The continuity of life is attributed to the faithful copying of DNA. It is therefore intriguing that some organisms chemically modify their DNA in ways that can lead to inheritable changes. Soon after the discovery that a fraction of cytosines are methylated in DNA of some eukaryotes—including vertebrates, plants, and some fungi—evidence began to accumulate that DNA methylation is inherently mutagenic. For example, in mammals, where methylation is almost exclusively found at cytosines immediately preceding guanines (“CpG” dinucleotides), it was noticed that methylated chromosomal regions are typically unexpectedly poor in CpGs and unexpectedly rich in TpGs and CpaAs, apparently as a result of C to T mutations at sites subject to methylation (Bird 1980). It was assumed that this reflected the fact that 5-methylcytosine, like cytosine, is somewhat unstable, spontaneously decomposing to 5-methyl-uridine (thymine) and TpGs, and that G–T mismatches are repaired less efficiently than C–G mismatches. Discovery of repeat-induced point mutation (RIP) in the filamentous fungus Neurospora crassa raised the possibility that DNA methyltransferases and similar enzymes may be directly responsible for such mutations (Selker 1998; Freitag et al. 2002). What positive functions does DNA methylation serve? Although the control and function of DNA methylation are both still hotly debated, it is becoming increasingly clear that DNA methylation is one of several mechanisms that defend the eukaryotic genomes from mischievous DNA. Neurospora crassa provided an early, and exceptionally clear, illustration of this and it continues to provide a convenient model to elucidate DNA methylation and other epigenetic processes.

The Methylation Landscape

Although Neurospora crassa was initially reported to have no methylated DNA (Antequera et al. 1984), observations in the 1980s revealed cytosine methylation in this model eukaryote. After methods became established to transform Neurospora, it was noticed that some transforming sequences are subject to de novo methylation, especially when they are repeated (Bull and Wootten 1984; Selker et al. 1987a; Orbach et al. 1988). At the same time, natural patches of methylation were found associated with several SS-rRNA pseudogenes (Selker and Stevens 1985; Margolin et al. 1998). No methylation has been reported in any of the hundreds of protein-coding genes that have been studied. Indeed, as discussed below, recent analyses suggest that most DNA methylation is associated with relics of transposons. We now know that ~1.5% of the cytosines in the DNA of N. crassa are methylated (Russell et al. 1987; Foss et al. 1993).

The first methylated region characterized in detail is the 1.4-kb zeta–eta (ζ–η) region (Selker and Stevens 1985; Selker et al. 1993b). This consists of a diverged tandem duplication of a 0.8-kb segment of DNA, including a 5S rRNA gene. Comparison of this region with the corresponding chromosomal region of strains lacking the duplication led to the idea that repeated sequences can somehow induce DNA methylation (Selker et al. 1985; 1987a; Grayburn and Selker 1989) and ultimately led to the discovery of the genome defense system that we named RIP (Selker et al. 1987b; Selker and Garrett 1988; Cambareri et al. 1989). Both the ζ–η region and the psi-63 (Ψ63) region, the second methylated region discovered in Neurospora (Metzenberg et al. 1985; Foss et al. 1993; Miao et al. 1994; Margolin et al. 1998), are products of RIP.

The Prototypical Genome Defense System: RIP

RIP detects duplicated sequences in the haploid genomes of special dikaryotic cells resulting from fertilization and then riddles both copies of the duplicated sequence with polarized transition mutations, C–G pairs are replaced with T–A pairs (Cambareri et al. 1989). RIP has a clear sequence preference (Cs preceding As are mutated most frequently), making it straightforward to recognize sequences that have been methylated by RIP (Margolin et al. 1998). In a single passage through the sexual cycle, up to ~30% of the G–C pairs in duplicated sequences can be changed to A–T pairs (Cambareri et al. 1991). After RIP, remaining cytosines are generally methylated, including those not in symmetrical sequences (e.g., CpgIs) (Selker et al. 1993a). This heterogeneous methylation can extend beyond the mutational hotspots—e.g., GpGIs are less methylated than CpGIs. It is now clear that RIP is an ancient defense mechanism that evolved independently in fungi, plants, and vertebrates.
tated region and even beyond the edge of the segment that was originally duplicated (Campariere et al. 1989; Selker et al. 1993a; Irelan and Selker 1997; Miao et al. 2000).

Results of genome-wide analyses of DNA methylation suggest that the vast majority of methylated residues are in relics of RIP (Galagan et al. 2003; Selker et al. 2003). Indeed, the only methylation not known to have resulted from RIP is that in the tandemly arranged DNA (Perkins et al. 1986a). A survey of methylated Neurospora sequences isolated by affinity chromatography using the methyl-binding domain (MBD) of McCP2 revealed clear evidence of RIP in 47 of 51 sequenced fragments (Selker et al. 2003). Analysis of these sequences revealed significant similarities with a variety of transposable elements identified in other organisms, including both retrotransposons (e.g., lolligad, Tad, Tecn, and DAB1) and DNA-type transposons (e.g., Dodo1, Dodo2, Dodo3, listless, dPunt, Punt3, and Nogo) (Selker et al. 2003). Similar information came from large-scale genome sequencing projects. About 10% of the genome is composed of repetitive DNA and most of this is identifiable as transposon-like sequences that have been riddled with mutations by RIP (Galagan et al. 2003). Not a single intact transposable element was identifiable, consistent with failures to detect transpositions in standard Neurospora strains (Kinsey and Hebler 1989). We conclude that RIP is an efficient genome defense mechanism. Does this mean that Neurospora cannot generate gene families by gene duplication? Nearly 20% of N. crassa genes are found in multigene families, but, interestingly, nearly all of the paralogs are sufficiently divergent and/or short that they should be invisible to RIP (Galagan et al. 2003). Thus, although Neurospora has gene families, RIP is presumably an obstacle to evolution of new genes through gene duplication (Galagan and Selker 2004).

DE NOVO AND MAINTENANCE METHYLATION IN NEUROSPORA

Laboratory experiments revealed that most products of RIP, as well as their unmethylated counterparts, are methylated (Selker et al. 1987a, 1993a; Cambareri et al. 1991; Singer et al. 1995; Miao et al. 2000). Similarly, predicted products of RIP identified in the genome based on their sequence composition are generally methylated (Galagan et al. 2003; Selker et al. 2003). In principle, such methylation could reflect signals for de novo methylation that work in vegetative cells. Native, unmethylated sequences are not susceptible, while many foreign sequences with resemblance to sequences mutated by RIP are (M. Freitag and E. Selker, unpubl.). This methylation appears loosely correlated with the copy number of transforming DNA. Although single-copy sequences are less frequently methylated than multicopy sequences (Pandit and Russo 1992; Romano and Macino 1992), we have found that sequences differ in their susceptibility to methylation in vegetative cells. Native, unmethylated sequences are not susceptible, while many foreign sequences with resemblance to sequences mutated by RIP are (M. Freitag and E. Selker, unpubl.). This methylation appears loosely correlated with the copy number of transforming DNA. Although single-copy sequences are less frequently methylated than multicopy sequences (Pandit and Russo 1992; Romano and Macino 1992; Selker et al. 1993b), multicopy sequences are not always methylated (Selker and Garrett 1988).

Observations on the specificity of methylation in Neurospora led to the “collapsed chromatin model,” in which DNA methylation is the fate of sequences that are inert (i.e., completely inactive [Selker 1990a]). One prediction of the model was that short sequences should not have the ability to trigger methylation, for example, when inserted into an active gene. To explore this and other possibilities, we developed efficient gene-targeting systems that allowed us to test the methylation potential of single copies of sequences integrated precisely and without excision sequences at a common chromosomal position (at the am locus on LG PR; Miao et al. 1994) or the his-locus on LG IR. Margolin et al. (1997). Using these systems, we demonstrated that mutations per se (e.g., numerous A to G mutations) do not trigger methylation and that products of RIP, such as the ζ-η region, effectively contain multiple, additive methylation signals, whose effects can spread hundreds of base pairs into flanking sequences. Fragments of the ζ-η region as short as 171 bp can trigger methylation (Selker et al. 1993b; Miao et al. 2000). We found that mutation density per se does not determine whether sequences become methylated and that neither A-T richness nor high densities of TpA dinucleotides, typical attributes of methylated sequences in Neurospora, are essential features of methylation signals. Nevertheless, both A-T richness and high densities of TpA dinucleotides appear to promote methylation in Neurospora. These and other findings from “transplanting” small fragments of genes, pieces of DNA mutated by RIP, and synthetic sequences led us to con-
clude that methylated sequences do not simply reflect the absence of signals that prevent methylation; they appar-
etly contain positive signals that trigger methylation (Miao et al. 2000).

To better define the nature of these signals, we devel-
oped a more sensitive assay to test the capacity of short (25–100-bp) synthetic oligonucleotides to trigger methyl-
ation at a specific locus (Tamaru and Selker 2003). Our
system used a his-5 targeting vector carrying a 100-bp ζ segment surrounded by a lightly RIP-mutated allele of the am gene. We demonstrated that this mosaic con-
struct does not trigger methylation itself at his-5, but does provide a sensitive context to assess the potential of vari-
ous sequences to induce methylation. A variety of ran-
don sequences consisting of only A and T residues trig-
ger methylation of nearby cytosines. Introduction of G:C pairs into the A:T-rich sequences was strongly in-
hibitory and both As and Ts were found to be required on the same strand to trigger significant methylation. Never-
theless, neither TpA nor ApT dinucleotides were essen-
tial. Tests of 20-, 25-, 40-, and 80-mer fragments of the most potent sequence, (TAAA)₅₆, showed that longer tracts of this sequence act as stronger signals. It seems possible that an unidentified “A:T-hook”-type protein mediates methylation in Neurospora. Consistent with this,
possibility, we found that Distamycin A, an analog of the A:T-hook motif, interferes with de novo methylation in
Neurospora (Tamaru and Selker 2003). Nevertheless,
some of our findings do not support this hypothesis. In particular, we found that sequences with only two bp A:T tracts (e.g., (CTA)n) can induce methylation. In addition,
using A:T-hook protein HMG-I (kindly provided by R. Reeves, Washington State University, Pullman), we
showed that the sequence preference for HMG-I binding does not simply correlate with the sequence preferences for de novo methylation. Additional work will be re-
quired to determine how relics of RIP, and similar degen-
erate sequences, are recognized to trigger methylation in vegetative cells of Neurospora.

FOREWARD GENETICS APPROACH TO ELUCIDATE CONTROL AND FUNCTION OF DNA METHYLATION

Neurospora is well suited to identify components of the methylation machinery by genetic approaches that do not rest on prior knowledge or preconceptions. We have
successfully used several approaches to identify mutants defective in methylation (dim), but have not yet saturated the Neurospora genome for such mutations. DNA meth-
ylation can silence genes in Neurospora (Rountree and
Selker 1997), allowing for the direct selection of muta-
tions that affect DNA methylation. Nevertheless, most
known dim mutants have been identified in other ways,
such as in “brute force” screens using Southern hy-
bridization to identify mutants that affect methylation (Foss et al. 1993, 1995). Two mutants, dim-2 and dim-5, abolish all detectable DNA methylation. Mapping and com-
plementation studies revealed that the dim-2 gene en-
codes a DNA methyltransferase (DMTase) (Kouzmintova
and Selker 2001), while dim-5 encodes a histone methyl-
transferase (HMTase) (Tamaru and Selker 2001).

Biochemical characterization of DIM-2 should be in-
teresting. This predicted 1454-amino-acid protein is re-
sponsible for all detected DNA methylation in Neu-
rospora, including methylation in a variety of sequence
contexts. Mutations preventing all methylation have not
been described in other eukaryotes, which may reflect the
fact that DNA methylation results from multiple
DMTases in most organisms that have been characterized
and that DNA methylation is essential in some organ-
isms. Although the DIM-2 carboxy-terminal domain in-
cludes all the well conserved motifs characteristic of
DMTases, its amino-terminal domain shows no marked
similarity to previously described proteins (Kouzmintova
and Selker 2001).

CONTROL OF DNA METHYLATION BY HISTONE METHYLATION

The discovery that dim-5 encodes a HMTase provided
the first indication that DNA methylation is regulated by
histone methylation, at least in some organisms. Bio-
chemical work showed that DIM-5 is specific for lysine 9 of
histone H3 and that it efficiently trimethylates this
residue in vitro and in vivo, unlike all previously charac-
terized HMTases (Tamaru and Selker 2001; Tamaru et al.
2003). In order to determine whether histone H3 is the
critical substrate for DIM-5 in vivo, we transformed a
Dim⁻ Neurospora strain with engineered histone H3 genes
carrying a mutation at lysine 9 that would preclude
methylation. Amino acid substitutions at lysine 9 dramat-
ically reduced DNA methylation, implicating histone H3
as the critical substrate for DNA methylation (Tamaru
and Selker 2001). Evidence that some DNA methylation
in Arabidopsis is also controlled by histone methylation
came soon thereafter (Jackson et al. 2002; Malagnac et
al. 2002). Interestingly, the critical mark in Arabidopsis ap-
ppears to be di-, rather than tri-, methyl-lysine 9 (Jackson
et al. 2004). Most recently, Jenewein and colleagues have
provided evidence in mice that trimethylation of lysine 9
of histone H3 directs DNA methylation to major satellite
repeats at pericentric heterochromatin (Lehnertz et al.
2003). The extent to which DNA methylation is con-
trolled by histones in eukaryotes is not yet known.

SEARCH FOR OTHER COMPONENTS OF DNA METHYLATION PATHWAY

In principle, forward genetics should reveal all compo-
nents of the DNA methylation pathway except for those
that are essential or redundant. However, use of reverse
genetics and biochemical approaches can facilitate the
identification of components in the methylation process.
For example, evidence from other systems that a hetero-
chromatin protein, HP1 (Eisenberg and Elgin 2000),
binds methylated lysine 9 of histone H3 (Bannister et al.
2001; Jacobs et al. 2001; Lachner et al. 2001), led us to
search for an HP1 homolog in Neurospora and to test its
possible involvement in DNA methylation by disrupting
the HP1 gene (hp1Δ) by RIP. We found that HP1 is indeed essential for DNA methylation in *Neurospora*, implying that this protein directs DM-2 to DNA associated with chromatin in which histone H3 is methylated at lysine 9 (Freitag et al. 2004a). 

Clues that the RNAi machinery may be involved in DNA methylation (Wassengegger et al. 1994; Aufsatz et al. 2002; Chan et al. 2004) prompted us to test if components of the *Neurospora* RNAi machinery, including apparent RNA-dependent RNA polymerases, dicers, argonaute homologs, and RISC have been identified by genetic or bioinformatic approaches (Gaigan et al. 2003). We found no evidence that any of the RNAi machinery is involved in initiation or maintenance of DNA methylation (Freitag et al. 2004b). In contrast, mutational analyses of histone deacetylase genes of *Neurospora* revealed that at least one of these genes is involved in DNA methylation in *Neurospora* (K. Smith et al., unpubl.), consistent with the observation that the histone deacetylase inhibitor Trichostatin A reduces DNA methylation in some chromosomal regions of this organism (Selker 1998).

**CONCLUSIONS**

Why control DNA methylation through histone methylation? In the last few years it has become increasingly clear that histones are more than structural proteins; they are informational molecules (Jenuwein and Allis 2001). Particular combinations of posttranslational modifications of histones, including phosphorylation, methylation, acetylation, ubiquitination, and ADP ribosylation, can influence the function of the associated DNA and the modifications can depend on the presence or absence of other modifications. For example, methylation of lysine 9 on histone H3 is inhibited by phosphorylation of serine 10 and methylation of lysine 4, which are associated with active sequences (Rea et al. 2000; E. Berge et al., unpubl.) and there are indications that histone acetylation also influences methylation of lysine 9 (Selker 1998; Nakayama et al. 2001; K. Smith et al., unpubl.). Thus histones are well suited to integrate information bearing on the DNA that they are associated with.

*Neurospora* provided the first example of a genome defense system, RIP, and has more recently revealed two additional genetic mechanisms related to RNA interference (RNAs) discovered in plants and animals that should also resist change in the genome. The first, “quelling,” is specific to the vegetative phase of the organism. It is essential for DNA methylation in *Neurospora* (K. Smith et al., unpubl.), consistent with the observation that the histone deacetylase inhibitor Trichostatin A reduces DNA methylation in some chromosomal regions of this organism (Selker 1998).

ACKNOWLEDGMENTS

I thank Michael Freitag, Kyesmit Adhvaryu, and Kristina Smith for comments on the manuscript and thank present and past members of my laboratory for their contributions to our research. I gratefully acknowledge funding from NIH (grant GM35690) and NSF (grant MCB0131383), which contributed to this work.

REFERENCES


**SELKER**


Genome Defense and DNA Methylation in *Neurospora*

E.U. SELKER

Access the most recent version at doi:10.1101/sqb.2004.69.119

This article cites 64 articles, 35 of which can be accessed free at:
http://symposium.cshlp.org/content/69/119.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to *Cold Spring Harbor Symposia on Quantitative Biology* go to:
http://symposium.cshlp.org/subscriptions