Cytosine modifications modulate the chromatin architecture of transcriptional enhancers

Elise A. Mahé, Thierry Madigou, Aurélien Sérandour, Maud Bizot, Stéphane Avner, Frédéric Chalmel, Gaëlle Palierne, Raphaël Métivier, Gilles Salbert

Cytosine modifications modulate the chromatin architecture of transcriptional enhancers

Elise A. Mahé,[1,2] Thierry Madigou,[1,2] Aurélien A. Sérandour,[3,5] Maud Bizot,[1,2] Stéphane Avner,[1,2] Frédéric Chalmel,[4] Gaëlle Palierne,[1,2] Raphaël Métivier,[1,2] and Gilles Salbert[1,2]

[1]*CNRS UMR6290, Equipe SP@RTE, Institut de Génétique et Développement de Rennes, Campus de Beaulieu, 35042 Rennes Cedex, France;* [2]*Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France;* [3]*EMBL, Heidelberg 69117, Germany;* [4]*Inserm U1085-IRSET, Université de Rennes 1, F-35042 Rennes, France*

Epigenetic mechanisms are believed to play key roles in the establishment of cell-specific transcription programs. Accordingly, the modified bases 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) have been observed in DNA of genomic regulatory regions such as enhancers, and oxidation of 5mC into 5hmC by Ten-eleven translocation (TET) proteins correlates with enhancer activation. However, the functional relationship between cytosine modifications and the chromatin architecture of enhancers remains elusive. To gain insights into their function, 5mC and 5hmC levels were perturbed by inhibiting DNA methyltransferases and TETs during differentiation of mouse embryonal carcinoma cells into neural progenitors, and chromatin characteristics of enhancers bound by the pioneer transcription factors FOXA1, MEIS1, and PBX1 were interrogated. In a large fraction of the tested enhancers, inhibition of DNA methylation was associated with a significant increase in monomethylation of H3K4, a characteristic mark of enhancer priming. In addition, at some specific enhancers, 5mC oxidation by TETs facilitated chromatin opening, a process that may stabilize MEIS1 binding to these genomic regions.

[Supplemental material is available for this article.]

Cell differentiation relies on the engagement of short genomic regions called transcriptional enhancers, acting at a distance from the transcription start site of their target genes. Enhancers are endowed with the capacity to bind numerous cell-specific and general transcription factors (TFs) in a temporal sequence which is partly dictated by the timely controlled expression of TFs and/or their regulated binding to chromatin through signaling pathways (Spitz and Furlong 2012; Heinz et al. 2015). In their inactive state, enhancers can be targeted by so-called "pioneer factors" (PFs) which bind condensed chromatin and favor transition of enhancers to a primed state, in turn facilitating binding of additional TFs and further activation (Iwafuchi-Doi and Zaret 2014). Several TFs, including FOXA1, SPI1, PBX1, CEBPB, the AP-1 complex, and MEIS1, have been qualified as PFs and play major roles in pluripotent cell differentiation during development and in adulthood (Heidt et al. 2007; Lupien et al. 2008; Ghisletti et al. 2010; Biddie et al. 2011; Magnani et al. 2011; Siersbæk et al. 2011). PF engagement during lineage commitment is thought to provoke changes in enhancer chromatin architecture. In agreement with this view, binding of SPI1 to enhancer chromatin during monocyte-to-osteoclast differentiation impacts cytosine methylation (5mC) at CpGs through the recruitment either of DNA methyltransferases (DNMTs) at active enhancers, or of Ten-eleven translocation (TET) enzymes, which catalyze 5mC oxidation into 5-hydroxymethylcytosine (5hmC) (Tahiliani et al. 2009), at inactive enhancers (de la Rica et al. 2013). 5hmC is a catalytic intermediate in a pathway of active DNA demethylation generating further oxidative derivatives (5-formylcytosine [5fC] and 5-carboxylcytosine [5caC]) subsequently replaced by cytosines through action of the base excision repair (BER) machinery (He et al. 2011; Maiti and Drohat 2011). However, due to the relative stability of 5hmC and its ability to bind specific proteins (Iurlaro et al. 2013; Spruijt et al. 2013; Bachman et al. 2014), 5hmCpGs may also act as signaling modules just as 5mCpGs do (Bird 2011). In this respect, high levels of 5hmC were found at selected enhancers prior to binding of the 3-aa loop extension-homeodomain (TALE-HD) proteins MEIS1 and PBX1 during differentiation of mouse P19 embryonal carcinoma cells (ECCs) into neural progenitors (Sérandour et al. 2012), suggesting that DNA hydroxymethylation at enhancers could favor MEIS1 and PBX1 engagement. Moreover, consistent with data showing that TF binding to regulatory regions lowers their CpG methylation levels (Stadler et al. 2011), we previously observed that FOXA1 recruitment to enhancers is associated with DNA demethylation during P19 ECC differentiation (Sérandour et al. 2011). Nonetheless, it remains unclear which steps of enhancer priming and activation are regulated through DNA demethylation, and in particular, if 5hmC plays a role per se in these processes. In order to gain insights into the role of cytosine modifications at PF-bound enhancers, we have here analyzed the impact of DNA methylation/demethylation processes on the chromatin structure of FOXA1- and TALE-HD-bound enhancers.

[5]Present address: INSERM U1232, CRCINA, Hôpital Nord CHU Nantes, 44800 Saint-Herblain, France
Corresponding author: gilles.salbert@univ-rennes1.fr
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.211466.116.

Results

P19 ECCs as a model-system to study enhancer activation by pioneer factors during NPC specification

Although different from the embryonic stem cell (ESC) model classically used to analyze neural progenitor cell (NPC) commitment over a 6-d process, P19 ECCs, after treatment with retinoic acid (RA), have the ability to generate NPCs showing a hydroxymethylome similar to the one observed in ESC-derived NPCs, in a rapid (2 d) and robust fashion (Supplemental Fig. S1A–C). Importantly, 82.5% of the regions gaining 5hmC during P19 cell differentiation overlap with hydroxymethylated regions from ESC-derived NPCs. Furthermore, 5hmC-positive regions from ECC-derived NPCs identify 28 out of the 48 enhancers, as demonstrated by Visel et al. (2007) to be active in neural tissues from E11.5 dpc mouse embryos (Supplemental Fig. S1D,E; http://enhancer.lbl.gov/). Similarly, 5hmC-positive regions from ESC-derived NPCs identify 22 out of these 48 enhancers. Moreover, it has been demonstrated that P19 ECC differentiation into NPCs relies on the induction of FOXA1, PBX1, and MEIS1, three transcription factors which are also induced by RA in mouse ESCs (Kashyap et al. 2013), and is enhanced by their overexpression (Qin et al. 2004; Tan et al. 2010; Dong et al. 2012; Yamada et al. 2013). These observations indicate (1) that the hydroxymethylome of ECC-derived NPCs is not aberrant, and (2) that ECC-derived NPCs constitute a highly relevant cellular model to study the activation of enhancers related to developmental programs controlled by pioneer factors.

In this context, we previously reported a major reconfiguration of the hydroxymethylome of P19 ECCs upon RA treatment (Sérandour et al. 2012), which we show here to be correlated with loss of pluripotency factors NANOG and POU5F1 and activation of the pioneer factor genes *FoxA1* and *Meis1* (Fig. 1A,B). MEIS1 is known to associate with PBX1 to form a prevalent TALE-HD complex (Chang et al. 1997) which is likely to control P19 differentiation (Qin et al. 2004; Yamada et al. 2013). Accordingly, data obtained by chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) indicate that MEIS1 and PBX1 may directly control the expression of NPC marker genes like *Tshz1*, *Pax6*, and *Irx3* (Fig. 1C). In addition to the role of MEIS1/PBX1 dimers, FOXA1 might also control P19 cell fate decision. Indeed, FOXA1 is rapidly induced in response to RA (Fig. 1B) and is known to stimulate NPC commitment in P19 cells (Tan et al. 2010; Dong et al. 2012). To investigate the relationship between FOXA1 and TALE-HDs, we mapped FOXA1 binding sites in P19-derived NPCs. A significant fraction of FOXA1 sites were also enriched in MEIS1 and PBX1 (Fig. 1D–G), suggesting an interplay between these pioneer factors in mediating chromatin activation at specific sites such as within the *HoxB* cluster during NPC lineage commitment.

Pioneer factor-bound enhancers are enriched in 5hmC

Primed enhancers are characterized by the presence of histone H3 monomethylated on lysine 4 (H3K4me1) and the absence of acetylated H3K27 (H3K27ac), whereas active enhancers present both marks (for review, see Calo and Wysocka 2013). We previously observed that, during RA-induced differentiation of P19 ECCs into NPCs, differentially hydroxymethylated regions (DhMRs) are detected at enhancers either being silenced for down-DhMRs or undergoing activation for up-DhMRs (Sérandour et al. 2012). In order to obtain a more complete view of these dynamic processes at the full sets of up- and down-DhMRs, we first analyzed the temporal variations of H3K4me1 and H3K27ac histone modifications genome-wide by ChIP-seq, together with those of 5hmC and 5mC by hydroxymethylated or methylated DNA immunoprecipitation (hMeDIP-seq and MeDIP-seq, respectively) during RA-induced differentiation of P19 ECCs. In addition, chromatin opening was assessed by formaldehyde assisted isolation of regulatory elements (FAIRE) (Giresi et al. 2007). For both types of DhMRs, the levels of H3K4me1 and chromatin opening followed the gradual changes observed for 5hmC, whereas 5mC levels showed only mild fluctuations over time (Fig. 2A for up-DhMRs; Supplemental Fig. S2A for down-DhMRs). When comparing called peaks, 51.6% and 34.5% of the hydroxymethylated regions overlapped with H3K4me1 and FAIRE peaks, respectively, whereas 57% overlapped with H3K27ac peaks. These observations suggest a strong correlation between 5mC oxidation and both enhancer priming and activation during ECC differentiation. We next analyzed FOXA1, MEIS1, and PBX1 engagement at DhMRs and observed that, consistent with an enrichment in PF binding motifs at enhancers gaining 5hmC during differentiation (Sérandour et al. 2012), PF binding peaked at up-DhMRs (Fig. 2B) but not at down-DhMRs (Supplemental Fig. S2B), and heat maps clearly identified PF-bound regions as being enriched in 5hmC (Fig. 2C). In addition, these three PFs shared a significant fraction of their binding sites (Fig. 1D–G), indicating that FOXA1 and TALE-HD proteins target a common subset of enhancers. Collectively, these data suggest that the recruitment of these critical regulators at enhancers could influence or be sensitive to cytosine modifications.

Cytosine modifications increase TALE-HD binding in vitro

TALE-HDs bind DNA through the formation of MEIS1/PBX1 heterodimers or through dimerization with HOX proteins (for review, see Moens and Selleri 2006). ChIP-seq of MEIS1 and PBX1 in RA-treated P19 ECCs demonstrated that most MEIS1-bound regions are also interacting with PBX1 (Supplemental Fig. S3A). Accordingly, analysis of the proteins associated with PBX1-bound chromatin in RA-differentiated P19 cells by ChIP coupled to mass spectrometry (ChIP-MS) revealed the presence of several HOX and TALE-HD proteins, including MEIS1 (Supplemental Fig. S3B; Supplemental Table 1). In accordance with a putative pioneer function of TALE-HDs, both MEIS1 and PBX1 from NPC nuclear extracts can be efficiently captured by an artificial probe containing 5mCpGs in a $TG(ACG)_4$ sequence (Spruijt et al. 2013). In addition, MEIS1 and PBX1 engage up-DhMRs subsequently to 5hmC occurrence during RA-induced P19 cell differentiation (Sérandour et al. 2012), suggesting that they could sense cytosine modifications rather than induce them. To address the possible influence of cytosine modifications on the binding of TALE-HD proteins to their recognition sites in vitro, pull-down analyses of cell extracts from differentiated P19 cells were then performed using probes containing the TGATTTACG PBX1/HOXA9 motif (Shen et al. 1997) in which the cytosine from each strand of the CpG dinucleotide was differentially modified (either methylated or hydroxymethylated), as well as probes containing the MEIS1/PBX1 motif TGATTGACAG (Chang et al. 1997) modified at underlined positions. For each probe, the specific binding of MEIS1/2 and PBX1/2/3 was first verified by supershift experiments (Supplemental Fig. S3C). Pull-down assays revealed that MEIS1 and PBX1/2/3/4 were better retained by the PBX1/HOXA9 probe containing specific combinations of modified cytosines (such as hemi-methylated, fully hydroxy- or methylated, or both methylated and hydroxymethylated probes) compared to the probe containing

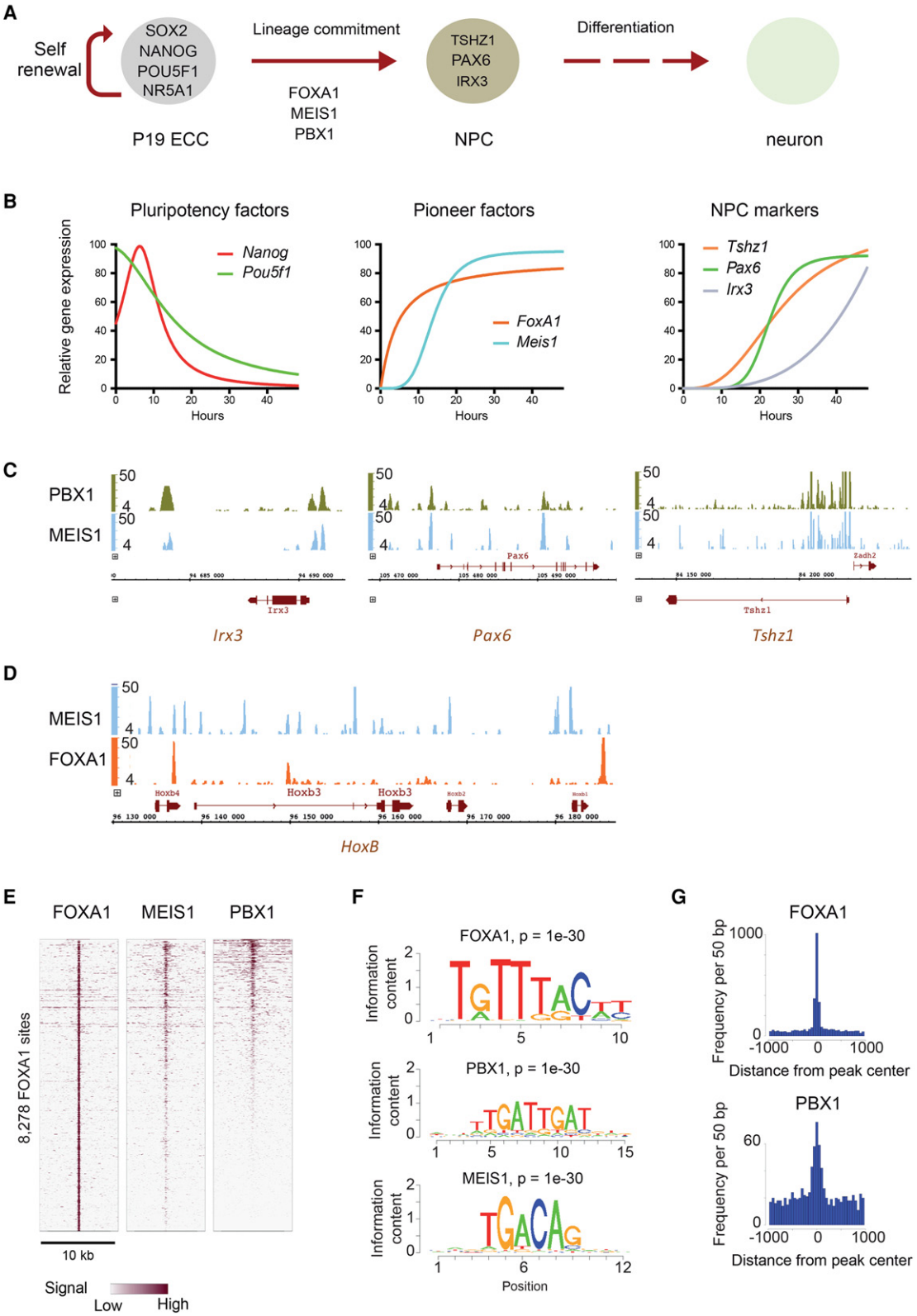

Figure 1. Pioneer factors and NPC differentiation. (*A*) Pluripotent P19 EC cells derive from a 6.5-dpc mouse embryo and can differentiate into neural progenitor cells (NPCs) when treated with retinoic acid (RA). Pluripotency factors expressed in these cells include POU5F1, NANOG, SOX2, and NR5A1. Under the influence of RA, FOXA1, MEIS1, and PBX1 are induced and control the commitment of P19 EC cells to NPCs which express marker genes like *Tshz1*, *Pax6*, and *Irx3*. (*B*) RT-qPCR analysis of marker gene expression during 48 h of NPC differentiation after addition of RA. Curves were fitted with the GraphPad Prism software. (*C*) NPC marker genes are direct targets of TALE-HDs. Integrated Genome Browser (IGB) screenshots showing MEIS1 and PBX1 ChIP-seq signals at *Irx3*, *Pax6*, and *Tshz1* loci in RA-treated cells. (*D*) IGB screenshot of FOXA1 and MEIS1 ChIP-seq signals from NPCs at the *HoxB* cluster, showing that FOXA1 and MEIS1 share common binding sites. (*E*) Heat maps showing FOXA1, MEIS1, and PBX1 ChIP-seq signals at 8278 FOXA1 sites in P19 cell-derived NPCs. (*F*) Logos of motifs enriched at FOXA1 binding sites. (*G*) Distribution of FOXA1 and PBX1 motifs in 50-bp windows centered on FOXA1 binding sites.

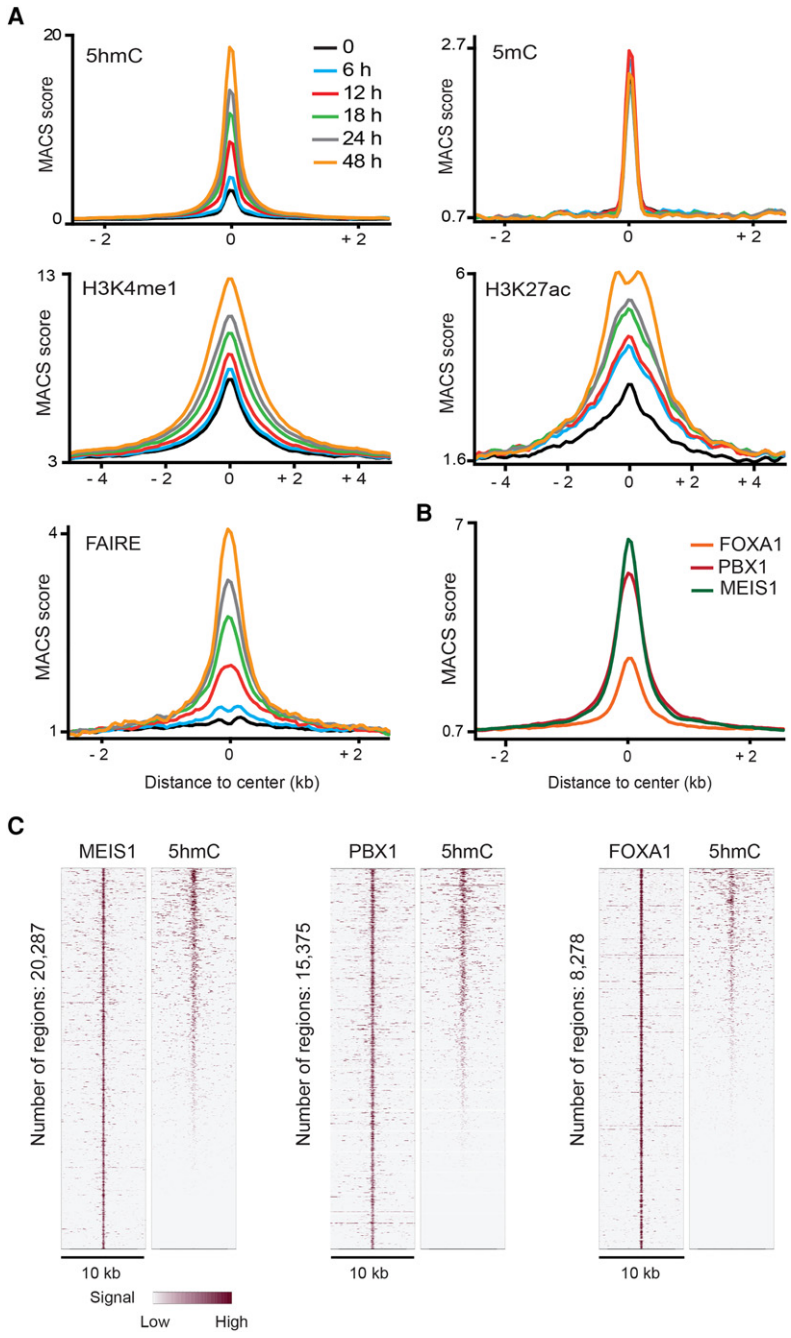

Figure 2. De novo hydroxymethylated enhancers bind pioneer factors driving NPC commitment. (*A*) Average kinetic profiles of 5hmC, 5mC, H3K4me1, H3K27ac, and chromatin opening (FAIRE) at 20,492 regions gaining 5hmC (up-DhMRs) at 48 h of P19 ECC differentiation. (*B*) Average profiles of FOXA1, PBX1, and MEIS1 binding at up-DhMRs at 48 h of differentiation. (*C*) Heat map representations of 5hmC signal at MEIS1, PBX1, and FOXA1 enriched regions in 48 h RA-treated P19 cells.

unmodified cytosines (Fig. 3A; Supplemental Fig. S3D). Likewise, MEIS1 and PBX1/2/3/4 proteins were better retained, although less markedly, by PBX1/MEIS1 probes containing modified cytosines (Fig. 3A). These results suggest that MEIS1 and PBX1/2/3/4 bind preferentially to sequences containing modified cytosines in vitro. This is consistent with data from Spruijt et al. (2013) indicating that MEIS1 and PBX1 are preferentially retained by fully methylated CpGs and extends their findings by showing that TALE-HDs also prefer CpGs where 5mC is combined with 5hmC on the other strand rather than to an unmodified cytosine.

We next evaluated the presence of 5hmCpGs at 20,287 MEIS1 binding sites in vivo by using single-CpG resolution data obtained by SCL-exo-seq in RA-treated P19 cells (Sérandour et al. 2016). This method allows the genome-wide identification of hydroxymethylated CpGs by a combination of chemical labeling (glucosylation followed by biotinylation), immobilization on streptavidin beads, and exonuclease digestion of the bead-trapped DNA molecules (Sérandour et al. 2016). As shown in Figure 3B, 5hmCpGs were enriched at bulk MEIS1 binding sites (see also Supplemental Fig. S4A). A cluster of 958 regions with centered 5hmCpGs was next recovered using the TMEV software (Saeed et al. 2003) in order to only consider regions with 5hmCpGs located within a few base pairs of MEIS1 binding sites. Sites from this cluster had, on average, a higher CpG density compared to bulk MEIS1 sites (Supplemental Fig. S4B). Heat maps, as well as profiles of MEIS1 ChIP-seq and SCL-exo signals (Fig. 3B–D), showed that 5hmCpGs accumulated within −15/+15 bp of the MEIS1 ChIP-seq peak summits (defined as 0 in the heat maps and the average profiles). The 958 sequences were next analyzed and a consensus sequence was defined with the CLC Sequence Viewer 6 software (Fig. 3D). Whereas 95 regions contained the TGACAG consensus MEIS1 site in close vicinity (within 2 to 20 bp) to 5hmCpGs, 29 regions did not include the consensus site but showed a TGACGG site, 24 of which had a hydroxymethylated CpG detected by SCL-exo (see examples in Supplemental Fig. S4C). De novo motif discovery from the 958 sequences was then implemented with both SeqPos (Fig. 3E; Supplemental Fig. S4D) and SEME (Supplemental Fig. S4E). SeqPos recovered a TGACG motif with a high probability, whereas the probability to find the MEIS1 motif was lower (Fig. 3E). The SEME tool ranked first the motif TGAC/TGGATGG (Supplemental Fig. S4E), which closely resembles the PBX1 motif TGATTGAT (Fig. 1F). In addition, despite being enriched in MEIS1 and PBX1 motifs, as shown by de novo motif discovery (Supplemental Fig. S4F), MEIS1 binding regions devoid of CpGs were characterized by a lower average MEIS1 ChIP-seq signal than the cluster of 958 MEIS1 sites with centered 5hmCpGs (Fig. 3F). Finally, the ability of MEIS1 to interact with a TGACGG sequence was assessed by competition pull-down experiments in which a fivefold excess of a

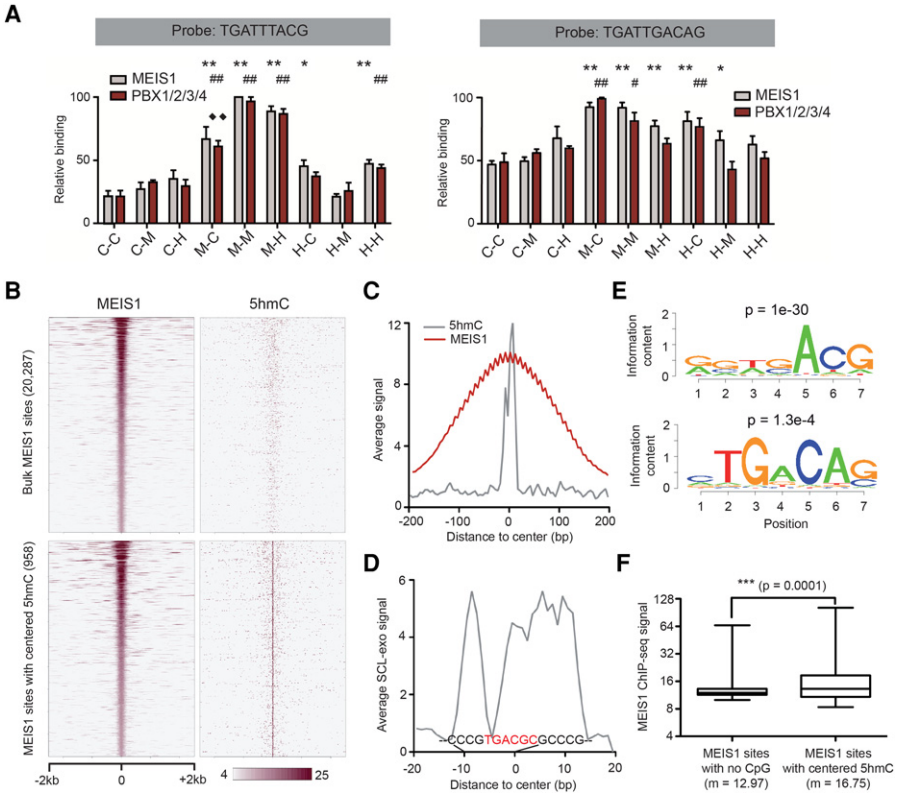

Figure 3. DNA binding by TALE-HD proteins is sensitive to cytosine modifications. (*A*) MEIS1 and PBX1/2/3/4 in vitro interaction with differentially modified synthetic PBX1/HOXA9 (*left* panel) and MEIS1/PBX1 (*right* panel) probes. Results are expressed according to the strongest binding equivalent to 100 (C: unmodified cytosine; M: methylated cytosine; H: hydroxymethylated cytosine; first letter: upper strand; second letter: lower strand). Statistical significance of differential binding to modified probes compared to the unmodified probe was tested with a Mann-Whitney test ($n = 5$; MEIS1 protein: [*] $P < 0.05$, [**] $P < 0.01$; PBX1/2/3/4: [#] $P < 0.05$, [##] $P < 0.01$). For PBX1 binding, ◆◆ corresponds to $P < 0.01$ between M-C and M-H conditions. (*B*) Heat maps of MEIS1 ChIP-seq and SCL-exo (5hmC) signals at 20,287 bulk MEIS1 binding sites (*upper* panels) and 958 MEIS1 sites showing a centered SCL-exo signal (*lower* panels) in P19 cell-derived NPCs. (*C*) Average MEIS1 ChIP-seq (red curve) and SCL-exo (gray curve) signals at the 958 MEIS1 binding sites showing a centered SCL-exo signal in P19 cell-derived NPCs. Note that oscillations in the MEIS1 profile are due to the lower resolution of the MEIS1 ChIP-seq .wig file compared to the SCL-exo .wig file. (*D*) Average SCL-exo signal over 40 bp for the cluster of 958 MEIS1 sites depicted in *B*. The consensus sequence determined from the aligned 958 sequences with the software CLC Sequence Viewer 6 is indicated on the graph. A motif related to the TGACAG consensus MEIS1 motif has been highlighted in red font. (*E*) Examples of de novo motifs found enriched (*P*-value is indicated above the logos) by the SeqPos motif tool from Cistrome. (*F*) Box plot analysis of MEIS1 ChIP-seq signal for MEIS1 sites with no CpGs within the ChIP-seq peaks ($n = 3628$) and the 958 MEIS1 sites with centered SCL-exo signal (m = mean ChIP-seq signal).

double-stranded oligonucleotide containing the TGACGG sequence was added to the PBX1/MEIS1 probe. Results indicated that indeed MEIS1 binds to the TGACGG site and that the fully methylated and fully hydroxymethylated sequences tend to show a stronger interaction with MEIS1 (Supplemental Fig. S4G). Collectively, these data suggest that MEIS1 can bind in vivo to TGACGG sequences including 5hmCpGs, although it might concern only a fraction of its target regions.

Cytosine modifications are required for the correct timing of neural progenitor differentiation

We next perturbed the DNA methylation and demethylation processes in P19 cells by using subtoxic concentrations of either 5-aza-2′-deoxycytidine (5-azadC), a hypomethylating agent which forms adducts with DNMTs (Jüttermann et al. 1994), or dimethyloxalylglycine (DMOG), a cell-penetrating derivative of N-Oxalylglycine (NOG) that inhibits Fe^{2+}/2-OG dioxygenases such as TET enzymes (Rose et al. 2011). As assessed by dot blot analysis, 5-azadC treatment reduced global 5mC levels in both undifferentiated and differentiated P19 cells and also limited 5hmC formation in RA-treated cells (Supplemental Fig. S5A). Treatment of P19 ECCs with DMOG did not impact global 5mC levels but inhibited significantly the increase in 5hmC observed during RA-induced differentiation (Supplemental Fig. S5B,C). We next evaluated the impact of the perturbation of 5mC and 5hmC levels on the production of mRNAs encoding pluripotency factors (POU5F1 and NANOG), neural progenitor markers (TSHZ1, PAX6, and IRX3), and cell fate determination TFs (FOXA1, MEIS1, and PBX1) during RA-induced P19 cell differentiation (Supplemental Fig. S6A). Although RA-mediated changes in expression of pluripotency factors and *Meis1* were not dramatically affected by 5-azadC and DMOG treatments, *Foxa1* and *Pbx1* mRNA levels did show some sensitivity to the drugs. However, such sensitivity was not observed at the protein level (Supplemental Fig. S6B). Notably, induction of *Pax6* and *Irx3* expression by RA was significantly delayed by 5-azadC and DMOG treatments (Supplemental Fig. S6A), suggesting that both DNMT and TET inhibition, although not affecting the repression of pluripotency factor genes, delayed differentiation of neural progenitors. This is consistent with the observed delayed induction of a subset of NPC marker genes during neural differentiation of *Tet2*$^{-/-}$ mouse ESCs (Hon et al. 2014). Since both RA-mediated loss of pluripotency and PF induction were not affected by the drugs while neural differentiation was delayed, these results indicate that our experimental conditions are suitable for the study of the role of DNA modifications at PF-bound enhancers.

Cytosine modifications control enhancer priming

In order to determine the role of cytosine modifications in enhancer priming and activation, we next analyzed the effects of 5-azadC or DMOG treatments on the levels of 5mC, 5hmC, H3K4me1, H3K27ac, and chromatin opening (FAIRE), as well as on MEIS1 and FOXA1 recruitment, at enhancers selected from genome-wide DIP- and ChIP-seq data according to (1) gain of 5hmC during differentiation, (2) enrichment in H3K4me1 and H3K27ac, and (3) recruitment of MEIS1 and/or FOXA1. Hence, 15 enhancers were selected, with five bound by MEIS1 but not by FOXA1 (enhancers M1 to M5), five recruiting both MEIS1 and FOXA1 (enhancers MF1

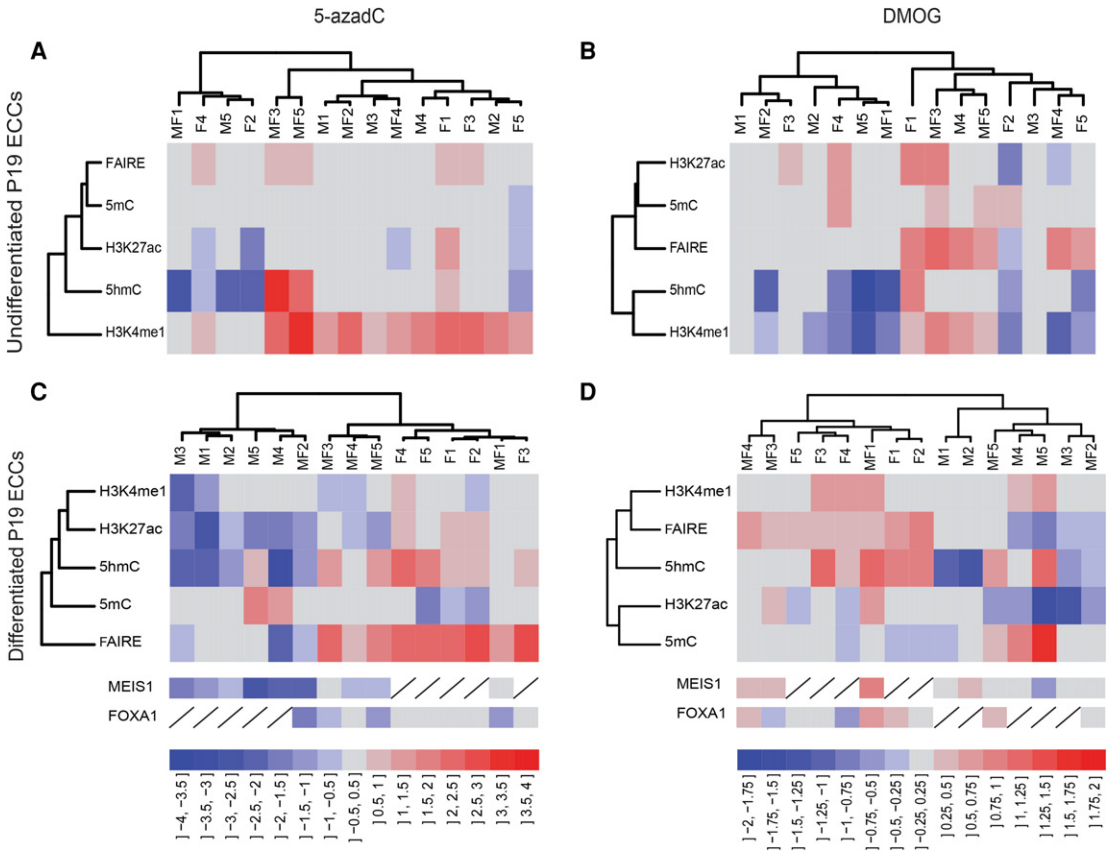

Figure 4. Cytosine modifications regulate enhancer priming. Dendrograms and heat maps from hierarchical clustering of enhancer features (*left* dendrogram) and the 15 selected enhancers studied (*top* dendrogram). (*A*,*B*): undifferentiated P19 ECCs. (*C*,*D*): differentiated neural progenitors. The color scale bars at the bottom of each column indicate $\log_2$-transformed fold-change between control cells and cells treated with the inhibitor, with a scale of 0.5 for 5-azadC (*A*,*C*) and 0.25 for DMOG (*B*,*D*). Variations in MEIS1 and FOXA1 recruitment are represented below heat maps in panels *C* and *D* but were not used for hierarchical clustering.

to MF5), and five bound by FOXA1 but not by MEIS1 (enhancers F1 to F5). As a control, we selected an enhancer enriched in H3K4me1 and H3K27ac and bound by both FOXA1 and MEIS1 but having no CpG in 400 bp flanking FOXA1 and MEIS1 ChIP-seq peak summits (enhancer "No CpG"). Notably, the chromatin features of this control enhancer were not affected by 5-azadC or DMOG treatments (Supplemental Fig. S7A,B). Since DMOG could also inhibit histone demethylases of the Jumonji family (Rose et al. 2011), we controlled whether DMOG treatment increased H3K27 methylation, potentially leading to inhibition of enhancers, and found that H3K27me3 levels did not increase in DMOG-treated cells for any of the selected enhancers (Supplemental Fig. S7C). These data indicate that, in our conditions, the observed effects of the DMOG treatment mainly reflected TET inhibition.

The chromatin features of these 15 enhancers were then analyzed by (h)MeDIP-, ChIP- and FAIRE-qPCR experiments. Results of these experiments are depicted in the histograms of Supplemental Figures. S8 and S9 and were analyzed by principal component analysis (PCA) with data expressed as fold change between control and inhibitor-treated samples from either undifferentiated or differentiated P19 cells (detailed results for PCA and hierarchical clustering on principal components [HCPC] are given in Supplemental Figs. S10, S11). Due to the low sensitivity of the MeDIP technique to detect slight changes in 5mC levels, we also interrogated the cytosine modification status of CpGs included in CCGG MspI sites whenever present in the selected enhancers, taking advantage of the fact that MspI cannot cut a C(5hm)CGG site after glucosylation of the 5hmC (eight out of 15 enhancers) (Supplemental Fig. S12). Results obtained with undifferentiated P19 cells showed that inhibition of both DNMTs (Fig. 4A) and TETs (Fig. 4B) mainly affected 5hmC and H3K4me1 levels. Remarkably, under DNMT inhibition, HCPC identified two clusters of enhancers: one characterized by a decrease in 5hmC at sites having significant levels of 5hmC in control condition (Supplemental Fig. S13: MF1, F4, M5, and F2), consistent with a decrease in the substrate for TET-mediated oxidation; and a second one grouping regions with increased levels of H3K4me1, indicating that 5mC prevents monomethylation of H3K4 (Fig. 4A). In addition, enhancers showing a decrease in 5hmC in the presence of DMOG also showed a decrease in H3K4me1 levels in undifferentiated cells (Fig. 4B). This is consistent with data showing that, in MCF-7 cells, binding of the H3K4 methyltransferase KMT2A (also known as MLL1) to the *TFF1* enhancer depends on DNA hypomethylation (Jeong et al. 2014) and that 5mC repels KMT2A in vitro (Spruijt et al. 2013).

Importantly, data obtained with differentiated P19 cells revealed a positive correlation between the levels of 5hmC and chromatin opening (Fig. 4C,D). Since inhibition of 5mC formation by 5-azadC did not impact chromatin opening (FAIRE) in undifferentiated cells despite an increase in H3K4me1 levels (Fig. 4A),

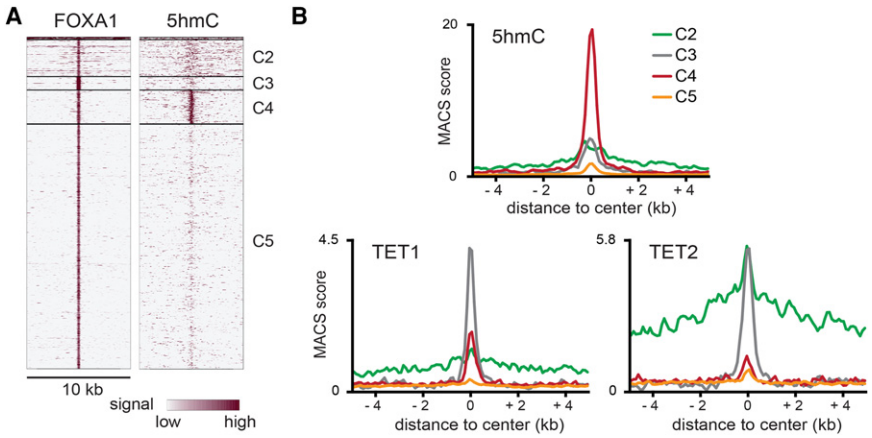

Figure 5. High levels of 5hmC correlate with low levels of TETs at FOXA1-bound enhancers. (*A*) Heat maps of FOXA1 and 5hmC signals in clusterized FOXA1 binding regions outside CpG islands. (*B*) Average profiles of 5hmC, TET1, and TET2 recruitment at the C2, C3, C4, and C5 clusters shown in *A*.

chromatin opening could be directly triggered by the presence of 5hmC itself and not by the loss of 5mC. This hypothesis is supported by data indicating that nucleosomes formed in vitro on a hydroxymethylated DNA template are unstable due to a lower interaction of DNA with H2A-H2B dimers (Mendonca et al. 2014). In order to compare the data obtained with DMOG, a compound affecting the enzymatic activity of TETs, and the effects of a reduction in TET proteins, P19 cells were transfected with siRNA targeting all three *Tets* (Supplemental Fig. S14). Since the siRNA targeting *Tet3* also reduced *Tet2* mRNA in RA-treated cells (Supplemental Fig. S14), 5hmC levels and chromatin opening were assessed in *Tet1* and *Tet2* siRNA conditions only (Supplemental Fig. S15). Although *Tet* mRNA levels were only partially reduced by the siRNAs, data indicated that, depending on the enhancer, TET depletion either reduced or increased 5hmC levels and the observed changes paralleled those induced by DMOG for six out of eight tested enhancers. As in DMOG-treated cells, the FAIRE signal at M1 and M2 enhancers did not vary, although 5hmC was reduced in siRNA conditions. Other enhancers showing either a decrease (M5) or an increase (MF1, F1) in FAIRE signal under DMOG treatment were also affected in the siRNA experiment. Indeed, M5 showed a significant reduction of the FAIRE signal in *Tet2* siRNA conditions in the presence of RA, whereas an increase in FAIRE signal was evidenced at MF1 and F1 in undifferentiated cells depleted in TET1. Hence, these siRNA experiments indicate that, at specific enhancers, an elevation in 5hmC levels can correlate with an elevation in FAIRE signal.

Interestingly, TET inhibition by DMOG (Fig. 4D) or knockdown by siRNAs (Supplemental Fig. S15) led either to a decrease or to an increase in 5hmC levels, depending on the analyzed enhancer, a phenomenon which has also been observed in human ECCs transfected with *TET* siRNAs (Putiri et al. 2014). Such an increase in 5hmC levels when TETs are inhibited could, in principle, occur at enhancers where 5hmC superoxidation in 5fC and 5caC takes place at a high rate, most probably because TET enzymes have a higher residency time at these particular enhancers. Based on their increased ability to gain 5hmC in the presence of DMOG, enhancers bound by FOXA1 seemed to be more prone to superoxidation than MEIS1-only enhancers (Fig. 4D). In agreement with this hypothesis, clustering according to FOXA1 and 5hmC levels of FOXA1-bound enhancers genome-wide revealed that 5hmC and TET levels were inversely correlated for three clusters out of four (Fig. 5), suggesting that high TET residency may indeed be associated with high 5hmC superoxidation. Hence, FOXA1 binding to enhancers might favor the engagement of TET enzymes involved in 5hmC oxidation. Accordingly, it was recently demonstrated that FOXA1 interacts with TET2 in prostate cancer cells (Takayama et al. 2015) and that 5hmC oxidation in human ECCs is mainly dependent on TET2 (Putiri et al. 2014). Similarly, MEIS1 binding sites were clustered according to their 5hmC and MEIS1 levels, and the clusters were analyzed for TET1 and TET2 engagement (Supplemental Fig. S16). Clusters C2 and C3 presented low and high 5hmC levels, respectively, and were similarly associated with TET1, whereas C2 was enriched in TET2 and, interestingly, presented the highest enrichment in FOXA1. These data, together with results from DMOG and siRNA experiments, strongly suggest that FOXA1-only and FOXA1/MEIS1-bound enhancers are more likely to undergo 5hmC superoxidation through TET2 recruitment than enhancers bound only by MEIS1.

Although MEIS1 engagement was decreased at eight out of the 10 MEIS1-bound enhancers (statistically significant only for four of them) in the 5-azadC condition (Fig. 4C), the impact of DMOG treatment on its recruitment was variable but had a tendency to follow changes in chromatin opening levels at seven out of 10 enhancers (Fig. 4D). These observations are consistent with a model (Supplemental Fig. S17) in which MEIS1 binding could be stabilized by 5mC/5hmC (in agreement with in vitro data shown in Fig. 3A) and facilitated by chromatin opening. Accordingly, the rapid increase in 5hmC levels at MEIS1 binding regions was systematically followed by chromatin opening and MEIS1 recruitment (Fig. 6), and a significant fraction of these sites (Fig. 6, clusters 1+2, including 1032 regions out of 3366) showed chromatin opening before MEIS1 engagement. In turn, binding of MEIS1 may favor acetylation of H3K27, as suggested by the correlation between these two events (Fig. 4; Supplemental Figs. S8,S9). All together, these results indicate that 5mC removal is required for enhancer priming and that 5hmC per se is functionally linked to chromatin opening events allowing recruitment of MEIS1.

Discussion

Whereas FOXA1 is known to bind nucleosomes (Sekiya et al. 2009), the data presented herein suggest that MEIS1 engagement at enhancers requires chromatin opening in correlation with TET-dependent 5mC oxidation. MEIS1-bound enhancers are particularly enriched in 5hmC, and we show evidence that MEIS1 can bind to sites including 5hmCpGs. Hence, 5mC oxidation could favor MEIS1 binding to DNA by promoting chromatin opening, suggesting that MEIS1 might poorly bind to nucleosomal DNA. Interestingly, the related TALE-HD protein PBX1, which dimerizes with MEIS1 in P19-derived NPCs, is enriched in pull-down experiments with naked methylated DNA but not with methylated DNA wrapped around histones (Bartke et al. 2010). These in vitro

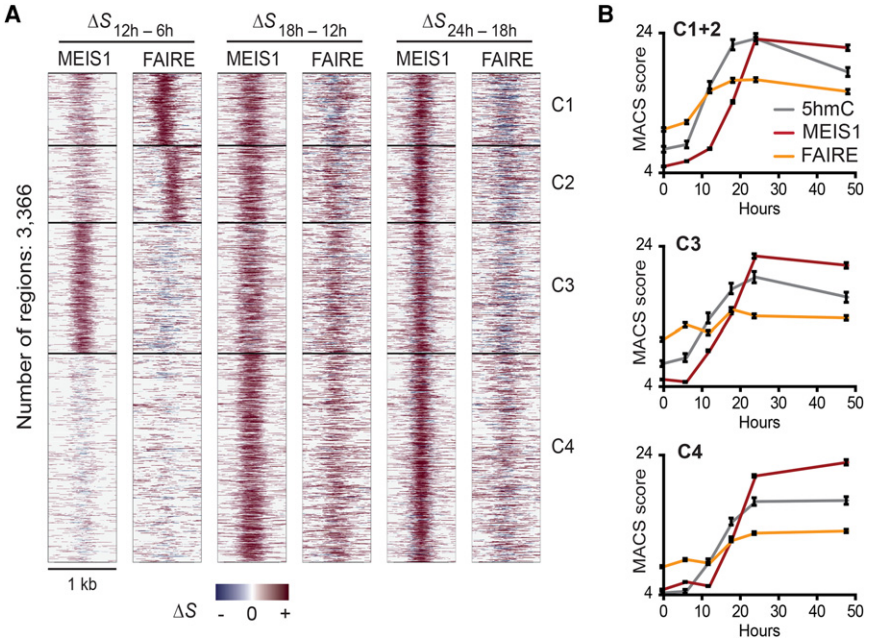

Figure 6. Gain of 5hmC at MEIS1-bound enhancers is followed by chromatin opening and MEIS1 binding. (*A*) Heat maps of MEIS1 and FAIRE differential signals (Δ*S*) during P19 cell differentiation in clusterized MEIS1 binding regions. Blue: decrease; red: increase. (*B*) Dynamic average profiles of 5hmC, chromatin opening, and MEIS1 recruitment at clusters shown in *A*.

observations are in agreement with a model in which 5hmC first favors nucleosome instability in vivo and then stabilizes PBX1/MEIS1 on DNA at CpG-containing recognition sites. Nonetheless, we cannot exclude that MEIS1 binds actually to sites with unmodified cytosines after complete DNA demethylation. Oxidative and classical hairpin bisulfite sequencing of DNA recovered from MEIS1-bound immunoprecipitated chromatin could eventually bring the answer to this question, but purifying the required amount of ChIPed DNA appears challenging.

By inhibiting DNA methylation, we observed a remarkable increase in H3K4me1 at a majority of tested enhancers, strongly suggesting that 5mC is a repellant for H3K4 methyltransferases in vivo. In accordance with this idea, KMT proteins are directly repelled by 5mC in vitro (Birke et al. 2002; Bach et al. 2009; Spruijt et al. 2013), suggesting that oxidation of 5mC could be sufficient for allowing KMT recruitment to CpG-containing DNA sequences through their CXXC domain and methylation of H3K4. In agreement with this hypothesis, our kinetic ChIP-seq data as well as the PCA analysis of DMOG-treated P19 ECCs indicated a strong correlation between monomethylation of H3K4 and the presence of 5hmC. Interestingly, the recently released structure of the human KMT2A CXXC domain in complex with CpG-containing DNA (PDB, 4NW3) reveals that the CXXC domain is anchored to both cytosines of the CpG by two residues, namely K_{1186} and Q_{1187}, which establish contacts with the fifth carbon of the cytosines. Hence, the presence of a methyl group or a hydroxymethyl group on the fifth carbon would be predicted to destabilize the interaction between the CXXC domain of KMT2A and DNA, as observed in vitro by Spruijt et al. (2013). Accordingly, sites which gain 5hmC in the presence of DMOG (most likely because of a reduced oxidation of 5hmC), and thus which have lower occurrence of a complete DNA demethylation process, should have reduced levels of H3K4me1. Our data actually show the opposite, since four out of the five enhancers gaining H3K4me1 in DMOG + RA treated cells also gained 5hmC (Fig. 4D), suggesting that the hydroxymethyl group might favor additional interactions with the CXXC domain of KMT2A.

Conversely to what was observed for H3K4me1, our data indicate that the presence of 5hmC itself at enhancers could facilitate chromatin opening, leading to the hypothesis that this modified base plays a role in the destabilization or eviction of nucleosomes. Still, it cannot be excluded that chromatin opening could occur through the action of remodelers or histone chaperones that would be brought to enhancers by TET enzymes and would not require erasure of cytosine modifications per se. In that case, 5hmC would merely reflect the presence of a TET enzyme and would not, stricto sensu, participate in chromatin opening. If this hypothesis was correct, then blocking superoxidation with DMOG should not lead to an increase in FAIRE signal. Nonetheless, such an increase was noted for five out of the seven enhancers that experienced a gain in 5hmC in DMOG + RA-treated cells, indicating that 5hmC might indeed be directly involved in the regulation of chromatin opening. Interestingly, apart from PBX1, PBX2, MEIS1, MEIS2, and PKNOX2 TALE-HD proteins, and their HOX partners, we identified PBX1-bound chromatin as being enriched in components of the chromatin remodeling complexes NuRD and FACT through ChIP-MS (Supplemental Table 1). NuRD is known to be associated with both active and repressed chromatin and can be targeted by MECP2, MBD2, and MBD3 CpG-binding proteins (Whyte et al. 2012; Günther et al. 2013; Shimbo et al. 2013). Whereas neither MECP2 nor MBD2 were identified in the PBX1 ChIP-MS experiment, MBD3 was (Supplemental Table 1). These data are consistent with the observed engagement of NuRD and MBD3 at 5hmC-enriched chromatin regions (Yildirim et al. 2011; Baubec et al. 2013). Nonetheless, MBD3 is still found at these regions in *Dnmt* triple-KO mouse ES cells, suggesting that 5hmC does not play a direct role in MBD3 targeting (Baubec et al. 2013). The presence of FACT at PBX1-bound enhancers indicates that the chromatin specificities of these sites could also impact FACT engagement. FACT facilitates H2A/H2B eviction from nucleosomes (Dinant et al. 2013), a process which relies on its ability to lower H2B interaction with DNA (Hondele et al. 2013). Interestingly, FACT activity can be enhanced by poly-ADP ribosylation of its subunit SPT16 by PARP1 (Dinant et al. 2013), a protein also detected by ChIP-MS at PBX1-bound chromatin (Supplemental Table 1). Hence, TET engagement at enhancers could ultimately lead to FACT recruitment and activation. Collectively, this information suggests that cytosine hydroxymethylation by TETs may influence nucleosome stability in a direct or indirect manner. This is consistent with reports showing (1) that nucleosomes found at hydroxymethylated CTCF binding sites are more sensitive to MNase digestion (Teif et al. 2014), and (2) that nucleosomes formed in vitro on a hydroxymethylated DNA template are unstable due to a lower interaction of DNA with H2A-H2B dimers (Mendonca et al. 2014). In addition,

a recent study described the engagement of FACT at a POU5F1-bound enhancer in ES cells and its role in nucleosome eviction (Shakya et al. 2015). Interestingly, in this study, DMOG reduces the engagement of FACT at this POU5F1-bound enhancer, suggesting that FACT recruitment depends either on histone demethylation, or on DNA hydroxymethylation, or on both (Shakya et al. 2015).

Collectively, our data further validate the concept that DNA demethylation is required for the proper regulation of enhancer priming and activation during cell differentiation and provide additional clues for understanding how nonpromoter DNA (hydroxy)methylation can participate in the cell-specific control of gene activity.

Methods

Cell culture

P19.6 mouse embryonal carcinoma cells were grown and differentiated as described (Sérandour et al. 2012). To inhibit DNA methyltransferases, cells were plated in a medium containing 20 nM of 5-aza-2-deoxycytidine (5-azadC, Sigma) for 24 h. The medium was then renewed and cells were cultured in the presence of 20 nM 5-azadC for 48 h with or without RA treatment. Inhibition of TET proteins was performed by treating P19 cells with 0.5 mM of dimethyloxalylglycine (DMOG, Cayman Chemical). Due to DMOG instability, the medium was renewed every 24 h. *Tet1-*, *2-*, and *3*-specific 27mer siRNA duplexes (Trilencer-27 siRNA) and universal negative siRNA duplex were purchased from Origene Technologies. Transfections of P19 cells with siRNAs at a final concentration of 10 nM were performed in Opti-MEM medium using Oligofectamine (Invitrogen) as recommended by the manufacturer. After 6 h of transfection, the medium was renewed and cells were cultured for 48 h with or without RA treatment.

Chromatin and DNA immunoprecipitation

ChIP experiments were carried out as previously described (Sérandour et al. 2011). ChIP followed by high-throughput sequencing (ChIP-seq) experiments were performed as in Schmidt et al. (2009). DNA immunoprecipitation protocols are available on the EpiGeneSys website (http://www.epigenesys.eu/en/protocols): "Methylated DNA Immunoprecipitation (MeDIP)" (Protocol 33 - Michaël Weber and Dirk Schübeler) and "Hydroxymethylated DNA Immunoprecipitation (hMeDIP) in mammalian cells" (Protocol 64 - Elise Mahé, Gilles Salbert).

Formaldehyde assisted isolation of regulatory elements

Formaldehyde assisted isolation of regulatory elements was performed as previously described (Eeckhoute et al. 2009).

Sequencing and bioinformatics

The (h)MeDIP, ChIP, FAIRE, and input libraries were prepared using the TruSeq ChIP Sample Prep kit (ref. IP-202-1012, Illumina). The libraries were sequenced by HiSeq following the manufacturer's protocol. Alignment of reads, conversion of extracted reads to .wig signal files, and peak calling were performed as previously described (Sérandour et al. 2012). Blacklisted regions available from UCSC (http://genome.ucsc.edu) were removed from each .wig signal file. CpG island (CGi) coordinates were also obtained from UCSC. Signals were normalized in order to compare (h)MeDIP-, FAIRE-, or ChIP-seq signals from different time points in time-course experiments. Proper normalization requires both calibrating the highest peaks and accounting for noise heterogeneity. For MEIS1, which is an induced transcription factor, signals within .wig files were calibrated according to the number of mapped reads. For other signals, assuming that the highest peaks remained, on average, at similar levels, we selected 15,000 positions with the highest signal for each time point sample, determined their mean signal value, and chose as a reference the sample with the highest mean signal. Then, all signals within a given .wig file were calibrated by multiplication by a factor set to $\alpha = \bar{s}_{ref}/\bar{s}_i$, where $\bar{s}_{ref}$ is the mean value of the reference signal and $\bar{s}_i$ the average value of signal s_i, in their respective top 15,000 positions. In order to reduce artifacts linked to noise heterogeneity between samples, we first defined for each sample an upper noise threshold τ as being twice the mean value of the signal for all positions of the corresponding .wig file. A peak-calling algorithm (Sérandour et al. 2012) was next used to identify peaks with a signal higher than τ in at least three consecutive positions. Called peaks were then extended by 100 bp on each side to include positions corresponding to the sides of the peaks even though they had an initial signal below τ. Finally, the original signal present within these extended regions was retrieved to generate normalized noise-free .wig files. Heat maps were generated by the Cistrome web-platform (http://cistrome.org/ap/; Liu et al. 2011). Motif search was run online through the SeqPos tool from Cistrome, and the analysis of motif distribution in genomic regions was done using CentDist (http://biogpu.ddns.comp.nus.edu.sg). TET1 ChIP-seq and SCL-exo data are from the NCBI GEO GSM941681 and GSE70635 data sets.

Pull-down assay

Differentially modified 20-bp synthetic oligonucleotides (Sigma), biotinylated on the forward strand and containing either the PBX1/HOXA9 (TGATTTACG) or the PBX1/MEIS1 (TGATTGACAG) recognition sites, were used as probes for pull-down assays. For each probe, cytosines from each CR dinucleotide (R = A or G) were unmodified (C), methylated (M), or hydroxymethylated (H). DNA/protein complexes were purified with streptavidin-coupled magnetic beads (Invitrogen) after incubation in binding buffer (Hepes 20 mM pH 7.9, 50 mM KCl, 1 mM DTT, 2.5 mM $MgCl_2$, 5 µg/µL BSA, 10% glycerol) with 10, 20, or 40 µg of extracts from P19 cells treated for 48 h with RA. Protein binding was then analyzed by Western blot. Signals were quantified using the Quantity One 1-D Analysis Software from Bio-Rad. Competition assays were run with the biotinylated PBX1/MEIS1 probe and a fivefold excess of differentially modified double-stranded oligonucleotides of the following sequence: TACACTGATTGACGGTGCAT.

ChIP-MS

ChIP followed by mass spectrometry analysis (ChIP-MS) was performed in triplicate as described in Mohammed et al. (2013). Samples were analyzed by the Proteomics core facility of the Cambridge Institute (Cancer Research UK). The nuclear fraction from 5×10^7 P19 cells cross-linked with formaldehyde was extracted and sonicated. Immunoprecipitation was performed with 100 µL of magnetic beads (Dynal) prebound either with 20 µg of PBX1 antibody or with 20 µg of control IgG. After several washes of the beads, the bead-bound proteins were digested with trypsin. The beads were then incubated overnight at 37°C to allow the salvage of the digested peptides, which were then analyzed by nano-LC-MS/MS. Mass spectrometry was performed using an LTQ Velos-Orbitrap mass spectrometer (Thermo Scientific) coupled to an Ultimate RSLCnano-LC system (Dionex). Results were analyzed with the Scaffold Viewer software with protein threshold and peptide threshold set at 90%. The mass spectrometry proteomics data

have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino et al. 2016) partner repository with the data set identifier PXD006158.

Chemiluminescent EMSA

EMSA experiments were performed with nuclear extracts (NEs) prepared from P19 cells treated for 48 h with 1 µM RA. Two micrograms of NE were incubated with binding buffer (Hepes 20 mM pH 7.9, 10 mM KCl, 1 mM DTT, 2.5 mM $MgCl_2$, 10% glycerol), 1 µg poly(dI-dC), and 150 fmol of the biotinylated probes used in pull-down assays. For supershift assays, NEs were first pre-incubated with the antibody of interest. EMSAs were performed using the "GelshiftTM Chemiluminescent EMSA" kit (Active Motif) following the manufacturer's instructions.

Dot blot assay

Genomic DNA was extracted using a DNA extraction kit (DNeasy Blood and Tissue kit; Qiagen) with a supplementary step of RNase digestion. Three hundred nanograms, 150 ng, and 75 ng of each sample DNA was spotted to a nylon membrane (Hybond N, GE Healthcare) previously soaked in 2× SSPE solution (0.3 M NaCl, 20 mM sodium phosphate, 2 mM EDTA) and inserted in the dot blot apparatus (SCIE-PLAS). DNA was cross-linked to the membrane by 15 min exposure to UV light before blotting. After revelation, membranes were stained with methylene blue to control the amount of DNA spotted. Signals were quantified using the Quantity One 1-D Analysis Software from Bio-Rad.

RNA preparation and reverse transcription

RNA was isolated by TRIzol extraction (Invitrogen) and reverse-transcribed using M-MLV reverse transcriptase (Invitrogen) and Pd(N)6 random hexamers (Amersham Pharmacia Biosciences).

Epimark 5hmC and 5mC analysis

Methylation or hydroxymethylation status of selected CpGs included in a MspI CCGG restriction site was verified using the EpiMark 5hmC and 5mC Analysis kit (New England Biolabs) according to the manufacturer's protocol. Briefly, genomic DNA was first treated by T4 beta-glucosyltransferase (T4-BGT) in order to add a glucose moiety to the hydroxyl group of 5hmC, and then digested by either MspI or HpaII restriction enzymes. Both enzymes recognize a CCGG sequence, but HpaII cleaves only unmodified sites, whereas MspI cleaves 5mC or 5hmC but not 5ghmC. Approximate quantification of cytosine modifications was obtained by real-time PCR. Specific oligonucleotides were designed whenever those designed for ChIP-qPCR were not appropriate. Analysis and calculations were performed according to the kit protocol.

Principal component analysis and clustering

The R software (R Core Team 2013), with the FactoMineR package (Lê et al. 2008), was used to perform principal component analysis and hierarchical clustering. Briefly, chromatin mark variations across several genomic regions were investigated by PCA in which variables were genomic regions and individuals were chromatin marks. Data were sequentially scaled to unit variance and then projected onto orthogonal components. The resulting eigenvalues (based on both the covariance and correlation matrices) were used to retain the first components accounting for over 90% of the variance in the data. Finally, the coordinates of the individuals and variables on the selected principal components were used for the hierarchical classification.

Real-time PCR (qPCR)

Real-time PCR was performed using SYBR Green Master Mix (Bio-Rad, France) on a Bio-Rad CFX96 machine with 39 cycles of amplification. All primers were purchased from Sigma and are listed in the Supplemental Materials and Methods section.

Statistical analyses

Statistical analyses were performed using GraphPad Prism software. Statistical significance was determined using the Mann-Whitney test.

Data access

All sequencing data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) database under accession number GSE82314. PBX1 ChIP-MS data are accessible at the PRIDE Archive database from the European Bioinformatics Institute (https://www.ebi.ac.uk/pride/) under accession number PXD006158.

Acknowledgments

We thank C. Le Péron and A. Laurent for critical reading of the manuscript. We also thank the "Human and Environmental Genomics" platform (Rennes) for Illumina sequencing. We are thankful to C. D'Santos and Chris Taylor from the Proteomics Core Facility of the Cambridge Institute (Cancer Research UK) for running the MS experiments. This work was supported by the CNRS, the University of Rennes I, and the Association de la Recherche contre le Cancer. E.A.M. was supported by a PhD fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche.

Author contributions: E.A.M. planned and ran experiments, analyzed the data, and wrote the manuscript. T.M., A.A.S., M.B., and G.P. ran experiments. S.A. ran bioinformatic analyses. F.C. ran the PCA analyses of the ChIP-qPCR experiments. R.M. designed and ran ChIP-qPCR experiments and analyzed the data. G.S. planned the study, analyzed the data, and wrote the manuscript.

References

Bach C, Mueller D, Buhl S, Garcia-Cuellar MP, Slany RK. 2009. Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2. *Oncogene* **28:** 815–823.

Bachman M, Uribe-Lewis S, Yang X, Williams M, Murrell A, Balasubramanian S. 2014. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat Chem* **6:** 1049–1055.

Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. 2010. Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* **143:** 470–484.

Baubec T, Ivánek R, Lienert F, Schübeler D. 2013. Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell* **153:** 480–492.

Biddie SC, John S, Sabo PJ, Thurman RE, Johnson TA, Schiltz RL, Miranda TB, Sung MH, Trump S, Lightman SL, et al. 2011. Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding. *Mol Cell* **43:** 145–155.

Bird A. 2011. The dinucleotide CG as a genomic signaling module. *J Mol Biol* **409:** 47–53.

Birke M, Schreiner S, García-Cuéllar MP, Mahr K, Titgemeyer F, Slany RK. 2002. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res* **30:** 958–965.

Calo E, Wysocka J. 2013. Modification of enhancer chromatin: what, how, and why? *Mol Cell* **49:** 825–837.

Chang CP, Jacobs Y, Nakamura T, Jenkins NA, Copeland NG, Cleary ML. 1997. Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. *Mol Cell Biol* **17:** 5679–5687.

de la Rica L, Rodríguez-Ubreva J, García M, Islam AB, Urquiza JM, Hernando H, Christensen J, Helin K, Gómez-Vaquero C, Ballestar E. 2013. PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome Biol* **14:** R99.

Dinant C, Ampatziadis-Michailidis G, Lans H, Tresini M, Lagarou A, Grosbart M, Theil AF, van Cappellen WA, Kimura H, Bartek J, et al. 2013. Enhanced chromatin dynamics by FACT promotes transcriptional restart after UV-induced DNA damage. *Mol Cell* **51:** 469–479.

Dong D, Meng L, Yu Q, Tan G, Ding M, Tan Y. 2012. Stable expression of FoxA1 promotes pluripotent P19 embryonal carcinoma cells to be neural stem-like cells. *Gene Expr* **15:** 153–162.

Eeckhoute J, Lupien M, Meyer CA, Verzi MP, Shivdasani RA, Liu XS, Brown M. 2009. Cell-type selective chromatin remodeling defines the active subset of FOXA1-bound enhancers. *Genome Res* **19:** 372–380.

Ghisletti S, Barozzi I, Mietton F, Polletti S, De Santa F, Venturini E, Gregory L, Lonie L, Chew A, Wei CL, et al. 2010. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* **32:** 317–328.

Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD. 2007. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res* **17:** 877–885.

Günther K, Rust M, Leers J, Boettger T, Scharfe M, Jarek M, Bartkuhn M, Renkawitz R. 2013. Differential roles for MBD2 and MBD3 at methylated CpG islands, active promoters and binding to exon sequences. *Nucleic Acids Res* **41:** 3010–3021.

He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, et al. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333:** 1303–1307.

Heidt AB, Rojas A, Harris IS, Black BL. 2007. Determinants of myogenic specificity within MyoD are required for noncanonical E box binding. *Mol Cell Biol* **27:** 5910–5920.

Heinz S, Romanoski CE, Benner C, Glass CK. 2015. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol* **16:** 144–154.

Hon GC, Song CX, Du T, Jin F, Selvaraj S, Lee AY, Yen CA, Ye Z, Mao SQ, Wang BA, et al. 2014. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol Cell* **56:** 286–297.

Hondele M, Stuwe T, Hassler M, Halbach F, Bowman A, Zhang ET, Nijmeijer B, Kotthoff C, Rybin V, Amlacher S, et al. 2013. Structural basis of histone H2A-H2B recognition by the essential chaperone FACT. *Nature* **499:** 111–114.

Iurlaro M, Ficz G, Oxley D, Raiber EA, Bachman M, Booth MJ, Andrews S, Balasubramanian S, Reik W. 2013. A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. *Genome Biol* **14:** R119.

Iwafuchi-Doi M, Zaret KS. 2014. Pioneer transcription factors in cell reprogramming. *Genes Dev* **28:** 2679–2692.

Jeong KW, Andreu-Vieyra C, You JS, Jones PA, Stallcup MR. 2014. Establishment of active chromatin structure at enhancer elements by mixed-lineage leukemia 1 to initiate estrogen-dependent gene expression. *Nucleic Acids Res* **42:** 2245–2256.

Jüttermann R, Li E, Jaenisch R. 1994. Toxicity of 5-aza-2′-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proc Natl Acad Sci* **91:** 11797–11801.

Kashyap V, Laursen KB, Brenet F, Viale AJ, Scandura JM, Gudas LJ. 2013. RARγ is essential for retinoic acid induced chromatin remodeling and transcriptional activation in embryonic stem cells. *J Cell Sci* **126:** 999–1008.

Lê S, Josse J, Husson F. 2008. FactoMineR: an R package for multivariate analysis. *J Stat Softw* **25:** 1–18.

Liu T, Ortiz JA, Taing L, Meyer CA, Lee B, Zhang Y, Shin H, Wong SS, Ma J, Lei Y, et al. 2011. Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol* **12:** R83.

Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W, Carroll JS, Liu XS, Brown M. 2008. FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132:** 958–970.

Magnani L, Ballantyne EB, Zhang X, Lupien M. 2011. PBX1 genomic pioneer function drives ERα signaling underlying progression in breast cancer. *PLoS Genet* **7:** e1002368.

Maiti A, Drohat AC. 2011. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem* **286:** 35334–35338.

Mendonca A, Chang EH, Liu W, Yuan C. 2014. Hydroxymethylation of DNA influences nucleosomal conformation and stability in vitro. *Biochim Biophys Acta* **1839:** 1323–1329.

Moens CB, Selleri L. 2006. Hox cofactors in vertebrate development. *Dev Biol* **291:** 193–206.

Mohammed H, D'Santos C, Sérandour AA, Ali HR, Brown GD, Atkins A, Rueda OM, Holmes KA, Theodorou V, Robinson JL, et al. 2013. Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. *Cell Rep* **3:** 342–349.

Putiri EL, Tiedemann RL, Thompson JJ, Liu C, Ho T, Choi JH, Robertson KD. 2014. Distinct and overlapping control of 5-methylcytosine and 5-hydroxymethylcytosine by the TET proteins in human cancer cells. *Genome Biol* **15:** R81.

Qin P, Haberbusch JM, Zhang Z, Soprano KJ, Soprano DR. 2004. Pre-B cell leukemia transcription factor (PBX) proteins are important mediators for retinoic acid-dependent endodermal and neuronal differentiation of mouse embryonal carcinoma P19 cells. *J Biol Chem* **279:** 16263–16271.

R Core Team. 2013. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rose NR, McDonough MA, King ON, Kawamura A, Schofield CJ. 2011. Inhibition of 2-oxoglutarate dependent oxygenases. *Chem Soc Rev* **40:** 4364–4397.

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, et al. 2003. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34:** 374–378.

Schmidt D, Wilson MD, Spyrou C, Brown GD, Hadfield J, Odom DT. 2009. ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. *Methods* **48:** 240–248.

Sekiya T, Muthurajan UM, Luger K, Tulin AV, Zaret KS. 2009. Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA. *Genes Dev* **23:** 804–809.

Sérandour AA, Avner S, Percevault F, Demay F, Bizot M, Lucchetti-Miganeh C, Barloy-Hubler F, Brown M, Lupien M, Métivier R, et al. 2011. Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome Res* **21:** 555–565.

Sérandour AA, Avner S, Oger F, Bizot M, Percevault F, Lucchetti-Miganeh C, Palierne G, Gheeraert C, Barloy-Hubler F, Péron CL, et al. 2012. Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Res* **40:** 8255–8265.

Sérandour AA, Avner S, Mahé EA, Madigou T, Guibert S, Weber M, Salbert G. 2016. Single-CpG resolution mapping of 5-hydroxymethylcytosine by chemical labeling and exonuclease digestion identifies evolutionarily unconserved CpGs as TET targets. *Genome Biol* **17:** 56.

Shakya A, Callister C, Goren A, Yosef N, Garg N, Khoddami V, Nix D, Regev A, Tantin D. 2015. Pluripotency transcription factor Oct4 mediates stepwise nucleosome demethylation and depletion. *Mol Cell Biol* **35:** 1014–1025.

Shen WF, Rozenfeld S, Lawrence HJ, Largman C. 1997. The Abd-B-like Hox homeodomain proteins can be subdivided by the ability to form complexes with Pbx1a on a novel DNA target. *J Biol Chem* **272:** 8198–8206.

Shimbo T, Du Y, Grimm SA, Dhasarathy A, Mav D, Shah RR, Shi H, Wade PA. 2013. MBD3 localizes at promoters, gene bodies and enhancers of active genes. *PLoS Genet* **9:** e1004028.

Siersbæk R, Nielsen R, John S, Sung MH, Baek S, Loft A, Hager GL, Mandrup S. 2011. Extensive chromatin remodelling and establishment of transcription factor 'hotspots' during early adipogenesis. *EMBO J* **30:** 1459–1472.

Spitz F, Furlong EE. 2012. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* **13:** 613–626.

Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C, Münzel M, Wagner M, Müller M, Khan F, et al. 2013. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* **152:** 1146–1159.

Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, et al. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480:** 490–495.

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, et al. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324:** 930–935.

Takayama K, Misawa A, Suzuki T, Takagi K, Hayashizaki Y, Fujimura T, Homma Y, Takahashi S, Urano T, Inoue S. 2015. TET2 repression by androgen hormone regulates global hydroxymethylation status and prostate cancer progression. *Nat Commun* **25:** 8219.

Tan Y, Xie Z, Ding M, Wang Z, Yu Q, Meng L, Zhu H, Huang X, Yu L, Meng X, et al. 2010. Increased levels of FoxA1 transcription factor in

pluripotent P19 embryonal carcinoma cells stimulate neural differentiation. *Stem Cells Dev* **19:** 1365–1374.

Teif VB, Beshnova DA, Vainshtein Y, Marth C, Mallm J-P, Höfer T, Rippe K. 2014. Nucleosome repositioning links DNA (de)methylation and differential CTCF binding during stem cell development. *Genome Res* **24:** 1285–1295.

Visel A, Minovitsky S, Dubchak I, Pennacchio LA. 2007. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res* **35:** D88–D92.

Vizcaino JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, et al. 2016. 2016 update of the PRIDE database and related tools. *Nucleic Acids Res* **44:** D447–D456.

Whyte WA, Bilodeau S, Orlando DA, Hoke HA, Frampton GM, Foster CT, Cowley SM, Young RA. 2012. Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature* **482:** 221–225.

Yamada T, Urano-Tashiro Y, Tanaka S, Akiyama H, Tashiro F. 2013. Involvement of crosstalk between Oct4 and Meis1a in neural cell fate decision. *PLoS One* **8:** e56997.

Yildirim O, Li R, Hung JH, Chen PB, Dong X, Ee LS, Weng Z, Rando OJ, Fazzio TG. 2011. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* **147:** 1498–1510.

Received June 16, 2016; accepted in revised form April 5, 2017.

Cytosine modifications modulate the chromatin architecture of transcriptional enhancers

Elise A. Mahé, Thierry Madigou, Aurélien A. Sérandour, et al.

Genome Res. 2017 27: 947-958 originally published online April 10, 2017
Access the most recent version at doi:10.1101/gr.211466.116