NA methylation is a key mechanism used by higher eukaryotes to regulate gene expression. The addition of a methyl group to carbon atom number 5 within cytosine bases in DNA is known to repress the transcription of genes into messenger RNA molecules, thus reducing the production of the proteins coded by these genes. Most methylation occurs at CpG dinucleotides—cytosines that are paired with guanines—and these often cluster together to form CpG islands in the promoter regions of genes. In the late 1990s, it was discovered that transcription was repressed when methyl CpG binding proteins were recruited to methylated CpG islands (Hendrich and Bird, 1998).

Subsequent studies have confirmed that the binding of these proteins throughout the genome is proportional to the density of DNA methylation (Baubec et al., 2013), and have identified additional proteins with a high affinity for methylated CpG sites (reviewed in Defossez and Stancheva, 2011). Moreover, in recent years, other screening approaches based mainly on mass spectrometry have revealed that more proteins bind to methylated DNA than previously thought (Mittler et al., 2009; Bartke et al., 2010; Bartels et al., 2011; Spruijt et al., 2013). Now, in eLife, Heng Zhu and co-workers at the Johns Hopkins University School of Medicine—including Shaohui Hu as first author—use a high-throughput screening method to show that many human transcription factors also interact with genomic DNA sequences containing methylated CpG sites (Hu et al., 2013).

To this end, the Johns Hopkins researchers made use of a published protein microarray consisting of 1,321 transcription factors and 210 co-factors (Hu et al., 2009). Hu et al. incubated the array with 154 distinct human promoter sequences, each of which contained at least one methylated CpG dinucleotide. Their results revealed that 150 (97%) of the 154 methylated human promoter sequences showed specific binding to at least one protein on the microarray. Moreover, of the 1531 proteins, 47 (3%) showed binding to methylated cytosines within the promoters. Most of the proteins bound to methylated DNA in a sequence-dependent manner; however, a minority bound to many different methylated DNA probes, indicating that binding can sometimes occur independent of DNA sequence (Figure 1).

A number of transcription factors, including KLF4—a recently identified methyl-CpG binding
protein (Spruijt et al., 2013)—interacted with methylated sequences that did not resemble their known consensus DNA binding motifs. Using a technique based on electrophoresis, Hu et al. showed that KLF4 binds methylated and non-methylated DNA in a non-competitive manner: this suggests that different domains of the protein may be responsible for each type of binding, which they duly confirmed using molecular modeling and mutagenesis studies.

The Johns Hopkins researchers then mined published ChIP-seq data from stem cells to identify the target DNA sequences of KLF4, and compared these with data on genome-wide DNA methylation. Strikingly, KLF4 binding appears to be bimodal in nature throughout the genome, with 38% of KLF4 binding sites showing less than 20% methylation, and 48% showing methylation levels over 80%. Finally, Hu et al. used ChIP-bisulfite sequencing, which makes it possible to determine the methylation status of each cytosine within a target DNA sequence, to confirm that KLF4 also binds to both methylated and non-methylated DNA in vivo.

Hu et al. only profiled a small fraction of the complete human methylome for interactions with transcription factors; further proteins capable of binding genomic methyl CpG sequences surely await identification. The same holds true for interactions with methylated non-CpG sequences such as methyl-CpA (cytosine adjacent to adenine), which are fairly abundant in embryonic stem cells (Ramsahoye et al., 2000; Lister et al., 2009). To determine the physiological relevance of these interactions, it will be important to deduce the affinity with which proteins bind these sequences compared to their known targets; initial experiments along these lines are presented in the current eLife paper. Furthermore, recent evidence suggests that non-methylated CpG islands recruit activator proteins, many of which contain a CXXC motif (reviewed in Long et al., 2013). The transcription factor microarray approach used by the Johns Hopkins team, combined with quantitative mass spectrometry-based technology (Spruijt et al., 2013), could thus be used to identify the complete cellular complement of proteins that bind specifically to non-methylated CpG islands.

Finally, this study and other recently published papers force us to reconsider the mechanism(s) via which CpG methylation regulates transcription. Although DNA methylation is generally considered to be a repressive epigenetic modification, experiments presented by Hu et al. suggest that in some cases, methylation of a given promoter sequence can result in activation of transcription. Moreover, other work has revealed a temporal uncoupling of DNA methylation and transcriptional repression during Xenopus embryogenesis (Bogdanovic et al., 2011). Further experiments are therefore required to determine whether the functional readout of CpG methylation is affected by the repertoire and abundance of different DNA methylation ‘readers’ acting at any given time in a cell or a developing organism.

Michiel Vermeulen is at the Department of Molecular Cancer Research, University Medical Center Utrecht, Utrecht, Netherlands
m.vermeulen-3@umcutrecht.nl

Competing interests: The author declares that no competing interests exist.

Published 03 September 2013

References
Bartels SJ, Spruijt CG, Brinkman AB, Jansen PW, Vermeulen M, Stunnenberg HG. 2011. A SILAC-based screen for Methyl-CpG binding proteins identifies RBP-J as a DNA methylation and sequence-specific
binding protein. PLOS ONE 6:e25884. doi: 10.1371/journal.pone.0025884.
Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. 2010. Nucleosome-interacting proteins regulated by DNA and Histone methylation. Cell 143:470–84. doi: 10.1016/j.cell.2010.10.012.
Baubec T, Ivanek R, Lienert F, Schubeler D. 2013. Methylation-dependent and -independent genomic targeting principles of the MBD protein family. Cell 153:480–92. doi: 10.1016/j.cell.2013.03.011.
Bogdanovic O, Long SW, van Heeringen SJ, Brinkman AB, Gomez-Skarmeta JL, Stunnenberg HG, et al. 2011. Temporal uncoupling of the DNA methylome and transcriptional repression during embryogenesis. Genome Res 21:1313–27. doi: 10.1101/gr.114843.110.
Defossez PA, Stancheva I. 2011. Biological functions of methyl-CpG-binding proteins. Prog Mol Biol Trans Sci 101:377–98. doi: 10.1016/B978-0-12-387685-0.00012-3.
Hendrich B, Bird A. 1998. Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 18:6538–47.
Hu S, Wan J, Su Y, Song Q, Zeng Y, Nguyen HN, et al. 2013. DNA methylation presents distinct binding sites for human transcription factors. eLife 2:e00726. doi: 10.7554/eLife.00726.
Hu S, Xie Z, Onishi A, Yu X, Jiang L, Lin J, et al. 2009. Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. Cell 139:610–22. doi: 10.1016/j.cell.2009.08.037.
Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462:315–22. doi: 10.1038/nature08514.
Long HK, Blackledge NP, Klose RJ. 2013. ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. Biochem Soc trans 41:727–40. doi: 10.1042/BST20130028.
Mittler G, Butter F, Mann M. 2009. A SILAC-based DNA protein interaction screen that identifies candidate binding proteins to functional DNA elements. Genome Res 19:284–93. doi: 10.1101/gr.081711.108.
Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. 2000. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc Natl Acad Sci USA 97:5237–42. doi: 10.1073/pnas.97.10.5237.
Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C, et al. 2013. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. Cell 152:1146–59. doi: 10.1016/j.cell.2013.02.004.