mechanism that leaves a gene with an "imprint" from its mother or father. Analysis of endogenous imprinted genes isolated in the last 2 years supports previous links between imprinting and DNA methylation. Of particular interest is the fact that these new findings suggest that imprinting may have evolved as an extension of the host defense role that DNA methylation plays against invading organisms.

Genetic analyses in mice predict that only a small number of genes are imprinted (*2*); although these genes likely act during embryonic development (*3*, *4*), it is not clear whether imprinted genes share a common function. Indeed, the recent identification of four imprinted genes in the mouse suggests they do not. *Igf2* (*5*) and *Igf2r* (*6*) encode insulin-like growth factor type 2 and its receptor; *H19* (*7*) encodes an embryonic RNA of unknown function; and *Snrpn* (*8*) encodes part of a ribonucleoprotein that catalyzes RNA splicing in the brain. *Igf2* and *Snrpn* are exclusively paternally expressed, while *Igf2r* and *H19* are exclusively maternal. In all cases the repressed locus shows a complete absence of messenger RNA. This means that the cellular transcription machinery must be able to discriminate between the maternal and paternal gene copy. Because inbred mice, which are genetically identical at all loci, are used for these experiments, this discrimination cannot be due to nucleotide sequence differences, but must be due to some form of parent-specific modification that affects the ability of a gene to be transcribed.

A simple model to explain how imprinted genes are modified in a parent-specific manner proposes that the modification is added during gametogenesis [see figure and (*9*)]. This is the only period when the maternal and paternal genomes are separate and can be subjected to differing influences. According to this model there are at least two steps involved: recognition of a sequence element, also known as an "imprinting box" (*5*), at the gene locus and modification of this sequence by an imprinting factor. Any gene containing an imprinting box is subject to parent-specific modification by the imprinting factor during gametogenesis. Since some genes are paternally expressed and others are maternal, an additional complication is to propose the existence of maternal and paternal variants of imprinting boxes and imprinting factors. Many of the features proposed for this model are based on analyses of endogenous imprinted genes (*5–8*) and imprinted "transgenes" [DNA constructs containing eukaryotic and prokaryotic sequences (*10*)] in mice, and together the data can be used to describe four properties of the imprinting factor and the imprinting box. First, addition of the imprinting factor to the imprinting



A model for genetic imprinting. The maternal imprinting box (MIB) of a gene on a maternally inherited chromosome has been modified by the oocyte imprinting factor (OIF). The same MIB remains unmodified on a paternally inherited chromosome. The opposite situation occurs with a gene containing a paternal imprinting box (PIB), modified by a spermatogenic imprinting factor (SIF) on the paternally inherited chromosome but not modified on the maternal. The OIF and the SIF are removed from the chromosomes of early germ cells, producing amnesic chromosomes that have erased their imprint. After sex determination, germ cells differentiate either into oocytes or spermatozoa. Oocytes produce OIF and spermatozoa produce SIF that interact with their respective imprinting boxes, leaving a chromosome once more with a memory of its mother or father.

The author is at the Research Institute of Molecular Pathology, Dr. Bohr-Gasse 7, A-1030 Vienna, Austria.

box is reversible; second, the factor changes transcription; third, the factor modifies the imprinting box in gametogenesis and is erased in the gametes of the next generation; and finally, the factor is heritable in a chromosome-specific manner in the diploid embryo. These properties closely resemble the behavior of DNA methylation in mammals.

DNA methylation, like gene imprinting, is reversible, associated with changes in transcription, subject to fluctuations during gametogenesis and development, and heritable in a chromosome-specific manner (*11*). Although all of these features suggest a link between methylation and the imprinting factor, the strongest link has been demonstrated by the use of transgenic mice. In a few of these mice the foreign transgene becomes methylated in a parent-specific manner in the gamete and diploid cells of the embryo and, subsequently, the modification is erased and reestablished upon passage through the germ line (*10*). In one exceptional case (*12*), methylation is associated with repression of the *TG.A* transgene. A drawback to these experiments is that these transgenes contain prokaryotic DNA sequences that are foreign to the mouse, and thus their modification may not accurately reflect what happens to endogenous imprinted genes. DNA methylation in both prokaryotes and eukaryotes is considered to be primarily a system to neutralize invading, foreign DNA (*13,14*): in mice retroviral genomes introduced into the early embryo are reliably methylated (*15*). Thus, methylation of transgenes may simply reflect a host defense response. This concern has now been addressed by analyzing the methylation patterns of endogenous imprinted genes.

Is DNA methylation the imprinting factor? If the imprinting factor for endogenous genes is DNA methylation, then the methylation profile of an imprinted gene should follow a pattern in which parent-specific methylation is present in the mature gamete and is maintained in the diploid embryo and adult as long as imprinted gene expression persists. The *Igf2r* gene (*16*) contains a discrete region located in a 5' intron that is methylated with a developmental profile exactly as predicted above. These results strongly suggest that methylation is the imprinting factor for endogenous genes as well as for transgenes. Earlier work describing methylation of an expressed imprinted *TG.A* transgene (*12*) showed that the repressed locus was methylated, while the expressed locus was unmodified. In complete contrast to these results, the methylated region in *Igf2r* is on the expressed (maternal) locus, but the repressed (paternal) locus is not modified. Thus in *Igf2r*, methylation is associated with expression, and the locus is both maternally imprinted and maternally expressed. This situation can result in maternal-specific expression if *Igf2r* transcription depends on either methylation-sensitive repressors or methylation-sensitive activators (*16*). In the *Igf2* locus, parent-specific methylation of the expressed (paternal) locus was identified (*17*), but only in late embryos and adults. A more detailed study of three imprinted genes (*Igf2r*, *Igf2*, and *H19*) confirms that differential methylation occurs at all three loci in late embryos and adults, but suggests that gamete methylation may act as a focus for later parent-specific methylation that occurs after implantation (*18*). Thus, so far only methylation of the *Igf2r* gene (*16*) and the *TG.A* transgene (*12*) clearly meets the requirements for the imprinting factor, and, although there are clear indications that methylation is involved in imprinting *Igf2* and *H19*, the final verdict is not yet in. The exciting possibility now exists to test directly the involvement of methylation in imprinting by using mice that lack DNA methyltransferase. Mutant mice that have been generated by gene inactivation after homologous recombination in embryonic stem cells (*19*) do not survive to birth, but embryos could be analyzed for the expression of imprinted genes. Because methylation of the *Igf2r* gene is associated with expression of the locus (*16*) and methylation of the *TG.A* transgene is associated with repression of the locus (*12*), I would predict that embryos lacking DNA methyltransferase will also lack *Igf2r* transcripts but will show an increase in *TG.A* expression.

Why are imprinted genes methylated? The methylation profile of two imprinted genes, *Igf2r* and the *TG.A* transgene, suggests that the imprinting factor in mammals is methylation. What then is the nature of the imprinting box at these two genes and are the two boxes similar? At the *Igf2r* locus a discrete region is methylated (*16*), while at the *TG.A* transgene locus a larger region containing plasmid, retroviral, and eukaryotic sequences is methylated (*12*). Although we have not yet defined the sequences that specify the imprinting box, the ability of retroviruses to become methylated de novo in the early embryo (*15*) suggests the following: The *TG.A* transgene becomes methylated as part of the host defense function of DNA methyltransferase because it contains sequences that are recognized as foreign DNA, and the *Igf2r* gene is methylated by the same system because it contains an imprinting box that looks like foreign DNA. If this is the case, genomic imprinting may be the missing link that bridges the evolution of DNA methyltransferase from an agent for host defense to a mechanism that regulates gene expression. By extension, this implies that the origins of gene imprinting lie in an existing biochemical system that serves to neutralize foreign invading DNA. The methylation profile of *Igf2r* and the *TG.A* transgene also suggests the curious prospect that imprinting and the host defense function for methylation may occur only in the maternal germ line. Evidence to support this point of view is that genes that are imprinted by methylation on the maternal chromosome can be either maternally expressed [*Igf2r* (*16*)] or maternally repressed [*TG.A* transgene (*12*)]. In addition, in all cases of parent-specific transgene methylation, it is always the maternal locus that is methylated, never the paternal (*9*). A paternal imprinting factor may not be required (although the paternal imprinting system may just differ from the maternal one). Whatever the outcome concerning the mechanism of imprinting, the role of DNA methylation in mammals is a topic that generates a certain amount of heat, especially among colleagues whose favorite organisms have chosen to eschew the benefits of reprogramming their inherited genetic information. A suggestion that imprinting may have evolved in mammalian oocytes as an extension of the host defense role of DNA methylation is just a little extra fuel to keep the pot boiling.

References and Notes

1. M. Monk and A. Surani, Eds., *Genomic Imprinting* (*Dev. Suppl.*) (The Company of Biologists, Ltd., Cambridge, United Kingdom, 1990).
2. C. V. Beechey, B. M. Cattanach, A. G. Searle, *Mouse Genome* **87**, 64 (1990).
3. J. McGrath and D. Solter, *Cell* **37**, 179 (1984).
4. M. A. H. Surani, S. C. Barton, A. Burling, *Nature* **308**, 548 (1984).
5. T. M. DeChiara, E. J. Robertson, A. Efstratiadis, *Cell* **64**, 849 (1991).
6. D. P. Barlow *et al.*, *Nature* **349**, 84 (1991).
7. M. S. Bartolomei, S. Zemel, S. M. Tilghman, *ibid.* **351**, 153 (1991).
8. S. E. Leff *et al.*, *Nature Genet.* **2**, 259 (1992).
9. M. A. Surani, *Curr. Opin. Genet. Dev.* **1**, 241 (1991).
10. _____ W. Reik, N. Allen, *Trends Genet.* **4**, 59 (1988).
11. A. P. Bird, *Cell* **70**, 5 (1992).
12. J. R. Chaillet, T. F. Vogt, D. R. Beier, P. Leder, *ibid.* **66**, 77 (1991).
13. T. H. Bestor, *Philos. Trans. R. Soc. London B* **326**, 179 (1990).
14. W. Doerfler, *Biol. Chem. Hoppe-Seyler* **372**, 557 (1991).
15. D. Jähner and R. Jaenisch, *Mol. Cell. Biol.* **5**, 2212 (1985).
16. R. Stöger, P. Kubicka, C-G. Liu, T. Kafri, A. Razin, H. Cedar, D. P. Barlow, *Cell*, **73**, 1 (1993).
17. H. Sasaki *et al.*, *Genes Dev.* **6**, 1843 (1992).
18. H. Cedar and A. Razin, personal communication.
19. E. Li, T. H. Bestor, R. Jaenisch, *Cell* **69**, 915 (1992).

20. I thank my "imprinting" co-workers R. Stöger, N. Schweifer, and P. Kubicka for contributing to the ideas expressed here; H. Cedar and A. Razin for communicating results prior to publication; and L. Ballou, M. Birnstiel, B. Hogan, T. Schuster, and E. Wagner for editing the manuscript.