View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Current Biology, Vol. 12, 65–68, January 8, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S0960-9822(01)00622-4

Genome-Wide Profiling of DNA Methylation Reveals Transposon Targets of CHROMOMETHYLASE3

Rachel Tompa,[1] Claire M. McCallum,[1] Jeffrey Delrow,[1] Jorja G. Henikoff,[1] Bas van Steensel,[2] and Steven Henikoff[1,3]
[1] Howard Hughes Medical Institute and
Fred Hutchinson Cancer Research Center
Seattle, Washington 98109
[2] Swammerdam Institute for Life Sciences
University of Amsterdam
Amsterdam
The Netherlands

Summary

DNA methylation has been implicated in a variety of epigenetic processes, and abnormal methylation patterns have been seen in tumors [1–3]. Analysis of methylation patterns has traditionally been conducted either by using Southern analysis after cleavage with methyl-sensitive restriction endonucleases or by bisulfite sequencing [4]. However, neither method is practical for analyzing more than a few genes. Here, we describe a simple technique for genome-wide mapping of DNA methylation patterns. Fragmentation by a methyl-sensitive restriction endonuclease is followed by size fractionation and hybridization to microarrays. We demonstrate the utility of this method by characterizing methylation patterns in *Arabidopsis* methylation mutants. This analysis reveals that CHROMOMETHYLASE3 (CMT3) [5], which was previously shown to maintain CpXpG methylation [6, 7], preferentially methylates transposons, even when they are present as single copies within the genome. Methylation profiling has potential applications in disease research and diagnostic screening.

Results and Discussion

Chromomethylases (CMTs) are chromodomain-containing cytosine DNA methyltransferase homologs [5, 8, 9]. Chromodomains are thought to be chromatin interaction modules that mediate chromatin inheritance. However, the molecular mechanisms whereby different chromodomains interact with chromatin appear to differ between classes of chromodomains. For example, the N-terminal chromodomain of HP1 binds methylated lysine 9 of histone H3 [10], the HP1 C-terminal "shadow" chromodomain binds a pentameric peptide target [11], and the chromodomain of MOF1 binds RNA [12]. This latter observation, together with evidence that DNA methylation in plants may be guided by double-stranded RNA [13], has led to the idea that CMT chromodomains are involved in the recognition of target sites by an RNAi-mediated process [14].

The CMT3 chromomethylase appears to be a key plant DNA methyltransferase, because *Arabidopsis cmt3* mutant plants are deficient in CpXpG methylation [6, 7], and similar results have been reported for a maize homolog [15]. Furthermore, *cmt3* mutants abolish epigenetic silencing at both the SUPERMAN and PAI loci in *Arabidopsis* [6, 7]. However, it is not known what is responsible for the preference of CMT3 for particular CpXpG target sites. Therefore, we sought to assay for CMT3 targets in an unbiased way to determine if there is an underlying pattern of target specificity.

To identify CMT3 targets genome wide, we adapted a fractionation procedure that we had previously introduced for chromatin profiling [16]. DNA from cells of interest is isolated, digested with a methyl-sensitive enzyme, and size fractionated (Figure 1). Experimental and control fragments of <2.5 kb are fluorescently labeled with either Cy3 or Cy5 and hybridized to a DNA microarray. In the present study, DNA was purified from *A. thaliana chromomethylase3-2* (*cmt3-2*) null mutant [5] or *methyltransferase1* (*met1*) antisense mutant [17] and corresponding wild-type (WT) tissue, and samples were digested with MspI, a restriction endonuclease that preferentially cleaves CCGG but is blocked by methylation of the outer cytosine. The DNA was size fractionated, and labeled fragments were hybridized to a DNA microarray. Fragments from mutant DNAs were labeled with Cy5, and fragments from WT DNA were labeled with Cy3 fluorescent dyes. The resulting Cy3:Cy5 ratio reveals changes in methylation between the WT and mutant plants; a log ratio significantly lower than 0 indicates relative hypermethylation in the mutant, whereas a log ratio greater than 0 indicates relative hypomethylation. The detection of significant log ratios provides a qualitative assessment of differential methylation between experimental and control samples; sites that are found to have altered methylation can then be analyzed further using conventional methods. Our microarray contained 240 single-copy fragments selected at random throughout the genome. An additional 144 loci that have been previously implicated in epigenetic phenomena were included as well.

We found eight loci displaying significant changes in CpXpG methylation in *cmt3* (Figure 2A). All eight loci have positive log ratios, indicating that *cmt3* mutant plants are hypomethylated at CpXpG sites for these loci. Our genome-wide sampling extends previous observations in which *cmt3* plants were shown to have decreased, but not increased, levels of CpXpG methylation at particular sites [6, 7]. In contrast, we detected significant loci with either positive or negative log ratios in *met1* plants (Figure 2B), in support of previous work showing both hypomethylation and hypermethylation [18].

In *cmt3* plants, four of the eight hypomethylated loci (loci 2–5, Figure 2A and Table 1) were from the set of randomly chosen fragments. Notably, all 4 were determined to be retrotransposons, both LTR and non-LTR, whereas only 10 out of the 240 random fragments were identified as transposable elements, based on manual scrutiny of blastx [19] search results. The inference that retrotransposons are targeted by CMT3 is consistent

[3]Correspondence: steveh@fhcrc.org

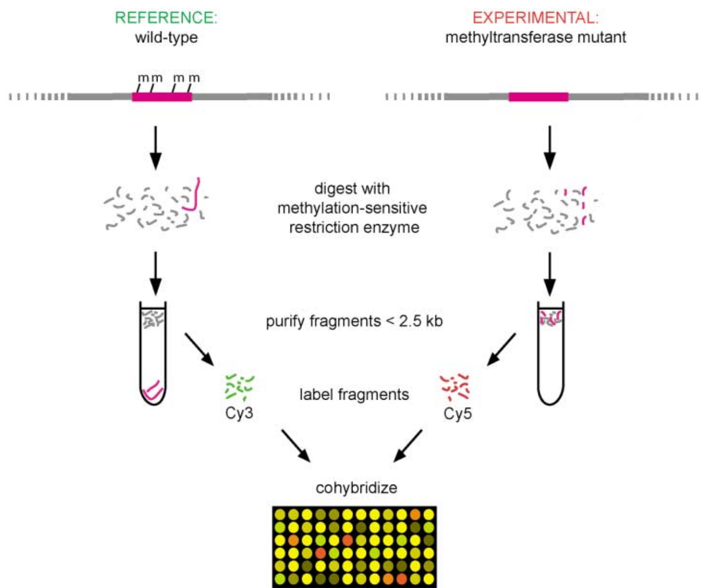

Figure 1. Schematic Diagram of the Methylation-Profiling Procedure Described in the Text

Methylated bases (m) that would protect a region (magenta) from restriction digestion in the reference sample are indicated. In the experiments described here, genomic DNA was isolated from *Arabidopsis thaliana* leaves and stems using the FastDNA kit (Bio101). Samples containing 20–30 μg DNA were digested for 1–2 hr with MspI (330 units, GIBCO-BRL). Digested DNA samples were size fractionated on 11-ml sucrose gradients (5%–30%) by ultracentrifugation, as previously described [16]. Gradient fractions containing DNA fragments smaller than 2.5 kb, as determined by agarose gel electrophoresis, were pooled and concentrated by isopropanol precipitation. DNA samples (1 μg) were labeled with Cy3- or Cy5-dCTP (Amersham Pharmacia) by random priming, mixed, and cohybridized to microarrays [22]. After washing, arrays were spun dry in a centrifuge and scanned using a GenePix 4000 fluorescent scanner (Axon Instruments). To obtain the 240 random single-copy loci, nonoverlapping 1-kb segments were selected at random from a database consisting of available contigs for *A. thaliana* chromosomes 1, 3, and 5 and complete chromosomes 2 and 4 from GenBank (August 2, 2000). The Primer3 program running on a SUN Solaris workstation [23] was applied with default settings to select 17-bp candidate primers with $T_m > 50°C$ that encompass 700 ± 25 bp fragments. To stringently avoid redundant loci, each fragment was then searched against the database using blastn (word size = 20, match score = 1, mismatch score = −50, gap open penalty = 50, gap extension penalty = 50, and no filtering) [19], reporting only those hits with E value < 0.15, except for the first hit to itself. Custom oligonucleotide primers were obtained in 96-well microtiter plates from MWG-Biotech and were used for PCR amplification with Ex-Taq polymerase (Fisher/Panvera Labs) [16]. Microarray construction and hybridization protocols were modified from those described elsewhere [24]. Microarrays were constructed by mechanically spotting PCR amplification products onto poly-lysine-coated microscope slides using an OmniGrid high-precision robotic gridder (GeneMachines). All loci were present in duplicate on the microarray, and each experiment was repeated three times.

with the activation of transcription of certain retrotransposons in *cmt3* mutant individuals [6]. Furthermore, each of these four loci is present in the genome in only a single copy, a consequence of our blastn filtering to exclude repetitive sequences from the microarray. We conclude that targeting of transposons by CMT3 does not require them to be repetitive in the genome.

Three of the remaining four significant loci included both the promoter and ORF of Athila, a repeated retrotransposon found near centromeres [20], and the 180-bp centromeric repeat. Both repeats have previously been shown to be hypomethylated in *cmt3*, based on enhanced cleavage by MspI restriction endonuclease detected using Southern analysis [6]. The other signifi-

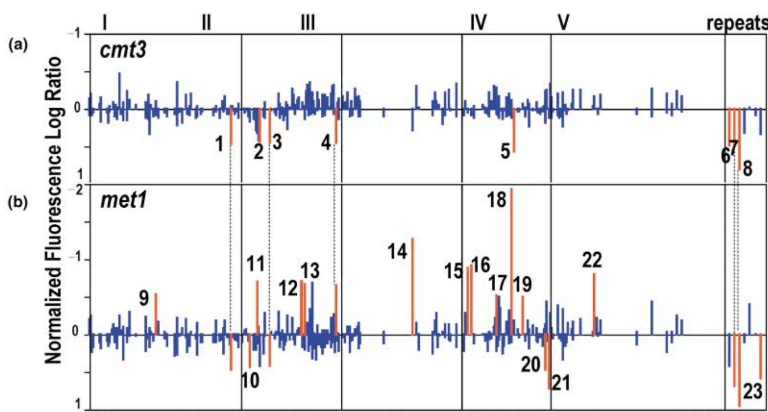

Figure 2. Chromosomal Maps of Cy3:Cy5 Ratios of Hybridizations with Genomic Probes from *cmt3-2* and *met1* Mutants Versus Wild-Type, WT

The 384 loci are represented by their approximate position along the chromosomes (I–V), and their relative heights are the normalized log fluoresence ratios. Loci indicated by red bars show significant differences from the mean log ratio of 0. Numbers 1–8 refer to loci listed in Table 1. p values were determined using CyberT [25], with window size = 101 and Bonferroni correction = 0.05. Only loci in which mean log ratios were greater than 0.4 or less than −0.4, and in which both duplicates were independently characterized as significant by CyberT, were classified as significant. Changes reported as significant did not depend on dye-labeling orientation, because reversing Cy3/Cy5 labeling between sample pairs gave similar ratios between mutant and WT (data not shown). Prior to use, *cmt3-2* was backcrossed to its parental ecotype, No-0. No-0 served as the wild-type *CMT3*$^+$ control. *cmt3-2,* which truncates the CHROMOMETHYLASE3 protein within the catalytic domain [5], behaves as a null in complementation tests, and it completely suppresses the *clk* phenotype [6]. Based on the frequency of homozygous methylation mutations discovered in a mutational screen for suppressors of *clk-st* [6], we estimate a maximum probability of 0.0005 that a background suppressor mutation, in addition to *cmt3-2*, contributes to the observed differences. *met1* is an antisense line [17] whose parental ecotype, C24, served as the wild-type *MET1*$^+$ control.

Table 1. Loci Displaying Significant Methylation Changes in *cmt3* and *met1* Mutant Plants

Locus	Annotation	Methylation in $cmt3^{-/-}$	P Value	Methylation in $met1^{-/-}$	P Value	GenBank ID
1	Rap2.1 promoter	hypo	0.016	hypo	0.002	AC083835 (99137–98479)
2	Copia-like retrotransposon	hypo	0.004	–	0.011	AC006219 (29200–29899)
3	Tal1-like non-LTR retroelement	hypo	0.008	hypo	0.005	AC007063 (10926–11647)
4	Non-LTR retroelement	hypo	0.034	hyper	0.049	AC002387 (31095–31781)
5	Copia-like retrotransposon	hypo	0.002	–	0.228	AL161554 (179012–179717)
6	180-bp centromeric repeat	hypo	0.048	–	0.034	AC011621 (15014–15133)
7	Athila ORF	hypo	0.018	hypo	0.006	X81801 (2091–2775)
8	Athila promoter	hypo	0.013	hypo	0.004	X81801 (328–1008)
9	Cryptochrome 2 apoprotein			hyper	0.014	AC000104 (45542–46262)
10	Unknown			hypo	0.008	AC006955 (70121–70828)
11	Non-LTR retroelement			hyper	0.020	AC007730 (9338–10032)
12	Ser/Thr protein kinase			hyper	0.022	AC005312 (36146–36851)
13	Unknown			hyper	0.008	AC002334 (41794–42468)
14	5s rDNA			hyper	0.008	AL356013 (144–430)
15	Oxidoreductase homolog			hyper	0.030	AF069442 (31380–30736)
16	Hypothetical protein			hyper	0.038	AF069441 (44085–43402)
17	Disease-resistance protein			hyper	0.018	AL035528 (76686–77360)
18	Probable retroelement			hyper	0.001	AL080254 (11415–12107)
19	Hypothetical protein			hyper	0.014	AL049480 (31772–32462)
20	Unknown			hypo	0.002	AL035709 (31497–31991)
21	Unknown			hypo	0.004	AL35708 (3056–3629)
22	Hypothetical protein			hyper	0.004	AC005964 (22084–22788)
23	CHV telomere repeat			hypo	0.018	AP000737 (2048–2690)

Loci of *cmt3* and *met1* are numbered in Figure 2. P values are means of duplicate determinations.

cant locus, Rap2.1, is a member of the APETALA2 family that has a retroelement at its 5′ end [21]. Rap2.1 was identified in a screen for loci that become silenced in allopolyploid hybrids [21]. Thus, seven of the eight CMT3 targets have retrotransposon features, consistent with the idea that an RNAi-based process is involved in target recognition [14]. No comparable transposon bias was seen for MET1 targets that were not also targeted by CMT3 (Table 1, bottom).

Interestingly, five hypomethylated loci in *cmt3* plants were also shown to be affected by *met1* (Figure 2). A non-LTR retrotransposon was hypomethylated in *cmt3* and hypermethylated in *met1*, whereas the other four were hypomethylated in both *cmt3* and *met1* (Table 1). Previous studies of CpXpG methylation in *A. thaliana* have shown that *met1* plants display genome-wide hypomethylation at CpG sites but dense hypermethylation of non-CpG sites at selected loci [6]. Our results extend these findings to a genome-wide sample and identify additional sites acted upon by both MET1 and CMT3.

One feature of our method is flexibility in site specificity that is made possible by the use of different restriction endonucleases for digestion of probe DNA on the same batch of microarrays. In the experiments described here, we utilized MspI, which detects (^{me}C)CGG. Other restriction endonucleases may be used to investigate other specificities. For example, digestion with HpaII would allow detection of C(^{me}C)GG, and digestion with Sau3AI would allow detection of GAT(^{me}C)N. It should be noted that the nonrandom distribution of restriction sites means that occasional loci will not be detected with certain restriction endonucleases. For example, the SUPERMAN locus lies within a >4-kb region that has no CCGG sites (although there are 18 GATC sites). Thus, although the SUPERMAN promoter region is known to be hypomethylated in $cmt3^{-}$ plants, it was not detected on our array using MspI digestion, because all probe fragments were smaller than the largest SUPERMAN MspI fragments. However, the use of different restriction endonucleases in preparing probes should result in more regions of altered methylation being detected.

In conclusion, methylation profiling is robust and reliable, as is illustrated by the fact that our data are in agreement with previous reports. This method should be applicable on a large scale to DNA methylation alterations in a variety of situations. For example, abnormal patterns of methylation have been seen in tumors [3], and so methylation profiling might be used for cancer diagnosis.

Acknowledgments

We thank Luca Comai, Harmit Malik, Paul Talbert, Danielle Vermaak, and Kami Ahmad for helpful discussions. Array construction and hybridizations were performed at the Fred Hutchinson Center Microarray Facility. Funding for this work was provided by the Howard Hughes Medical Institute and a grant from the National Institutes of Health (GM29009).

Received: October 3, 2001
Revised: October 25, 2001
Accepted: October 30, 2001
Published: January 8, 2002

References

1. Martienssen, R., and Colot, V. (2001). DNA methylation and epigenetic inheritance in plants and filamentous fungi. Science *293*, 1070–1074.
2. Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089–1093.
3. Robertson, K.D. (2001). DNA methylation, methyltransferases, and cancer. Oncogene *20*, 3139–3155.
4. Thomassin, H., Oakeley, E.J., and Grange, T. (1999). Identifica-

tion of 5 methylcytosine in complex genomes. Methods *19*, 465–475.
5. McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. (2000). Targeted screening for induced mutations. Nat. Biotechnol. *18*, 455–457.
6. Lindroth, A.M., et al., and Jacobson, S.E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science *292*, 2077–2080.
7. Bartee, L., Malagnac, F., and Bender, J. (2001). Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. Genes and Dev. *15*, 1753–1758.
8. Henikoff, S., and Comai, L. (1998). A DNA methyltransferase homolog with a chromodomain exists in multiple forms in *Arabidopsis*. Genetics *149*, 307–318.
9. Rose, T.M., Schultz, E.R., Henikoff, J.G., Pietrokovski, S., McCallum, C.M., and Henikoff, S. (1998). Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. Nucleic Acids Res. *26*, 1628–1635.
10. Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074–1080.
11. Smothers, J.F., and Henikoff, S. (2000). The HP1 chromo shadow domain binds a consensus peptide pentamer. Curr. Biol. *10*, 27–30.
12. Akhtar, A., Zink, D., and Becker, P.B. (2001). Chromodomains are protein-RNA interaction modules. Nature *407*, 405–409.
13. Wassenegger, M. (2000). RNA-directed DNA methylation. Plant Mol. Biol. *43*, 203–220.
14. Matzke, M., Matzke, A.J.M., and Kooter, J.M. (2001). RNA: guiding gene silencing. Science *293*, 1080–1083.
15. Papa, C.M., Springer, N.M., Muszynski, M.G., Meeley, R., and Kaeppler, S.M. (2001). Maize chromomethylase *Zea* methyltransferase2 is required for CpNpG methylation. Plant Cell *13*, 1919–1928.
16. van Steensel, B., Delrow, J., and Henikoff, S. (2001). Chromatin profiling using targeted DNA adenine methyltransferase. Nat. Genet. *27*, 304–308.
17. Finnegan, E.J., Peacock, W.J., and Dennis, E.S. (1996). Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. Proc. Natl. Acad. Sci. USA *93*, 8449–8454.
18. Kishimoto, N., et al., and Finnegan, E.J. (2001). Site specificity of the *Arabidopsis* METI DNA methyltransferase demonstrated through hypermethylation of the superman locus. Plant Mol. Biol. *46*, 171–183.
19. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.
20. Pelissier, T., Tutois, S., Deragon, J.M., Tourmente, S., Genestia, S., and Picard, G. (2001). Athila, a new retroelement from *Arabidopsis thaliana*. Plant Mol. Biol. *29*, 441–452.
21. Comai, L., Tyagi, A.P., Winter, K., Holmes-Davis, R., Reynolds, S.H., Stevens, Y., and Byers, B. (2000). Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. Plant Cell *12*, 1551–1568.
22. Pollack, J.R., Perou, C.M., Alizadeh, A.A., Eisen, M.B., Pergamenschikov, A., Williams, C.F., Jeffrey, S.S., Botstein, D., and Brown, P.O. (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat. Genet. *23*, 41–46.
23. Rozen S., and Skaletsky, H.J. (1998). Primer3 (http://www.genome.wi.mit.edu/genome_software/other/primer3.html).
24. DeRisi, J.L., Iyer, V., and Brown, P.O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. Science. *278*, 680–686.
25. Baldi, P., and Long, A.D. (2001). A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes. Bioinformatics *17*, 509–519.