The Inheritance of Methylation Patterns in Vertebrates

Michael H. Wigler
Cold Spring Harbor Laboratory
P.O. Box 100
Cold Spring Harbor, New York 11724

The genomic DNAs of most organisms contain modified bases. In vertebrates, 5-methyl-cytosine (^mC) is the only modified base; it results from enzymatic transfer of the methyl group of S-adenosyl-methionine to cytosine residues in DNA. Most but not all ^mC occurs in the dinucleotide 5'-CpG, and in mammals and birds approximately 50–70% of all such dinucleotides are modified.

Considerable interest in DNA methylation has been created by increasing evidence linking methylation patterns to patterns of gene expression; this subject has been recently reviewed elsewhere (Razin and Riggs, Science 210, 604–610, 1980). The actual distribution of ^mC within specific genes can be probed with the use of bacterial restriction endonucleases, such as Hpa II and Hha I, that will not cleave recognition sequences containing ^mCpG.

Tissue-specific differences in methylation patterns have been noted (Waalwijk and Flavell, NAR 5, 4631–4641, 1978; Mandel and Chambon, NAR 7, 2081–2103, 1979; McGhee and Ginder, Nature 280, 419–420, 1979; van der Ploeg and Flavell, Cell 19, 947–958, 1980) and strong correlations exist between methylation and transcriptional inactivity of integrated viral genomes (Desrosiers et al., PNAS 76, 3839–3843, 1979; Cohen, Cell 19, 653–662, 1980; Sutter and Doerfler, PNAS 77, 253–256, 1980). Experiments with DNA-mediated gene transfer have suggested a causal link between methylation and inhibition of gene expression (Pollack et al., PNAS 77, 6463–6467, 1980; Wigler et al., Cell 24, 33–40, 1981). Recent experiments involving oocyte microinjection and in vitro transcription with DNA molecules methylated in vitro also confirm this conclusion (recently reported at the 1981 Annual Genetics Meeting in Koln, Germany). Agents that can disrupt DNA methylation in vivo can cause diverse effects such as alterations in the pathway of differentiation and reactivation of genes residing in the inactive X chromosome (Taylor and Jones, Cell 17, 771–779, 1979; Mohandas et al., Science 211, 393–396, 1981).

If, in fact, methylation can modulate gene expression, it is important to understand the factors that determine methylation patterns in the cells of vertebrate organisms. It was hypothesized that a methylation pattern, once established in somatic cells, could become inherited in progeny cells (Holiday and Pugh, Science 187, 226–232, 1975; Riggs, Cytogen. Cell Gen. 14, 9–25, 1975). Because the CpG dinucleotide, which bears most of the vertebrate methylation, is a simple palindrome, methylation on one strand could direct the methylation on a newly replicated strand through the action of a "maintenance" methylase that recognizes only hemimethylated sites. In this manner, the organism would have at least one mechanism for the stable somatic inheritance of methylation patterns. Such phenomena as maintenance of X chromosome inactivation and the stability of the differentiated phenotype could be explained in this way.

In recent years, evidence has been obtained in favor of this model. Bird (JMB 118, 49–60, 1978) studied the distribution of methylation in the ribosomal genes of Xenopus laevis red blood cells. These genes exist in a highly methylated state. They are virtually resistant to digestion with Hpa II, although a few unmethylated Hpa II sites are randomly distributed within these genes. The sensitivity to Hpa II digestion of denatured and reannealed ribosomal genes indicated that virtually all methylated sites were symmetrically methylated, thus suggesting the action of a maintenance methylase of the type postulated above. The amplified ribosomal genes of Xenopus, however, are not methylated (Dawid et al., JMB 51, 341–360, 1970; Bird and Southern, JMB 118, 27–47, 1978), which suggests that the specificity for methylation does not reside in the DNA sequence immediately flanking a potential site. Similarly, the endogenous mouse mammary tumor virus sequences are heavily methylated but the integrated sequences arising by horizontal infection are not methylated (Cohen, loc. cit.).

Finally, direct evidence for the passive maintenance of methylation patterns comes from two recent studies that have utilized the techniques of DNA-mediated transformation in cultured mouse cells (Pollack et al., loc. cit.; Wigler et al., loc. cit.). DNA molecules were methylated in vitro with the bacterial modification enzyme M-Hpa II, which methylates the internal cytosines of the $\frac{CpCpGpG}{GpGpCpC}$ sequence (Mann and Smith, NAR 4, 4211–4221, 1977). Such DNAs and their unmethylated controls were then transferred into growing cells, and the state of methylation in progeny strands was examined many generations later by Hpa II digestion and Southern blotting. These studies were performed with four distinct DNA molecules: bacterial plasmid pBR322, bacteriophage φX174 RF DNA, the cloned herpes thymidine kinase gene and the cloned chicken thymidine kinase gene. In summary, it was found that there was evidence of inheritance of methylation at all the Hpa II sites of φX174 and most of the sites in chicken thymidine kinase. In some cases inheritance was observed in cells transformed with

pBR322. There was no evidence for inheritance at the Hpa II sites of the herpes thymidine kinase gene and at one of the Hpa II sites of the chicken thymidine kinase gene. In these two cases, however, there was selection for expression of the thymidine kinase genes, and this may have selected against retention of methylation at these sites. Although complete fidelity of inheritance was not observed in any instance, it is clear that cells do have mechanisms for the inheritance of methylation patterns. These studies do not rule out the possibility, and in fact suggest, that additional factors may influence the probability of inheritance.

Although studies with DNA-mediated transfer on unmethylated DNA indicate that de novo methylation is infrequent (Pollack et al., loc. cit.; Wigler et al., loc. cit.), many examples can be cited in which unmethylated sequences introduced into vertebrate cells can subsequently become methylated (Desrosiers et al., loc. cit.; Sutter and Doerfler, loc. cit.; Pollack et al., loc. cit.). Moreover, the methylation of integrated retroviral DNA makes it clear that heritable methylation patterns can be established for new sequences acquired in the germline (Cohen, loc. cit.). The significance of these examples of de novo methylation is not clear. No one, however, has yet found sites methylated in somatic cells that are not also methylated in sperm DNA (Mandel and Chambon, loc. cit.; van der Ploeg and Flavell, loc. cit.). One form of the inheritance model is that a master pattern of methylation exists in the germline cells, subsequently in the fertil-

ized zygote, maintained intact in the germline for the next generation but passed on in varying degrees to the somatic tissues. Specific loss of methylation could be directed in either of two ways: by the action of a demethylating enzyme (although no such activity has been found); or by a programmed omission of methylation at a specific site following DNA replication. Some random loss of methylation would also be expected to result from an occasional failure of the maintenance methylase. This would result in a variable methylation pattern, which is frequently observed in somatic tissues.

How then is the initial methylation pattern established in the germline, and what is the unit of inheritance? We have not answered these questions at present, although the means to do so are at hand. For the somatic cells, as we have discussed, the unit of inheritance appears to be the individual methylation site, which also appears to be its own genetic determinant. Evidence has been presented by Bird et al. (Cell 17, 889–891, 1979) that the genome of animal cells can be divided into methylated and unmethylated domains, making it likely that the unit of change in methylation pattern would constitute a new contiguous stretch of methylated sequences. The genetic factors that determine the boundaries of these domains and their stability during evolution remain a mystery. Rapidly arising individual differences in methylation patterns could provide a rich source for genetic variation within a species.