Published in final edited form as:
Neuron. 2014 August 20; 83(4): 823–838. doi:10.1016/j.neuron.2014.07.013.

Dnmt3a regulates global gene expression in olfactory sensory neurons and enables odorant-induced transcription

Bradley M. Colquitt[1,2], **Eirene Markenscoff-Papadimitriou**[1,2], **Rachel Duffié**[2], and **Stavros Lomvardas**[2,*]

[1]Neuroscience Graduate Program, University of California, San Francisco, CA, 94158, USA

[2]Department of Anatomy, University of California, San Francisco, CA 94158, USA

Summary

During differentiation, neurons exhibit a reorganization of DNA modification patterns across their genomes. The *de novo* DNA methyltransferase Dnmt3a is implicated in this process, but the effects of its absence have not been fully characterized in a purified neuronal population. To better understand how DNA modifications contribute to neuronal function, we performed a comprehensive analysis of the epigenetic and transcriptional landscapes of Dnmt3a-deficient mature olfactory sensory neurons (mOSNs), the primary sensory neurons of the olfactory epithelium. Dnmt3a is required for both 5mC and 5hmC patterning within accessible genomic regions, including hundreds of neurodevelopmental genes and neural enhancers. Loss of Dnmt3a results in the global disruption of gene expression via activation of silent genes and reduction of mOSN-expressed transcripts. Importantly, the DNA modification state and inducibility of odorant-activated genes is markedly impaired in Dnmt3a knockouts, suggesting a crucial role for this enzyme in establishing an epigenetic landscape compatible with neuronal plasticity.

Introduction

The sculpting of DNA modification patterns during cellular differentiation is essential to the construction of cellular identity. Once thought to be largely static in somatic tissues after embryogenesis, the patterning of cytosine modifications is now appreciated to be dynamic in certain cellular and genomic contexts. In particular, 5-hydroxymethylcytosine (5hmC), an oxidized derivative of 5-methylcytosine (5mC), is highly enriched in neurons (Globisch et al., 2010; Kriaucionis and Heintz, 2009; Ruzov et al., 2011), increases in abundance during neurogenesis (Hahn et al., 2013; Szulwach et al., 2011), and is localized to gene bodies, regions upstream of transcription start sites (TSSs), and enhancer elements (Colquitt et al., 2013; Hahn et al., 2013; Mellén et al., 2012; Song et al., 2011; Szulwach et al., 2011). To date, it is unclear how neuron-specific 5hmC patterning is established during the transition

*Correspondence to: stavros.lomvardas@ucsf.edu.

from neuronal progenitor to differentiated neuron. Recent studies (Hahn et al., 2013; Szulwach et al., 2011) noted that the increase of 5hmC during neurodevelopment is not strictly accompanied by a reduction of 5mC, suggesting that *de novo* DNA methylation and cytosine oxidation, mediated by the Tet family of enzymes (Wu and Zhang, 2011), are coupled during this developmental period. Moreover, substantial *de novo* 5mC patterning was recently found in cortical forebrain neurons relative to non-neuronal cell types (Lister et al., 2013).

Dnmt3a, one of two *de novo* DNA methyltransferases in mammals, is expressed in neural precursor cells and neurons during late embryogenesis and in post-mitotic neurons in the postnatal central nervous system (CNS) (Feng et al., 2005). Its ablation specifically within the CNS (Nguyen et al., 2007) recapitulates the early death (3–4 weeks) seen in mice with constitutive loss of the enzyme (Okano et al., 1999). In addition, its loss within post-mitotic neurons is associated with deficits in long-term potentiation, learning, and memory (Feng et al., 2010). Recent work identified a role for Dnmt3a in both the repression and facilitation of gene expression in neural stem cells (Wu et al., 2010), suggesting that the enzyme's effects extend beyond establishing repressive 5mC. Here, we tested the hypothesis that Dnmt3a-mediated *de novo* DNA methylation contributes to 5hmC patterning within neurons and is necessary to define neuronal regulatory and transcriptional states. To explore this model *in vivo*, we analyzed the relationships between Dnmt3a-dependent 5mC and 5hmC patterning, chromatin accessibility, histone modifications, and transcription within both the mouse main olfactory epithelium (MOE) and the resident sensory neurons of the tissue, mature olfactory sensory neurons (mOSNs).

We find that Dnmt3a is required for a restricted fraction of 5mC and 5hmC patterning within mOSNs that is strongly associated with highly accessible regions, including enhancer elements and transcription start sites. We also identify a set of regions exhibiting broad but precisely defined Dnmt3a-dependent 5mC and 5hmC patterning, high accessibility, and an association with neurodevelopmental genes. Interestingly, these Dnmt3a-dependent regions display intermediate levels of 5mC and 5hmC in Dnmt3a heterozygotes, indicating that levels of available Dnmt3a strongly influence DNA modification patterning. Moreover, the absence of Dnmt3a in mOSNs results in widespread transcriptional alterations including the upregulation of repressed genes and modest downregulation of expressed genes. Finally, in the absence of Dnmt3a, immediate early genes display significant reductions of 5mC and 5hmC levels and have compromised odorant-induced gene expression. These data support a model in which Dnmt3a-mediated DNA modification patterning is necessary to establish both repressive and active transcriptional states within neurons and to generate the full set of gene responses associated with neuronal plasticity.

Results

We find that *Dnmt3a* transcription increases along the developmental lineage of olfactory sensory neurons, from horizontal basal cells (HBCs) – the multipotent stem cells of the tissue (Leung et al., 2007) – through globose basal cells (GBCs) – the neuronal progenitors of mOSNs (Caggiano et al., 1994) – and finally to mOSNs (Fig. 1A), in agreement with previous work (MacDonald et al., 2005; Watanabe et al., 2006). Notably, this developmental

increase is similar to the enrichment of 5hmC along the mOSN developmental path we described in an earlier study (Colquitt et al., 2013). In contrast, *Dnmt3b*, the other mammalian *de novo* DNA methyltransferase, is weakly expressed in HBCs but is not expressed within GBCs or mOSNs. Thus, Dnmt3a is the primary source of *de novo* 5mC within mOSNs. In three-week old Dnmt3a wildtype (WT) mice Dnmt3a protein is most abundant in the immature neuronal stage, between GBCs and mOSNs, but is absent from basal stem and apical non-neuronal layers (Fig. 1B and Fig. S1A). Importantly, we do not find significant alterations in Dnmt3a KO MOEs of the expression of the other DNA methyltransferases (except for a 2-fold increase of *Dnmt3b* expression ($p = 0.02$) which still remains at 5% of the expression level of *Dnmt3a* in WT) or of the three Tet members (Fig. S1B). In addition, mOSN differentiation is largely unaffected in Dnmt3a KO MOEs, as determined by unaltered percentages of mOSNs and GBCs within the tissue (Fig. S1C), unaffected mitosis rates (Fig. S1D), and only mild increases in apoptosis rates (0.3% of total cells in WT to 0.9% in KOs, $p = 0.11$, Fig. S1D and E). To determine the effects of the absence of Dnmt3a on global 5mC and 5hmC levels, we immunostained Dnmt3a WT and KO MOEs with antibodies specific for the two modified bases (Fig. 1C). Interestingly, we find no significant difference in 5mC levels along the basal-apical axis ($p = 0.39$, Student's t-test, N=3). However, we detect a significant reduction of 5hmC in Dnmt3a KO MOEs within the mOSN layer ($p = 0.03$, Students t-test, N=3), despite initially reaching levels similar to those in WT within more basal layers (Fig. 1C).

These results indicate that Dnmt3a is necessary to provide some component of *de novo* 5hmC patterning during the transition from neuronal progenitor to neuron. To determine what genomic regions exhibit Dnmt3a-dependent 5mC and 5hmC patterning, we isolated mOSNs via fluorescence-activated cell sorting (FACS) from the MOEs of three-week old *Omp*-(*Olfactory marker protein*)-IRES-EGFP and Dnmt3a WT, Het, and KO mice and performed DNA immunoprecipitation against 5mC and 5hmC coupled to deep sequencing (DIP-seq, Table S1). The specificity of these antibodies has been extensively demonstrated by us and other laboratories (Colquitt et al., 2013; Ito et al., 2010; Williams et al., 2011). We used two methods to identify genomic regions that contain Dnmt3a-dependent 5mC and 5hmC – peak analysis using a small peak size (500 basepairs, bp) and differential enrichment analysis in broader windows (5–50 kilobases, kb) – and found that the loss of Dnmt3a affects the modification status of a restricted set of genomic regions. For the first method, biological variation was controlled for by intersecting differential peak sets from two independent biological replicates. For the second, variation across biological replicates is directly modeled during differential enrichment identification (Anders and Huber, 2010). Both sets of Dnmt3a-dependent 5mC (N=955) and 5hmC (N=5,767) 500 bp regions constitute only a small fraction of the total number of 5mC (N=91,859) and 5hmC-enriched (N=166,572) regions in mOSNs, 1% and 3% respectively. A reciprocal analysis using identical peak cutoffs indicates that only two regions significantly gain 5mC and one region gains 5hmC in Dnmt3a KO relative to WT mOSNs. This specificity is readily seen in a representative ~1.5 Mb locus in which the great majority of 5mC and 5hmC patterning is unaffected by the loss of Dnmt3a except for a few sparsely distributed regions (Fig. 1D). Genome-wide, 5mC and 5hmC levels are highly correlated between Dnmt3a genotypes; however, 5mC and 5hmC levels in Het and KO mOSNs are reduced in regions that are

highly modified in WT mOSNs (linear model slopes < 1, Fig. 1E). We find little overlap between Dnmt3a-dependent 5mC and 5hmC peak sets (Fig. 1F), reflecting their complementary distribution. However, regions containing Dnmt3a-dependent 5mC display mildly reduced 5hmC levels and vice versa (Fig. 1G). Relative to total 5mC and 5hmC peak sets, Dnmt3a-mediated 5mC peaks are localized to CpG islands (CGIs) while Dnmt3a-mediated 5hmC peaks are associated not only with CGIs, but also regions upstream of TSSs, and conserved intergenic regions (Fig. 1H). In addition, both Dnmt3a-dependent 5mC and 5hmC are enriched at intergenic MOE enhancers, described in more detail in Figure 5.

Using the second differential enrichment method, we find large domains (average size ~9 kb) of Dnmt3a-dependent 5mC (N=1,162) and 5hmC (N=1,805) patterning that are strongly associated with gene regions (2025 of 2549, p = 4E-58, Fisher's exact-test). These domains, which we refer to as "Dnmt3a-dependent blocks", are significantly associated with gene regions that are implicated in neuronal projection development and function (e.g. *Ephrin* family members, *Robo1* and *Robo2*, and amyloid beta precursor protein (*App*)), cell adhesion (*Pcdha* cluster and *Dscaml1*), and neurodevelopmental transcriptional regulation (*Foxg1*, *Otx2*, *Olig2*, and *Klf5*, Fig. 2A and Table S2). Of these, approximately two hundred are transcription factors. Interestingly, 94 of the 1944 associated genes are entirely contained within the region of Dnmt3a-dependency. Dnmt3a-dependent blocks are highly conserved, containing significantly higher vertebrate phyloP scores (median±SD, 0.11±0.17) than flanking regions (0.09±0.08, p = 0.02, two-sided Wilcoxon) or an equal number of random genomic regions (0.05±0.11, p = 8E-6, Fig. 2B). Similarly, these blocks are significantly more accessible to DNase I cleavage, a hallmark of active genomic sites, in Dnmt3a WT MOEs (median RPKM±SD, 0.67±0.31) relative to flanking (0.55±0.18, p < 2E-16) and random regions (0.35±0.36, p < 2E-16, Fig. 2C).

To integrate these Dnmt3a-dependent blocks into the broader epigenetic landscape of the MOE, we performed ChIP-seq against H3K4me1 (a mark associated with genes and enhancer elements), H3K27ac (TSSs and active enhancer elements), and H3K27me3 (repressed loci) using native chromatin preparations isolated from the MOEs of three-week old Dnmt3a WT mice. Although the MOE is composed of multiple cell types, ~45% of the tissue consists of mOSNs making it the dominant cell type in this tissue (Fig. S1C). We hierarchically clustered Dnmt3a-dependent blocks by their average H3K27ac and H3K27me3 levels, producing three clusters (enriched for H3K27ac only, H3K27me3 only, or both modifications, Fig. 2D). This grouping by histone modifications resulted in a partial segregation of 5mC and 5hmC levels (Fig. 2D): Regions enriched for H3K27ac (cluster 1, example in Fig. 2E, N=1,219) are enriched for 5hmC while those enriched for H3K27me3 (cluster 2, example in Fig. 2E, N=317) are correspondingly enriched for 5mC. Interestingly, a substantial fraction of these regions (N=913) are enriched for both histone modifications and possess high levels of both DNA modifications (cluster 3, examples in Fig. 2F and Figs. S2A–D). Notably, the set of genes associated with this dually modified cluster is enriched for genes implicated in neurodevelopment including several transcription factors – *Neurog1*, *Dlx5*, *Lhx2*, and *Emx2* – known to be important for OSN maturation or olfactory receptor expression (Table S2) (Hirota and Mombaerts, 2004; Nicolay et al., 2006). In addition, DNase I accessibility within Dnmt3a WT MOEs is elevated precisely within the area of

H3K27ac, H3K27me3, and Dnmt3a-dependent 5mC and 5hmC enrichment (Fig. 2E, F and Figs. S2A–D). This dual presence of histone modifications associated with activation and repression within single loci may be due to developmental heterogeneity within the tissue or may reflect differential monoallelic epigenetic states.

Interestingly, 5mC and 5hmC levels within Dnmt3a-dependent blocks exhibit an intermediate reduction in Dnmt3a Het mOSNs (Fig. 3A and B). This observation could result from an admixture of highly modified and unmodified sequences, as might be generated by tissue or allelic heterogeneity, or by sequences with partially reduced methylation across individual alleles, which would indicate that the DNA modification patterning in these affected regions is sensitive to Dnmt3a dosage. To distinguish between these two models, we performed bisulfite sequencing using gDNA from Dnmt3a WT, Het, and KO MOEs on four gene regions that contain Dnmt3a-dependent 5mC or 5hmC (*Egr1, Nr4a2, Fos*, and the *Dlx5/6* intergenic region), one that contains Dnmt3a-independent 5mC and 5hmC (*Hspa1a*), and three that have enrichment for neither modification (TSSs of *Egr1, Nr4a2*, and *Hspa1a*, Fig. 3C, Table S3). As bisulfite sequencing cannot distinguish between 5mC and 5hmC (Huang et al., 2010), these data provide a conflated measure of 5mC and 5hmC status. In agreement with the DIP-seq data, Dnmt3a-dependent 5(h)mC exhibits intermediate levels in Dnmt3a Het MOE relative to WT and KO MOEs (Fig. 3D). Moreover, individual sequencing clones from Dnmt3a Het MOEs display a partial loss of modification and are not in either fully modified or unmodified states (Figs. 3D–F and Figs S3A). In particular, this partial modification loss does not appear to be stereotyped across clones, consistent with a model in which Dnmt3a protein is limiting and acts by a non-processive mechanism (Gowher and Jeltsch, 2001). Although we detect little 5(h)mC in the CH context, known to be generated by Dnmt3a (Gowher and Jeltsch, 2001; Meissner et al., 2005; Ramsahoye et al., 2000), we do observe Dnmt3a-dependent 5(h)mC within the CA context at the Dlx5/6 intergenic region (roughly corresponding to a previously characterized enhancer (Zerucha et al., 2000)).

Our finding that Dnmt3a-dependent patterning is strongly associated with gene elements and with accessible transcription factor domains suggests that the enzyme generally modifies accessible regions of the neuronal genome. A recent study described a positive relationship between chromatin accessibility and neuronal DNA methylation in the CH dinucleotide context (Lister et al., 2013), which was hypothesized to be mediated by Dnmt3a. To examine this association, we compared 5mC and 5hmC levels in Dnmt3a WT and KO mOSNs with DNase I accessibility in Dnmt3a WT MOEs. In the absence of Dnmt3a, regions flanking DNase I hypersensitive sites in the MOE exhibit significant depletions of 5mC (Fig. 4A, mean log2(KO/WT)±SD RPKM = −0.09±0.52, p=1E-21, two-sided Wilcoxon) and 5hmC (−0.15±0.39, p=2E-48), and Dnmt3a-dependent 5mC and 5hmC peak sets significantly overlap with the set of accessible regions (Fig. 4B, p=1E-9 and 3E-259, respectively, Fisher's exact test). Moreover, global chromatin accessibility positively correlates with Dnmt3a-dependent 5mC (10 kb windows, Pearson R=0.22, p < 1E-4, permutation test) and 5hmC (R=0.36, p < 1E-4) patterning (Figs. 4C and D). In a previous study (Colquitt et al., 2013), we overexpressed Tet3 in mOSNs (Tet3-tg) and found that regions with high levels of 5hmC in WT mOSNs exhibited depletions of the mark in Tet3-tg

mOSNs, consistent with the role of Tet3 in progressive oxidation of modified cytosines (Wu and Zhang, 2011). Incorporating these data with the Dnmt3a data here, we find a positive correlation between Tet3-tg/wildtype differential 5hmC and Dnmt3a-dependent 5hmC (Pearson R = 0.46, p < 1E-4, Figs. 4C and D), suggesting that Dnmt3a operates on genomic loci that are also sensitive to neuronal Tet-mediated oxidation.

As noted above, Dnmt3a-dependent 5mC and 5hmC are significantly associated with potential MOE enhancers (Fig. 1H), defined as intergenic H3K4me1 peaks. We find positive correlations at these regions between flanking 5hmC levels and both H3K27ac levels (R=0.33, p < 1E-4, permutation test) and DNase hypersensitivity (R=0.22, p < 1E-4, Fig. 5A) but only weak correlations between these features and flanking 5mC levels (R=−0.04 and 0.07, respectively, p < 1E-4 for both). To further investigate the relationship between enhancer activity and 5hmC levels, we compared enhancer 5hmC levels and the expression of nearby genes along the mOSN developmental lineage. We clustered the 5000 potential enhancers with the highest 5hmC variance across HBCs, GBCs, and mOSNs by their 5hmC levels (data from Colquitt et al., 2013), producing 8 clusters each representing a developmental 5hmC path (Fig. 5B) and computed the median expression levels of nearby genes. In support of an active role for 5hmC in enhancer function, we find a striking correspondence between developmental enhancer 5hmC patterning and nearby gene expression.

Enhancers have been previously classified as ‘poised’ (competent for activation, enriched for H3K4me1) or ‘active’ (engaged with its target gene, enriched for both H3K4me1 and H3K27ac) (Creyghton et al., 2010). To generate a similar classification for MOE enhancers, H3K27ac peaks were intersected with intergenic H3K4me1 peaks, yielding H3K27ac-negative H3K4me1 peaks (‘poised’) and H3K27ac-positive peaks (‘active’). We find that 5mC levels in Dnmt3a KO mOSNs are reduced relative to WT flanking enhancers regardless of classification (Fig. 5C). Similarly, 5hmC is depleted from the flanking regions of both poised and active enhancers in Dnmt3a KOs, although the depletion is strongest at regions flanking active enhancers. As with DNase hypersensitive sites in general, 5hmC in Dnmt3a Het mOSNs is present at intermediate levels flanking both poised and active enhancers.

The reduction of 5mC and 5hmC levels at active and poised enhancers in the absence of Dnmt3a may indicate that the enzyme modifies all enhancers or only those that are active within neural tissues. To resolve this, we computed Dnmt3a KO versus WT 5mC and 5hmC ratios at VISTA enhancers, functionally validated regulatory elements whose tissue activities have been determined by mouse reporter transgenics (Visel et al., 2007), and found that those with Dnmt3adependent 5mC or 5hmC patterning are strongly associated with activity in the nose and other neural or face tissues (forebrain, midbrain, hindbrain, and neural tube, Fig. 5D). Conditioning these enhancers by their tissue activity indicates that, although most enhancer types exhibit reduced 5mC and 5hmC in Dnmt3a KO mOSNs, those with neural activity display most substantial reductions (Fig. 5E). Moreover, VISTA enhancers with at least two-fold reduction of 5hmC are significantly associated with homeodomain transcription factors implicated in neurodevelopment (Table S4).

To determine if Dnmt3a-dependent 5mC and 5hmC patterning is critical for the proper regulation of neuronal gene expression programs, we prepared RNA-seq libraries from FACS-isolated Dnmt3a WT, Het, and KO mOSNs with two biological replicates for each condition (Tables S1 and S5). Overall, reads per kilobase per million reads (RPKM) values from each genotype are highly correlated (Fig. S4A). To assess global gene expression, we averaged RPKM values across genotypes and replicates to obtain a transcriptional index. Plotting expression fold change against this index, we find a striking bidirectional change of expression in Het and KO relative to WT mOSNs that is dependent on transcriptional output (Fig. 6A and B). On average, lowly expressed or silent genes (as determined by Gaussian-based clustering of RPKMs) are transcriptionally upregulated in KO mOSNs (mean log2(KO/WT RPKM)±SD = 0.9±1.1) while highly expressed genes are weakly downregulated (−0.3±0.4). Moreover, genes with an absolute KO-WT fold change of at least 2-fold exhibit an intermediate disruption of expression levels in Het mOSNs, in concordance with Dnmt3a dosage response of 5mC and 5hmC levels (Fig. 6C). To assess how Dnmt3a-sensitive genes are regulated during mOSN development, we clustered these gene sets by their expression in Dnmt3a wildtype mOSNs, GBCs, and HBCs, using previously published datasets (Fig. S4B and C) (Colquitt et al., 2013; Magklara et al., 2011) and find that downregulated genes (N=387) are expressed in mOSNs (Fig. S4B) while upregulated genes (N=1,845) are typically expressed in HBCs, GBCs, or not at all in the MOE (Fig. S4C).

To independently verify these effects, and to gauge if the general effects seen in the RNA-seq data are due to normalization artifacts, we performed RT-qPCR using RNA isolated from FAC-sorted Dnmt3a WT, Het, and KO mOSNs (N=3, each genotype). We tested the expression of eleven genes predicted to be downregulated, four genes predicted to be upregulated, and four predicted to remain the same according to KO/WT fold changes from the RNA-seq datasets (Fig. 6D). Of these genes, the expression of seven downregulated genes, one upregulated gene, and no unchanged genes displayed significant dependencies on Dnmt3a genotype ($p < 0.05$, one-way ANOVA). The remaining three upregulated genes exhibit a trend toward increased expression in Dnmt3a-deficient mOSNs by RT-qPCR. Furthermore, expression estimates from the RNA-seq and RT-qPCR datasets are strongly correlated (Fig. 6E, R=0.76, p=1E-4, Pearson correlation test). Finally, *in situ* hybridizations (ISH) using probes against genes that are expressed in mOSNs yet exhibit reduced expression in Dnmt3a KOs by RNA-seq (*Omp* and *Lrrc23*) display reduced expression (Fig. 6F and S4D).

As described above, Dnmt3a-dependent 5mC and 5hmC patterning is strongly associated with gene elements. To determine if modification changes correlate with transcriptional alterations, we examined the relationships between genic 5mC and 5hmC levels and expression. Grouping genes by Dnmt3a WT mOSN expression quartiles indicates that 5mC and 5hmC levels are largely maintained across genes (Fig. S4E). However, TSS-flanking 5mC is decreased in Dnmt3a KO mOSNs at moderately expressed genes. In contrast, 5hmC levels are steadily reduced across the gene body (with a bias toward the 5’ end) with increasing levels of expression (Fig. S4E). To determine how the alteration of 5mC and 5hmC levels across genes correlates with changes in expression, we computed correlation values between 5mC or 5hmC KO/WT fold-change across genes and the corresponding

expression fold-change (Fig. 6G). After adjusting for multiple comparisons, we find that Dnmt3a-dependent 5mC levels surrounding the TSS are negatively correlated with expression changes, consistent with its repressive role. In contrast, Dnmt3a-dependent 5hmC levels across the gene body are positively correlated with expression changes in Dnmt3a KO mOSNs while those present at the TSS are negatively correlated, consistent with a dual regulatory function (Colquitt et al., 2013; Wu et al., 2011). Ordering average 5mC levels at the TSS and 5hmC levels at the gene body by KO/WT expression fold-change demonstrates striking differences between the two modifications (Fig. 6H): 5mC levels are mildly reduced in Dnmt3a KO mOSNs at the TSSs of upregulated genes, while 5hmC levels display Dnmt3a dosage-sensitive depletions within downregulated genes. Two representative upregulated (*Utg2a1*) and downregulated (*Kirrel*) genes (RT-qPCR expression levels found in Fig. 6D) provide examples of this relationship (Fig. 6I and J): 5mC levels are reduced at the TSS of *Utg2a1* and 5hmC levels are reduced within the gene body of *Kirrel*.

Hierarchical clustering of Dnmt3a-dependent block gene expression demonstrates a similar bidirectional disruption of transcription (Fig. S4F). The two main fold-change expression clusters (upregulated, cluster 1; down-regulated, cluster 2) are not significantly enriched for particular gene ontology terms with respect to each other, but, as with the full set of genes, lowly expressed genes tend to exhibit increased expression while active genes tend to display reduced expression in Dnmt3a KO mOSNs (Fig. S4G). Moreover, in agreement with the association of 5mC with repressed and 5hmC with active transcriptional states, 5mC is enriched in regions associated with cluster 1 (repressed and upregulated) while 5hmC is enriched in regions associated with cluster 2 (active and downregulated, Fig. S4G).

A previous study observed increased levels of the repressive histone modification H3K27me3 at the TSSs of genes downregulated in neural stem cells in the absence of Dnmt3a (Wu et al., 2010). This increase was argued to result from the inappropriate spread of Polycomb complex activity, which generates H3K27me3, onto TSSs that are normally protected by Dnmt3a-mediated 5mC. To determine if a similar increase occurs within the olfactory epithelium, we performed H3K27me3 ChIP-seq using Dnmt3a KO MOEs. In contrast to the previous report, we do not observe an increase of H3K27me3 levels at genes with reduced expression in KO mOSNs and instead detect a modest depletion of H3K27me3 at the TSSs of genes with increased expression (Fig. S4H, median RPKM±SD: WT = 0.37±1.33, KO = 0.31±1.09, p = 1E-4, Wilcoxon two-sided).

These expression defects led us to investigate transcriptional responses to odorant-induced neuronal activation in Dnmt3a-deficient animals. In particular, we hypothesized that the mild transcriptional effects observed in Dnmt3a KO mOSNs would be exacerbated upon gene induction. Initially, we exposed Dnmt3a WT mice to a concentrated mixture of previously characterized odorants (1:1:1 amyl acetate:acetophenone:octanal, am:ac:oc) (Mombaerts, 2004; Rubin and Katz, 1999) and assayed transcriptional levels in their MOEs relative to those of mice exposed to water (Fig. 7A). We detected maximal induction of the immediate early genes *Fos*, *Egr1*, and *Nr4a1* after one hour exposure (Fig. 7B and S5A). The induction of *Egr1* occurs broadly in the mOSN layer (Fig. 7C). To globally assay odorant-induced transcription, we prepared two replicate RNA-seq libraries from MOEs of Dnmt3a WT, Het, and KO mice that had been exposed to one hour of am:ac:oc or water

(Tables S1 and S6). Across genotypes, we identified 23 genes that were significantly upregulated and one gene that was downregulated (FDR<0.05) in response to odorant exposure (Table S7). Among these genes was the transcription factor *Nr4a2*, which strikingly exhibits both depletions of 5mC and 5hmC levels in its gene body and the reduction of odorant-induced expression in Dnmt3a KOs (Fig. 7D). In contrast, *Pcdh10*, which is induced to wild type levels in the Dnmt3a KO (Fig. 7D and S5B), has wild type levels of DNA modifications in the KO OSNs (Fig. 7D). On average, the fold transcriptional response to odorant of the set of 23 genes is significantly reduced in Dnmt3a KOs (median log2(Het/WT) RPKM±SD: −0.93±1.30, p = 1E-3, two-sided Wilcoxon, Fig. S5C), and their gene bodies are significantly depleted of both 5mC (p = 4E-5) and 5hmC (p = 2E-4) (Fig. S5D). Interestingly, although odorant-induced transcription levels are reduced in Dnmt3a Hets (Fig. 7E, median log2(Het/WT) RPKM±SD: −0.22±0.29, p = 3E-3, paired two-sided Wilcoxon) and KOs (−0.94±1.23, p = 7E-4), basal levels in both genotypes remain unchanged (Het: −0.08±0.46, p = 0.16 ; KO: 0.03±0.39, p = 0.84).

This association between reduced 5mC, 5hmC, and reduced gene activation generalizes across odorant-induced genes (Fig. 7F and G). Dnmt3a KO/WT levels of 5mC and (to a lesser extent) 5hmC in the bodies of odorant-induced genes are positively correlated with KO/WT fold-change odorant-induced RNA levels (5mC: Pearson's R = 0.45, p = 0.02, permutation test; 5hmC: R = 0.16, p=0.21). Strikingly, unlike the mild global reduction of expressed genes in Dnmt3a KO mOSNs (log2(KO/WT RPKM) of −0.3), we find a median −0.9 log2 reduction of odorant-induced expression in Dnmt3a KO relative to WT (p = 2E-3, Wilcoxon two-sided), indicating that activity-dependent responses are particularly sensitive to the loss of the enzyme (Fig. 7H). Several lines of evidence argue against these effects being a result of a general disruption of odorant detection. The loss of Dnmt3a does not affect the protein levels of several markers of mOSN identity, including the major mOSN signaling components (Adcy3 and Gnal), the mOSN axon guidance molecule Cntn4, and the glutamate vesicular transporter Slc17a6 (Fig. S5E). Similarly, glomerulus formation of the most highly expressed olfactory receptor, *Olfr1507*, is unaffected by the absence of Dnmt3a (Fig. S5F). Finally, phosphorylation of the S6 ribosomal subunit, a marker for neuronal activity (Knight et al., 2012), is induced to similar levels in both Dnmt3a WT and KO animals exposed to am:ac:oc for one hour (Fig. S5G).

Discussion

The patterning of 5-methylcytosine and 5-hydroxymethylcytosine during neurodevelopment is dynamic, and it is likely that these changes inform the development and function of neurons. Our comprehensive analysis of how the lack of Dnmt3a, a *de novo* DNA methyltransferase strongly implicated in neurodevelopment, influences the epigenetic and transcriptional states of mOSNs provides insight into how DNA modifications contribute to defining the neuronal state. Although *de novo* DNA methyltransferase activity is completely absent in Dnmt3a KO mOSNs, the vast majority of 5mC and 5hmC patterning is unaffected by the loss of the enzyme. Instead, Dnmt3a is required to pattern defined components of the 5mC and 5hmC landscapes of mOSNs that are strongly associated with accessible regions of the genome. Recent work described a global build up of mCH within neurons in the forebrain (Lister et al., 2013) and dentate gyrus (Guo et al., 2013) that parallels Dnmt3a

expression. Moreover, in the dentate gyrus methylation in this dinucleotide context but not the CG context is partially dependent on Dnmt3a. The apparent stability of much of mOSN 5mC patterning in the absence of Dnmt3a could be due to relatively low levels of mCH within this peripheral neuronal population or to the poor sensitivity of the antibody-based approach used here in detecting low levels of modification. Moreover, we cannot exclude that the antibody we used in these experiments has significantly lower affinity for mCH vs mCG. Genome-wide bisulfite sequencing will be needed to resolve this issue. In addition, it will be important to assay DNA modification patterning in GBCs (OSN progenitors) in the absence of Dnmt3a to specifically identify sites of Dnmt3a activity during differentiation. Furthermore, because the *Dnmt3a* mutant allele is constitutive, we cannot rule out possible compensatory effects of Dnmt3b, the other mammalian *de novo* DNA methyltransferase, during earlier developmental timepoints, including in HBCs, the multipotent stem cell population of MOE. However, both Dnmt1 and Dnmt3b expression is largely unaltered in the MOEs of Dnmt3a KO mice, making it unlikely that the persistence of 5mC or 5hmC in this tissue is due to atypical activities of the other methyltransferases. Nonetheless, the constitutive mutation makes it difficult to formally rule out the influence of other factors, such as the alteration of epigenetic patterning during early development, that may have indirectly produced the observations presented here.

Interestingly, Dnmt3a is required to fully generate the elevated 5hmC levels established during mOSN development. This dependency suggests that *de novo* 5mC in neurons is targeted by the Tet oxidation machinery. In support of this model, we also find that Dnmt3a-dependent 5mC and 5hmC patterning is positively correlated with Tet3-sensitive 5hmC patterning. In neurons, 5hmC occurs nearly exclusively on cytosines in CG dinucleotides while *de novo* 5mC occurs broadly in multiple contexts (Lister et al., 2013). This suggests that Dnmt3a, which has been shown to provide both CG and CH methylation in non-neural systems (Meissner et al., 2005; Ramsahoye et al., 2000), generates a 5mC platform on which neuronal Tet activity can provide elevated levels of 5hmC. Alternatively, 5mC oxidation may occur broadly across cytosine contexts but with less processivity in CG contexts, resulting in higher levels of 5hmC, than on those in CH contexts. Recent structural data and *in vitro* enzymatic assays of human TET2 indicate that the Tet enzymatic domain specifically oxidizes 5mC within the CG context, in support of the former model (Hu et al., 2013). However, these assays have not been extended to mouse Tet3, which is the dominant Tet family member in mOSNs (Colquitt et al., 2013), or to Tet1.

We describe a set of broad domains that contain Dnmt3a-dependent 5mC and 5hmC patterning, which we term Dnmt3a-dependent blocks. These regions are strongly associated with key genomic loci, including genes and regulatory elements associated with neuronal development. In particular, approximately 200 transcription factors are embedded within or near these Dnmt3a-dependent blocks. The spatial correspondence between accessibility and Dnmt3a-mediated patterning suggests either that Dnmt3a is only able to modify open genomic regions or that Dnmt3a itself establishes zones of chromatin accessibility. This patterning is reminiscent of recently characterized ‘super-enhancers’, which are defined by the extensive occupancy of transcription factors and structural proteins implicated in mediating enhancer-promoter contacts as well as the enrichment of active histone

modifications (Hnisz et al., 2013; Whyte et al., 2013). Although the function of these regions is unclear, super-enhancers are strongly associated with transcription factors involved in the specification of cell identity. Moreover, much like Dnmt3a-dependent blocks, super-enhancers often span transcription factor loci, thus establishing large, permissive regulatory domains across these regions. These likenesses suggest that Dnmt3a-mediated DNA modification patterning is an additional feature of super-enhancers; however, what role this patterning plays in defining the function of these enhancers remains to be determined.

The finding that 5mC and 5hmC levels depend on Dnmt3a dosage, evidenced by intermediate 5mC and 5hmC levels in heterozygotes, argues that either Dnmt3a protein level is limiting or that Dnmt3a-modified sites are engaged in a methylation-demethylation cycle, rendering them vulnerable to fluctuations in Dnmt3a activity. Interestingly, we find that these DNA modification effects are propagated to intermediate levels of gene expression in Dnmt3a Het mOSNs. This is surprising considering standard models of DNA methylation function in which the methylation of key CGs alters the affinity of protein trans factors. Instead, these data support a model in which the density of 5mC and its derivative 5hmC – and not the modification of key positions – is the functionally relevant output of Dnmt3a-mediated patterning. Indeed, the mosaicism of 5mC or 5hmC in Dnmt3a heterozygotes, the distributive activity of the enzyme (Gowher and Jeltsch, 2001), as well as its ability to methylate cytosines in multiple contexts suggest that Dnmt3a stochastically methylates targeted sites. This behavior may permit the fine-tuning of gene expression levels through intermediate DNA modification patterning.

In keeping with such fine regulation, the absence of Dnmt3a in mOSNs causes bidirectional alterations to gene expression profiles, through the upregulation of repressed genes and downregulation of expressed genes. Traditionally associated with transcriptional repression, DNA methylation is increasingly appreciated as a bifunctional modification depending on its genomic context and interpretation. In particular, gene body methylation in *Arabidopsis thaliana* (Lister et al., 2008; Zilberman et al., 2008), the mammalian X chromosome (Hellman and Chess, 2007), and embryonic stem cells (Lister et al., 2009) is positively correlated with expression. Similarly, although 5hmC at the TSS is negatively correlated with transcription in neurons, 5hmC levels flanking TSSs and within gene bodies generally positively correlate, and to a greater degree than 5mC, with transcriptional output (Colquitt et al., 2013; Hahn et al., 2013; Mellén et al., 2012; Song et al., 2011; Szulwach et al., 2011). We observed in previous work that overexpressing Tet3 in mOSNs resulted in either the down- or up-regulation of transcription depending on whether 5hmC was decreased or increased, respectively, in gene bodies (Colquitt et al., 2013). Accordingly, reduced gene body 5hmC in Dnmt3a KO mOSNs correlates with mildly reduced transcriptional output, consistent with 5hmC playing a facilitating role in transcription. Here, despite the reduction of 5hmC levels in the absence of *de novo* DNA methyltransferase activity, the bulk of 5hmC patterning in mOSNs is unaltered, indicating that most of this patterning is stable in this neuronal population. This strongly suggests that neuronal 5hmC is not exclusively an intermediate in Tet-mediated demethylation but is itself stably maintained in neurons. Recent characterizations of proteins that bind 5hmC, including MBD3 (Yildirim et al.,

2011), MeCP2 (Mellén et al., 2012), and a diverse array of other factors (Spruijt et al., 2013), suggest that the modified base is actively interpreted and may directly influence gene regulation.

Finally, our finding that odorant-induced transcriptional responses require Dnmt3a, combined with the observation that substantial *de novo* 5mC occurs in neurons around the time of synaptogenesis (Lister et al., 2013), argue that *de novo* neuronal DNA modification patterns are critical to regulate activity-sensitive transcriptional programs. This connection is supported by experiments indicating that MeCP2 (methyl-CpG binding protein 2), which binds both 5mC and 5hmC in neurons (Mellén et al., 2012), is phosphorylated by neuronal activity (Cohen et al., 2011; Zhou et al., 2006), indicating that a neuron's activity state is communicated to one of the primary interpreters of the epigenome. Furthermore, DNA modification patterns associated with activity-dependent genes are sensitive to neuronal activity in the hippocampus (Guo et al., 2011a, 2011b). Interestingly, despite significant reductions of 5mC and 5hmC levels on activity-regulated genes in the absence of odorant stimulation, basal transcription levels are unaffected by Dnmt3a deletion. Transcriptional perturbations of these genes only become apparent with gene induction. This relationship suggests that a Dnmt3a-mediated DNA modification signature on these activity-regulated genes is required for enhanced transcription upon neuronal stimulation. Taking into account recent observations suggesting unexpected plasticity in the peripheral olfactory system (Abraham et al., 2014; Kass et al., 2013), our data may provide insight into the molecular underpinnings of activity-induced transcription in olfactory neurons. In effect, Dnmt3a may establish an epigenetic landscape that prepares the neuronal genome for future transcriptional responses, supporting a crucial role for DNA modifications in neuronal plasticity.

Experimental Procedures

Animal care and use. Mice were treated in compliance with the rules and regulations of IACUC under protocol approval number AN099395-01. Dnmt3a constitutive null mice (Okano et al., 1999) were obtained from MGI (2182412).

Fluorescence activated cell sorting. Mature OSNs were isolated from the MOEs of OMP-ires-GFP knockin animals (Shykind et al., 2004). Single cell suspensions from 1–10 MOEs were made using the Papain dissociation system (Worthington): MOE were dissected into Earl's Buffered Saline Solution (EBSS), minced in papain dissociation solution (Worthington), and incubated at 37°C for 45 minutes. Cells were washed once in 1:10 inhibitor solution and once in PBS. For a given assay (DIP-seq or RNA-seq), 2E5 to 1E6 cells were used. Population purity was assayed by resorting a small fraction of the sorted cells. Purities of ~90–95% were routinely achieved.

DNA immunoprecipitation (DIP)-seq. Genomic DNA was isolated using the DNeasy Blood and Tissue Kit (Qiagen) from the cell populations described in the main text, prepared for paired-end Illumina sequencing using 75–200 ng gDNA and standard protocols, and immunoprecipitated as in (Weber et al., 2005) with minor modifications. See Supplemental Experimental Procedures for details.

Chromatin immunoprecipitation (ChIP)-seq. Nuclei were isolated from the MOEs of three-week old Dnmt3a wildtype and knockout mice and native ChIP was performed as described (Magklara et. al, 2011) using anti-H3K27ac (Millipore, cma309), anti-H3K4me1 (Abcam, ab8895), and anti-H3K27me3 (Millipore, 07-449) antibodies. For library preparation, ChIP DNA was sonicated for 120–180 seconds on a Covaris S220 and prepared for sequencing using the Ovation Ultralow Library Kit (Nugen).

Odorant exposure. Mice were divided into two clean cages containing no bedding and placed into separate dark cabinets. After 10 minutes in these cages, a 4 cm×4 cm piece of Whatman paper with either 80 μL of water or a 1:1:1 mixture of pure amyl acetate:acetophenone:octanal (all purchased from Sigma-Aldrich) was placed at one end of the cage. Details for the timing of odorant exposure and associated RNA analysis are given in the Extended Experimental Procedures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Sequencing data is deposited in GEO under accession GSE52464. This work was supported by the NIH (DP2 OD006667, R01MH091661), NSF (Graduate Research Fellowship), and the Rett Syndrome Trust. BC designed the research project, performed most experiments, analyzed the data, and wrote the manuscript. EMP performed the DNase-seq and several ChIP-seq assays. RD performed the bisulfite sequencing and analysis. SL designed the research project and wrote the manuscript.

References

Abraham NM, Vincis R, Lagier S, Rodriguez I, Carleton A. Long term functional plasticity of sensory inputs mediated by olfactory learning. Elife. 2014; 3

Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010; 11:1–28.

Caggiano M, Kauer JS, Hunter DD. Globose basal cells are neuronal progenitors in the olfactory epithelium: a lineage analysis using a replication-incompetent retrovirus. Neuron. 1994; 13:339–352. [PubMed: 8060615]

Cohen S, Gabel HW, Hemberg M, Hutchinson AN, Sadacca LA, Ebert DH, Harmin DA, Greenberg RS, Verdine VK, Zhou Z, et al. Genome-wide activity-dependent MeCP2 phosphorylation regulates nervous system development and function. Neuron. 2011; 72:72–85. [PubMed: 21982370]

Colquitt BM, Allen WE, Barnea G, Lomvardas S. Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. Proc. Natl. Acad. Sci. U. S. A. 2013; 110:14682–14687. [PubMed: 23969834]

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. U. S. A. 2010; 107:21931–21936. [PubMed: 21106759]

Feng J, Chang H, Li E, Fan G. Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system. J. Neurosci. Res. 2005; 79:734–746. [PubMed: 15672446]

Feng J, Zhou Y, Campbell SL, Le T, Li E, Sweatt JD, Silva AJ, Fan G. Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. Nat. Rev. Neurosci. 2010; 13:423–430.

Globisch DD, Münzel MM, Müller MM, Michalakis SS, Wagner MM, Koch SS, Brückl TT, Biel MM, Carell TT. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. PLoS One. 2010; 5:e15367–e15367. [PubMed: 21203455]

Gowher H, Jeltsch A. Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. J. Mol. Biol. 2001; 309:1201–1208. [PubMed: 11399089]

Guo JU, Ma DK, Mo H, Ball MP, Jang M-H, Bonaguidi MA, Balazer JA, Eaves HL, Xie B, Ford E, et al. Neuronal activity modifies the DNA methylation landscape in the adult brain. Nat. Neurosci. 2011a; 14:1345–1351. [PubMed: 21874013]

Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, Zhong C, Hu S, Le T, Fan G, et al. Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. Nat. Neurosci. 2013; 17:215–222. [PubMed: 24362762]

Guo JUJU, Su YY, Zhong CC, Ming GGL, Song HH. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell. 2011b; 145:423–434. [PubMed: 21496894]

Hahn MA, Qiu R, Wu X, Li AX, Zhang H, Wang J, Jui J, Jin S-G, Jiang Y, Pfeifer GP, et al. Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. Cell Rep. 2013; 3:291–300. [PubMed: 23403289]

Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. Science. 2007; 315:1141–1143. (80-.). [PubMed: 17322062]

Hirota J, Mombaerts P. The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory sensory neurons. Proc. Natl. Acad. Sci. U. S. A. 2004; 101:8751–8755. [PubMed: 15173589]

Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-André V, Sigova AA, Hoke HA, Young RA. Super-Enhancers in the Control of Cell Identity and Disease. Cell. 2013; 155:934–947. [PubMed: 24119843]

Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, Shi YG, Zhu J, Wang P, Xu Y. Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. Cell. 2013; 155:1545–1555. [PubMed: 24315485]

Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS One. 2010; 5:e8888. [PubMed: 20126651]

Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature. 2010; 466:1129–1133. [PubMed: 20639862]

Kass MD, Rosenthal MC, Pottackal J, McGann JP. Fear learning enhances neural responses to threat-predictive sensory stimuli. Science. 2013; 342:1389–1392. [PubMed: 24337299]

Knight ZA, Tan K, Birsoy K, Schmidt S, Garrison JL, Wysocki RW, Emiliano A, Ekstrand MI, Friedman JM. Molecular profiling of activated neurons by phosphorylated ribosome capture. Cell. 2012; 151:1126–1137. [PubMed: 23178128]

Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science. 2009; 324:929–930. (80-.). [PubMed: 19372393]

Leung CT, Coulombe PA, Reed RR. Contribution of olfactory neural stem cells to tissue maintenance and regeneration. Nat. Rev. Neurosci. 2007; 10:720–726.

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell. 2008; 133:523–536. [PubMed: 18423832]

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo Q-M, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009; 462:315–322. [PubMed: 19829295]

Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, et al. Global Epigenomic Reconfiguration During Mammalian Brain Development. Science. 2013; 341:1237905. (80-.). [PubMed: 23828890]

MacDonald JL, Gin CSY, Roskams AJ. Stage-specific induction of DNA methyltransferases in olfactory receptor neuron development. Dev. Biol. 2005; 288:461–473. [PubMed: 16288735]

Magklara A, Yen A, Colquitt BM, Clowney EJ, Allen W, Markenscoff-Papadimitriou E, Evans ZA, Kheradpour P, Mountoufaris G, Carey C, et al. An epigenetic signature for monoallelic olfactory receptor expression. Cell. 2011; 145:555–570. [PubMed: 21529909]

Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res. 2005; 33:5868–5877. [PubMed: 16224102]

Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Neuron. 2012; 151:1417–1430.

Mombaerts P. Genes and ligands for odorant, vomeronasal and taste receptors. Nat. Rev. Neurosci. 2004; 5:263–278. [PubMed: 15034552]

Nguyen S, Meletis K, Fu D, Jhaveri S, Jaenisch R. Ablation of de novo DNA methyltransferase Dnmt3a in the nervous system leads to neuromuscular defects and shortened lifespan. Dev. Dyn. 2007; 236:1663–1676. [PubMed: 17477386]

Nicolay DJ, Doucette JR, Nazarali AJ. Transcriptional regulation of neurogenesis in the olfactory epithelium. Cell. Mol. Neurobiol. 2006; 26:803–821. [PubMed: 16708285]

Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999; 99:247–257. [PubMed: 10555141]

Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc. Natl. Acad. Sci. U. S. A. 2000; 97:5237–5242. [PubMed: 10805783]

Rubin BD, Katz LC. Optical imaging of ddorant representations in the mammalian olfactory bulb. Neuron. 1999; 23:499–511. [PubMed: 10433262]

Ruzov A, Tsenkina Y, Serio A, Dudnakova T, Fletcher J, Bai Y, Chebotareva T, Pells S, Hannoun Z, Sullivan G, et al. Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. Cell Res. 2011; 21:1332–1342. [PubMed: 21747414]

Shykind BM, Rohani SC, O'Donnell S, Nemes A, Mendelsohn M, Sun Y, Axel R, Barnea G. Gene switching and the stability of odorant receptor gene choice. Cell. 2004; 117:801–815. [PubMed: 15186780]

Song C-X, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, Li Y, Chen C-H, Zhang W, Jian X, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat. Biotechnol. 2011; 29:68–72. [PubMed: 21151123]

Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PWTC, Bauer C, Münzel M, Wagner M, Müller M, Khan F, et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. Cell. 2013; 152:1146–1159. [PubMed: 23434322]

Szulwach KE, Li X, Li Y, Song C-X, Wu H, Dai Q, Irier H, Upadhyay AK, Gearing M, Levey AI, et al. 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. Nat. Neurosci. 2011; 14:1607–1616. [PubMed: 22037496]

Visel A, Minovitsky S, Dubchak I, Pennacchio LA. VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic Acids Res. 2007; 35:D88–D92. [PubMed: 17130149]

Watanabe DD, Uchiyama KK, Hanaoka KK. Transition of mouse de novo methyltransferases expression from Dnmt3b to Dnmt3a during neural progenitor cell development. Neuroscience. 2006; 142:727–737. [PubMed: 16973295]

Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schübeler D. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat. Genet. 2005; 37:853–862. [PubMed: 16007088]

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell. 2013; 153:307–319. [PubMed: 23582322]

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PAC, Rappsilber J, Helin K. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature. 2011; 473:343–348. [PubMed: 21490601]

Wu HH, Zhang YY. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. Genes Dev. 2011; 25:2436–2452. [PubMed: 22156206]

Wu H, Coskun V, Tao J, Xie W, Ge W, Yoshikawa K, Li E, Zhang Y, Sun YE. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. Science. 2010; 329:444–448. (80-.). [PubMed: 20651149]

Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev. 2011; 25:679–684. [PubMed: 21460036]

Yildirim O, Li R, Hung J-H, Chen PB, Dong X, Ee L-S, Weng Z, Rando OJ, Fazzio TG. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell. 2011; 147:1498–1510. [PubMed: 22196727]

Zerucha T, Stühmer T, Hatch G, Park BK, Long Q, Yu G, Gambarotta A, Schultz JR, Rubenstein JLR, Ekker M. A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain. J. Neurosci. 2000; 20:709–721. [PubMed: 10632600]

Zhou Z, Hong EJ, Cohen S, Zhao W-N, Ho H-YH, Schmidt L, Chen WG, Lin Y, Savner E, Griffith EC, et al. Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. Neuron. 2006; 52:255–269. [PubMed: 17046689]

Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature. 2008; 456:125–129. [PubMed: 18815594]

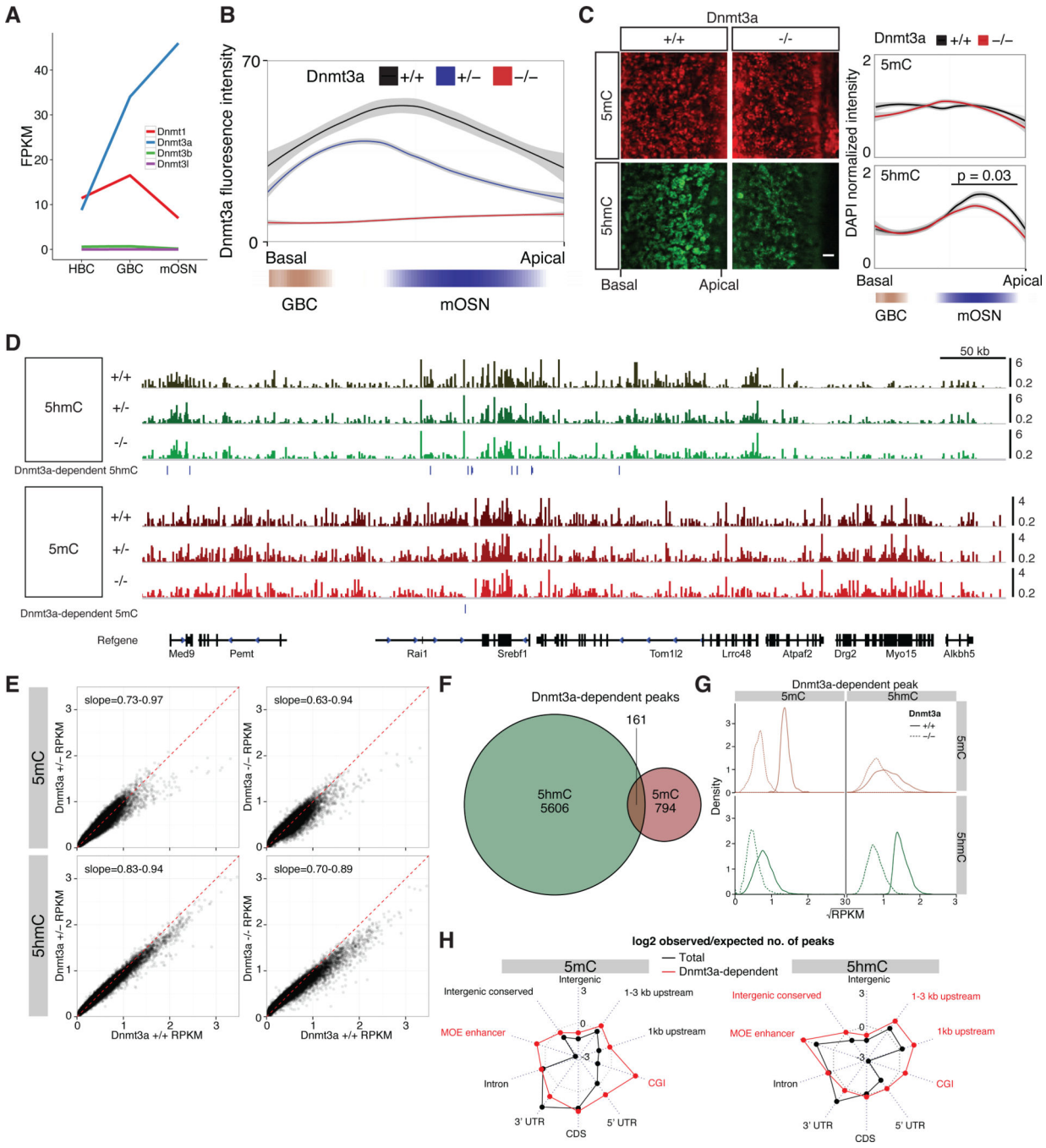

Figure 1.

Dnmt3a is necessary for a defined fraction of neuronal 5mC and 5hmC patterning. (A) RNA-seq fragments per kilobase per million (FPKM) of *Dnmt* family members in horizontal basal cells (HBCs, multipotent stem cells), globose basal cells (GBCs, neuronal progenitors), and mature olfactory sensory neurons (mOSNs, primary sensory neurons). Data from Colquitt et al., 2013. (B) Basal-apical quantification of Dnmt3a immunoreactivity in *Dnmt3a* +/+, +/−, and −/− MOEs. Lines are average intensity across three biological replicates, error is 95% confidence interval. Positions of GBCs and mOSNs determined

using Neurog1-EGFP and OMP-EGFP MOEs, respectively. (C) *Left*, Immunofluorescence of 5mC and 5hmC in *Dnmt3a* +/+ and −/− MOE. *Right*, Quantification of 5mC and 5hmC intensity along basal-apical axis. Lines are average intensity across three biological replicates, error is 95% confidence interval. Scale bar is 20 μm. (D) Representative locus showing 5mC and 5hmC DIP-seq tracks from Dnmt3a WT, Het, and KO mOSNs. Regions with significantly depleted 5mC or 5hmC are indicated below each modification set. (E) Scatter plots of Dnmt3a 5mC and 5hmC levels in Dnmt3a WT, Het, and KO mOSNs. Each point is the average RPKM within non-overlapping genome-wide 10 kilobase (kb) windows. Shown is a random 10% subset of the data. Range of linear slopes across biological replicates is given for each comparison. (F) Overlap between Dnmt3a-dependent 5mC and 5hmC peak sets. (G) Levels of 5mC and 5hmC in either Dnmt3a-dependent 5mC or 5hmC peak sets. RPKM values have been square-root transformed for visual clarity. (H) Genomic feature-peak intersections represented as the ratio of the observed number of peaks that intersect with a given feature set over the expected number of peaks (modeled as peaks randomly positioned across the genome). Shown are intersections for total and Dnmt3a-dependent 5mC or 5hmC peak sets. Red labels indicate features with statistically greater proportional number of Dnmt3a-dependent peaks than total peaks ($p < 0.01$, Fisher exact test). See also Figure S1.

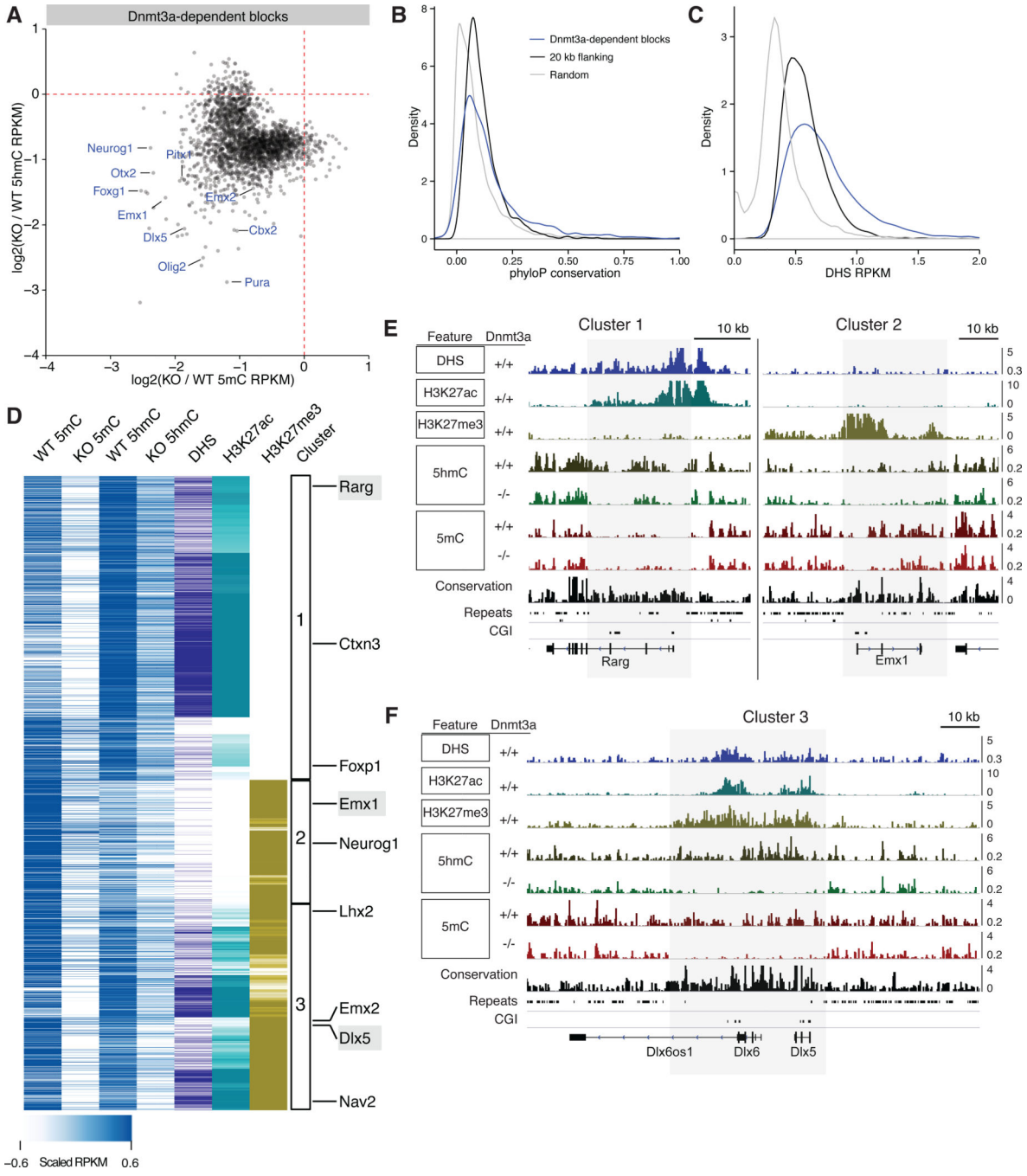

Figure 2.

Dnmt3a is required for 5mC and 5hmC patterning within genomic blocks associated with key neurodevelopmental genes. (**A**) log2(KO/WT RPKM) of average 5mC and 5hmC RPKM within the union of the 5–50 kb regions containing Dnmt3a-dependent 5mC or 5hmC patterning, ‘Dnmt3a-dependent blocks’ (N=2,549). Genes of interest that intersect with these regions are indicated. (**B**) Densities of average vertebrate phyloP conservation across Dnmt3a-dependent blocks, 20 kb flanking both sides of each block, or an equal number of random genomic regions. (**C**) Densities of average DNase I hypersensitivity

(DHS) RPKM in Dnmt3a WT MOEs across the same regions in (**B**). (**D**) Hierarchical clustering of the union of Dnmt3a-dependent 5mC or 5hmC regions by average H3K27ac and H3K27me3 levels. Displayed are average 5mC and 5hmC RPKM in Dnmt3a WT and KO mOSNs and average DHS, H3K27ac and H3K27me3 RPKM in Dnmt3a WT MOE. Each track is input-subtracted and scaled to a common range. Numbered boxes indicate positions of three broad clusters from hierarchical clustering. Representative genes associated with each cluster are indicated at far right. (**E**) Tracks of representative genes found in cluster 1 (*Rarg*) and cluster 2 (*Emx1*). CGIs, CpG islands. Shaded boxes denote regions containing Dnmt3a-dependent 5mC and/or 5hmC. Y-axis is RPKM. (**F**) As (**E**) with representative genes found in cluster 3 (*Dlx5/6*). See also Figure S2.

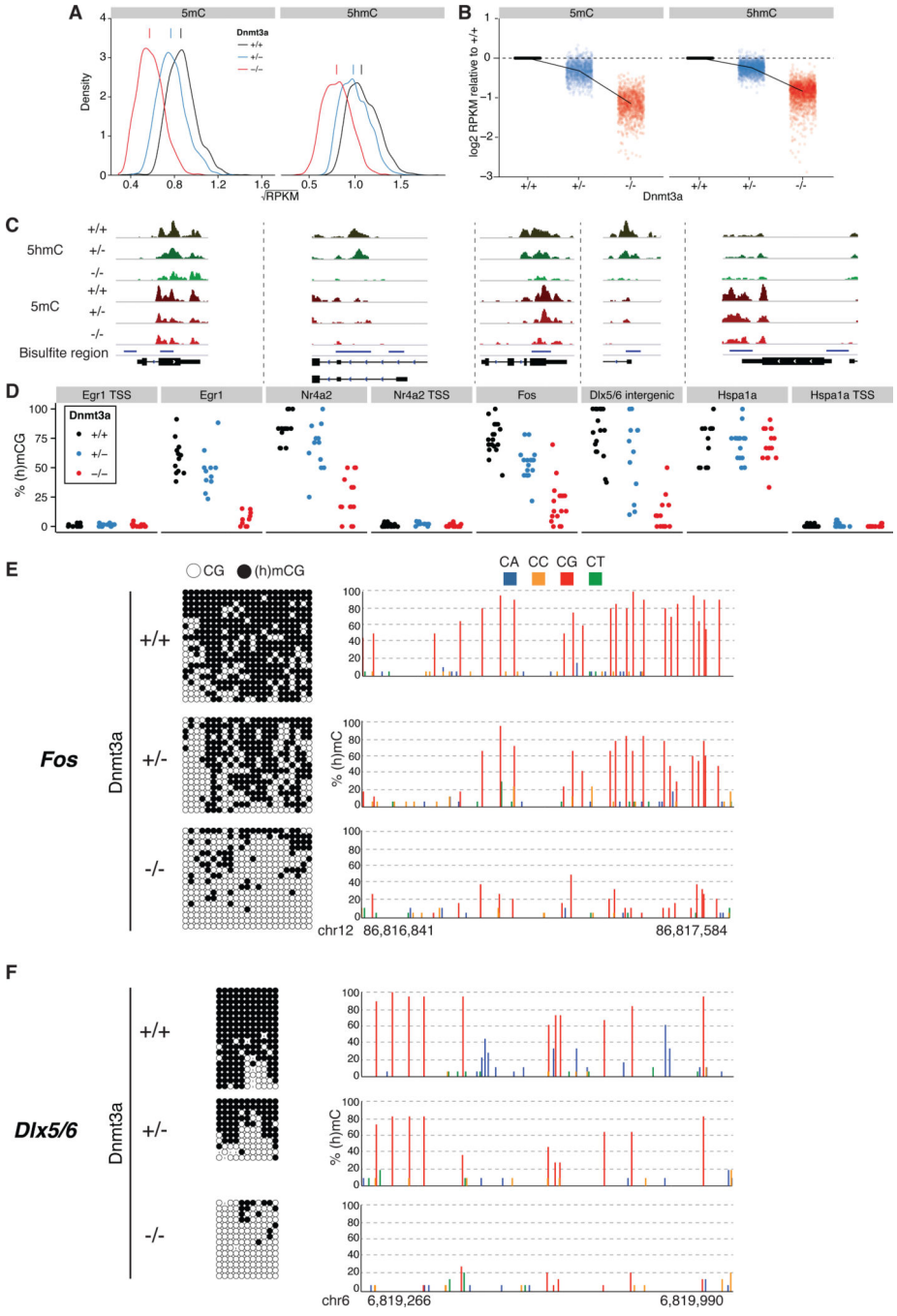

Figure 3.

5mC and 5hmC patterning in Dnmt3a-dependent regions is sensitive to Dnmt3a heterozygosity. (**A**) Densities of Dnmt3a WT, Het, and KO 5mC and 5hmC square-root RPKM (transformed for better visualization) within Dnmt3a-dependent blocks. Ticks indicate distribution medians. (**B**) Average 5mC and 5hmC RPKM normalized to WT within Dnmt3a-dependent blocks. Lines intersect with genotype medians. (**C**) 5hmC and 5mC DIP-seq tracks from Dnmt3a WT, Het, KO mOSNs. Regions analyzed by bisulfite sequencing are indicated below each track. (**D**) Summary of percent modified CGs for each tested locus.

Each point is the average CG modification percentage across individual sequencing clones. (**E**) *Left*, diagram depicting 5(h)mC status of individual CGs (columns) across sequencing clones (rows) for the *Fos* locus. *Right*, percent modification of cytosines across sequencing clones within all four dinucleotide contexts. (**F**) As (**E**) for the *Dlx5/6* intergenic region. See also Figure S3.

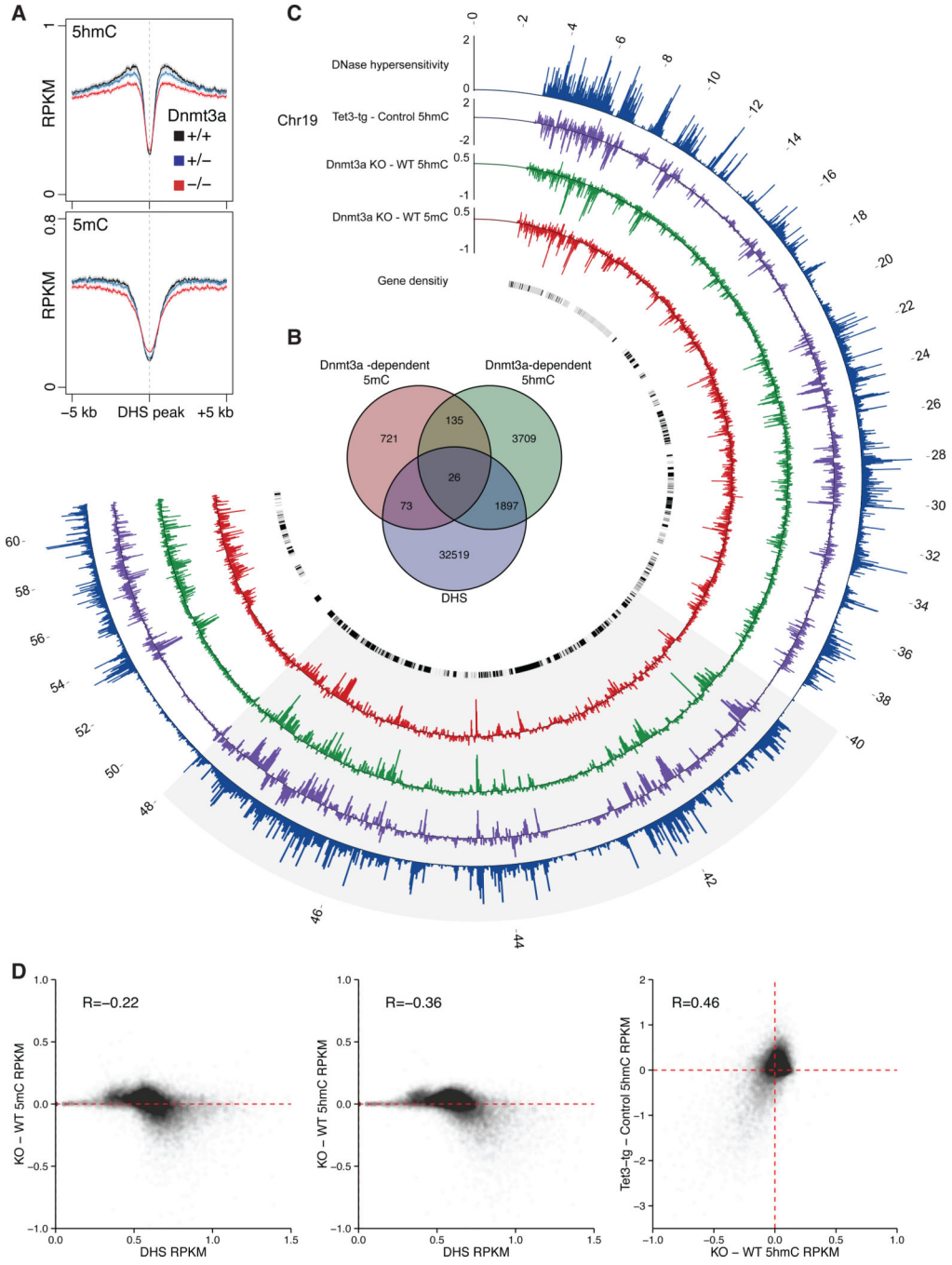

Figure 4.

Dnmt3a-dependent 5mC and 5hmC patterning occurs in accessible chromatin. (**A**) Average 5hmC and 5mC levels spanning MOE Dnmt3a WT DNase hypersensitive sites (N=12,419). Error is bootstrapped 95% confidence intervals. (**B**) Overlap of Dnmt3a-dependent 5mC and 5hmC peak sets with DHS regions. (**C**) Chromatin accessibility and Dnmt3a-dependent DNA modification patterning across chromosome 19. *Outside row*, MOE Dnmt3a WT DNase accessibility. *Second row from outside*, Difference between 5hmC levels in Tet3-tg mOSNs (OMP-tTA×tetO-Tet3-GFP) and control mOSNs (OMP-GFP). Data from Colquitt

et al., 2013. Dnmt3a KO minus WT 5hmC (*third row from outside)* and 5mC (*fourth row from outside*) levels in mOSNs. 5mC and 5hmC values are averages across biological replicates. Y-axes in RPKM. (**D**) Dnmt3a KO – WT 5mC (*left*) and 5hmC (*center*) RPKM versus Dnmt3a WT DNase I hypersensitivity (DHS). (*right*) Tet3-tg – Control mOSN 5hmC RPKM (Colquitt et al., 2013) versus Dnmt3a KO – WT 5hmC RPKM.

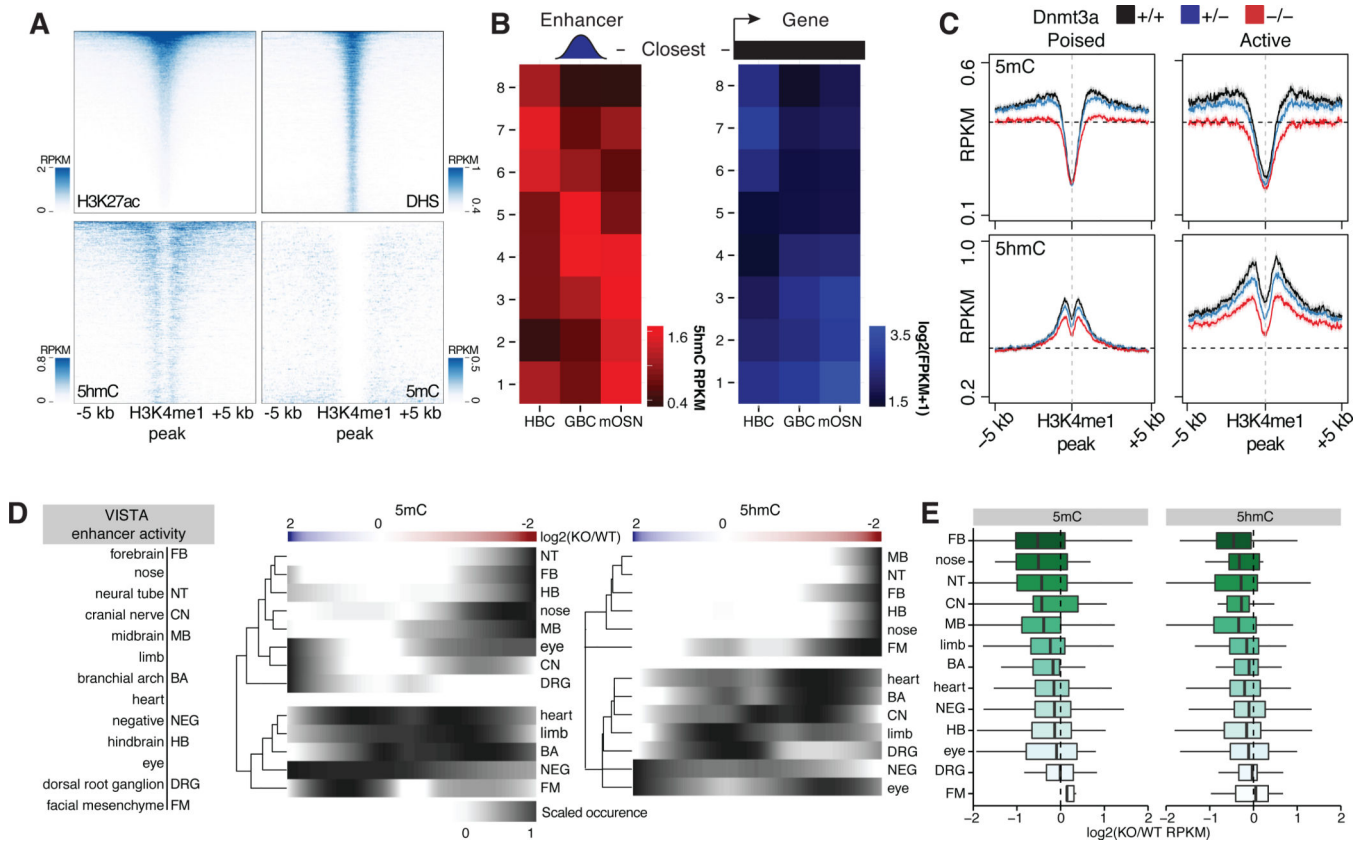

Figure 5.

Dnmt3a is required for 5mC and 5hmC patterning at neuronal enhancers. (**A**) MOE H3K27ac, MOE DNase hypersensitivity (DHS), mOSN 5hmC, and mOSN 5mC levels over MOE intergenic H3K4me1 peaks ordered by H3K27ac levels. (**B**) Analysis of the relationship between enhancer 5hmC levels (*left*) and the expression of closest genes (within 50 kb, *right*) within HBCs, GBCs, and mOSNs. (**C**) Average 5mC and 5hmC levels flanking 'poised' and 'active' enhancers (defined in text) in mature olfactory sensory neurons (mOSNs) from Dnmt3a +/+, +/−, and −/− MOEs. Error bars are bootstrapped 95% confidence intervals. (**D**) Association of enhancer activity with Dnmt3a-dependent 5mC and 5hmC. Shown are matrices of VISTA enhancer activities (rows) ordered by 5mC (*left*) or 5hmC (*right*) log2(Dnmt3a KO/WT RPKM) (columns). Each row was then loess smoothed and normalized, and the matrix was hierarchically clustered. (**E**) Distributions of average 5mC and 5hmC log2(KO/WT RPKM) across VISTA enhancers conditioned by tissue activity.

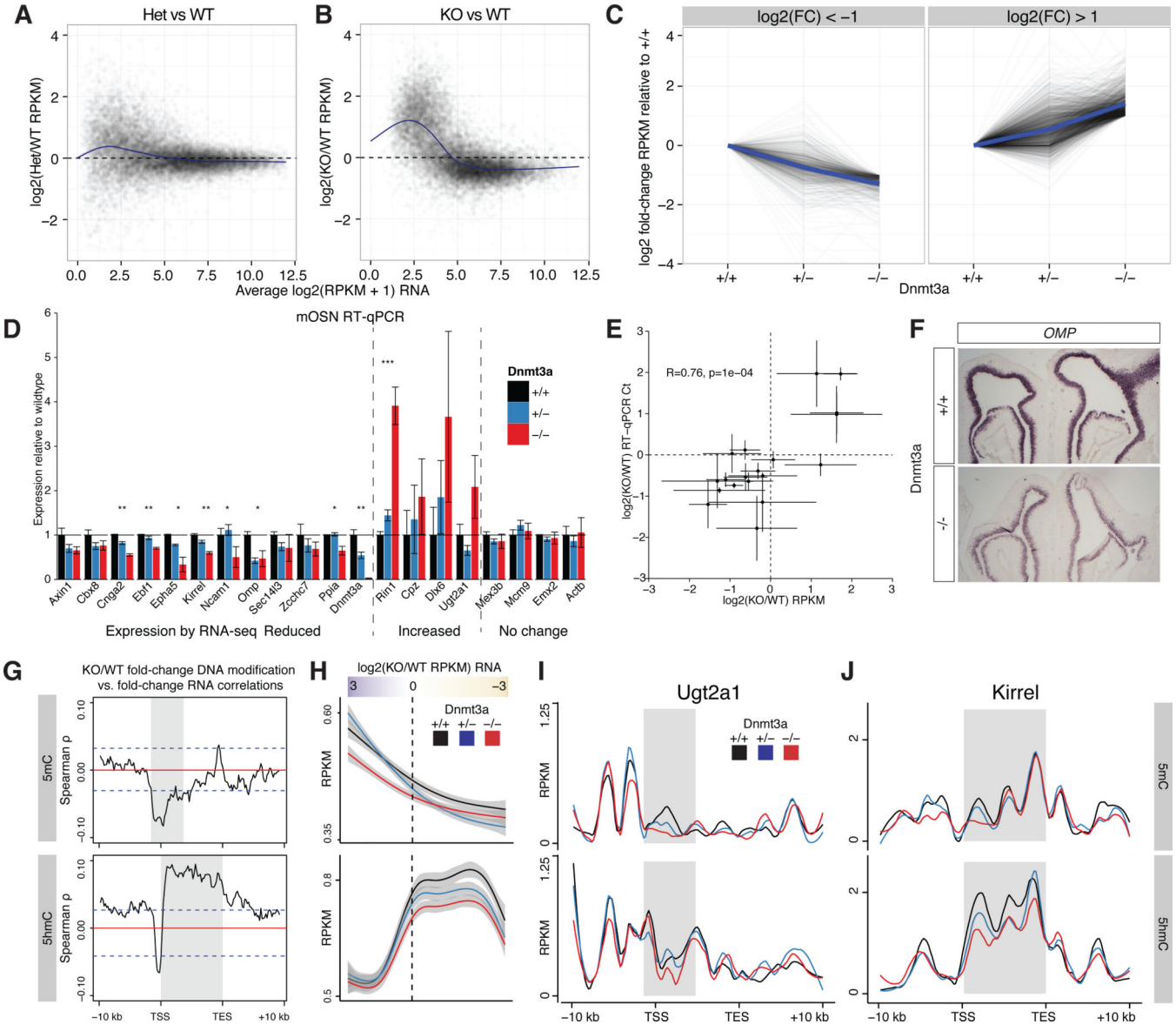

Figure 6.

Loss of Dnmt3a results in the bidirectional disruption of transcription in mOSNs. (**A, B**) Fold-change expression in Dnmt3a Het (**A**) and KO (**B**) mOSNs relative to Dnmt3a WT mOSNs. *X-axis*, average log2(RPKM+1) across Dnmt3a WT and Het (**A**) or WT and KO (**B**) for each gene. *Y-axis*, log2 fold-change for each gene. (**C**) RNA log2(RPKM+1) values in Dnmt3a WT, Het, and KO mOSNs normalized to WT. Displayed are values from genes that have an absolute log2 fold change > 1. (**D**) RT-qPCR using RNA from Dnmt3a WT, Het, and KO mOSNs. Values are normalized to *Tubb5* expression then to WT for each gene. *p<0.05, **p<0.01, ***p<0.001, one-way ANOVA, N=3 each genotype. (**E**) Comparison of fold-change expression estimates by RT-qPCR (*y-axis*) and RNA-seq (*x-axis*). Error bars are SEM for RT-qPCR and the range of two replicates for RNA-seq. (**F**) *In situ* hybridizations (ISH) of Dnmt3a WT and KO coronal MOE sections using probe against *Omp*. (**G**) Spearman rho correlations between Dnmt3a KO/WT fold-change 5mC or 5hmC at each

gene position and fold-change gene expression. Dashed blue lines indicate rho at which the FDR-adjusted p-value < 0.01. Shaded regions indicate positioned used for analysis in (**H**). (**H**) Relationships between average TSS 5mC and gene body 5hmC levels and KO/WT fold-change expression. Lines are loess-smoothed averages, error is 95% confidence interval. (**I** and **J**) Dnmt3a WT, Het, and KO 5mC and 5hmC levels across an upregulated gene (*Ugt2a1*, **I**) and a downregulated gene (*Kirrel*, **J**). Shades regions correspond to those in (**G**). See also Figure S4.

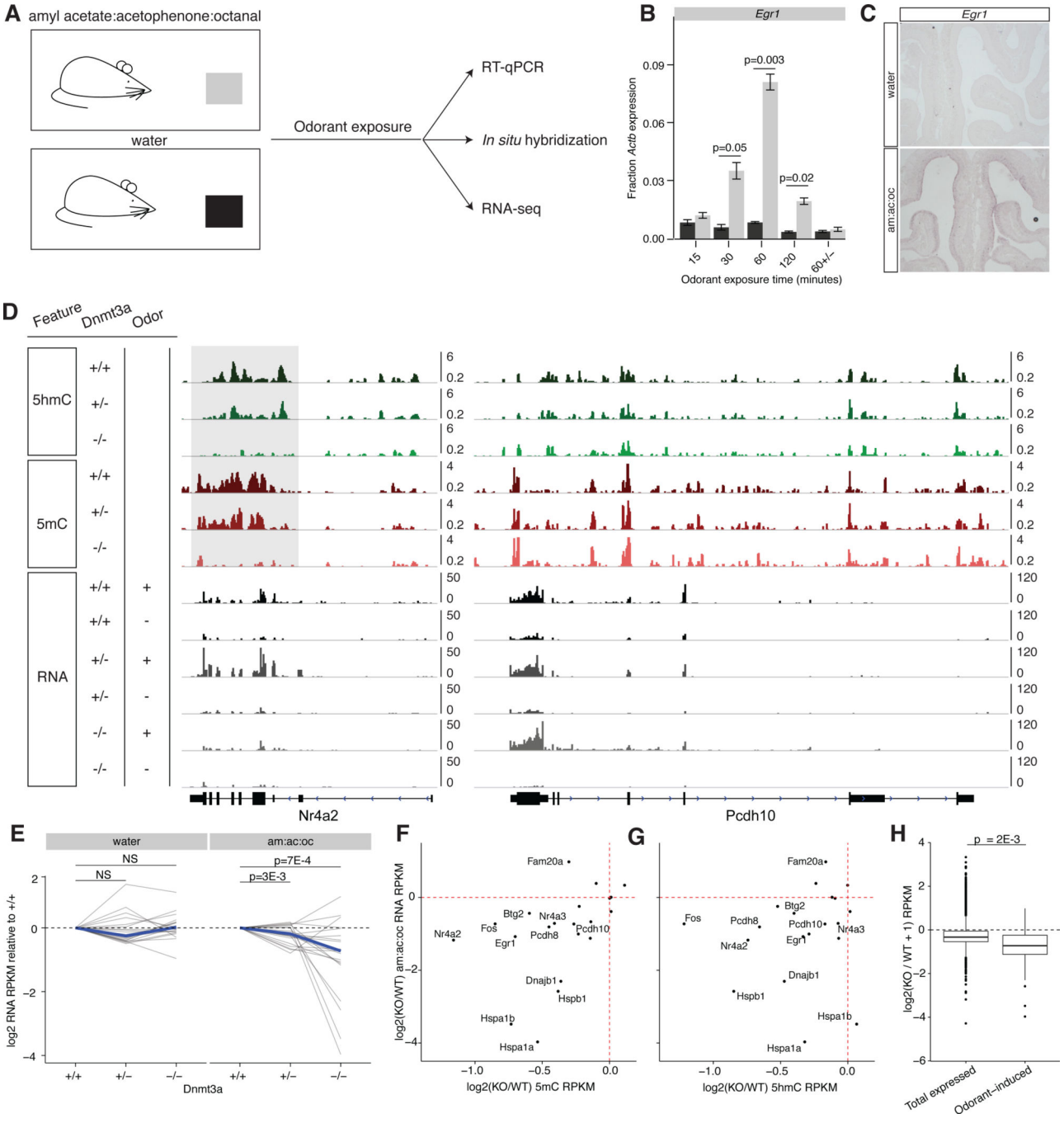

Figure 7.

Loss of Dnmt3a disrupts DNA modification states and inducibilities of odorant-sensitive genes. (**A**) Schematic of odorant exposure experiments. (**B**) RT-qPCR of immediate early gene *Egr1* expression in MOEs from 3-week old wildtype mice that had been exposed to either water or 1:1:1 amyl acetate:acetophenone:octanal (am:ac:oc) for 15, 30, 60, or 120 minutes or exposed for 60 minutes then allowed to recover in a new cage for 60 minutes (60+/−). Expression levels are expressed as the fraction of *Actb* expression ± SEM. P-values from Student's t-test. (**C**) *In situ* hybridizations against *Egr1* on coronal MOE sections from

Dnmt3a WT mice exposed to either water or am:ac:oc for one hour. (**D**) 5hmC, 5mC, and RNA tracks of two representative odorant-induced genes, *Nr4a2* and *Pcdh10*. 1 hour exposure to am:ac:oc ('odor +') or water ('odor −'). (**E**) Wildtype-normalized RPKM values of 23 odorant-induced genes in either water or am:ac:oc conditions. Blue bars intersect with the median of each genotype. Values are the averages across two biological replicates. (**F, G**) Gene-body 5mC (**F**) and 5hmC (**G**) loss is correlated with reduced odorant-induced expression. *X-axis*, average log2(KO/WT RPKM) 5mC or 5hmC across gene-bodies; *y-axis*, log2(KO/WT) RPKM of induced RNA expression of odorant-induced genes. (**H**) Comparison of KO vs. WT fold-change for total mOSN expressed genes (as defined in text) and odorant-induced genes. P-value from Wilcoxon two-sided. See also Figure S5.