GENETICS

Reprogramming of DNA methylation at NEUROD2-bound sequences during cortical neuron differentiation

Maria A. Hahn[1], Seung-Gi Jin[2], Arthur X. Li[3], Jiancheng Liu[4], Zhijun Huang[2], Xiwei Wu[5], Byung-Wook Kim[1], Jennifer Johnson[2], Adrienne-Denise V. Bilbao[2], Shu Tao[5], Jacob A. Yim[1], Yuman Fong[1], Sandra Goebbels[6], Markus H. Schwab[6,7], Qiang Lu[4]*†, Gerd P. Pfeifer[2]*†

The characteristics of DNA methylation changes that occur during neurogenesis in vivo remain unknown. We used whole-genome bisulfite sequencing to quantitate DNA cytosine modifications in differentiating neurons and their progenitors isolated from mouse brain at the peak of embryonic neurogenesis. Localized DNA hypomethylation was much more common than hypermethylation and often occurred at putative enhancers within genes that were upregulated in neurons and encoded proteins crucial for neuronal differentiation. The hypomethylated regions strongly overlapped with mapped binding sites of the key neuronal transcription factor NEUROD2. The 5-methylcytosine oxidase ten-eleven translocation 2 (TET2) interacted with NEUROD2, and its reaction product 5-hydroxymethylcytosine accumulated at the demethylated regions. NEUROD2-targeted differentially methylated regions retained higher methylation levels in *Neurod2* knockout mice, and inducible expression of NEUROD2 caused TET2-associated demethylation at its in vivo binding sites. The data suggest that the reorganization of DNA methylation in developing neurons involves NEUROD2 and TET2-mediated DNA demethylation.

INTRODUCTION

Neurogenesis in the embryonic mouse brain initiates in the ventricular zone (VZ) and sub-ventricular zone. Neuroepithelial cells differentiate into radial glial cells, a fate-restricted neural progenitor cell (NPC) population that can give rise to young neurons directly or through intermediate progenitor cells. Embryonic day 14.5 (E14.5) and E15.5 represent the peak stages of neurogenesis in the mouse cortex. This is a critical step at which proliferating NPCs transition from proliferation to differentiation (*1*–*6*).

This transition is initiated by a number of transcription factor changes, e.g., down-regulation of NPC-specific factors and engagement of neuronal differentiation–promoting factors that cooperate with epigenetic modifications present on DNA and chromatin (*1*, *2*, *4*, *5*). These epigenetic processes are likely important for achieving and maintaining the differentiated state, leading to the establishment of neuronal cell identity (*2*, *4*, *5*, *7*–*9*). Further commitment steps lead to the formation of the different types of cortical neurons.

5-Methylcytosine (5mC), formed primarily at CpG sequences in mammals, is an epigenetic DNA modification linked to the control of gene expression, the faithful execution of certain developmental pathways, and long-term maintenance of cellular memory (*10*–*14*). When found at promoters or enhancers, 5mC is generally viewed as a repressive DNA modification incompatible with gene expression. Genome-wide and gene-specific 5mC patterns undergo marked changes during mammalian development (*15*) and also change as part of cell differentiation steps. For example, in vitro differentiation of embryonic stem (ES) cells into neuronal progenitors involves de novo methylation of hundreds of gene promoters and also the formation of hypomethylated DNA regions (*16*–*18*). However, it is largely unknown to what extent and with what sequence specificity DNA methylation patterns are modulated in vivo during key neurodevelopmental transitions.

We previously characterized 5mC and 5-hydroxymethylcytosine (5hmC) patterns during neuronal differentiation using genome-wide mapping approaches (*19*). These in vivo studies have shown that the modified cytosine base 5hmC, which is produced by oxidation of 5mC by ten-eleven translocation (TET) proteins, increases along gene bodies of neuronal differentiation–associated genes in differentiating neurons along with a loss of the Polycomb mark H3K27me3 (*19*). To compare NPCs and their derived neurons, we used a dual-reporter labeling strategy by using transgenic mice in which NPCs are labeled with green fluorescent protein (GFP) expressed from the *Nes* (Nestin) promoter, and differentiated neurons are labeled with red fluorescent protein (RFP) expressed from the *Dcx* (doublecortin) promoter. This in vivo system allows the efficient parallel purification of NPCs and daughter neurons from the same mouse brain (*19*, *20*).

In the present study, we used whole-genome bisulfite sequencing (WGBS) to comprehensively interrogate DNA methylation patterns in NPCs and in neurons at single-base resolution. We observed hundreds of specific DNA hypomethylation events during this critical step of neurogenesis in mice. Most of these methylation-reprogramming events were targeted to genomic binding sites of the basic helix-loop-helix (bHLH) transcription factor NEUROD2, a critical neuronal differentiation factor, and were associated with up-regulation of many neuronal function–related genes that harbor these hypomethylated differentially methylated regions (DMRs). Loss- and gain-of-function studies provide further support for a role of NEUROD2 in DNA methylation remodeling.

[1]Department of Surgery, City of Hope, Duarte, CA 91010, USA. [2]Center for Epigenetics, Van Andel Research Institute, Grand Rapids, MI 49503, USA. [3]Department of Information Sciences, City of Hope, Duarte, CA 91010. [4]Department of Developmental and Stem Cell Biology, City of Hope, Duarte, CA 91010, USA. [5]Department of Molecular and Cellular Biology, City of Hope, Duarte, CA 91010, USA. [6]Max Planck Institute of Experimental Medicine, Department of Neurogenetics, D-37075 Göttingen, Germany. [7]Cellular Neurophysiology and Center for Systems Neuroscience (ZSN), Hannover Medical School, 30625 Hannover, Germany.
*These authors contributed equally to this work as senior authors.
†Corresponding author. Email: gerd.pfeifer@vai.org (G.P.P.); qlu@coh.org (Q.L.)

RESULTS

Global DNA methylation changes during neurogenesis

The goal of our study was to obtain detailed, genome-wide information on DNA methylation changes at single-base resolution during the transition from proliferating NPCs to differentiated neurons in the embryonic mouse brain. We made use of a double-transgenic system, in which progenitors are labeled with GFP expressed from the nestin (*Nes*) promoter, and differentiated neurons are marked with RFP under the control of the *Dcx* promoter (*20*). The use of a differentiation reporter in conjunction with a progenitor cell–specific promoter helps alleviate the problem of carryover of GFP from a primitive cell to progeny, thus allowing effective copurification of NPCs and daughter neurons from the brain (see scheme in Fig. 1A). We isolated DNA from NPCs and neurons from E15.5 mouse brain and subjected it to WGBS. The bisulfite-based approach detects the sum of 5mC and 5hmC as modified bisulfite-resistant cytosines (*21*, *22*). Thus, we were scoring events that reflect either de novo methylation (formation of 5mC and/or 5hmC from C) or loss of modified cytosines (DNA demethylation; change of 5mC and/or 5hmC into C). DNA demethylation or de novo methylation may represent critical steps during neuronal differentiation, but the extent and sequence characteristics of these processes in vivo are unknown. The data from the two WGBS biological replicates of NPCs and neurons were highly correlated with $R = 0.96$ and $R = 0.95$, respectively (fig. S1, A and B). The total percentage of cytosines at CpG sequences that were modified to 5mC/5hmC (subsequently referred to as "methylated") ranged from 77.4 to 80.1%, with no significant difference between NPCs and neurons. We also determined the percentage of methylated cytosines at CHG and CHH sequences to assess the contribution of non-CpG methylation, which is found in certain mammalian cell types including neuronal cells (*23*, *24*). These values were only between 0.4 and 0.7%, which is close to the background level of remaining unconverted cytosines achieved by use of the bisulfite sequencing technique. The data show that DNA methylation at this developmental stage of mouse neurogenesis is mostly limited to CpG sequences.

At CpG islands and at promoters, most of CpG dinucleotide positions were close to unmethylated, with a methylation level of 0 to 10%, as expected (fig. S1, C and D). In contrast, methylation levels increased to 60 to 100% in gene bodies and in intergenic regions (fig. S1, E and F). There was generally no major difference in CpG methylation profiles between NPCs and neurons in the different genomic compartments (fig. S1, C to F).

Hypomethylated DMRs in gene bodies

We used the Bioconductor package Dispersion Shrinkage for Sequencing (DSS) (*25*) to identify DMRs between NPCs and neurons. This analysis identified 1306 hypomethylated DMRs in neurons but only 26 hypermethylated DMRs, indicating that hypomethylation in neurons is 50 times more common than hypermethylation (Fig. 1B and table S1).

Figure 1C shows an example of a genomic region about 200 kb upstream of the *Sox5* gene on chromosome 6, where several neighboring CpGs become hypomethylated in neurons. These hypomethylated DMRs were generally no more than a few hundred base pairs in length. More than half of the hypomethylated DMRs were in intragenic regions, very few in promoters, and the rest were in intergenic areas (Fig. 1D). The very few hypermethylated DMRs were most common in intergenic genomic regions. Less than 10% of the hypomethylated intragenic DMRs were in exons, placing most of them into intronic sequences (Fig. 1E). There were only two hypermethylated DMRs that occurred in exons. The composite profiles revealed that hypomethylated DMRs were associated with a localized gain of 5hmC according to 5hmC immunoprecipitation (Fig. 1F), suggesting a likely role of 5mC oxidation in the DNA demethylation process. Gene ontology (GO) analysis of genes with intragenic hypomethylated DMRs showed enrichment of terms related to cell adhesion, neuron projection development, and axon formation (Fig. 1G). The most highly enriched category of "cell adhesion" included adhesion molecules encoded by the genes *Cadm1*, *Clstn2*, *App*, *Negr1*, *Ptprk*, *Ptprm*, *Magi1*, *Nrxn3*, *Sdk1*, *Sdk2*, *Nrxn1*, *Btbd9*, *Ctnna2*, *Lsamp*, *Reln*, *Dst*, *Dscaml1*, *Spock1*, *Cdh4*, *Rgmb*, *Pvrl1*, *Sorbs1*, *Nfasc*, *Stab2*, *Cdh13*, *Col19a1*, *Col14a1*, and *Ntm*. Many of these genes encode cell surface proteins critical for axon formation, neuron migration, axon guidance, and synapse formation.

We also inspected hypomethylated DMRs that occur within 10 kb upstream of the transcription start sites (TSSs). This analysis showed an enrichment of terms such as developmental maturation and metal ion transport (fig. S2A). Genes with DMRs within 10 kb downstream of the transcription end sites (TESs) were associated with p53 signaling and cell cycle regulation (fig. S2B).

Hypomethylation at putative enhancer regions

The presence of hypomethylated DMRs in introns suggested that they might be associated with enhancers. Figure 2 shows the enrichment profile of the active enhancer mark H3 lysine 27 acetylation (H3K27Ac) at the hypomethylated DMRs and browser views of these intragenic regions that become hypomethylated in neurons. Because of limited cell numbers obtainable from NPCs for comprehensive chromatin immunoprecipitation sequencing (ChIP-seq), we used published ChIP-seq data for H3K27Ac obtained from mouse forebrain at E14.5 (*26*), which contains ~85% neurons. This dataset gave the strongest signals and was used in quantitative comparisons. Very similar H3K27Ac profiles were also seen in an independent dataset from E16.5 mouse cortex (*27*), which has more than 90% differentiated neurons, and in Encyclopedia of DNA Elements (ENCODE) data from E15.5 forebrain, suggesting that these profiles predominantly reflect enhancers in differentiated neurons. This conclusion is also supported by the strong mirror image of H3K27Ac and DNA methylation levels that we observed in neurons, where H3K27Ac generally coincided with DNA sequences having low methylation. The profile of H3K27Ac over the center of hypomethylated DMRs and their flanking sequences show a strong accumulation of the H3K27Ac enhancer mark at the center of hypomethylated DMRs (Fig. 2, A and B). Examples of H3K27Ac-associated enhancer regions that undergo DNA hypomethylation in neurons are shown for the genes *Auts2* (autism susceptibility candidate 2, a regulator of neuronal migration) and *Itpr1* (inositol 1,4,5-trisphosphate receptor, which acts as a calcium release channel) (Fig. 2, C and D).

Using microarray data for NPCs and neurons, we asked whether genes with intragenic hypomethylated DMRs undergo changes in gene expression. Of the genes associated with intragenic hypomethylation, for which expression data were available, 194 became at least 1.5-fold activated and 108 genes became at least 1.5-fold repressed in neurons (fig. S3A; see table S2 for statistical analysis). In total, 29% of the intragenic hypomethylated DMRs were colocalized with H3K27Ac-marked putative enhancers (fig. S3B), and this colocalization was more frequent in activated genes compared to repressed

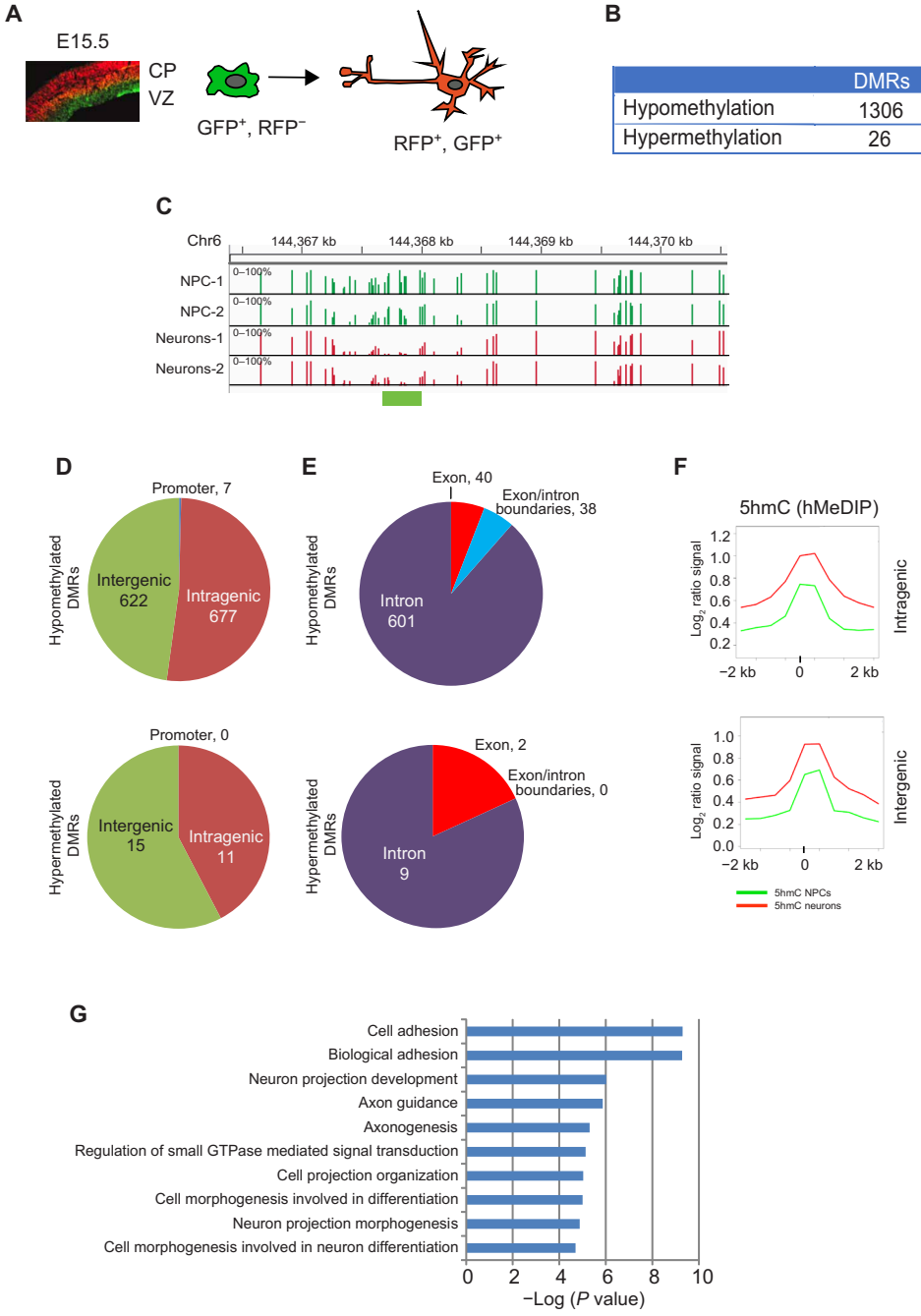


Fig. 1. Hypomethylated and hypermethylated DMRs and their genomic distribution. (**A**) Outline of the experimental approach to purify NPCs and early differentiating neurons. NPCs were marked with *Nes*-GFP and neurons were marked with *Dcx*-RFP. CP, cortical plate. (**B**) Numbers of DMRs identified. Hypomethylated DMRs are much more frequent than hypermethylated ones. (**C**) Example of neighboring CpG sites that undergo synchronous hypomethylation (green bar indicates the DMR). This region is ~200 kb upstream (5′) of the promoter of the *Sox5* gene. (**D**) Genomic distribution of hypomethylated and hypermethylated DMRs in intragenic, intergenic, and promoter regions. (**E**) Hypomethylated DMRs accumulate in introns and are much more frequent than hypermethylated DMRs. (**F**) Hypomethylated intragenic and intergenic DMRs in neurons are characterized by 5hmC peaks defined by hmeDIP. Green, NPCs; red, neurons. (**G**) Gene ontology analysis of genes with intragenic hypomethylated DMRs. GTPase, guanosine triphosphatase.

genes (fig. S3, D and E). In neurons, intragenic H3K27Ac-associated DNA hypomethylation was associated with stronger gene activation, in contrast to intragenic hypomethylated sites, which were not colocalized with H3K27Ac (Fig. 2E). Because enhancer gene pairs are difficult to assign, we cannot exclude the possibility that an intragenic putative enhancer regulates a more distant gene. In contrast to the intragenically hypomethylated genes, the genes (n = 11) carrying intragenic hypermethylated areas were predominantly repressed in neurons (fig. S3F). In intergenic genomic areas, there were 622 hypomethylated DMRs (Fig. 1D). A smaller fraction of these intergenic hypomethylated DMRs also colocalized with sequences carrying H3K27 acetylation (fig. S3C).

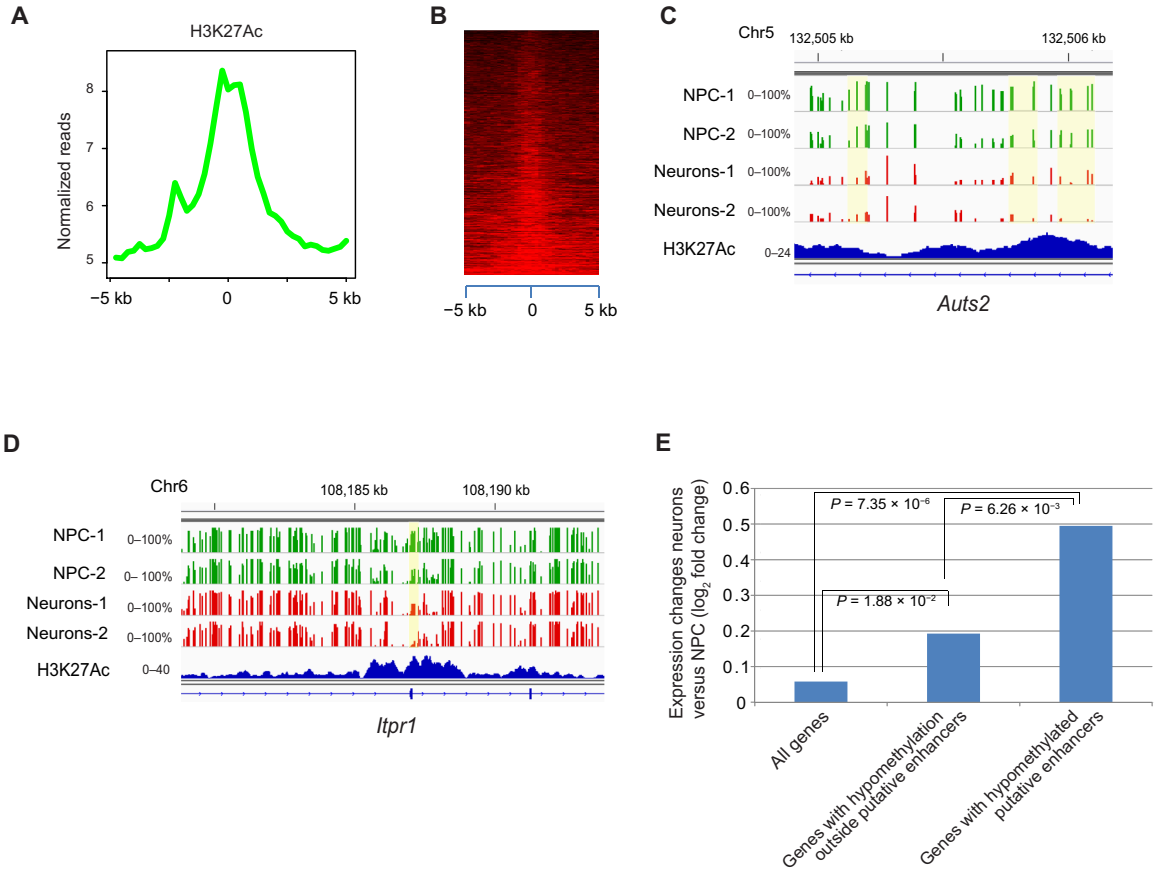


Fig. 2. Enrichment of hypomethylated DMRs at putative active enhancer regions. (**A**) H3K27Ac profile distribution at centered hypomethylated DMRs. Hypomethylated DMRs were analyzed along with 5 kb of their flanking DNA sequences. (**B**) Heatmap of H3K27Ac distribution at centered hypomethylated DMRs. High read counts are shown in red. (**C** and **D**) Intragenic sequences undergo hypomethylation in neurons (hypomethylated DMRs are shaded in yellow) and are localized near H3K27Ac-marked regions (blue). (C) *Auts2* gene. (D) *Itpr1* gene. (**E**) Genes with hypomethylated intragenic putative enhancers are more strongly activated than genes with hypomethylation not linked to putative enhancers or than all genes combined. Statistical significance was determined using *t* tests.

NEUROD2-linked DNA hypomethylation during neurogenesis

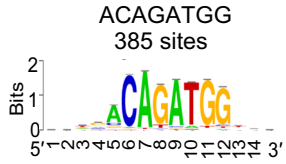
Since the hypomethylated DMRs tended to localize to putative enhancer regions, we then conducted a de novo transcription factor motif finding analysis to identify potential regulatory elements that become hypomethylated during neuronal differentiation. This analysis revealed a highly significant enrichment of two E-box–type motif classes (Fig. 3A). One motif, 5′ACAGATGG, was assigned by the JASPAR database to the bHLH transcription factors neurogenic differentiation factor 1/2 (NEUROD1/2), twist-related protein 2 (TWIST2), T-cell acute lymphocytic leukemia protein 1 (TAL1), transcription factor 3 (TCF3), basic helix-loop-helix family member a15 (BHLHA15), and several others (Fig. 3A). The other enriched motif was 5′CATATG, which was assigned to NEUROD1/2, basic helix-loop-helix family member e22 (BHLHE22), atonol homologue 1 (ATOH1), and a few others. Some of these factors showed only very low expression levels in NPCs and neurons, including *Myod1*, *Atoh1*, *Tal1*, *Bhlhe23*, *Pou5f1*, and *Myog*. Expression of a number of other candidates was down-regulated in neurons, including *Neurog1* and *Neurog2* (Fig. 3B). The most significantly up-regulated genes were *Neurod2* and *Bhlhe22*, which reached very high expression levels in neurons (Fig. 3B), consistent with expression analysis in mice (*28*, *29*). *Neurod1* was also up-regulated. BHLHE22/BHLHB5 is a transcription factor required for neuron differentiation and neuronal circuit assembly and functions primarily as a transcriptional repressor (*30*, *31*). NEUROD2 is an important neurogenic differentiation factor and transcriptional activator (*28*, *32*–*36*). Broad areas, ~30 kb in length, of H3K27 acetylation are found at the *Neurod2* gene in E14.5 mouse cortex (Fig. 3C) but not at the *Neurod1* or *Bhlhe22* genes. This feature has been suggested to be a determinant of so-called "super-enhancers" and may be a hallmark of cell identity genes often encoding important lineage-specific differentiation factors (*37*).

We then compared available ChIP-seq data for NEUROD2 in E14.5 mouse cortex (*38*), which has ~85% neurons, with our hypomethylated DMRs (Fig. 4, A to C, and fig. S4, A to F). According to a heatmap of NEUROD2 binding at the hypomethylated DMRs and surrounding areas, a large fraction of these DMRs colocalized with NEUROD2 binding (Fig. 4C). A peak identification algorithm identified the NEUROD2 peaks in 51% (343 of 677) of intragenic hypomethylated DMRs and in 54% (334 of 622) of intergenic hypomethylated DMRs. Examples for the colocalization of NEUROD2 peaks and hypomethylated DMRs are shown in Fig. 4 and fig. S4, and data for intragenic regions are summarized in table S3.

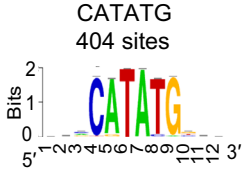
A number of important neuronal differentiation genes showed the feature of NEUROD2-associated DNA hypomethylation, including *Nfasc* (Fig. 4A) and *Efnb2* (fig. S4), and a number of others

A

Motif 1 evalue = 1.3×10^{-29}
NEUROD1/2, TWIST2, TAL1:TCF3, BHLHA15, ATOH1, TCF12, NEUROG1/2, MYOG, MYOD1



Motif 2 evalue = 1.3×10^{-18}
NEUROD1/2, BHLHE22, ATOH1, OLIG1/2/3, BHLHA15, TWIST2, NEUROG1/2, BHLHE23, POU3F4, POU5F1B, TAL1:TCF3



B

Gene	NPC $\log_2$	Neurons $\log_2$	Neurons versus NPC $\log_2$	FDR
Bhlhe22	9.12	12.29	3.17	5.5×10^{-5}
Neurod2	11.34	12.92	1.58	6.0×10^{-5}
Neurod1	10.08	11.83	1.76	2.19×10^{-5}
Myod1	6.61	7.16	0.56	>0.1
Bhlha15	9.20	9.69	0.49	>0.1
Atoh1	6.93	7.40	0.48	>0.1
Tal1	7.04	7.44	0.40	>0.1
Bhlhe23	6.83	7.23	0.40	>0.1
Twist2	7.46	7.84	0.38	>0.1
Pou5f1	6.49	6.57	0.09	>0.1
Olig3	7.57	7.50	−0.06	>0.1
Myog	6.78	6.54	−0.24	>0.1
Tcf12	12.23	11.63	−0.60	4.4×10^{-2}
Pou3f4	9.68	8.56	−1.12	3.8×10^{-3}
Neurog2	12.87	10.75	−2.13	1.6×10^{-4}
Neurog1	10.20	7.27	−2.93	1.5×10^{-4}
Tcf3	10.67	7.65	−3.01	6.3×10^{-5}
Olig1	8.81	5.70	−3.11	1.9×10^{-4}
Olig2	8.45	3.23	−5.22	1.7×10^{-3}

C

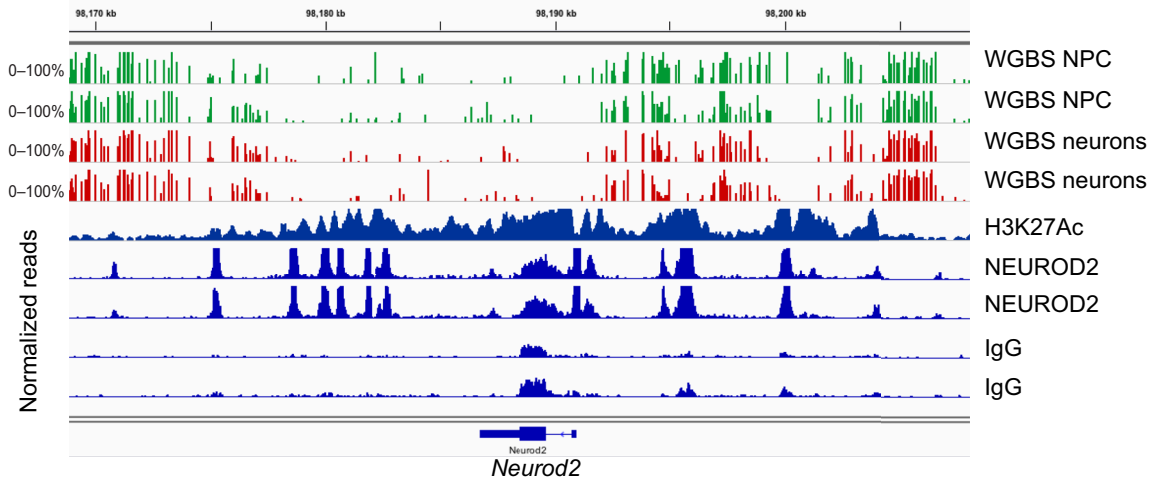


Fig. 3. De novo motif finding analysis for hypomethylated DMRs and candidate transcription factors. (**A**) Hypomethylated DMRs were subjected to motif analysis using Regulatory Sequence Analysis Tools. Two specific motifs were enriched with high significance, motif 1 (5′ACAGATGG3′) and motif 2 (5′CATATG3′). Other motifs had e values of 0.001 or larger. The most significantly scoring candidate DNA-binding factors are listed for each motif. (**B**) Expression levels and changes in expression of candidate transcription factor genes between neurons and NPCs was determined by microarray analysis. (**C**) The *Neurod2* gene in E14.5 embryonic mouse brain is broadly covered by H3K27Ac and has multiple NEUROD2-occupied binding sites. FDR, false discovery rate. IgG, immunoglobulin G.

such as *Nav2*, *Myt1l*, and *Dlg2* (table S3). In contrast, the numerous other genomic H3K27Ac-enriched sequences without NEUROD2 peaks only very rarely were associated with loss of DNA methylation in neurons. GO analysis for intragenic hypomethylated NEUROD2-associated DMRs showed a strong enrichment for terms of neuron differentiation, including neuron projection development and axonogenesis for the up-regulated genes (n = 116) (fig. S4G). For the fewer NEUROD2-linked hypomethylation-associated genes that were down-regulated in neurons (n = 58), a number of other terms were moderately enriched (fig. S4H). This observation suggests that NEUROD2 may function as a transcriptional repressor for some genes or, alternatively, that NEUROD2 regulates a more distant gene in those cases.

A role for TET-mediated 5mC oxidation in neuronal DNA demethylation

DNA demethylation at NEUROD2 motifs may be linked to 5mC oxidation by TET proteins. Hypomethylated DMRs are associated with broader and more intensive increases of 5hmC in neurons in the presence of NEUROD2 versus regions without this transcription factor (Fig. 4D). This is true for both intragenic and intergenic hypomethylated DMRs, suggesting that NEUROD2 binding at the DNA demethylation sites is linked to 5mC oxidation. Examples for 5hmC accumulation at specific genes are shown in fig. S5. For most hypomethylated DMRs in neurons, the WGBS data do not indicate complete "demethylation." Therefore, the remaining signals (also at sequences immediately flanking the hypomethylated sites) may be

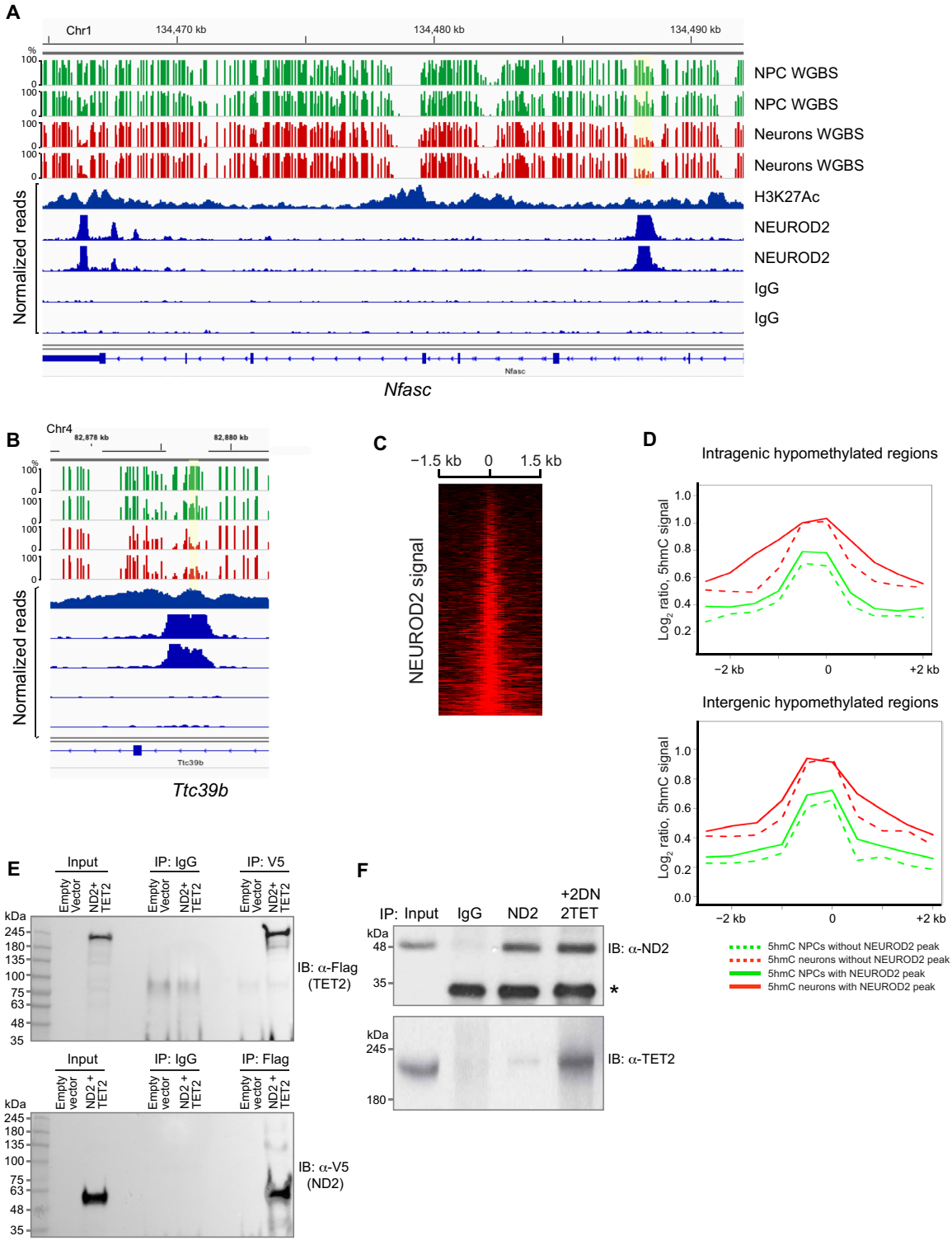


Fig. 4. Features of NEUROD2-associated DNA hypomethylation. (**A** and **B**) Intragenic DNA hypomethylation in neurons coincides with NEUROD2 peaks in E14.5 brain. Several examples are shown as genome browser views. (A) *Nfasc*; (B) *Ttc39b*. The identified hypomethylated DMRs are shaded in yellow. See figs. S4 and S5 for more examples. FDR, false discovery rate. (**C**) Heatmap distribution for NEUROD2 signal around centered hypomethylated DMRs (−1.5 to +1.5 kb relative to the center of the hypomethylated DMR). We sorted the heatmap by the number of NEUROD2 reads at the center of hypomethylated DMRs. Red color indicates a high number of reads, and black indicates a low number of reads or no reads. (**D**) Distribution of 5hmC signal around NEUROD2 peak-centered hypomethylated DMRs and at hypomethylated DMRs that do not have NEUROD2 peaks. Green, NPCs; red, neurons; solid lines, with NEUROD2 peaks; dotted lines, without NEUROD2 peaks. Data for intragenic and intergenic regions are shown separately. Note that the 5hmC enrichment increases during the transition of NPCs to neurons and is broader in neurons at NEUROD2 peaks. (**E**) Co-IP of NEUROD2 (ND2) and TET2 in 293 T cells. IB, immunoblotting. (**F**) Co-IP of NEUROD2 (ND2) and TET2 in E15.5 embryo mouse brain. Asterisk (*) denotes IgG light chain.

derived from 5hmC, which, in the case of enhancer demethylation, can perhaps be viewed as an intermediate base of the demethylation process rather than as a stable DNA base. The E-box motifs recognized by NEUROD2 do not contain a CpG site themselves (Fig. 3A). However, several CpGs in close proximity to the motifs and the NEUROD2 peaks undergo DNA hypomethylation. To further confirm an interaction between NEUROD2 and TET proteins, we focused on TET2, which has been linked to DNA demethylation at enhancer-like sequences in other cell types (*39*–*44*). Using coexpression of tagged TET2 and NEUROD2 proteins (Fig. 4E) or coimmunoprecipitation (co-IP) of the endogenous proteins in E15.5 mouse cortical tissue (Fig. 4F), we observed a clear interaction between NEUROD2 and TET2, further supporting a role of TET2-mediated DNA demethylation activity at NEUROD2-bound sequences.

We then proceeded to test whether methylation of NEUROD2-bound sequences interferes with gene expression from a linked promoter. We cloned the promoters of the genes *Tiam2* and *Rgmb* into the pGL3-basic luciferase vector. NEUROD2-bound sequences located within less than 15 kb downstream of the TSS of these genes were inserted in an unmethylated form or in a CpG-methylated form downstream of the luciferase gene, and the activity was measured. We found that CpG methylation decreased the expression from both promoters in two different NEUROD2-expressing cell lines (fig. S6).

Reduced DNA demethylation in *Neurod2* knockout mice

To test a functional role of NEUROD2 in DNA demethylation, we obtained DNA from E16.5 cortices of wild-type and *Neurod2* homozygous or heterozygous knockout mice (*28*), which contain >90% differentiated neurons at this stage of development. This DNA was converted with sodium bisulfite, and several NEUROD2-targeted genomic regions were analyzed by deep sequencing. The percent methylation at individual CpGs was determined (Fig. 5). A hypomethylated DMR in the gene *Nfasc* (Fig. 5A) showed 5 to 15% higher methylation levels at almost all individual CpG sites in the *Neurod2* knockout mice compared to wild-type or heterozygous mice ($P < 0.05$) (Fig. 5B). A similar effect was seen at several CpG sites in the genes *Gabbr2* and *Satb2* (fig. S7, A and B). The bisulfite sequencing analysis allows analysis of individual DNA molecules (Fig. 5C and fig. S7, C and D). We found that for the *Nfasc* locus, there were either mostly unmethylated or mostly methylated DNA strands in both wild-type and knockout mice (fig. S7, C and D). However, in the *Neurod2*-deficient mice, the fraction of unmethylated strands was significantly reduced with a concomitant increase of fully methylated DNA strands (Fig. 5C and fig. S7, C and D). These data suggest that, although there is a degree of cellular heterogeneity with respect to methylation state and likely neuronal subtypes, loss of NEUROD2 reduces the extent of DNA hypomethylation in differentiated neurons.

NEUROD2 induces DNA demethylation

We wanted to know whether NEUROD2 is able to change the methylation state at its target sites in cells in which it is not normally expressed. We chose the mouse P19 embryonal carcinoma cell line, which has previously been used to induce neuronal differentiation with bHLH transcription factors including NEUROD2 (*45*, *46*). We expressed NEUROD2 from a doxycycline-inducible lentiviral vector (Fig. 6A). Fifteen days after doxycycline treatment, the NEUROD2-expressing cells expressed neuronal marker genes, including *Mapt*, *Tubb3*, and several others (fig. S8, A and B). These cells also expressed increased levels of *Tet2* and *Tet3* but only low and non-inducible levels of *Tet1*. We chose several genes with intragenic NEUROD2 peaks in the brain for DNA methylation analysis. These target loci also harbored identical NEUROD2 peaks in P19 cells as observed with overexpressed NEUROD2 (*46*). The NEUROD2-targeted sequences were highly methylated (>80%) in P19 cells without NEUROD2 induction (Fig. 6, B to E, and fig. S8, C to G). After doxycycline treatment, NEUROD2-bound regions became demethylated. The most substantially demethylated sequences were observed for the genes *Aff3*, *Nfasc*, *Cog7*, and *Zhx2* ($P < 0.0015$; Fisher's exact test, two-tailed). Significant demethylation was also observed at the genes *Rgmb*, *Wwc1*, *Glis2*, and *Tiam2* (Fig. 6 and fig. S8). Next, we analyzed expression of those genes that underwent NEUROD2-induced demethylation in P19 cells with and without NEUROD2 induction. We observed that the genes *Aff3*, *Nfasc*, *Zhx2*, *Rgmb*, *Itpr2*, *Cog7*, *Wwc1*, *Glis2*, and *Tiam2* were all induced after doxycycline treatment (fig. S8B).

To test an involvement of the 5mC oxidase in the demethylation process, we used short hairpin–mediated RNA (shRNA) knockdown to deplete *Tet2* from P19 cells. After induction of NEUROD2 expression with doxycycline, we tested the methylation of several NEUROD2 target regions and the expression of the associated genes *Aff3*, *Zhx2*, and *Cog7* (fig. S9). We observed higher methylation levels at the target sequences in the *Tet2* knockdown cells (fig. S9A). Expression of *Aff3* and *Cog7* was reduced but not that of the *Zhx2* gene (fig. S9B), which may be regulated primarily by other mechanisms.

Together, this set of data shows that NEUROD2 has the capacity to induce DNA demethylation near its target sites in vivo concomitant with induction of gene expression and supports a role of this factor in remodeling DNA methylation in neurons.

DISCUSSION

Our data show that loss of DNA methylation during cortical neuron differentiation occurs at a set of defined localized regions targeted by the bHLH transcription factor NEUROD2 but is less common at other places in the neuronal genome. The data suggest a critical role of NEUROD2 in epigenome remodeling during neuron development in vivo.

Previous studies have analyzed DNA methylation changes in vitro after differentiation of ES cells toward the neuronal lineage. Whereas numerous methylation changes were observed during the initial differentiation of ES cells to NPCs, fewer changes seemed to occur when NPCs undergo final differentiation to neurons in this system (*17*, *18*). Our study is the first comprehensive analysis of DNA cytosine methylation changes occurring during the key cell differentiation steps at early stages of neuron formation in the mouse brain in vivo.

DNA methylation changes occurring at the initial stages of neuronal differentiation from NPCs to young neurons are restricted to specific sequences. They almost never occur at CpG islands or at gene promoters. These data are consistent with our previous lower resolution analysis, in which we observed that activation of neuronal differentiation genes is not frequently accompanied by changes in 5mC but rather by a loss of the Polycomb mark H3K27me3 and by deposition of 5hmC along gene bodies of the same genes (*19*). This conversion of 5mC to 5hmC would not be detected by bisulfite sequencing. However, bisulfite sequence analysis can detect the conversion of 5mC/5hmC to cytosine, which we found to occur with a strong preference for targeting intragenic sequences. This intragenic hypomethylation was commonly associated with H3K27Ac-marked

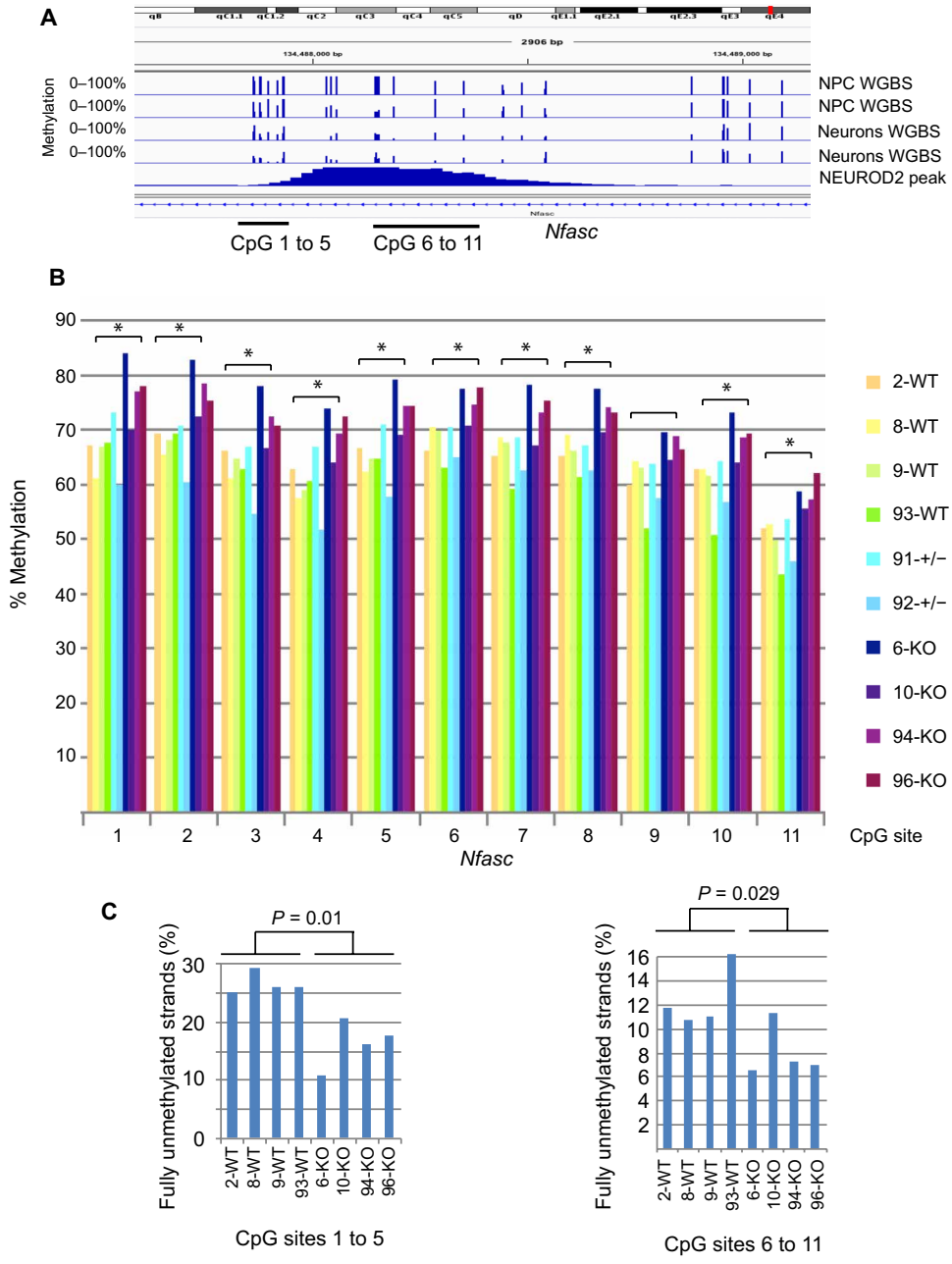


Fig. 5. DNA methylation changes in *Neurod2* knockout mice. (**A**) Genomic view of a hypomethylated NEUROD2-bound region in the *Nfasc* gene. WGBS data are shown for NPCs and neurons along with the associated NEUROD2 peak. (**B**) Analysis of CpG methylation by bisulfite deep sequencing at the 11 CpGs of the *Nfasc* gene as indicated in (A). The analysis was conducted at E16.5 with four wild-type (WT), two heterozygous, and four homozygous *Neurod2* knockout (KO) embryos. The percent CpG methylation level is shown for each CpG. Measurements were taken from distinct samples and for each CpG site. A significantly higher methylation level was observed in the *Neurod2* knockout mice compared to wild-type mice at all CpG sites (*P* values between 0.01 and 0.02; *t* test, two-sided, indicated by asterisks) except for site no. 9 ($P = 0.06$). The mouse genotypes are indicated by the color code. (**C**) The percentage of fully unmethylated DNA strands is decreased in *Neurod2* knockout cortex both at CpG sites 1 to 5 ($P = 0.01$; *t* test, two-sided) and at sites 6 to 11 ($P = 0.029$; *t* test, two-sided) compared to wild-type mice ($n = 4$, each genotype, each sample measured individually). These data are based on single-molecule analysis as shown in fig. S7.

regions (Fig. 2) and with activation of the resident genes, many of which are important for neuronal differentiation (Fig. 2E, fig. S3A, and table S3). We used sequence motif analysis and found that hypomethylated genomic sequences were enriched for two consensus binding sites for bHLH transcription factors. The most interesting motif we identified was for NEUROD2, and a large fraction (more than half) of all intragenic hypomethylated DMRs coincided with strong NEUROD2 ChIP-seq peaks (Fig. 4 and figs. S4 and S5). For intergenic regions, a similar fraction of the hypomethylated sites were associated with NEUROD2. NEUROD2 is one factor that has been used for in vitro differentiation of pluripotent stem cells into neurons (*33*) and hence may function as a key regulator required

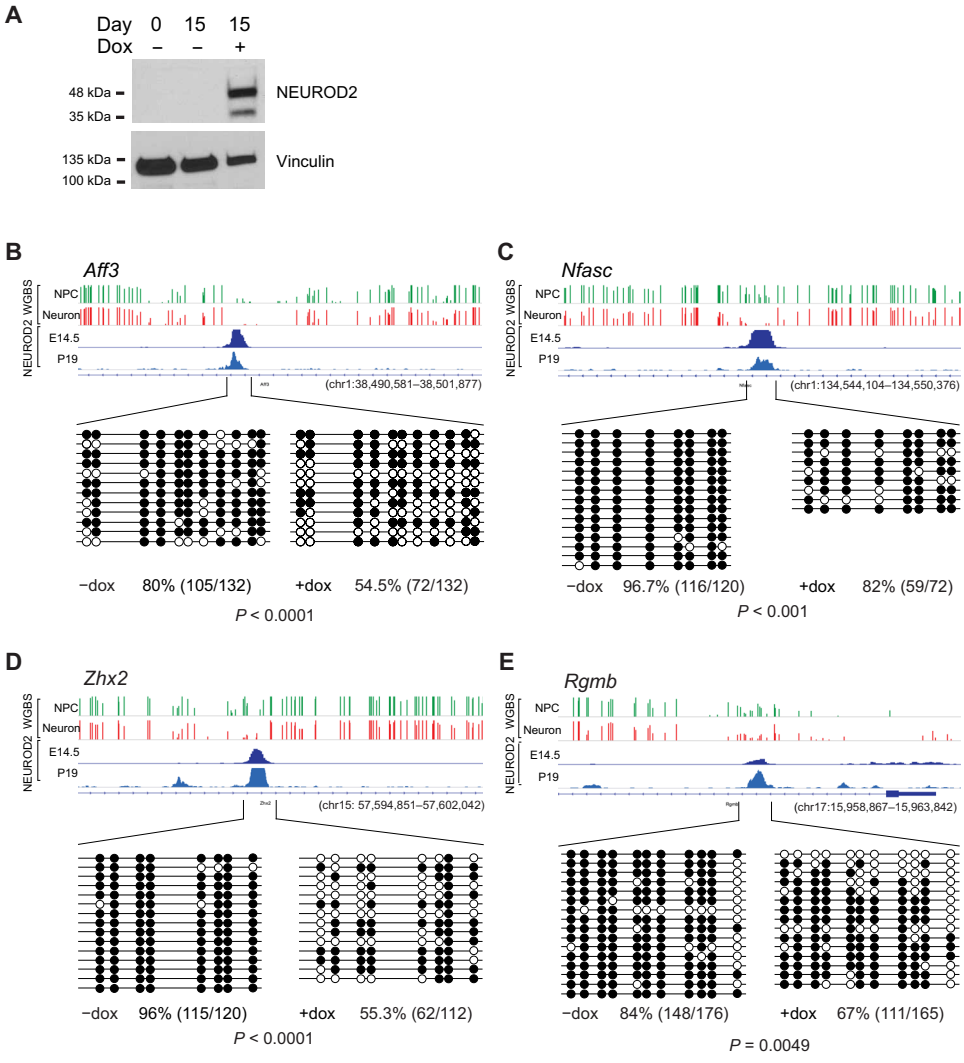


Fig. 6. NEUROD2 induces DNA demethylation. NEUROD2 was expressed from a doxycycline-inducible lentiviral vector in P19 cells. (**A**) Induction of NEUROD2 with doxycycline (Dox) is shown by Western blot. Vinculin is a loading control. (**B** to **E**) DNA methylation analysis at NEUROD2 target sites. The browser views show WGBS data from NPCs and neurons, mapping data for NEUROD2 in E14.5 cortex, and mapping data for NEUROD2 in P19 cells. The brackets indicate the genomic regions analyzed by bisulfite sequencing. The numbers and percentages of methylated and unmethylated CpG sites (black and white circles) in absence (–dox) and presence (+dox) of doxycycline are indicated. *P* values were determined by Fisher's exact test (two-tailed). (B) *Aff3*, (C) *Nfasc*, (D) *Zhx2*, and (E) *Rgmb*.

for the effective differentiation of neurons. Using single-cell expression profiling of brain-derived neurons, *Neurod2* was identified as a relatively early neuronal transcript during corticogenesis in mice, which rises when NPC markers such as *Sox2* are declining (*34*). *Neurod2* is important for normal central nervous system development and function and for survival of specific neuronal subtypes (*36*). It plays a role in regulating the anatomical and physiological maturation of thalamocortical connections (*35*) and, together with its homolog NEUROD6, is involved in neocortical projection neuron development and corpus callosum formation (*28*). The *Neurod2* gene is surrounded by multiple NEUROD2 peaks, suggesting effective autoregulation of the gene within a super-enhancer–like structure (Fig. 3C). *Bhlhe22*, a gene coding for a protein with the same DNA binding specificity, is also a NEUROD2 target and may therefore be downstream of NEUROD2.

Although the core factor binding sites we identified do not contain a CpG sequence (Fig. 3A), we observed that CpGs in closest proximity to the binding sites become hypomethylated. It is likely that access of transcription factors to their DNA sequences is determined by chromatin structure. Most methylated genomic regions are thought to be associated with inaccessible chromatin. Thus, mammalian cells have developed specialized mechanisms to keep CpG-rich regulatory sequences free of DNA methylation. There are about a dozen proteins that contain a CXXC zinc finger structure, a unique domain that recognizes unmethylated CpG islands (*47*). Several of these proteins may participate in preserving the unmethylated state of these sequences, including the TET 5mC oxidases that could remove inadvertently introduced 5mC bases (*48*, *49*). However, enhancers have lower CpG density and may not always be kept DNA methylation free. It is known that TET-mediated 5mC oxidation preferentially takes place at enhancer regions in several cell types examined (*24*, *39*–*41*, *44*, *50*–*54*). Although so far not well studied, a developmental transition or a differentiation switch may require the recruitment of a TET activity to enhancers to ensure effective

DNA demethylation. A few specific transcription factors associated with enhancer demethylation have been identified. Examples include the lineage-specific pioneer factors forkhead box A1/2 (FOXA1/2), GATA-binding transcription factor 4 (GATA4) (*55*–*57*), paired box transcription factor 7 (PAX7) (*58*), the stem cell transcription factor spalt-like transcription factor 4A (SALL4A) (*59*), and estrogen receptor (*43*). We observed that TET2 interacts with NEUROD2 and that the TET reaction product 5hmC accumulates at NEUROD2 binding sites that undergo demethylation in vivo (Fig. 4, D to F, and fig. S5), suggesting that a TET-mediated mechanism at putative enhancers is in play that leads to enhancer hypomethylation and activation in neurons. TET2 and TET3 play an important role in neuronal differentiation (*19*). We have not identified TET3 (long isoform) binding to the enhancer regions that become demethylated in neurons (*48*), and a suitable TET2 antibody was not available for ChIP-seq analysis. It would be difficult to determine whether these DNA demethylation events are DNA replication independent or replication dependent because differentiation may proceed through intermediate progenitor cells, which involves one round of DNA replication.

Our experiments with *Neurod2* knockout mice suggest that NEUROD2 is, at least in part, responsible for the DNA demethylation events. One possibility is that DNA demethylation is a consequence of NEUROD2 binding and enhancer activation, as suggested for DNA binding proteins in general (*16*). However, there are undoubtedly several other transcription factors that become engaged and DNA bound during the transition of progenitor cells toward neuronal differentiation, and yet, we observed DNA demethylation most commonly at NEUROD2-bound sites. Another scenario in which some other factor binding event near NEUROD2 consensus sites causes DNA demethylation, which is then followed by NEUROD2 binding, seems less likely because the demethylated sites strictly coincided with NEUROD2 peaks. The initial binding of a transcription factor that is insensitive to CpG methylation and insensitive to repressive chromatin, a so-called "pioneer factor," is likely a key initial step leading to enhancer demethylation and activation to promote expression of specific genes as part of the in vivo neuronal differentiation cascade. We examined other published ChIP-seq datasets from embryonic mouse brain or from ectopic expression experiments for the candidate neuronal pioneer factors achaete-scute homolog 1 (ASCL1) (*60*), NEUROD1 (*61*), and neurogenin 2 (NEUROG2) (*62*) but did not identify an overlap with our demethylated sites in neurons. Because our data show that hypomethylation events are precisely targeted to numerous NEUROD2 binding sites genome wide, we propose that NEUROD2 is a key factor endowed with DNA methylation remodeling features. We confirmed these functions in P19 cells, in which induction of NEUROD2 expression led to demethylation at its binding sites and induction of gene expression, which was diminished after *Tet2* knockdown (Fig. 6 and figs. S8 and S9).

Future experiments will be necessary to identify transcription factors that may cooperate with NEUROD2, e.g., as heterodimerization partners (TCF3/4) or as functional homologs (NEUROD6), for achieving the most efficient DNA demethylation of critical neuron-specific enhancers. A better understanding of the epigenome remodeling processes that occur during neuron differentiation in vivo will help in the development of more efficacious reprogramming approaches that can be used for the conversion of other cell types into functional neurons useful for regenerative medicine.

MATERIALS AND METHODS

Purification of NPCs and neurons

We used a double-transgenic reporter system to purify NPCs and neurons from E15.5 mouse cortex (*20*). We followed protocols for cell purification as reported previously (*19*). In this strategy, we used *Nes*-GFP mice and *Dcx*-RFP C57BL/6J mice to obtain littermate embryos that were positive for both GFP and RFP and also RFP single-positive embryos. These embryos were used for obtaining GFP^+RFP^- cells (NPCs) and RFP^+ cells (neurons), respectively, from the same mouse litters. All animal procedures were approved by the City of Hope's Institutional Animal Care and Use Committee.

Whole-genome bisulfite sequencing

Two biological replicates were used for NPCs and neurons, respectively. To generate the whole-genome libraries with bisulfite-converted DNA, neuronal cell DNA (1.2 μg each) was sonicated to approximately 150–base pair (bp) DNA fragments. Then, DNA was end-repaired by using the End-It DNA End-Repair Kit (Epicentre) and linker-ligated with T4 DNA ligase [New England BioLabs (NEB)]. The ligated DNA was bisulfite-converted by using the EZ DNA Methylation Gold Kit according to the manufacturer's instructions (Zymo Research) and amplified with Pfu Turbo DNA polymerase (Agilent). Paired-end sequencing was performed using a HiSeq2500 Illumina instrument. Between 220 and 300 million aligned paired-end reads were obtained for each sample. Data have been deposited to the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) databases (accession no. GSE101090).

Manual bisulfite sequencing

Two micrograms of each genomic DNA was converted with sodium bisulfite using the EZ DNA Methylation Gold Kit (Zymo Research), according to the manufacturer's instruction. Bisulfite-modified DNAs were amplified using target-specific primers (table S4). For sequence analysis, the polymerase chain reaction (PCR) products obtained after bisulfite conversion were purified using QIAquick PCR Purification Kit (Qiagen) and were then ligated into the pCR4 TA cloning vector (Thermo Fisher Scientific). Sixteen colonies for each cloned sample were sequenced and evaluated.

Targeted bisulfite deep sequencing in *Neurod2*-deficient mice

Generation of *Neurod2* knockout mice was previously described (*28*). DNA from E16.5 brain cortex of wild-type, $Neurod2^{-/-}$, and $Neurod2^{+/-}$ mice was converted with sodium bisulfite using the EZ DNA Methylation Gold Kit (Zymo Research). NEUROD2 peak-targeted regions were amplified with PCR primers specific to the sequences of interest. The primer sequences are available upon request. PCR products were ligated to linkers for next-generation sequencing using the KAPA Hyper Prep Kit (KAPA Biosystems). Ligated PCR products were sequenced on an Illumina HiSeq2500 instrument and aligned to the mouse genome. Statistical differences were calculated for each CpG site using the two-tailed *t* test with unequal variance.

ChIP-seq and 5hmC data analysis

Published data for H3K27Ac mapping in E14.5 mouse forebrain were downloaded from the GEO database (accession no. GSE52386) (*26*). The H3K27Ac-covered regions were defined as >100-bp sequences with continuous base coverage of five or more reads. Peaks that were less than 250 bp apart were merged into larger peaks.

Mapping data for 5hmC distribution genome wide in NPCs and neurons were obtained previously (*19*).

NEUROD2 ChIP-seq data from E14.5 mouse cortices were obtained from published datasets (GEO dataset: GSE67539) (*38*). NEUROD2 peaks were identified using customized R scripts and Bioconductor packages. NEUROD2 peaks, based on SRR1951389 and SRR1951390 files, were first identified individually using the readGAlignmentsFromBam function. Then, the common regions between these two sequence files were identified. A heatmap for NEUROD2 signal distribution across hypomethylated sites was generated on the basis of the regions ±1.5 kb from the centers of hypomethylated DMRs. Each genomic location was divided into 30 bins. The profile of NEUROD2 was obtained from SRR1951389. The heatmap was sorted by the sum of the number of reads at each analyzed region (3 kb in total).

Bioinformatics analyses of DNA methylation data

The analysis pipelines were implemented using R statistical language, except GO analysis, which was performed using the DAVID tool (https://david-d.ncifcrf.gov/). RefSeq genes were downloaded from the University of California, Santa Cruz mm9 genome annotation database.

The generation of composite profiles has been described (*19*, *63*). Promoters were defined as from −1.5 to +0.5 kb relative to the TSS. Gene body (intragenic) regions were defined as ranging from 0.5 kb downstream of the TSS to the TES. Methylation changes were classified as intergenic if they did not overlap with promoters or gene bodies.

Identification of genes with DNA hypomethylation and hypermethylation in promoters, intragenic sequences, and intergenic sequences

To identify DMRs, we used the Bioconductor package DSS (http://bioconductor.org/packages/release/bioc/html/DSS.html) (*25*). We used default parameters with a minimum length of DMRs of 50 bp, minimum number of CpGs of 4, a *P* value threshold of 0.02, and a minimum methylation difference of 0.3. The methylation levels of all CpG sites within each DMR were retrieved for each sample in the NPC and neuron group. The chromosomal locations of all DMRs were retrieved using DSS and matched with genomic features, changes in gene expression, presence of H3K27Ac, and NEUROD2 peaks.

GO analysis

GO and tissue-specific expression analysis of gene sets were performed using DAVID (*64*).

Motif-finding analysis

De novo motif discovery analysis was performed using Regulatory Sequence Analysis Tools (*65*), and the discovered motifs were matched to the JASPAR transcription factor database (*66*).

Gene expression data

The microarray gene expression data of neurons and NPCs were generated previously using standard procedures (GEO dataset: GSE22946) (*19*, *20*). False discovery ratio analysis of the microarray data was performed with the limma package at default setting.

Co-IP assay

We cotransfected 293T cells with Flag-tagged *Tet2* and V5-tagged *Neurod2* expression vectors. Forty-eight hours after transfection, cells were lysed, and immunoprecipitation was carried out with anti-Flag, anti-V5, or control immunoglobulin G (IgG) antibodies. The immunoprecipitated proteins were analyzed by Western blotting with anti-Flag or anti-V5 antibodies.

Mouse embryo brain cortices at E15.5 days were used for endogenous co-IP. We used the Active Motif Universal Magnetic Co-IP Kit (no. 54002) to perform the co-IP. Antibodies used were rabbit IgG isotype control (Abcam, ab171870), anti-NEUROD2 (Abcam, ab109406), and anti-TET2 (Proteintech, no. 21207-1-AP). The lysate for the co-IP was precleared using rabbit IgG control antibody for 30 min and gently agitated at 4°C. Immune complexes were collected on Dynabeads Protein G from Invitrogen (10003D) for 30 min. Proteins were separated on SDS gels and subjected to Western blotting.

Expression of NEUROD2 in P19 cells

A doxycycline-inducible lentiviral vector for expression of mouse NEUROD2 was constructed using the Gateway cloning system (Thermo Fisher Scientific) in plasmid pCW57.1 (a gift from D. Root; Addgene plasmid no. 41393) and packaged in 293FT cells.

P19 cells (American Type Culture Collection, CRL-1825) were cultured in minimum essential medium α with 7.5% bovine calf serum and 2.5% fetal bovine serum and infected in media containing the lentiviral particles. Lack of mycoplasma contamination was routinely verified. The infected cells were selected by treatment with puromycin (0.5 μg/ml) for 5 days. Expression of NEUROD2 was induced by treatment of the cells with doxycycline (1 μg/ml). The cells were replated onto poly-L-lysine/laminin–coated plates and maintained in neurobasal media containing B27 supplement and 10 μM vitamin C 3 days after doxycycline treatment. Induction of NEUROD2 protein expression was confirmed by Western blot analysis. The antibodies used for Western blotting were anti-NEUROD2 from Abcam (ab109406) and anti-vinculin from Cell Signaling Technology (no. 13901). Fifteen days after the beginning of the doxycycline treatment, the cells were harvested for DNA and RNA isolations. DNA was converted with sodium bisulfite for DNA methylation analysis. For quantitative reverse transcription (RT)–PCR analysis of neuronal marker genes, *Tet1/2/3* genes, and NEUROD2 target genes (*Aff3*, *Nfasc*, *Zhx2*, *Rgmb*, *Itpr2*, *Cog7*, *Wwc1*, *Glis2*, and *Tiam2*), we used TaqMan probes (table S4).

A set of knockdown vectors for *Tet2* containing four unique 29-mer shRNA (GCCACAGAGACTCAACGGTTATCAGGCTT, GAACAGCATCAGAATGATTG-TGGCTCACC, CTGC-GTTTCATCCAGTCTCTTGCTGAGAA, and GATACTCCT-GGTGAACAAAGTCAGAATGG) expression plasmids was obtained from OriGene (catalog no. TL513547), and the pool of shRNA plasmids was transfected into the NEUROD2-inducible P19 cells prepared as described above. Noneffective 29-mer scrambled shRNA cassette-containing plasmid was used as a control. Two days after transfection, GFP-positive cells were sorted by fluorescence-activated cell sorting. The expression of NEUROD2 was then induced as described above, and the cells were harvested for DNA and RNA analyses.

Luciferase assays

The Rgmb promoter region (chr17:15963232–15964104) was PCR-amplified using primer pairs (forward, 5′-CTTCCTCCTTTCAGG-CCAGC-3′; reverse, 5′-TACGGACCTCGGTGTCATCT-3′). The *Tiam2* promoter region (chr17:3326263–3326948) was PCR-amplified using primer pairs (forward, 5′-CGCGCGCTTTGTAGTTTGTG-3′;

reverse, 5′-AGAGAGAATGCGAAACCGCT-3′). The *Rgmb* NEUROD2-bound putative enhancer region (chr17:15961298–15961700) was PCR-amplified using primer pairs (forward, 5′-GCGTGTCTTGTACTTTCAGCC; reverse, GCAGACTTCGCGCACATTTA-3′). The *Tiam2* NEUROD2-bound region (chr17:3338720–3339246) was PCR-amplified using primer pairs (forward, 5′-GCCACAGGTGATCCAGAAAC-3′; reverse, 5′-CGACTCTGTGGCCATACCTT-3′). The NEUROD2 target PCR fragments were then treated with the DNA methyltransferase M.SssI (NEB, M0226M), according to the manufacturer's instructions, to obtain fully methylated DNA fragments. For each gene, the promoter region and the methylated or unmethylated NEUROD2 target DNA fragments were ligated into the pGL3-basic luciferase reporter vector (Promega, E1751).

One hundred thousand P19 cells or Neuro-2a cells were seeded into 24-well plates before transfection. To induce *NeuroD2* expression, P19 cells were treated with doxycycline, and Neuro-2a cells were infected with a lentiviral *Neurod2* expression vector for 24 hours before transfection. We transfected 47.5 ng of pGL3 luciferase reporter vectors (with gene-specific promoter and methylated or unmethylated NEUROD2 target sites) and 2.5 ng of internal control *Renilla* luciferase reporter vector (pRL-CMV, Promega, Madison, WI) into the cells. We harvested the cells 48 hours after transfection. All transfections were carried out at least in three independent experiments and in triplicate. Firefly and *Renilla* luciferase activities were assayed with the Dual Luciferase Assay Kit (Promega) according to the manufacturer's instructions. The firefly luciferase activities were normalized relative to *Renilla* activity.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/10/eaax0080/DC1

Fig. S1. Global analysis of WGBS data.
Fig. S2. GO analysis of genes associated with hypomethylated DMRs upstream of the TSS and downstream of the TES.
Fig. S3. Genes with intragenic hypomethylated DMRs preferentially become activated in neurons.
Fig. S4. Many intragenic hypomethylated DMRs are NEUROD2 targets.
Fig. S5. Accumulation of 5hmC at NEUROD2 binding sites undergoing DNA demethylation.
Fig. S6. Luciferase assays with methylated and unmethylated NEUROD2-targeted regions.
Fig. S7. DNA methylation changes at NEUROD2 target genes in *Neurod2* knockout mice.
Fig. S8. Analysis of NEUROD2 target genes in P19 cells.
Fig. S9. Analysis of NEUROD2 target genes in P19 cells after *Tet2* knockdown.
Table S1. DMRs between neurons and NPCs identified by the DSS method.
Table S2. Differential expression of genes with intragenic hypomethylated DMRs.
Table S3. NEUROD2 peaks at H3K27Ac-enriched hypomethylated DMRs within genes.
Table S4. Primers for bisulfite sequencing and RT-PCR.

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. B. Martynoga, D. Drechsel, F. Guillemot, Molecular control of neurogenesis: A view from the mammalian cerebral cortex. *Cold Spring Harb. Perspect. Biol.* **4**, a008359 (2012).
2. B. Yao, K. M. Christian, C. He, P. Jin, G. L. Ming, H. Song, Epigenetic mechanisms in neurogenesis. *Nat. Rev. Neurosci.* **17**, 537–549 (2016).
3. E. Taverna, M. Götz, W. B. Huttner, The cell biology of neurogenesis: Toward an understanding of the development and evolution of the neocortex. *Annu. Rev. Cell Dev. Biol.* **30**, 465–502 (2014).
4. L. C. Greig, M. B. Woodworth, M. J. Galazo, H. Padmanabhan, J. D. Macklis, Molecular logic of neocortical projection neuron specification, development and diversity. *Nat. Rev. Neurosci.* **14**, 755–769 (2013).
5. S. Lodato, P. Arlotta, Generating neuronal diversity in the mammalian cerebral cortex. *Annu. Rev. Cell Dev. Biol.* **31**, 699–720 (2015).
6. J. H. Lui, D. V. Hansen, A. R. Kriegstein, Development and evolution of the human neocortex. *Cell* **146**, 18–36 (2011).
7. Z. Wang, B. Tang, Y. He, P. Jin, DNA methylation dynamics in neurogenesis. *Epigenomics* **8**, 401–414 (2016).
8. M. Fasolino, Z. Zhou, The cucial role of DNA methylation and MeCP2 in neuronal function. *Genes* **8**, E141 (2017).
9. J. Hsieh, X. Zhao, Genetics and epigenetics in adult neurogenesis. *Cold Spring Harb. Perspect. Biol.* **8**, a018911 (2016).
10. M. M. Suzuki, A. Bird, DNA methylation landscapes: Provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465–476 (2008).
11. D. Schübeler, Function and information content of DNA methylation. *Nature* **517**, 321–326 (2015).
12. M. Kim, J. Costello, DNA methylation: An epigenetic mark of cellular memory. *Exp. Mol. Med.* **49**, e322 (2017).
13. C. M. Rivera, B. Ren, Mapping human epigenomes. *Cell* **155**, 39–55 (2013).
14. P. Sen, P. P. Shah, R. Nativio, S. L. Berger, Epigenetic mechanisms of longevity and aging. *Cell* **166**, 822–839 (2016).
15. M. Iurlaro, F. von Meyenn, W. Reik, DNA methylation homeostasis in human and mouse development. *Curr. Opin. Genet. Dev.* **43**, 101–109 (2017).
16. M. B. Stadler, R. Murr, L. Burger, R. Ivanek, F. Lienert, A. Schöler, E. Nimwegen, C. Wirbelauer, E. J. Oakeley, D. Gaidatzis, V. K. Tiwari, D. Schübeler, DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).
17. F. Mohn, M. Weber, M. Rebhan, T. C. Roloff, J. Richter, M. B. Stadler, M. Bibel, D. Schübeler, Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* **30**, 755–766 (2008).
18. M. J. Ziller, R. Edri, Y. Yaffe, J. Donaghey, R. Pop, W. Mallard, R. Issner, C. A. Gifford, A. Goren, J. Xing, H. Gu, D. Cacchiarelli, A. M. Tsankov, C. Epstein, J. L. Rinn, T. S. Mikkelsen, O. Kohlbacher, A. Gnirke, B. E. Bernstein, Y. Elkabetz, A. Meissner, Dissecting neural differentiation regulatory networks through epigenetic footprinting. *Nature* **518**, 355–359 (2015).
19. M. A. Hahn, R. Qiu, X. Wu, A. X. Li, H. Zhang, J. Wang, J. Jui, S. G. Jin, Y. Jiang, G. P. Pfeifer, Q. Lu, Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Rep.* **3**, 291–300 (2013).
20. J. Wang, H. Zhang, A. G. Young, R. Qiu, S. Argalian, X. Li, X. Wu, G. Lemke, Q. Lu, Transcriptome analysis of neural progenitor cells by a genetic dual reporter strategy. *Stem Cells* **29**, 1589–1600 (2011).
21. Y. Huang, W. A. Pastor, Y. Shen, M. Tahiliani, D. R. Liu, A. Rao, The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLOS ONE* **5**, e8888 (2010).
22. S. G. Jin, S. Kadam, G. P. Pfeifer, Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res.* **38**, e125 (2010).
23. R. Lister, E. A. Mukamel, J. R. Nery, M. Urich, C. A. Puddifoot, N. D. Johnson, J. Lucero, Y. Huang, A. J. Dwork, M. D. Schultz, M. Yu, J. Tonti-Filippini, H. Heyn, S. Hu, J. C. Wu, A. Rao, M. Esteller, C. He, F. G. Haghighi, T. J. Sejnowski, M. M. Behrens, J. R. Ecker, Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 1237905 (2013).
24. C. Luo, M. A. Lancaster, R. Castanon, J. R. Nery, J. A. Knoblich, J. R. Ecker, Cerebral organoids recapitulate epigenomic signatures of the human fetal brain. *Cell Rep.* **17**, 3369–3384 (2016).
25. H. Feng, K. N. Conneely, H. Wu, A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. *Nucleic Acids Res.* **42**, e69 (2014).
26. A. S. Nord, M. J. Blow, C. Attanasio, J. A. Akiyama, A. Holt, R. Hosseini, S. Phouanenavong, I. Plajzer-Frick, M. Shoukry, V. Afzal, J. L. R. Rubenstein, E. M. Rubin, L. A. Pennacchio, A. Visel, Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* **155**, 1521–1531 (2013).
27. A. N. Malik, T. Vierbuchen, M. Hemberg, A. A. Rubin, E. Ling, C. H. Couch, H. Stroud, I. Spiegel, K. K. H. Farh, D. A. Harmin, M. E. Greenberg, Genome-wide identification and characterization of functional neuronal activity-dependent enhancers. *Nat. Neurosci.* **17**, 1330–1339 (2014).
28. I. Bormuth, K. Yan, T. Yonemasu, M. Gummert, M. Zhang, S. Wichert, O. Grishina, A. Pieper, W. Zhang, S. Goebbels, V. Tarabykin, K. A. Nave, M. H. Schwab, Neuronal basic helix-loop-helix proteins Neurod2/6 regulate cortical commissure formation before midline interactions. *J. Neurosci.* **33**, 641–651 (2013).
29. M. H. Schwab, S. Druffel-Augustin, P. Gass, M. Jung, M. Klugmann, A. Bartholomae, M. J. Rossner, K. A. Nave, Neuronal basic helix-loop-helix proteins (NEX, neuroD, NDRF): Spatiotemporal expression and targeted disruption of the NEX gene in transgenic mice. *J. Neurosci.* **18**, 1408–1418 (1998).
30. P. S. Joshi, B. J. Molyneaux, L. Feng, X. Xie, J. D. Macklis, L. Gan, Bhlhb5 regulates the postmitotic acquisition of area identities in layers II-V of the developing neocortex. *Neuron* **60**, 258–272 (2008).
31. S. E. Ross, A. E. McCord, C. Jung, D. Atan, S. I. Mok, M. Hemberg, T. K. Kim, J. Salogiannis, L. Hu, S. Cohen, Y. Lin, D. Harrar, R. R. McInnes, M. E. Greenberg, Bhlhb5 and Prdm8 form a repressor complex involved in neuronal circuit assembly. *Neuron* **73**, 292–303 (2012).

32. F. Chen, J. T. Moran, Y. Zhang, K. M. Ates, D. Yu, L. A. Schrader, P. M. Das, F. E. Jones, B. J. Hall, The transcription factor NeuroD2 coordinates synaptic innervation and cell intrinsic properties to control excitability of cortical pyramidal neurons. *J. Physiol.* **594**, 3729–3744 (2016).
33. S. K. Goparaju, K. Kohda, K. Ibata, A. Soma, Y. Nakatake, T. Akiyama, S. Wakabayashi, M. Matsushita, M. Sakota, H. Kimura, M. Yuzaki, S. B. H. Ko, M. S. H. Ko, Rapid differentiation of human pluripotent stem cells into functional neurons by mRNAs encoding transcription factors. *Sci. Rep.* **7**, 42367 (2017).
34. L. Telley, S. Govindan, J. Prados, I. Stevant, S. Nef, E. Dermitzakis, A. Dayer, D. Jabaudon, Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex. *Science* **351**, 1443–1446 (2016).
35. G. Ince-Dunn, B. J. Hall, S. C. Hu, B. Ripley, R. L. Huganir, J. M. Olson, S. J. Tapscott, A. Ghosh, Regulation of thalamocortical patterning and synaptic maturation by NeuroD2. *Neuron* **49**, 683–695 (2006).
36. J. M. Olson, A. Asakura, L. Snider, R. Hawkes, A. Strand, J. Stoeck, A. Hallahan, J. Pritchard, S. J. Tapscott, NeuroD2 is necessary for development and survival of central nervous system neurons. *Dev. Biol.* **234**, 174–187 (2001).
37. W. A. Whyte, D. A. Orlando, D. Hnisz, B. J. Abraham, C. Y. Lin, M. H. Kagey, P. B. Rahl, T. I. Lee, R. A. Young, Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
38. E. Bayam, G. S. Sahin, G. Guzelsoy, G. Guner, A. Kabakcioglu, G. Ince-Dunn, Genome-wide target analysis of NEUROD2 provides new insights into regulation of cortical projection neuron migration and differentiation. *BMC Genomics* **16**, 681 (2015).
39. G. C. Hon, C. X. Song, T. Du, F. Jin, S. Selvaraj, A. Y. Lee, C. A. Yen, Z. Ye, S. Q. Mao, B. A. Wang, S. Kuan, L. E. Edsall, B. S. Zhao, G. L. Xu, C. He, B. Ren, 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol. Cell* **56**, 286–297 (2014).
40. F. Lu, Y. Liu, L. Jiang, S. Yamaguchi, Y. Zhang, Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev.* **28**, 2103–2119 (2014).
41. K. D. Rasmussen, G. Jia, J. V. Johansen, M. T. Pedersen, N. Rapin, F. O. Bagger, B. T. Porse, O. A. Bernard, J. Christensen, K. Helin, Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes Dev.* **29**, 910–922 (2015).
42. J. L. Sardina, S. Collombet, T. V. Tian, A. Gómez, B. Di Stefano, C. Berenguer, J. Brumbaugh, R. Stadhouders, C. Segura-Morales, M. Gut, I. G. Gut, S. Heath, S. Aranda, L. Di Croce, K. Hochedlinger, D. Thieffry, T. Graf, Transcription factors drive Tet2-mediated enhancer demethylation to reprogram cell fate. *Cell Stem Cell* **23**, 727–741.e9 (2018).
43. L. Wang, P. A. Ozark, E. R. Smith, Z. Zhao, S. A. Marshall, E. J. Rendleman, A. Piunti, C. Ryan, A. L. Whelan, K. A. Helmin, M. A. Morgan, L. Zou, B. D. Singer, A. Shilatifard, TET2 coactivates gene expression through demethylation of enhancers. *Sci. Adv.* **4**, eaau6986 (2018).
44. J. Yamazaki, J. Jelinek, Y. Lu, M. Cesaroni, J. Madzo, F. Neumann, R. He, R. Taby, A. Vasanthakumar, T. Macrae, K. R. Ostler, H. M. Kantarjian, S. Liang, M. R. Estecio, L. A. Godley, J. P. J. Issa, TET2 mutations affect non-CpG island DNA methylation at enhancers and transcription factor-binding sites in chronic myelomonocytic leukemia. *Cancer Res.* **75**, 2833–2843 (2015).
45. M. H. Farah, J. M. Olson, H. B. Sucic, R. I. Hume, S. J. Tapscott, D. L. Turner, Generation of neurons by transient expression of neural bHLH proteins in mammalian cells. *Development* **127**, 693–702 (2000).
46. A. P. Fong, Z. Yao, J. W. Zhong, Y. Cao, W. L. Ruzzo, R. C. Gentleman, S. J. Tapscott, Genetic and epigenetic determinants of neurogenesis and myogenesis. *Dev. Cell* **22**, 721–735 (2012).
47. H. K. Long, N. P. Blackledge, R. J. Klose, ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. *Biochem. Soc. Trans.* **41**, 727–740 (2013).
48. S. G. Jin, Z. M. Zhang, T. L. Dunwell, M. R. Harter, X. Wu, J. Johnson, Z. Li, J. Liu, P. E. Szabó, Q. Lu, G. L. Xu, J. Song, G. P. Pfeifer, Tet3 reads 5-Carboxylcytosine through Its CXXC domain and is a potential guardian against neurodegeneration. *Cell Rep.* **14**, 493–505 (2016).
49. K. Williams, J. Christensen, K. Helin, DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep.* **13**, 28–35 (2011).
50. O. Bogdanovic, A. H. Smits, E. de la Calle Mustienes, J. J. Tena, E. Ford, R. Williams, U. Senanayake, M. D. Schultz, S. Hontelez, I. van Kruijsbergen, T. Rayon, F. Gnerlich, T. Carell, G. J. Veenstra, M. Manzanares, T. Sauka-Spengler, J. R. Ecker, M. Vermeulen, J. L. Gómez-Skarmeta, R. Lister, Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nat. Genet.* **48**, 417–426 (2016).
51. E. L. Putiri, R. L. Tiedemann, J. J. Thompson, C. Liu, T. Ho, J. H. Choi, K. D. Robertson, Distinct and overlapping control of 5-methylcytosine and 5-hydroxymethylcytosine by the TET proteins in human cancer cells. *Genome Biol.* **15**, R81 (2014).
52. A. A. Serandour, S. Avner, F. Oger, M. Bizot, F. Percevault, C. Lucchetti-Miganeh, G. Palierne, C. Gheeraert, F. Barloy-Hubler, C. L. Péron, T. Madigou, E. Durand, P. Froguel, B. Staels, P. Lefebvre, R. Métivier, J. Eeckhoute, G. Salbert, Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Res.* **40**, 8255–8265 (2012).
53. S. Orlanski, V. Labi, Y. Reizel, A. Spiro, M. Lichtenstein, R. Levin-Klein, S. B. Koralov, Y. Skversky, K. Rajewsky, H. Cedar, Y. Bergman, Tissue-specific DNA demethylation is required for proper B-cell differentiation and function. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5018–5023 (2016).
54. C. W. Lio, J. Zhang, E. González-Avalos, P. G. Hogan, X. Chang, A. Rao, Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *eLife* **5**, e18290 (2016).
55. Y. A. Yang, J. C. Zhao, K. W. Fong, J. Kim, S. Li, C. Song, B. Song, B. Zheng, C. He, J. Yu, FOXA1 potentiates lineage-specific enhancer activation through modulating TET1 expression and function. *Nucleic Acids Res.* **44**, 8153–8164 (2016).
56. J. Donaghey, S. Thakurela, J. Charlton, J. S. Chen, Z. D. Smith, H. Gu, R. Pop, K. Clement, E. K. Stamenova, R. Karnik, D. R. Kelley, C. A. Gifford, D. Cacchiarelli, J. L. Rinn, A. Gnirke, M. J. Ziller, A. Meissner, Genetic determinants and epigenetic effects of pioneer-factor occupancy. *Nat. Genet.* **50**, 250–258 (2018).
57. K. S. Zaret, Pioneering the chromatin landscape. *Nat. Genet.* **50**, 167–169 (2018).
58. A. Mayran, K. Khetchoumian, F. Hariri, T. Pastinen, Y. Gauthier, A. Balsalobre, J. Drouin, Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate. *Nat. Genet.* **50**, 259–269 (2018).
59. J. Xiong, Z. Zhang, J. Chen, H. Huang, Y. Xu, X. Ding, Y. Zheng, R. Nishinakamura, G. L. Xu, H. Wang, S. Chen, S. Gao, B. Zhu, Cooperative action between SALL4A and TET proteins in stepwise oxidation of 5-methylcytosine. *Mol. Cell* **64**, 913–925 (2016).
60. D. S. Castro, B. Martynoga, C. Parras, V. Ramesh, E. Pacary, C. Johnston, D. Drechsel, M. Lebel-Potter, L. G. Garcia, C. Hunt, D. Dolle, A. Bithell, L. Ettwiller, N. Buckley, F. Guillemot, A novel function of the proneural factor Ascl1 in progenitor proliferation identified by genome-wide characterization of its targets. *Genes Dev.* **25**, 930–945 (2011).
61. A. Pataskar, J. Jung, P. Smialowski, F. Noack, F. Calegari, T. Straub, V. K. Tiwari, NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. *EMBO J.* **35**, 24–45 (2016).
62. A. Sessa, E. Ciabatti, D. Drechsel, L. Massimino, G. Colasante, S. Giannelli, T. Satoh, S. Akira, F. Guillemot, V. Broccoli, The Tbr2 molecular network controls cortical neuronal differentiation through complementary genetic and epigenetic pathways. *Cereb. Cortex* **27**, 3378–3396 (2017).
63. M. A. Hahn, X. Wu, A. X. Li, T. Hahn, G. P. Pfeifer, Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. *PLOS ONE* **6**, e18844 (2011).
64. W. Huang da, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
65. A. Medina-Rivera, M. Defrance, O. Sand, C. Herrmann, J. A. Castro-Mondragon, J. Delerce, S. Jaeger, C. Blanchet, P. Vincens, C. Caron, D. M. Staines, B. Contreras-Moreira, M. Artufel, L. Charbonnier-Khamvongsa, C. Hernandez, D. Thieffry, M. Thomas-Chollier, J. van Helden, RSAT 2015: Regulatory sequence analysis tools. *Nucleic Acids Res.* **43**, W50–W56 (2015).
66. A. Mathelier, O. Fornes, D. J. Arenillas, C. Y. Chen, G. Denay, J. Lee, W. Shi, C. Shyr, G. Tan, R. Worsley-Hunt, A. W. Zhang, F. Parcy, B. Lenhard, A. Sandelin, W. W. Wasserman, JASPAR 2016: A major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **44**, D110–D115 (2016).

Acknowledgments

Funding: This work was supported by NIH grants MH094599 (to Q.L. and G.P.P) and CA160965 (to G.P.P.) and by the Van Andel Research Institute. **Author contributions:** M.A.H., S.-G.J., Q.L., and G.P.P. conceived and designed the research. M.A.H., S.-G.J., J.L., Z.H., B.-W.K., J.J., A.-D.V.B., J.A.Y., Y.F., S.G., M.H.S., Q.L., and G.P.P. performed the experiments and analyzed the data. A.X.L., S.T., Z.H., and X.W. performed computational data analysis. M.A.H., Q.L., and G.P.P. interpreted the data and wrote the manuscript. All authors contributed to critical reading of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. Bisulfite sequencing data are available from the NCBI/GEO repository under accession no. GSE101090. Gene expression data are from GSE22946.

Submitted 14 February 2019
Accepted 13 September 2019
Published 23 October 2019
10.1126/sciadv.aax0080

Citation: M. A. Hahn, S.-G. Jin, A. X. Li, J. Liu, Z. Huang, X. Wu, B.-W. Kim, J. Johnson, A.-D. V. Bilbao, S. Tao, J. A. Yim, Y. Fong, S. Goebbels, M. H. Schwab, Q. Lu, G. P. Pfeifer, Reprogramming of DNA methylation at NEUROD2-bound sequences during cortical neuron differentiation. *Sci. Adv.* **5**, eaax0080 (2019).