



DNA Methylation and Epigenetic Inheritance in Plants and Filamentous Fungi

Robert A. Martienssen, *et al.*

Science **293**, 1070 (2001);

DOI: 10.1126/science.293.5532.1070

The following resources related to this article are available online at www.sciencemag.org (this information is current as of January 15, 2007):

A correction has been published for this article at:
<http://www.sciencemag.org/cgi/content/full/293/5537/1997>

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
<http://www.sciencemag.org/cgi/content/full/293/5532/1070>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/cgi/content/full/293/5532/1070#related-content>

This article **cites 67 articles**, 31 of which can be accessed for free:
<http://www.sciencemag.org/cgi/content/full/293/5532/1070#otherarticles>

This article has been **cited by** 142 article(s) on the ISI Web of Science.

This article has been **cited by** 26 articles hosted by HighWire Press; see:
<http://www.sciencemag.org/cgi/content/full/293/5532/1070#otherarticles>

This article appears in the following **subject collections**:
Botany
<http://www.sciencemag.org/cgi/collection/botany>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:
<http://www.sciencemag.org/help/about/permissions.dtl>

changes in the methylation machinery. ICF patients have mutations in the *DNMT3b* gene, which leads to undermethylation of satellite DNA and specific chromosomal decondensation. Rett syndrome patients have mutations in one of the MBDs (*MeCP2*) and therefore may not be able to interpret the methylation signal correctly (45). These two diseases suggest that methylation is not only needed to complete embryonic development but is also required for development after birth. The price for the requirements that methylation is essential in mammals seems to be a substantial increase in cancer risk.

The study of DNA methylation in mammals has been stimulated by the identification of the key enzymes that methylate DNA and their interactions with DNA and DNA binding proteins, as well as by the link between methylation and chromatin structure. Perhaps the major function of methylases in mammals is in the long-term silencing of noncoding DNA in the genome, which contains a very substantial portion of repetitive elements. Lack of methylation in promoters of essential genes allows them to be potentially active and to be regulated by other processes. On the other hand, we still do not know whether the specific patterns seen in the genes of differentiated cell types are involved in transcriptional control or whether they simply reflect altered chromatin states.

The DNA methylation field is currently in a state of high activity as the links between stable epigenetic states, chromatin structure,

and heterochromatinization begin to become clearer. Disruption of these links leads to substantial disease states, including chromosomal integrity, mental retardation, and cancer. Understanding how epigenetic states are established and maintained and developing strategies to modify them therapeutically is, therefore, likely to be an area of intense future research.

References and Notes

- E. Li, T. H. Bestor, R. Jaenisch, *Cell* **69**, 915 (1992).
- M. Okano, D. W. Bell, D. A. Haber, E. Li, *Cell* **99**, 247 (1999).
- T. Bestor, A. Laudano, R. Mattaliano, V. Ingram, *J. Mol. Biol.* **203**, 971 (1988).
- M. Okano, S. Xie, E. Li, *Nature Genet.* **19**, 219 (1998).
- F. Larsen, G. Gundersen, R. Lopez, H. Prydz, *Genomics* **13**, 1095 (1992).
- M. Gardiner-Garden, M. Frommer, *J. Mol. Biol.* **196**, 261 (1987).
- D. Takai, P. A. Jones, unpublished data.
- S. Delgado, M. Gomez, A. Bird, F. Antequera, *EMBO J.* **17**, 2426 (1998).
- G. Grigg, S. Clark, *Bioessays* **16**, 431 (1994).
- D. Macleod, J. Charlton, J. Mullins, A. P. Bird, *Genes Dev.* **8**, 2282 (1994).
- M. Brandeis *et al.*, *Nature* **371**, 435 (1994).
- I. G. Lin, T. J. Tomzynski, Q. Ou, C. L. Hsieh, *Mol. Cell Biol.* **20**, 2343 (2000).
- L. Han, I. G. Lin, C. L. Hsieh, *Mol. Cell Biol.* **21**, 3416 (2001).
- W. Reik, W. Dean, J. Walter, *Science* **293**, 1089 (2001).
- K. D. Robertson *et al.*, *Nature Genet.* **25**, 338 (2000).
- M. R. Rountree, K. E. Bachman, S. B. Baylin, *Nature Genet.* **25**, 269 (2000).
- F. Fuks, W. A. Burgers, N. Godin, M. Kasai, T. Kouzarides, *EMBO J.* **20**, 2536 (2001).
- G. L. Xu *et al.*, *Nature* **402**, 187 (1999).
- D. Jahner, R. Jaenisch, *Nature* **315**, 594 (1985).
- A. N. Magewu, P. A. Jones, *Mol. Cell Biol.* **14**, 4225 (1994).
- M. S. Turker, *Semin. Cancer Biol.* **9**, 329 (1999).
- J. K. Christman, G. Sheikhnejad, C. J. Marasco, J. R. Sufirin, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7347 (1995).
- D. Carotti, S. Funicello, F. Palitti, R. Strom, *Biochemistry* **37**, 1101 (1998).
- M. Fatemi, A. Hermann, S. Pradhan, A. Jeltsch, *J. Mol. Biol.* **309**, 1189 (2001).
- J. P. Issa *et al.*, *Nature Genet.* **7**, 536 (1994).
- P. L. Jones *et al.*, *Nature Genet.* **19**, 187 (1998).
- X. Nan *et al.*, *Nature* **393**, 386 (1998).
- J. Guy, B. Hendrich, M. Holmes, J. E. Martin, A. Bird, *Nature Genet.* **27**, 322 (2001).
- R. Z. Chen, S. Akbarian, M. Tudor, R. Jaenisch, *Nature Genet.* **27**, 327 (2001).
- B. Hendrich, J. Guy, B. Ramsahoye, V. A. Wilson, A. Bird, *Genes Dev.* **15**, 710 (2001).
- P. A. Jones, *Trends Genet.* **15**, 34 (1999).
- M. L. Gonzalgo *et al.*, *Cancer Res.* **58**, 1245 (1998).
- M. R. Rountree, E. U. Selker, *Genes Dev.* **11**, 2383 (1997).
- M. Constancia, W. Dean, S. Lopes, T. Moore, G. Kelsey, W. Reik, *Nature Genet.* **26**, 203 (2000).
- S. Eden *et al.*, *EMBO J.* **20**, 3518 (2001).
- A. T. Hark *et al.*, *Nature* **405**, 486 (2000).
- A. Yeivin, A. Razin, in *DNA Methylation: Molecular Biology and Biological Significance*, J. P. Jost, H. P. Saluz, Eds. (Birkhauser Verlag, Basel, Switzerland, 1993), vol. 64, pp. 523–568.
- J. A. Yoder, C. P. Walsh, T. H. Bestor, *Trends Genet.* **13**, 335 (1997).
- G. C. Prendergast, E. B. Ziff, *Science* **251**, 186 (1991).
- P. A. Jones *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6117 (1990).
- D. N. Cooper, H. Youssoufian, *Hum. Genet.* **78**, 151 (1988).
- W. M. Rideout III, G. A. Coetzee, A. F. Olumi, P. A. Jones, *Science* **249**, 1288 (1990).
- S. B. Baylin, J. G. Herman, J. R. Graff, P. M. Vertino, J. P. Issa, *Adv. Cancer Res.* **72**, 141 (1998).
- P. A. Jones, P. W. Laird, *Nature Genet.* **21**, 163 (1999).
- R. E. Amir *et al.*, *Nature Genet.* **23**, 185 (1999).
- R. J. W. O'Neill, M. J. O'Neill, J. A. Marshall Graves, *Nature* **393**, 68 (1998).
- R. W. Hendriks, H. Hinds, Z. Y. Chen, I. W. Craig, *Genomics* **14**, 598 (1992).
- P.A.J. and D.T. are supported by grant R01 CA82422 from the NCI.

REVIEW

DNA Methylation and Epigenetic Inheritance in Plants and Filamentous Fungi

Robert A. Martienssen* and Vincent Colot*†

Plants and filamentous fungi share with mammals enzymes responsible for DNA methylation. In these organisms, DNA methylation is associated with gene silencing and transposon control. However, plants and fungi differ from mammals in the genomic distribution, sequence specificity, and heritability of methylation. We consider the role that transposons play in establishing methylation patterns and the epigenetic consequences of their perturbation.

Epigenetic changes, so-called “epimutations,” occur because nucleotide sequence is not the only form of genetic information in

Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA.

*To whom correspondence should be addressed. E-mail: martiens@cshl.org (R.A.M.); colot@evry.inra.fr (V.C.).

†Permanent address: Unité de Recherches en Génomique Végétale (INRA-CNRS), 2 rue Gaston Crémieux, CP 5708, 91057 Evry Cedex, France.

the cell: Chromosomal proteins and DNA methylation can also be inherited, with important phenotypic consequences (1). In plants and filamentous fungi, genomic methylation is restricted mostly to transposons and other repeats (2, 3). In mammals, by contrast, coding sequences are methylated as well, except for the so-called CpG islands that often encompass the first exons of genes (4). This difference likely reflects the colonization of mammalian introns by transposons, and the

possibility of methylation spreading into flanking exons (3). Of course, the human genome contains far less exonic DNA (<2%) than transposons (>45%) (5), which thus contribute more to the level of cytosine methylation overall (4). Naïvely, when transposons lose methylation they become activated, while when genes gain extra methylation, they become silenced.

Some of the first epimutations were observed by B. McClintock, who noted that transposons underwent cycles of inactivity in maize. These “changes in phase” are associated with changes in DNA methylation and are distinct from “changes in state,” which are usually sequence rearrangements (6). Both defective and intact transposons can also be modified epigenetically (“preset”) after exposure to an active transposon (6).

These changes can be observed because transposon activity can influence neighboring pigment genes, in some cases through an outward reading promoter (7, 8). The influence of transposons on nearby genes has been documented in *Neurospora crassa* (9) and occurs in mouse, bacteria, yeast and *Drosophila* (10–13). Thus, changes in phase could have genome-wide consequences when defective elements reach high copy numbers (Fig. 1) (14). Epimutations in maize genes were first studied by R. A. Brink, and often result from allele interactions, known as paramutation (15). Genetically, paramutation resembles presetting, in that one allele (paramutagenic) can lead to a heritable change of expression of another allele (paramutable) of the same gene in heterozygotes. The presence of transposons near paramutable genes has revived the proposition that paramutation may be caused by transposons (16–19), though their role, if any, has yet to be established.

Silent transgenes were among the first examples of epimutations to be observed in plants other than maize (20), but more recently two examples of endogenous epimutations have been examined in *Arabidopsis*. Epialleles of the zinc-finger floral gene *SUPERMAN* (known as *clark kent* or *clk* alleles) are densely methylated in the transcribed portion of the gene and in a 65-bp hairpin-forming CpT microsatellite surrounding the transcriptional start site (21). Likewise, epialleles of the *PAI2* gene, which encodes an enzyme involved in tryptophan biosynthesis, are densely methylated and inactive, but only in strains containing an inverted and methylated duplication of the closely related gene *PAI1*. As with *SUP*, methylation at *PAI2* extends upstream of the transcriptional start site and ends in the 3' untranslated region (22, 23).

In fungi, transgene duplications are efficiently methylated by RIP (repeat induced point mutation) in *Neurospora crassa* (24) and MIP (methylation induced premeiotically) in *Ascobolus immersus* (25). De novo methylation is triggered in haploid nuclei, most likely by DNA pairing between repeated sequences, and is maintained vegetatively after the duplicated sequences segregate. Methylated genes behave as epimutant alleles (26, 27), although RIP leads principally to genetic mutations (24). Unlike in plants and animals, methylation-associated silencing has been shown to block transcript elongation in these fungi (3, 28, 29). In *Neurospora*, short TpA segments can act as portable signals for methylation (30), while in *Ascobolus*, transfer of DNA methylation between alleles is related to homologous recombination (31). As yet there is no evidence for either process in plants, but the implications for paramutation and transposon silencing are clear. Most transposons in fungi are heavily methylated

and inactivated, with the only exceptions being too short to be targeted by MIP or RIP (24, 32).

DNA Methyltransferases and Methylation Mutants

Cytosine methyltransferase enzymes catalyze the transfer of an activated methyl group from *S*-adenosyl methionine to the 5 position of the cytosine ring (5-Me-C). In mammals, two distinct types of methyltransferases, Dnmt1 and Dnmt3, have been functionally characterized (33). Dnmt1 is thought to be primarily involved in maintenance and Dnmt3 in establishment of genomic methylation patterns (33). In *Ascobolus*, the *masc2* gene was identified by its similarity to *Dnmt1*. The *masc2* gene product has methyltransferase activity in vitro, but a knockout *masc2* mutation displayed no obvious effect on MIP methylation or maintenance of genomic methylation patterns (34). Conversely, another predicted *Ascobolus* methyltransferase, *masc1*, did not display any in vitro activity, but knockout mutants were defective for MIP methylation (35). However, neither the single *masc1* mutant nor the *masc1 masc2* double mutant was defective for maintenance of genomic methylation patterns, suggesting that other *Ascobolus* methyltransferases remain to be identified (34, 35). In *Neurospora*, the *dim-2* methyltransferase was identified by positional cloning of a methylation-deficient *dim-2* mutation (36). The *dim-2* mutant can still perform the point mutations associated with RIP however (36). A second *Neurospora* gene, related to *masc1*, has been found by genome sequencing, and this gene product might control RIP (36).

Arabidopsis has at least 10 genes that could encode DNA methyltransferases, more than any other eukaryote sequenced so far (37). The MET class encodes genes related to the mammalian *Dnmt1* (38). *MET1* has been functionally characterized in a transgenic line where its expression is suppressed by an antisense-*MET1* transgene. Antisense-*MET1* plants have as little as 10% of the wild-type levels of 5-Me-C (39, 40). Mutants of the *MET1* gene (*ddm2* mutants) have also been identified by genetic screens for plants with reduced methylation of repeated sequences (41, 42). Three *MET1*-related genes, *MET2a*, *MET2b*, and *MET3*, remain to be functionally characterized (38). A second type of methyltransferase, the CMT “chromomethylase” class, is also related to Dnmt1 except that a novel chromo-domain amino acid motif is inserted between two of the canonical methyltransferase motifs (43). This class is unique to plants, and there are three related CMT genes in *Arabidopsis* (38). *CMT1* is predicted to be non-functional in all strains characterized (43), but *CMT3* has been shown to be functional (44–46). A third class of *Arabidopsis* methyltransferase is the “domain rearranged methyltrans-

ferase” or DRM class, which is most related to Dnmt3, except that the canonical methyltransferase motifs are organized in a novel order (47). Two predicted DRM genes, *DRM1* and *DRM2*, have not yet been characterized for function. However, it is attractive to think that, like the mammalian Dnmt3 enzymes, they might be involved in the establishment of methylation patterns. Finally, one gene (GenBank accession number At5g25480) resembles *Dnmt2*, a highly conserved but enigmatic putative mammalian methyltransferase gene, with homologs in *Schizosaccharomyces pombe* and *Drosophila*, and no known function (48).

In *Arabidopsis*, screens for mutants with demethylated DNA (*decrease in DNA methylation* or *ddm* mutants) were conducted in a similar manner to those in *Neurospora*, and multiple mutations were recovered at two loci (41). *DDM1* encodes a likely SNF2/SWI2 chromatin remodeling protein (49), while *DDM2* encodes *MET1* (42). Interestingly, when demethylated sequences from *ddm1* or antisense-*MET1* backgrounds are inherited independent from the methylation defect, they remain demethylated. This observation suggests that de novo methyltransferase activity in plants is quite low, and that mutants might have little phenotype (41). In mouse embryos, by contrast, genome demethylation and remethylation occur over a few cleavage divisions, and mutations in the de novo methylase *DNMT3* are responsible for the ICF syndrome in humans (immunodeficiency, centromere instability, and facial anomalies) (50, 51).

Epigenetic Consequences of DNA Methylation Mutants

In *Neurospora*, the *dim-2* mutation has no obvious phenotype under laboratory conditions despite genome-wide demethylation

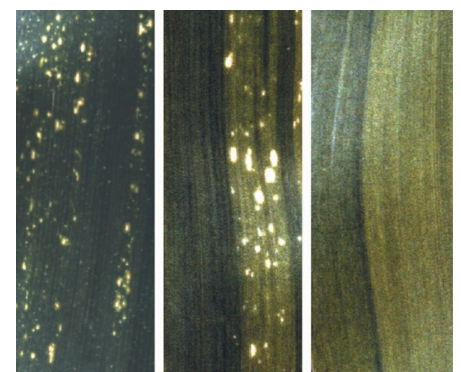


Fig. 1. Transposons in maize can co-ordinate the regulation of genes. Leaf samples from plants in which loss of transposon activity results in variegation. Unlinked mutations responsible for white lesions (left) or pale green pigmentation (right) are suppressed together in double mutant plants (middle). This is because both genes have been brought under the control of the same family of transposable elements (14).

(36). In *Ascobolus*, the *mascl* mutation also has no obvious phenotype during vegetative growth, but it blocks MIP methylation and confers sterility in *mascl* homozygous dikaryons (35). The effects of methylation defects are more conspicuous in plants. For example, in antisense-*MET1* strains of *Arabidopsis*, flowering time, fertility, leaf and floral morphology are all heritably altered (39, 40, 52). *ddm1* mutants display similar phenotypes, but often take many generations of inbreeding to manifest these changes (53). The molecular defects induced by both *ddm1* and antisense-*MET1* are varied. They include demethylation and ectopic expression of methylated genes, demethylation and movement of transposons, and local hypermethylation and silencing of genes.

Demethylation of a tandem repeat in the promoter accompanies the ectopic expression of the late flowering gene *fwa* in *ddm1* plants, and this results in a dominant epimutation (56). Similarly, *ddm1* targets a complex pathogen resistance gene cluster containing heavily methylated retrotransposons, leading to elevated expression in the dwarf epimutant *bal* (42). In a third example, a new allele of *dwarf4* arose in a single *ddm1* line. This mutation (*clam*) was caused by insertion of a CACTA transposable element, related to maize *Suppressor-Mutator* (55). In fact, several classes of transposons are demethylated and activated in *ddm1*. While both CACTA and MULE (*Mutator*-like) DNA transposons are mobilized, leading to a spectrum of new insertions (55, 56), heterochromatic retrotransposons related to *Athila* and *Tar17* do not appear to transpose in *ddm1* strains, despite being transcriptionally activated (57).

These seemingly disparate cases may be related by the nature of heterochromatin. Long arrays of tandem repeats may be responsible for heterochromatic gene silencing (58). By analogy, *fwa* and *bal* are associated with methylated repeats and may represent what B. McClintock called "cryptic" heterochromatin (59). Consistent with this view, pericentromeric and knob repeats, including many transposons, are also demethylated in *ddm1* (43), and this has led to the suggestion that DDM1 acts to enable access of DNA methyltransferases to heterochromatin (49). Thus, activation of both conspicuous and cryptic heterochromatin in *ddm1* mutants could lead to the variety of defects found.

Transgene silencing was relieved in both *ddm1* and *ddm2* (or *met1*) mutants, although only those transgenes that were silenced at the level of transcript initiation were uniformly activated in *ddm1* (61, 62). Screens for mutants that relieved this type of silencing recovered at least 12 mutations in 5 genes, including *DDM1* (57, 61, 63). *hog1* mutants resulted in a significant reduction in DNA methylation (63), but *mom1*, *sil1* and *sil2* mutants did not (57, 63).

These may encode chromatin modifying enzymes, a suggestion supported by cloning of the SNF2-related gene *MOM1* (57). DNA methylation mutants also affect endogenous epimutations. For example, *PAI2* silencing and methylation is relieved in *ddm1* (64), but only in strains without the *PAI1-PAI4* inverted duplication (65). Methylation of *PAI1-PAI4* was more reduced in *ddm2*, but this had less effect on methylation at *PAI2* (65).

Asymmetric Methylation and Antisense RNA

While methylation in mammalian genomes is generally restricted to CpG sequences, plant and fungal genomes have methylated cytosines in other contexts as well. The SUP and PAI epialleles, for example, have extensive asymmetric and CpXpG methylation (21, 23). Two genetic screens have been used to address the origin of this methylation and both have identified the chromomethylase gene *CMT3* (44, 45). Analysis of methylation patterns using bisulphite sequencing show that *cmt3* primarily affects CpXpG methylation and, furthermore, *CMT3* transgenes restore methylation in *cmt3* plants (45), apparently de novo, although some methylation is left intact and may be used as a mark for remethylation. This residual methylation must be generated by yet another methylase, and *CMT2*, *Dnmt2*-, and *Dnmt3*-like genes are good candidates. Intriguingly, antisense *CMT2* transgenes failed to inactivate the endogenous gene in over 50 transgenic plants (66), consistent with a requirement for *CMT2* in antisense silencing. Chromodomains in *Drosophila* direct proteins to heterochromatin and have been shown in one case to bind untranslated RNA (67). This has led to the speculation that chromomethylases might mediate RNA-directed DNA methylation (68, 69), which has been demonstrated to be predominantly asymmetric in plants (70, 71).

Thus, while DNA-DNA pairing is certainly possible (21, 72), antisense transcription of *SUP* and *PAI* might be responsible for the asymmetric methylation. In the case of *PAI1-PAI4*, the inverted duplication is tail-to-tail, and would be expected to generate a hairpin transcript (73). In the case of *SUP*, promoter methylation might be triggered by antisense transcription of the hairpin-forming CpT microsatellite at the start site of transcription, in either the endogenous or transgenic copies of the gene. Inverted repeats have been shown to be good targets for DNA methylation in *Arabidopsis*, and perhaps this is why (23). For similar reasons, RNA-mediated methylation is also an attractive model for paramutation and presetting in maize.

Paradoxically, methylation actually increases at *SUP* and *AG* in *met1* and *ddm1* mutants, leading to epimutant alleles at a very high frequency (52). Could this also be due to RNA-mediated methylation? Perhaps trans-

posons such as MITES (miniature invert repeat transposable elements), which reside in the 3' untranslated regions of many plant genes, initiate antisense transcription when they are activated in methylation mutants. Alternatively, short transcripts produced from CpT repeats elsewhere in the genome might target methylation at these genes. Both AG and SUP have repeats related to simple sequences found in other genes (52).

RNA-dependent DNA methyltransferases may serve to monitor transcribed transposable elements. Transposons generate dsRNA because of their random integration in the genome, and because they have promoters in their terminal repeats (74). For this reason, they appear to be major targets of RNA interference in *Caenorhabditis elegans* (75, 76). In plants, they may be targeted for methylation using RNA as a guide. This methylation would be maintained over generations, resulting in the differential methylation of transposons and genes. Asymmetric methylation has been found during early embryogenesis in mammals (77) when the components of RNA interference are also active (78). Perhaps mammalian exons are methylated by transcriptions from transposons in neighbouring introns. However, chromomethylases are not found in mammals or fungi, indicating some other methyltransferase must be responsible for de novo transposon methylation in these genomes.

In *Ascobolus* and *Neurospora*, candidate RNA-dependent DNA methyltransferases have not been found. This could explain why quelling (an RNA-mediated process) requires the continuous presence of a silencing transgene array (which may become methylated) but does not lead to methylation of the endogenous target gene (79). Nonetheless, asymmetric methylation is found in both fungi, suggesting that it is established and maintained by different mechanisms (24, 80).

A Role for Silencing and DNA Modification in Development

Mutants in the silencing genes *DDM1*, *MOM1*, *CMT3*, *HOG1*, *SIL1* and *SIL2*, have very minor effects on plant development as a whole. Even mutants in *MET1* have relatively mild phenotypes that can be explained by generation of epialleles elsewhere in the genome. Similarly, *dim-2* has no detectable phenotype in *Neurospora*. In mammalian cells, by contrast, differentiation of methyltransferase negative embryonic stem cells leads to rapid programmed cell death, and this may require p53 (81), which is absent from plants and fungi (37).

Although redundancy may explain the absence of strong phenotypes in plants, what can we learn about the biological role of DNA methylation, gene silencing, and chromatin modification from these mutants? It has been reported that the paternal genome is silenced

throughout much of early embryonic development in *Arabidopsis* (82). However, at least one gene used in this study is expressed from the paternal genome at the earliest detectable stage (83), indicating that transcriptional gene silencing is unlikely to be involved. Nonetheless, the *polycomb* genes *MEDEA/FIS1* and *FIE1*, which are required for seed development, are not expressed from paternal chromosomes in the endosperm. *fie1* is partially rescued by *anti-met1* transgenes, and there is some evidence that *ddm1* ameliorates *medea* after fertilization (84, 85). Consistent with these observations, anti-MET1 plants appear to be more tolerant of changes in parental dosage in the developing endosperm, a phenomenon related to imprinting (86).

As in *Drosophila*, mutants that destabilize endogenous gene repression have been found in plants. For example, *curly leaf* (a *polycomb* homolog), *pickle* (a *Mi-2*-like *SNF2* homolog) and *asymmetric leaves1* (a *myb* domain homolog) derepress homeotic gene expression and have strong phenotypic effects (87). Unlike *ddm1*, however, they do not affect genomewide methylation, supporting the notion that heterochromatin is the primary target of DDM1, rather than euchromatic genes.

Given that methylation is prevalent in heterochromatin, and genes affecting it have little developmental role in plants and fungi, what then is its function? An obvious candidate is transposon regulation. While it is clear that transposons can be highly detrimental, they typically accumulate in heterochromatin ensuring a minimal impact on gene expression. This is true both of *Arabidopsis* and *Neurospora*, though transposons are far more widely distributed in maize (37, 88, 89). Euchromatic insertions are observed in *Arabidopsis*, but they are only widespread in *ddm1* mutants (55, 56). This indicates that DDM1 might be required to target transposons to heterochromatin, and that it may be differently regulated, or have reduced function, in maize.

Telomeres and centromeres, which are uniquely eukaryotic chromosomal components, not only accumulate transposons but also share several properties with them, such as replication through reverse transcription in the case of telomeres and LINES. Retroelements are the largest single component of most eukaryotic genomes and although the analogy between the nucleus and a retroviral particle is perhaps far-fetched, the notion that centromeres and telomeres were derived from transposons is gaining some appeal.

Epigenetic Inheritance and Natural Variation

The expression level of a gene—rather than the sequence of the protein product—can often determine phenotypes that contribute to natural variation (90). Epigenetic modifications provide

a means of altering expression states. For example, methylation and silencing of the *cycloidea* gene of the plant *Linaria vulgaris* causes a flower morphology change. This methylation state arose approximately 250 years ago, and has been maintained by selection ever since (91). It has been proposed that epigenetic differences revealed by *ddm1* might contribute to the loss of fitness observed upon inbreeding (53). The fact that such differences are polymorphic between strains provides a potential explanation for hybrid vigor and suggests that transposons may be involved (92), although unlike epimutations, new transposon insertions are polymorphic both within and between strains (56). Epigenetic changes have also been implicated in the genome remodeling that occurs in allotetraploid plants, where two related but distinct diploid genomes are combined in a single nucleus. For example, in allotetraploids made between different diploid Brassicas, or between *Arabidopsis* and a related species *Cardaminopsis*, nucleolar dominance is observed. In this phenomenon, the rDNA genes in the tetraploid are expressed from only one of the two parental diploid genomes. The dominance can be reversed by the methyltransferase inhibitor compound 5-azacytidine, suggesting a role for methylation (93). Individual genes and transposon sequences can also be silenced from one of the two genomes in a tetraploid (94, 95). The application of genome-wide expression profiling to polyploids in wild-type and methylation-deficient mutant backgrounds may reveal the molecular basis for these epigenetic changes and contribute to our understanding of natural variation in plants.

References and Notes

- A. P. Wolffe, M. A. Matzke, *Science* **286**, 481 (1999).
- P. D. Rabinowicz *et al.*, *Nature Genet.* **23**, 305 (1999).
- V. Colot, J. L. Rossignol, *Bioessays* **21**, 402 (1999).
- J. A. Yoder, C. P. Walsh, T. H. Bestor, *Trends Genet.* **13**, 335 (1997).
- International Human Genome Sequencing Consortium, *Nature* **409**, 860 (2001).
- N. V. Fedoroff, in *Mobile DNA*, D. Berg, M. Howe, Eds. (American Society for Microbiology, Washington, DC, 1989), pp. 375–411.
- A. Barkan, R. A. Martienssen, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 3502 (1991).
- A. M. Settles, A. Baron, A. Barkan, R. A. Martienssen, *Genetics* **157**, 349 (2001).
- E. B. Cambareri, H. M. Foss, M. R. Rountree, E. U. Selker, J. A. Kinsey, *Genetics* **143**, 137 (1996).
- H. D. Morgan, H. G. Sutherland, D. I. Martin, E. Whitelaw, *Nature Genet.* **23**, 314 (1999).
- D. Berg, M. Howe, Eds., *Mobile DNA* (American Society for Microbiology, Washington, DC, 1989).
- T. Belenkaya *et al.*, *Mol. Gen. Genet.* **259**, 79 (1998).
- A. A. Kiger, S. Gigliotti, M. T. Fuller, *Genetics* **153**, 799 (1999).
- R. Martienssen, *Trends Genet.* **14**, 263 (1998).
- V. L. Chandler, W. B. Eggleston, J. E. Dorweiler, *Plant Mol. Biol.* **43**, 121 (2000).
- D. Schwartz, *Maize Genet. Coop. Newsl.* **63**, 42 (1989).
- R. Martienssen, *Curr. Biol.* **6**, 810 (1996).
- M. Matzke, A. Matzke, W. B. Eggleston, *Trends Plant Sci.* **1**, 586 (1996).
- J. E. Dorweiler *et al.*, *Plant Cell* **12**, 2101 (2000).
- M. A. Matzke, M. F. Mette, A. J. Matzke, *Plant Mol. Biol.* **43**, 401 (2000).
- S. E. Jacobsen, E. M. Meyerowitz, *Science* **277**, 1100 (1997).
- S. Melquist, B. Luff, J. Bender, *Genetics* **153**, 401 (1999).
- B. Luff, L. Pawlowski, J. Bender, *Mol. Cell* **3**, 505 (1999).
- E. U. Selker, *Trends Genet.* **13**, 296 (1997).
- J. L. Rossignol, G. Faugeron, *Experientia* **50**, 307 (1994).
- L. Rhounim, J. L. Rossignol, G. Faugeron, *EMBO J.* **11**, 4451 (1992).
- J. T. Ireland, E. U. Selker, *Genetics* **146**, 509 (1997).
- C. Barry, G. Faugeron, J. L. Rossignol, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4557 (1993).
- M. R. Rountree, E. U. Selker, *Genes Dev.* **11**, 2383 (1997).
- V. P. Miao, M. Freitag, E. U. Selker, *J. Mol. Biol.* **300**, 249 (2000).
- V. Colot, L. Maloisel, J. L. Rossignol, *Cell* **86**, 855 (1996).
- C. Goyon, J. L. Rossignol, G. Faugeron, *Nucleic Acids Res.* **24**, 3348 (1996).
- T. H. Bestor, *Hum. Mol. Genet.* **9**, 2395 (2000).
- F. Malagnac, A. Gregoire, C. Goyon, J. L. Rossignol, G. Faugeron, *Mol. Microbiol.* **31**, 331 (1999).
- F. Malagnac *et al.*, *Cell* **91**, 281 (1997).
- E. Kouzminova, E. U. Selker, *EMBO J.*, in press.
- The *Arabidopsis* Genome Initiative, *Nature* **408**, 796 (2000).
- E. J. Finnegan, K. A. Kovac, *Plant Mol. Biol.* **43**, 189 (2000).
- E. J. Finnegan, W. J. Peacock, E. S. Dennis, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8449 (1996).
- M. J. Ronemus, M. Galbiati, C. Ticknor, J. Chen, S. L. Dellaporta, *Science* **273**, 654 (1996).
- A. Vongs, T. Kakutani, R. A. Martienssen, E. J. Richards, *Science* **260**, 1926 (1993).
- E. Richards, personal communication.
- S. Henikoff, L. Comai, *Genetics* **149**, 307 (1998).
- A. M. Lindroth *et al.*, *Science* **292**, 2077 (2001).
- L. Bartee, F. Malagnac, J. Bender, *Genes Dev.* **15**, 1753 (2001).
- C. M. Papa, N. M. Springer, M. G. Muszynski, R. Meeley, S. M. Kaepler, *Plant Cell*, in press.
- X. Cao *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4979 (2000).
- A. Dong *et al.*, *Nucleic Acids Res.* **29**, 439 (2001).
- J. A. Jeddelloh, T. L. Stokes, E. J. Richards, *Nature Genet.* **22**, 94 (1999).
- M. Okano, D. W. Bell, D. A. Haber, E. Li, *Cell* **99**, 247 (1999).
- G. L. Xu *et al.*, *Nature* **402**, 187 (1999).
- S. E. Jacobsen, H. Sakai, E. J. Finnegan, X. Cao, E. M. Meyerowitz, *Curr. Biol.* **10**, 179 (2000).
- T. Kakutani, J. A. Jeddelloh, S. K. Flowers, K. Munakata, E. J. Richards, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12406 (1996).
- J. W. Soppe *et al.*, *Mol. Cell* **6**, 791 (2000).
- A. Miura *et al.*, *Nature* **411**, 212 (2001).
- T. Singer, C. Yordan, R. A. Martienssen, *Genes Dev.* **15**, 591 (2001).
- P. Amedeo, Y. Habu, K. Afsar, O. M. Scheid, J. Paszkowski, *Nature* **405**, 203 (2000).
- D. R. Dorer, S. Henikoff, *Genetics* **147**, 1181 (1997).
- B. McClintock, *Cold Spring Harbor Symp. Quant. Biol.* **16**, 13 (1951).
- CSHL/WUGSC/PEB *Arabidopsis* Sequencing Consortium, *Cell* **100**, 377 (2000).
- O. Mittelsten-Scheid, K. Afsar, J. Paszkowski, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 632 (1998).
- J. B. Morel, P. Mourrain, C. Béclin, H. Vaucheret, *Curr. Biol.* **10**, 1591 (2000).
- I. J. Furner, M. A. Sheikh, C. E. Collett, *Genetics* **149**, 651 (1998).
- J. A. Jeddelloh, J. Bender, E. J. Richards, *Genes Dev.* **12**, 1714 (1998).
- L. Bartee, J. Bender, *Nucleic Acids Res.* **29**, 2127 (2001).
- C. M. McCallum, L. Comai, E. A. Greene, S. Henikoff, *Nature Biotechnol.* **18**, 455 (2000).
- A. Akhtar, D. Zink, P. B. Becker, *Nature* **407**, 405 (2000).
- E. J. Finnegan, M. Wang, P. Waterhouse, *Curr. Biol.* **11**, R99 (2001).

69. Y. Habu, T. Kakutani, J. Paszkowski, *Curr. Opin. Genet. Dev.* **11**, 215 (2001).
70. M. Wassenegger, *Plant Mol. Biol.* **43**, 203 (2000).
71. M. A. Matzke, A. J. Matzke, J. M. Kooter, *Science* **293**, 1080 (2001).
72. J. Bender, *Trends Biochem. Sci.* **23**, 252 (1998).
73. E. U. Selker, *Cell* **97**, 157 (1999).
74. M. N. Raizada, M. I. Benito, V. Walbot, *Plant J.* **25**, 79 (2001).
75. R. F. Ketting, T. H. Haverkamp, H. G. van Luenen, R. H. Plasterk, *Cell* **99**, 133 (1999).
76. H. Tabara *et al.*, *Cell* **99**, 123 (1999).
77. B. H. Ramsahoye *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5237 (2000).
78. P. Svoboda, P. Stein, H. Hayashi, R. M. Schultz, *Development* **127**, 4147 (2000).
79. C. Cogoni *et al.*, *EMBO J.* **15**, 3153 (1996).
80. G. Faugeron, *Curr. Opin. Microbiol.* **3**, 144 (2000).
81. L. Jackson-Grusby *et al.*, *Nature Genet.* **27**, 31 (2001).
82. J. P. Vielle-Calzada, R. Baskar, U. Grossniklaus, *Nature* **404**, 91 (2000).
83. P. S. Springer, D. R. Holding, A. Groover, C. Yordan, R. A. Martienssen, *Development* **127**, 1815 (2000).
84. J. P. Vielle-Calzada *et al.*, *Genes Dev.* **13**, 2971 (1999).
85. R. Vinkenoog *et al.*, *Plant Cell* **12**, 2271 (2000).
86. S. Adams, R. Vinkenoog, M. Spielman, H. G. Dickinson, R. J. Scott, *Development* **127**, 2493 (2000).
87. M. Byrne, M. Timmermans, C. Kidner, R. Martienssen, *Curr. Opin. Plant Biol.* **4**, 38 (2001).
88. E. B. Cambareri, R. Aisner, J. Carbon, *Mol. Cell. Biol.* **18**, 5465 (1998).
89. P. SanMiguel *et al.*, *Science* **274**, 765 (1996).
90. R. Mauricio, *Nature Rev. Genet.* **2**, 370 (2001).
91. P. Cubas, C. Vincent, E. Coen, *Nature* **401**, 157 (1999).
92. R. Martienssen, *Curr. Opin. Genet. Dev.* **8**, 240 (1998).
93. Z. J. Chen, C. S. Pikaard, *Genes Dev.* **11**, 2124 (1997).
94. L. Comai *et al.*, *Plant Cell* **12**, 1551 (2000).
95. H. S. Lee, Z. J. Chen, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6753 (2001).
96. We thank E. Selker, E. Richards, V. Chandler, S. Kaeppler, S. Jacobsen, and J. Bender for communicating results prior to publication, two anonymous referees for suggestions for improvement, and our colleagues for many interesting discussions. R.M. and V.C. receive grant support from the NSF (DBI 1057338).

REVIEW

Translating the Histone Code

Thomas Jenuwein¹ and C. David Allis²

Chromatin, the physiological template of all eukaryotic genetic information, is subject to a diverse array of posttranslational modifications that largely impinge on histone amino termini, thereby regulating access to the underlying DNA. Distinct histone amino-terminal modifications can generate synergistic or antagonistic interaction affinities for chromatin-associated proteins, which in turn dictate dynamic transitions between transcriptionally active or transcriptionally silent chromatin states. The combinatorial nature of histone amino-terminal modifications thus reveals a "histone code" that considerably extends the information potential of the genetic code. We propose that this epigenetic marking system represents a fundamental regulatory mechanism that has an impact on most, if not all, chromatin-templated processes, with far-reaching consequences for cell fate decisions and both normal and pathological development.

Genomic DNA is the ultimate template of our heredity. Yet despite the justifiable excitement over the human genome, many challenges remain in understanding the regulation and transduction of genetic information (1). It is unclear, for example, why the number of protein-coding genes in humans, now estimated at ~35,000, only doubles that of the fruit fly *Drosophila melanogaster*. Is DNA alone then responsible for generating the full range of information that ultimately results in a complex eukaryotic organism, such as ourselves?

We favor the view that epigenetics, imposed at the level of DNA-packaging proteins (histones), is a critical feature of a genome-wide mechanism of information storage and retrieval that is only beginning to be understood. We propose that a "histone code" exists that may considerably extend the information potential of the genetic (DNA) code. We review emerging evidence that histone proteins and their associated covalent modifications contribute to a mechanism that can alter chromatin structure, thereby leading to inherited differences in transcriptional "on-off" states or to the stable propagation of chromosomes by defining a specialized higher order structure at centromeres. Under the assumption that a histone code exists, at least in some form, we discuss potential mecha-

nisms for how such a code is "read" and translated into biological functions.

Throughout this review, we have chosen epigenetic phenomena and underlying mechanisms in two general categories: chromatin-based events leading to either gene activation or gene silencing. In particular, we center our discussion on examples where differences in "on-off" transcriptional states are reflected by differences in histone modifications that are either "euchromatic" (on) or "heterochromatic" (off) (Fig. 1A). We also point out that, despite many elegant genetic and biochemical insights into chromatin function and gene regulation in the budding yeast *Saccharomyces cerevisiae*, some of the heterochromatic mechanisms (e.g., HP1-based gene silencing) discussed here do not exist in an obvious form in this organism. Thus, we will need to pursue other model systems, such as *Schizosaccharomyces pombe*, *Caenorhabditis elegans*, *Drosophila*, and mice, to "crack" the histone code.

Chromatin Template and Histone Code

In the nuclei of all eukaryotic cells, genomic DNA is highly folded, constrained, and compacted by histone and nonhistone proteins in a dynamic polymer called chromatin. For example, chromosomal regions that remain

transcriptionally inert are highly condensed in the interphase nucleus and remain cytologically visible as heterochromatic foci or as the "Barr body," which is the inactive X chromosome in female mammalian cells (2). The distinct levels of chromatin organization are dependent on the dynamic higher order structuring of nucleosomes, which represent the basic repeating unit of chromatin. In each nucleosome, roughly two superhelical turns of DNA wrap around an octamer of core histone proteins formed by four histone partners: an H3-H4 tetramer and two H2A-H2B dimers (3). Histones are small basic proteins consisting of a globular domain and a more flexible and charged NH₂-terminus (histone "tail") that protrudes from the nucleosome. It remains unclear how nucleosomal arrays containing linker histone (H1) then twist and fold this chromatin fiber into increasingly more compacted filaments leading to defined higher order structures.

Central to our current thinking is that chromatin structure plays an important regulatory role and that multiple signaling pathways converge on histones (4). Although histone proteins themselves come in generic or specialized forms (5), exquisite variation is provided by covalent modifications (acetylation, phosphorylation, methylation) of the histone tail domains, which allow regulatable contacts with the underlying DNA. The enzymes transducing these histone tail modifications are highly specific for particular amino acid positions (6, 7), thereby extending the information content of the genome past the genetic (DNA) code. This hypothesis predicts that (i) distinct modifications of the

¹Research Institute of Molecular Pathology (IMP) at the Vienna Biocenter, Dr. Bohrgasse 7, A-1030 Vienna, Austria. E-mail: jenuwein@nt.imp.univie.ac.at
²Department of Biochemistry and Molecular Genetics, University of Virginia Health Science Center, Charlottesville, VA 22908, USA. E-mail: allis@virginia.edu