nature genetics

# Targeted disruption of *DNMT1*, *DNMT3A* and *DNMT3B* in human embryonic stem cells

Jing Liao[1–3,5], Rahul Karnik[1–3,5], Hongcang Gu[1], Michael J Ziller[1–3], Kendell Clement[1–3], Alexander M Tsankov[1–3], Veronika Akopian[1–3], Casey A Gifford[1–3], Julie Donaghey[1–3], Christina Galonska[1–3], Ramona Pop[1–3], Deepak Reyon[4], Shengdar Q Tsai[4], William Mallard[1–3], J Keith Joung[4], John L Rinn[1–3], Andreas Gnirke[1] & Alexander Meissner[1–3]

© 2015 Nature America, Inc. All rights reserved.

**DNA methylation is a key epigenetic modification involved in regulating gene expression and maintaining genomic integrity. Here we inactivated all three catalytically active DNA methyltransferases (DNMTs) in human embryonic stem cells (ESCs) using CRISPR/Cas9 genome editing to further investigate the roles and genomic targets of these enzymes. Disruption of *DNMT3A* or *DNMT3B* individually as well as of both enzymes in tandem results in viable, pluripotent cell lines with distinct effects on the DNA methylation landscape, as assessed by whole-genome bisulfite sequencing. Surprisingly, in contrast to findings in mouse, deletion of *DNMT1* resulted in rapid cell death in human ESCs. To overcome this immediate lethality, we generated a doxycycline-responsive tTA-DNMT1* rescue line and readily obtained homozygous *DNMT1*-mutant lines. However, doxycycline-mediated repression of exogenous *DNMT1** initiates rapid, global loss of DNA methylation, followed by extensive cell death. Our data provide a comprehensive characterization of DNMT-mutant ESCs, including single-base genome-wide maps of the targets of these enzymes.**

DNA methylation is a heritable and reversible enzyme-mediated modification to DNA that is implicated in a diverse range of biological processes[1]. In mammals, there are three catalytically active DNMTs: DNMT1, which is largely responsible for the maintenance of DNA methylation over replication, and DNMT3A and DNMT3B, which generally perform *de novo* methylation of either unmethylated DNA or hemimethylated DNA to assist in maintenance[2]. Deletion of these enzymes in mice results in embryonic (*Dnmt1* and *Dnmt3b*) or postnatal (*Dnmt3a*) lethality, confirming their essential roles in development[3,4]. In line with knockout mouse models, mutations affecting the human DMNTs are linked to several diseases, including acute myeloid leukemia (AML) and immunodeficiency, centromere instability and facial anomalies (ICF) syndrome[5–8].

DNA methylation has been extensively mapped in human and mouse somatic cells as well as in ESCs[9–14]. Human ESC methylation patterns are most unique at hypomethylated regulatory elements that are enriched for binding of pluripotency-associated master regulators, such as OCT4, SOX2 and NANOG[14]. Additionally, both mouse and human ESCs display increased levels of non-CpG methylation that are not observed in most somatic lineages, consistent with the high activity of *de novo* methyltransferases in these cells[11,15–17].

Although deletion of *Dnmt1* is lethal in all dividing somatic cells[3,18–21], mouse ESCs are viable, despite global loss of DNA methylation. In fact, all three mouse DNMT genes can be removed from these cells without any deleterious effects in the undifferentiated state[22]. As such, mouse ESCs have become a powerful tool to study the role and function of the DNMT enzymes, allowing clarification of some of their specific targets and providing many general insights into the biology of DNA methylation[23]. No comparable efforts have been reported for human pluripotent stem cells, and loss-of-function studies have been limited to the depletion of *DNMT1* in the colon cancer cell line HCT116, which results in cell death[24,25] and therefore indicates a similar requirement for the maintenance of DNA methylation patterns in human cells. DNMT3B was reported to cooperate with DNMT1 to maintain methylation in HCT116 cells[26,27], and its depletion using RNA interference (RNAi) results in altered timing for neuronal differentiation and maturation[28]. Recently, a human ESC model for ICF syndrome with targeted disruption of *DNMT3B* was reported[29].

The limited number of human studies can likely be attributed to long-standing difficulties of genetic manipulation in the human ESC system rather than to lack of interest. However, new classes of genome-editing tools—transcription activator–like effector nucleases (TALENs) and clustered, regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems—have been applied for genomic editing in mammalian cells[30–34]. These techniques substantially enhance the targeting efficiency in human ESCs, allowing previously difficult studies to be undertaken[33,35].

Here we report the inactivation of all three DNMTs in human ESCs and present a detailed analysis of the DNA methylation changes in *DNMT1*, *DNMT3A* or *DNMT3B* knockout lines as well as in

[1]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. [2]Harvard Stem Cell Institute, Cambridge, Massachusetts, USA. [3]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts, USA. [4]Department of Pathology, Molecular Pathology Unit, Massachusetts General Hospital, Charlestown, Massachusetts, USA. [5]These authors contributed equally to this work. Correspondence should be addressed to A.M. (alexander_meissner@harvard.edu).

Received 1 August 2014; accepted 6 March 2015; published online 30 March 2015; doi:10.1038/ng.3258

tandem *DNMT3A* and *DNMT3B* double knockouts (*DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$). We further characterized the differentiation potential of the knockout lines using gene expression signatures, embryoid body differentiation and teratoma formation. We also used directed differentiation toward the endoderm lineage to assess the role of DNMT3A in this transition. Most notably, we were unable to obtain homozygous deletions in *DNMT1* without implementing a doxycycline-repressible rescue line, demonstrating that loss of DNMT1 is lethal. Taken together, our results highlight several unique aspects of DNA methylation biology in the context of human ESCs and provide controlled, tractable systems to dissect the function of DNMTs in precise detail.

## RESULTS

### Disrupting the catalytic domain of all three DNMTs

Of the three catalytically active DNMTs, *DNMT3B* is most highly expressed in undifferentiated ESCs (**Fig. 1a**) but also exhibited the most variation when examined across 25 pluripotent lines (**Fig. 1b**). As ESCs differentiate, *DNMT1* and *DNMT3A* remain expressed at comparable levels, whereas *DNMT3B* is strongly downregulated and switches to a catalytically inactive isoform (from isoform 3B1 (NM_006892) to isoform 3B3 (NM_175849); **Fig. 1a** and **Supplementary Fig. 1a**). Although inactive, this isoform can likely still form complexes with catalytically competent DNMT3A and/or DNMT3B to contribute to DNA methylation activity[36].

To better understand the role of the three DNMTs in human pluripotent stem cells and over cellular differentiation, as well as to study their specific target spectra at high resolution across the genome, we created targeted deletions for *DNMT1*, *DNMT3A* and *DNMT3B* individually and for both *DNMT3A* and *DNMT3B* in human ESCs. We selected the male line HUES64 for the following reasons: (i) it is on the US National Institutes of Health (NIH) registry and is generally available to researchers; (ii) it differentiates well into the three germ layers; (iii) it is karyotypically normal (**Supplementary Fig. 1b**) and grows well under standard culture conditions; and (iv) a substantial amount of publicly available transcriptional, epigenomic and transcription factor binding data have been generated for this line[9,37].

To avoid hypomorphic effects[24], we designed guide RNA (gRNA) sequences in the 5′-G(N)$_{19}$NGG-3′ framework targeting sequences encoding the catalytic domains for *DNMT1*, *DNMT3A* and *DNMT3B* (**Fig. 1c**). After picking and expanding individual clones, we confirmed the targeted disruption of the catalytic domains by Sanger sequencing (**Supplementary Fig. 1c,d**). All experiments yielded high rates of mutation ranging from 53 to 66%, with homozygous deletion occurring in 3–6% of the clones for *DNMT3A* and *DNMT3B* (**Supplementary Fig. 1c**). We derived double-knockout cells by targeting the already validated *DNMT3A*$^{-/-}$ clone (clone 139) with the *DNMT3B* gRNA and Cas9. To further validate the *DNMT3A* and *DNMT3B* knockouts, we performed quantitative PCR (qPCR) (**Fig. 1d**) and immunoblotting (**Fig. 1e**). Notably, although we obtained homozygous knockouts for *DNMT3A* and *DNMT3B*, both singly and in combination, we were unable to obtain any *DNMT1*$^{-/-}$ clones, despite the high (>50%) efficiency of deriving heterozygous lines.

### DNMT3-knockout ESCs remain pluripotent

All of our DNMT3-knockout clones grew well and were morphologically normal. We selected representative clones for each knockout (*DNMT3A*$^{-/-}$, clone 139; *DNMT3B*$^{-/-}$, clone 44; *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$, clone 21) and performed several assays

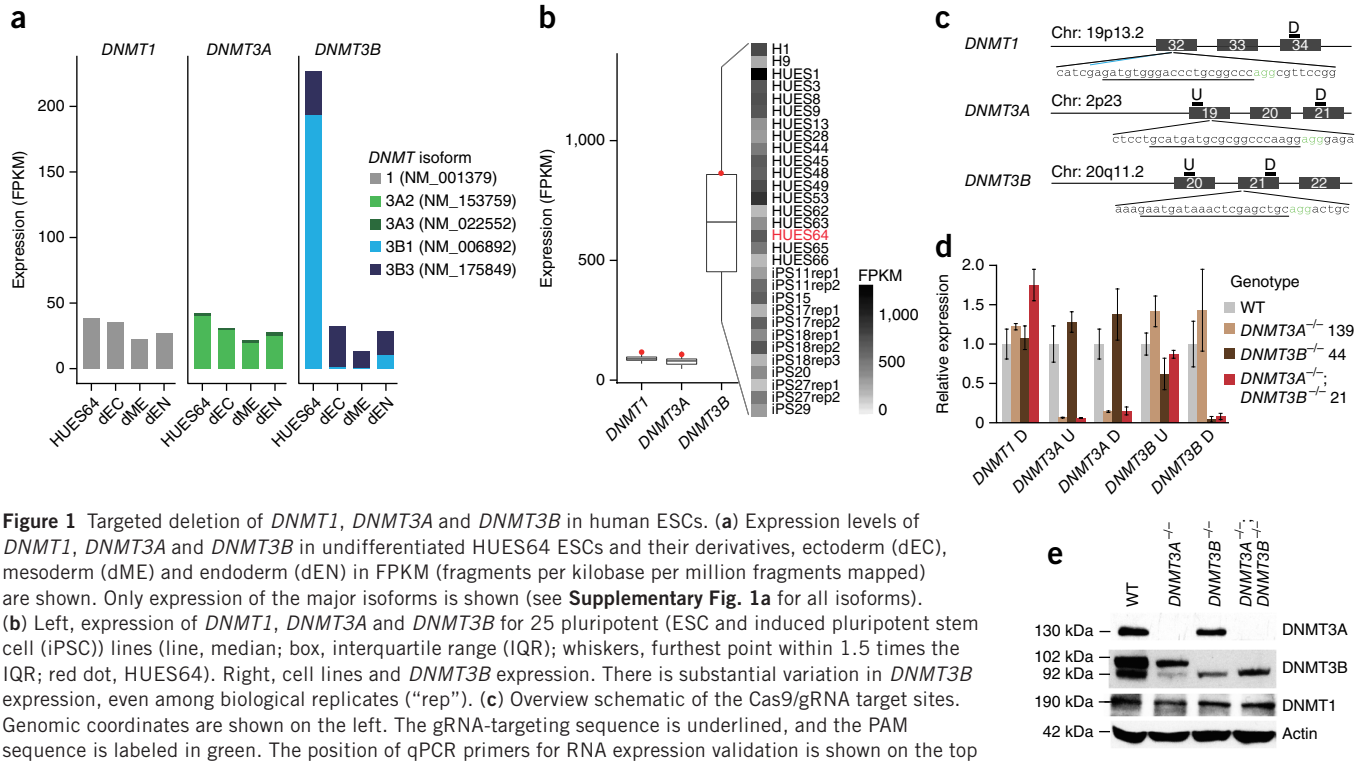



**Figure 1** Targeted deletion of *DNMT1*, *DNMT3A* and *DNMT3B* in human ESCs. (**a**) Expression levels of *DNMT1*, *DNMT3A* and *DNMT3B* in undifferentiated HUES64 ESCs and their derivatives, ectoderm (dEC), mesoderm (dME) and endoderm (dEN) in FPKM (fragments per kilobase per million fragments mapped) are shown. Only expression of the major isoforms is shown (see **Supplementary Fig. 1a** for all isoforms). (**b**) Left, expression of *DNMT1*, *DNMT3A* and *DNMT3B* for 25 pluripotent (ESC and induced pluripotent stem cell (iPSC)) lines (line, median; box, interquartile range (IQR); whiskers, furthest point within 1.5 times the IQR; red dot, HUES64). Right, cell lines and *DNMT3B* expression. There is substantial variation in *DNMT3B* expression, even among biological replicates ("rep"). (**c**) Overview schematic of the Cas9/gRNA target sites. Genomic coordinates are shown on the left. The gRNA-targeting sequence is underlined, and the PAM sequence is labeled in green. The position of qPCR primers for RNA expression validation is shown on the top of the exons. U, upstream; D, downstream. (**d**) RT-qPCR analysis for *DNMT1*, *DNMT3A* and *DNMT3B* in HUES64, *DNMT3A*$^{-/-}$ (139), *DNMT3B*$^{-/-}$ (44) and *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ (21) cells. Primer positions are shown in **c**, and sequences are provided in **Supplementary Table 4**. Error bars were generated from biological triplicates and represent one standard error. (**e**) Immunoblot analysis for DNMT3A, DNMT3B and DNMT1 in HUES64, *DNMT3A*$^{-/-}$ (139), *DNMT3B*$^{-/-}$ (44) and *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ (21) cells. For the DNMT3B antibody, the top band likely represents the major DNMT3B isoform in human ESCs (NM_006892), whereas the smaller band is likely isoform 3 (NM_175849) as reported in ref. 55, which was not targeted by our knockout strategy.

© 2015 Nature America, Inc. All rights reserved.

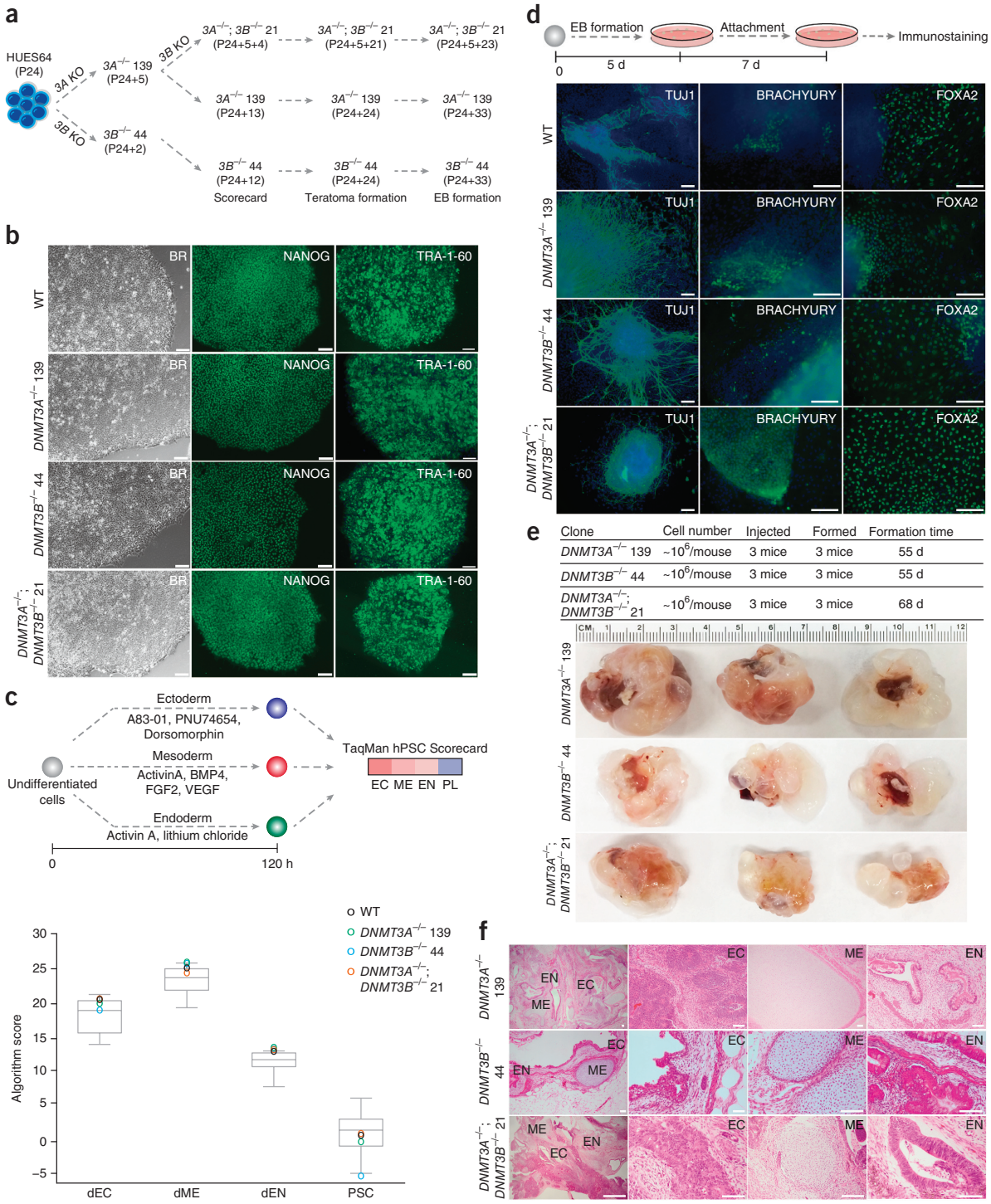



**Figure 2** Assessing the differentiation potential of the DNMT3 knockouts. (**a**) Overview of the clones and passage numbers used for the various assays. HUES64 cells at passage 24 (P24) were used to generate *DNMT3A* and *DNMT3B* single knockouts. The number after the plus sign represents the passage number after targeting. *DNMT3A*$^{-/-}$ clone 139 was passaged five times (P24+5) and then targeted using the *DNMT3B* gRNA. The *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ passage numbers were counted that point. *3A*, *DNMT3A*; *3B*, *DNMT3B*; EB, embryoid body. (**b**) Wild-type (WT) and *DNMT3A*$^{-/-}$, *DNMT3B*$^{-/-}$ and *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells show comparable morphology and expression of characteristic markers. Representative bright-field (BR) and pluripotency-associated marker (NANOG, TRA-1-60) immunostaining images are shown. 10× magnification; scale bars, 100 μm. (**c**) TaqMan hPSC Scorecard analysis of direct differentiation potential for wild-type, *DNMT3A*$^{-/-}$, *DNMT3B*$^{-/-}$ and *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells. Top, schematic overview of the conditions used for direct differentiation into ectoderm (EC), mesoderm (ME) and endoderm (EN). PL, pluripotent. Bottom, differentiation scores for wild-type, *DNMT3A*$^{-/-}$, *DNMT3B*$^{-/-}$ and *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells. The gray boxes represent the score range of 11 reference human PSC lines. (**d**) Top, schematic overview of embryoid body formation. Bottom, representative immunostaining images for TUJ1 (ectoderm), BRACHYURY (mesoderm) and FOXA2 (endoderm). 10× magnification; scale bars, 100 μm. (**e**) Top, summary of injection and teratoma formation results. Bottom, images of the nine teratomas collected. The ruler on top provides a size reference. (**f**) Representative images of hematoxylin and eosin staining of the teratomas from **e**. Neuroepithelium (ectoderm), cartilage with ossification center (mesoderm) and gut-like epithelium are indicated. Scale bars, 50 μm.

© 2015 Nature America, Inc. All rights reserved.

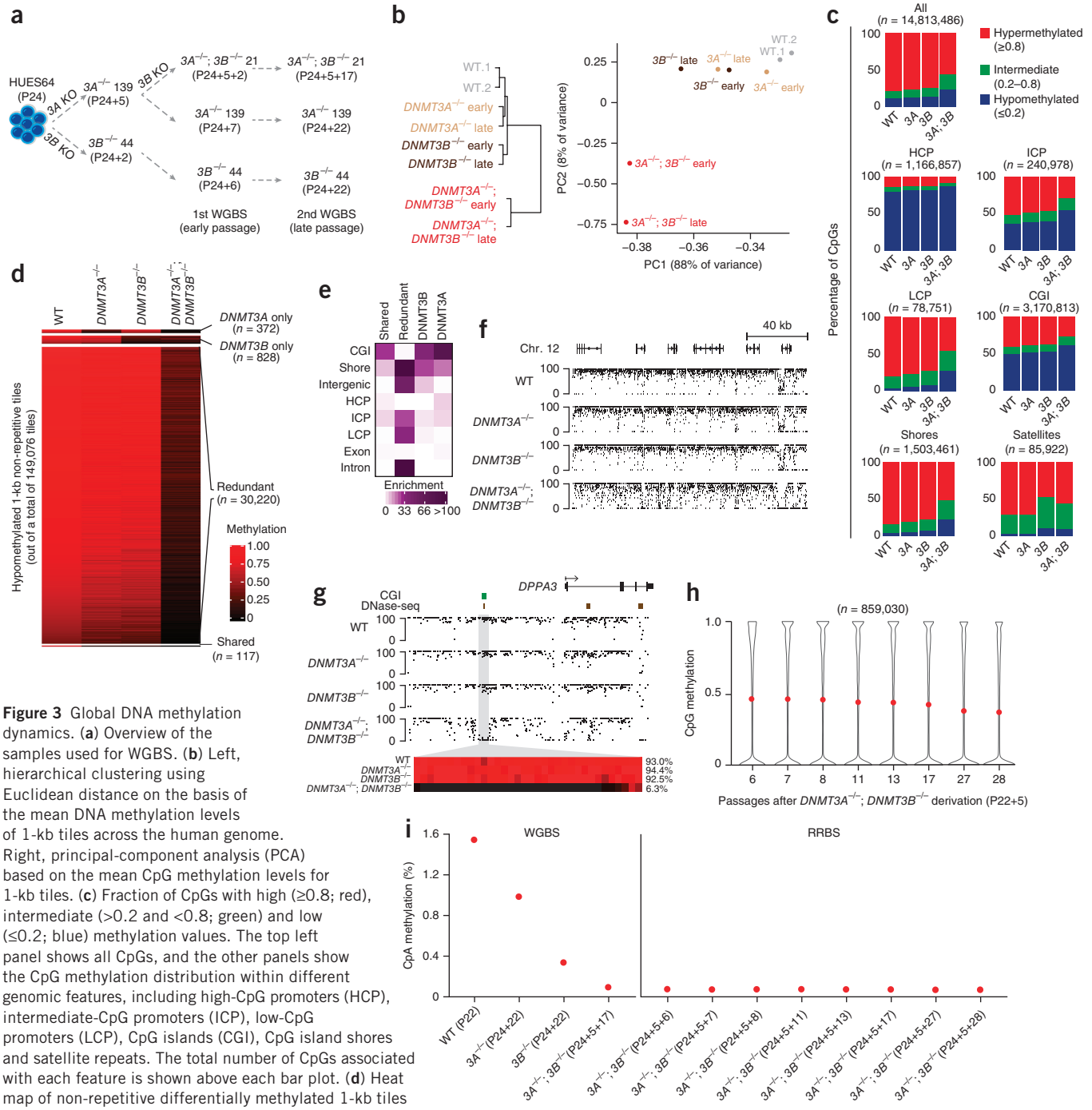



**Figure 3** Global DNA methylation dynamics. (**a**) Overview of the samples used for WGBS. (**b**) Left, hierarchical clustering using Euclidean distance on the basis of the mean DNA methylation levels of 1-kb tiles across the human genome. Right, principal-component analysis (PCA) based on the mean CpG methylation levels for 1-kb tiles. (**c**) Fraction of CpGs with high (≥0.8; red), intermediate (>0.2 and <0.8; green) and low (≤0.2; blue) methylation values. The top left panel shows all CpGs, and the other panels show the CpG methylation distribution within different genomic features, including high-CpG promoters (HCP), intermediate-CpG promoters (ICP), low-CpG promoters (LCP), CpG islands (CGI), CpG island shores and satellite repeats. The total number of CpGs associated with each feature is shown above each bar plot. (**d**) Heat map of non-repetitive differentially methylated 1-kb tiles ($q$ value < 0.05 and methylation difference greater than 0.4). (**e**) Enrichment ($-\log_{10}$ (hypergeometric $P$ value)) of genomic features in the target classes identified in **d**. (**f**) Genome browser tracks covering approximately 160 kb in the different lines. (**g**) Higher-resolution view of the *DPPA3* (*STELLA*) locus. The heat map below shows the DNA methylation values of individual CpGs within the gray region. The average DNA methylation value for the entire highlighted region is shown on the right. (**h**) CpG methylation levels over several passages of the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells. The red dot indicates mean methylation. (**i**) Global mean CpA methylation levels as assayed by WGBS and over several passages of the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells by RRBS.

to characterize these knockout lines, both in terms of their molecular phenotype and their differentiation potential (**Fig. 2a**). The clones were positive for expression of the pluripotency-associated markers NANOG and TRA-1-60 (**Fig. 2b**) and showed trilineage differentiation potential using the TaqMan hPSC Scorecard (**Fig. 2c** and **Supplementary Fig. 2**)[38]. We differentiated these clones into embryoid bodies, and the resulting cells were stained for markers of the three germ layers (**Fig. 2d**). Finally, we injected each clone into the kidney capsule of three independent immunocompromised (NOD-SCID) mice to test for teratoma formation (**Fig. 2e**) and confirmed the presence of cell types representing the three germ layers in several sections for each knockout clone (**Fig. 2f**). These results show that deletion in the DNMT3 enzymes, either individually or in combination, has no obvious effect on morphology, viability or initial differentiation potential in human ESCs. This observation is consistent with previous studies of the respective knockouts in mouse ESCs,

© 2015 Nature America, Inc. All rights reserved.

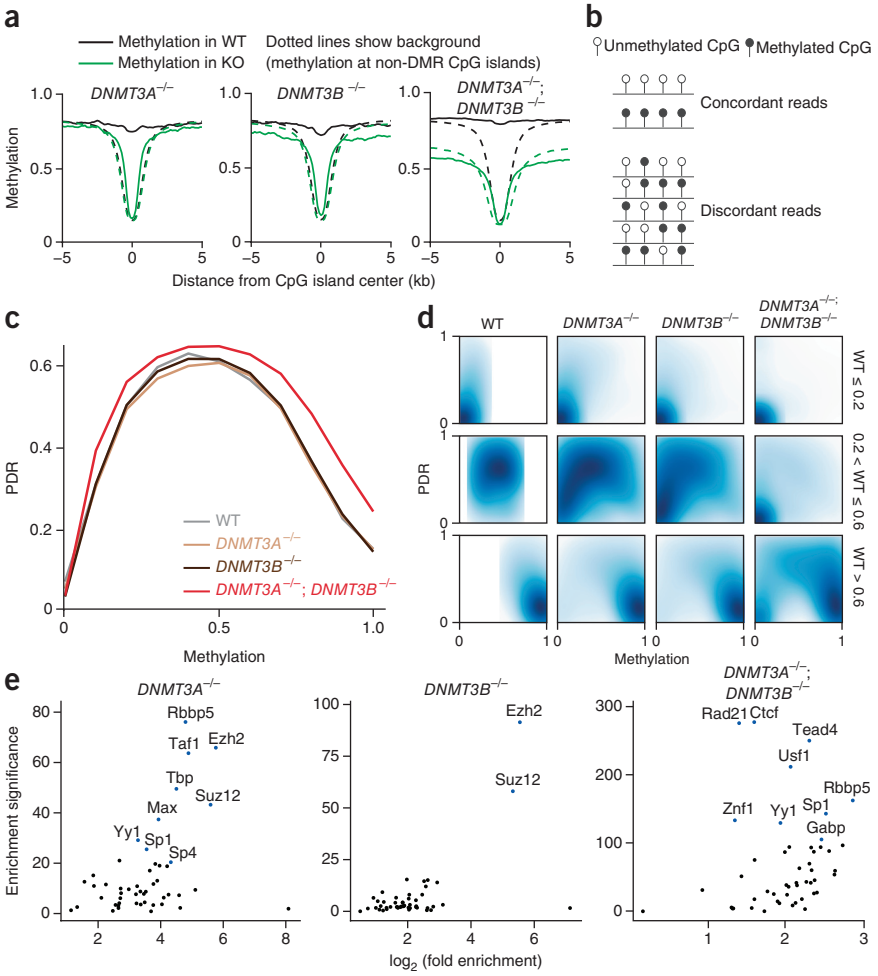



**Figure 4** Characterization of targets for DNMT3A and DNMT3B. (**a**) Composite plot of methylation around CpG islands in wild-type and knockout cells. The solid lines show the mean methylation of differentially methylated CpG islands, and the dashed lines show the mean methylation of all other CpG islands. KO, knockout. (**b**) Schematic for the definition of concordant and discordant reads. (**c**) PDR in wild-type, *DNMT3A*$^{-/-}$, *DNMT3B*$^{-/-}$ and *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells. (**d**) Two-dimensional density plots of PDR versus methylation for CpGs with low (≤0.2), medium (>0.2 and ≤0.6) and high (>0.6) levels of methylation in wild-type cells. (**e**) Enrichment of experimentally determined transcription factor binding sites in H1 ESCs generated by the Encyclopedia of DNA Elements (ENCODE) Project for DMRs in *DNMT3A*$^{-/-}$, *DNMT3B*$^{-/-}$ and *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells compared to wild-type cells. Enrichment significance was defined as the $-\log_{10}$ (*P* value) of the Fisher's exact test, with the background being non-repetitive 1-kb tiles that had at least 0.4 enrichment in wild-type cells.

where the individual knockouts could differentiate normally, although the mouse double knockouts could only generate teratomas at early passage numbers[4,39].

## Passaging results in progressive loss of DNA methylation

To characterize the effects of the DNMT3 knockouts on the human ESC methylome, we isolated genomic DNA from our mutant lines at an early (2–7) and late (17–22) passage and generated whole-genome bisulfite sequencing (WGBS) libraries (**Fig. 3a**). We used the methylation status of 6.9 million CpG sites that were covered at 5× in all 8 samples for the initial comparative analysis. Hierarchical clustering of the WGBS data showed that wild-type, *DNMT3A*$^{-/-}$ and *DNMT3B*$^{-/-}$ cells clustered closely together but remained clearly segregated nonetheless (**Fig. 3b**, left). Despite being derived from the *DNMT3A*$^{-/-}$ clone and therefore beginning in a DNMT3A-depleted background, the double knockout was clearly distinguishable, highlighting the fact that the effects of complete *de novo* methyltransferase ablation on global DNA methylation levels dominate those initiated by the original single knockout. We performed principal-component analysis to further assess the consequences of DNMT3 knockout and of passaging these lines (**Fig. 3b**, right). The first two principal components explained 96% of the variation between the samples and separated them into two main groups: the first group consisted of the wild-type human ESCs and the *DNMT3A* and *DNMT3B* single knockouts, whereas the second group included the early- and late-passage double knockouts.

We then analyzed the WGBS data from the later-passage samples, which captured 14.8 million CpG sites at 5× coverage in all 4 samples, to determine whether specific genomic features were particularly sensitive to the loss of *de novo* DNMT activity. We divided all the CpG sites into three classes according to their methylation values—high (≥0.8), intermediate (>0.2 and <0.8) or low (≤0.2)—and examined the relative proportion falling into each category (**Fig. 3c**). The fraction of regions with intermediate and low levels of methylation in the knockouts gradually increased with passage number, and was most apparent for the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ lines, a finding similar to those obtained from equivalent mouse knockout models (**Supplementary Fig. 3a**). Most genomic features showed a mild decrease in DNA methylation levels in the single knockouts and a more substantial change in the double knockouts, with the exception of satellites, which appeared particularly sensitive to the loss of *DNMT3B* (**Fig. 3c** and **Supplementary Fig. 3b–d**).

We next divided the genome into 1-kb tiles and compared the distribution of CpG methylation values within each tile across the samples (**Fig. 3d**). We defined differentially methylated regions (DMRs) as those tiles that were significantly different (weighted *t* test, *q* value < 0.05) between the wild-type cells and any of the knockouts and had an average methylation difference greater than 0.4, ignoring repetitive tiles. These DMRs could be divided into four classes: (i) DNMT3A-specific targets, (ii) DNMT3B-specific targets, (iii) shared targets that required both DNMT3A and DNMT3B; and (iv) redundant targets that required either DNMT3A or DNMT3B and were only lost in *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells. As suggested by the analysis for CpG levels above, almost all DMRs (~96%) were redundantly targeted by DNMT3A and DNMT3B and only lost methylation when both were ablated. For each class, we calculated the enrichment of different genomic features. Redundant targets were notably broad in scope and were enriched for intergenic regions, CpG island shores, promoters with intermediate and low CpG density, and introns (**Fig. 3e**). In contrast, the other three classes were largely enriched for CpG islands (**Supplementary Table 1**). These enrichments suggest that DNMT3A and DNMT3B act redundantly in regions of low CpG density but may be more selectively targeted to regions of high CpG density.

© 2015 Nature America, Inc. All rights reserved.

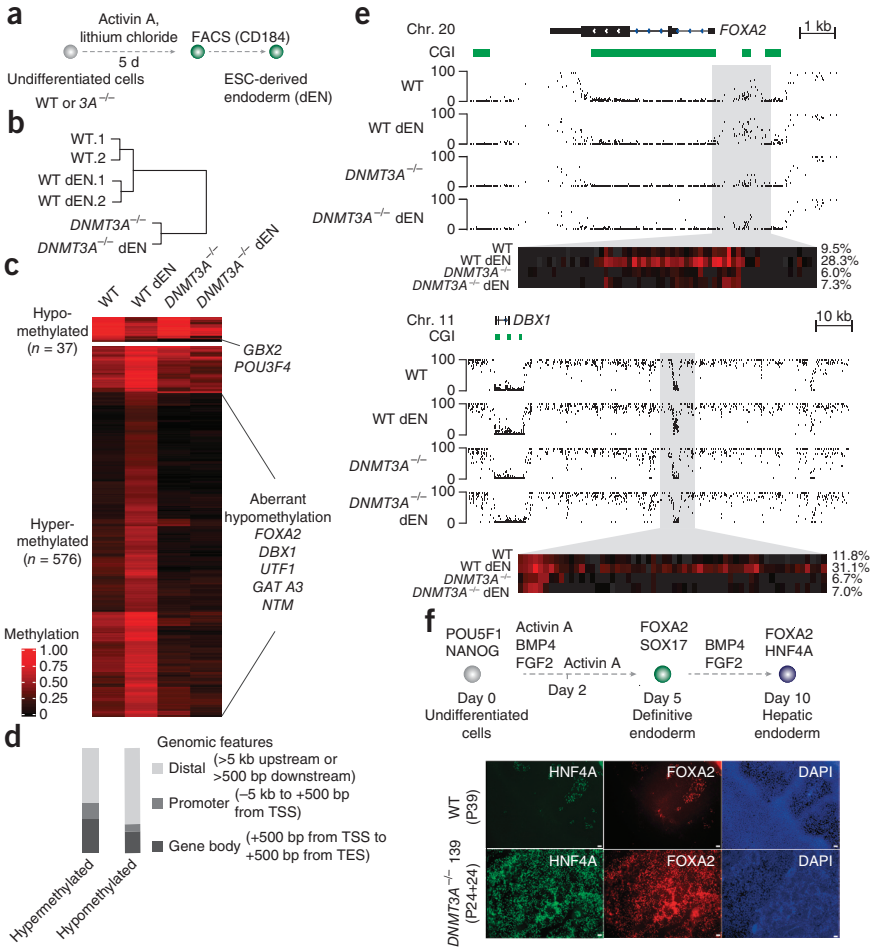



**Figure 5** Effect of *DNMT3A* deletion on endoderm differentiation. (**a**) Schematic for the generation of CD184+ dENs. (**b**) Global analysis for the WGBS data of undifferentiated and dEN cells from wild-type (P23 and P26) and *DNMT3A*$^{-/-}$ (P24+13) cells. Hierarchical clustering based on the mean DNA methylation levels of 1-kb tiles across the human genome using Euclidean distance is shown. 1 and 2 represent two biological replicates. The passage numbers of the replicates were within two passages of each other. (**c**) Heat map of methylation levels (black, 0; red, 1) in wild-type, wild-type dEN, *DNMT3A*$^{-/-}$ and *DNMT3A*$^{-/-}$ dEN cells of DMRs that change methylation significantly ($q$ value < 0.05 and methylation difference of at least 0.2) between wild-type and wild-type dEN cells. (**d**) Classification of endoderm DMRs into promoters, gene bodies and distal regions. TSS, transcription start site; TES, transcription end site. (**e**) Genome browser tracks of approximately 10-kb and 100-kb regions highlight global and local methylation levels for *FOXA2* and *DBX1*. The heat map below shows the DNA methylation values of individual CpGs within the highlighted region. The average DNA methylation value for the entire highlighted region is shown on the right. (**f**) Hepatoblast differentiation of wild-type (P39) and *DNMT3A*$^{-/-}$ (P24+24) cells. Top, schematic of hepatoblast differentiation. Bottom, representative images from two individual biological replicates are shown. 4× magnification of immunostaining images for the hepatoblast markers FOXA2 and HNF4A for HUES64 (wild type) and *DNMT3A*$^{-/-}$ (clone 139) cells are shown. DAPI, 4′, 6-diamidino-2-phenylindole. Scale bars, 100 µm.

The individual knockouts showed a mild, global reduction in methylation levels, an effect that became slightly more pronounced in the double knockouts (**Fig. 3f**). In contrast to this limited global decrease, several loci within the redundant target set, such as *DPPA3* (also known as *STELLA*), showed complete loss of methylation in the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells (**Fig. 3g**) . Another interesting pattern was seen at the *NANOG* promoter, which seemed to be sporadically targeted by both *de novo* DNMTs, leading to low-level methylation in the upstream region that was lost only in the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells (**Supplementary Fig. 3c**). NANOG is highly expressed, but the low-level methylation in wild-type cells might reflect its known transcriptional heterogeneity within the ESC population.

We continued passaging of the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells for 22 passages and profiled methylation over time using reduced-representation bisulfite sequencing (RRBS). In the absence of *de novo* DNMT activity, CpG methylation was gradually lost as a result of the imperfect fidelity of DNMT1 maintenance activity (44.5 to 35.4%; **Fig. 3h**). Using an approximated doubling time of 28.8 h, we estimated the fidelity of DNMT1 to be between 99.7 and 99.97%, with some noteworthy differences across genomic features (**Supplementary Fig. 3b**). In particular, features that were highly methylated and/or required continued repression by DNA methylation, such as transposable elements and satellite repeats[40], seemed to exhibit maintenance that was slightly superior to that of other genomic features.

Lastly, we investigated the global level of non-CpG methylation, which is known to be present in mouse[15] and human[11,17] ESCs. Consistent with our previous knockdown-based experiments[17], we found that *DNMT3B* depletion had a stronger effect than loss of *DNMT3A* (**Fig. 3i**, left). This finding could, however, simply be related to the comparatively high levels of *DNMT3B* in undifferentiated ESCs (**Fig. 1a,b**). In the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells, non-CpG methylation was rapidly reduced to background levels (**Fig. 3i**, right), confirming that continuous *de novo* DNMT activity is required for non-CpG methylation in dividing cells, whereas CpG methylation can still largely be propagated by DNMT1, with a slow, gradual decline due to imperfect maintenance over successive cell divisions (**Fig. 3h**).

## Shared and unique targets of DNMT3A and DNMT3B

As noted above, we found a strong enrichment of CpG islands in the set of redundant DNMT3A and DNMT3B targets (**Fig. 3e**). To better understand this observation and gain a global perspective on the effect of DNA methylation on these targeted CpG islands and their neighboring regions (**Supplementary Fig. 4a**), we generated composite plots (**Fig. 4a**) and aggregate heat maps (**Supplementary Fig. 4b**) for the mean methylation levels across 25,489 CpG islands that were experimentally verified to be hypomethylated in somatic tissues[41]. We saw a clear distinction in behavior between CpG islands that were hypomethylated in the knockouts and the background set of CpG islands that were hypomethylated in both wild-type and knockout lines. CpG islands in the former set are presumably specific targets of *de novo* methylation activity, as they became hypomethylated in the knockouts to canonical CpG island methylation levels. In the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells, we observed not only the focal loss

© 2015 Nature America, Inc. All rights reserved.

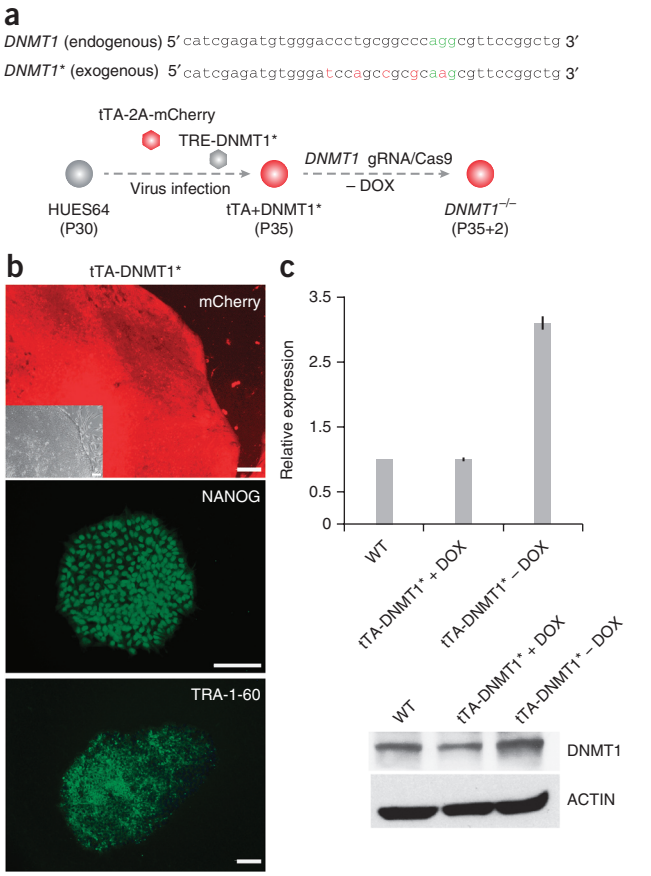



**Figure 6** *DNMT1*-knockout strategy and targeting efficiency. (**a**) Strategy to obtain *DNMT1* homozygous-knockout clones. Top, we created mutations within the exogenous *DNMT1* sequence that would prevent targeting by our gRNA but not alter the encoded protein (**Supplementary Fig. 6d**). An asterisk is used to distinguish the exogenous, modified *DNMT1* from endogenous *DNMT1*. Bottom, the *DNMT1* rescue line was established by infecting HUES64 cells with two separate lentiviruses (tTA-2a-mCherry and TRE-DNMT1*) at passage 30. *DNMT1* was then targeted as before into the tTA-DNMT1* line where exogenous *DNMT1** was expressed (no doxycycline (DOX)). (**b**) Representative bright-field image, mCherry expression and immunostaining for the pluripotent markers NANOG and TRA-1-60 for rescue line tTA-DNMT1*. Scale bars, 100 µm. (**c**) *DNMT1* expression levels in our *DNMT1** Tet-Off system. Top, RT-qPCR analysis for *DNMT1*. Primers amplify both endogenous and exogenous *DNMT1*. Error bars were generated from three biological replicates and represent one standard error. Bottom, immunoblotting analysis for DNMT1 under the control of doxycycline. The antibody can detect both endogenous and exogenous DNMT1.

of methylation at the center of the island but also a general decrease in methylation at CpG island shores. These results are consistent with the enrichment results and suggest that DNMT3A and DNMT3B are specifically targeted to a subset of CpG islands.

Our sequencing-based data allowed us to investigate the DNA methylation state of neighboring CpGs and classify them into concordant or discordant patterns (**Fig. 4b**). We calculated the proportion of discordant reads (PDR) as the number of discordant reads divided by the total number of reads covering a CpG and plotted the global distribution of this measure against CpG methylation[42]. In contrast to the single-knockout cells, the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells showed a global increase in PDR (**Fig. 4c**), which suggests that DNMT3A and DNMT3B have similar mechanisms of action and are likely to act redundantly to counteract global methylation loss due to imperfect DNMT1 fidelity. As expected, regions with low levels of methylation in wild-type cells had low PDR, and these regions remained hypomethylated with low PDR in the knockouts (**Fig. 4d**, top). Regions with intermediate levels of methylation showed decreased methylation and lower PDR in knockout cells, with especially striking hypomethylation in the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells (**Fig. 4d**, middle). One exception were imprinting control regions that are known to be variable in ESCs but generally maintained stable, intermediate levels of methylation even in the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells (**Supplementary Fig. 4c**). Finally, regions of high methylation had low PDR in wild-type cells and the single-knockout cells. These regions lost methylation specifically in the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells, with a corresponding increase in PDR (**Fig. 4d**, bottom).

To further assess the functional role of the DMRs found above, we calculated the degree of overlap of the DMRs in the three knockout lines with published transcription factor binding sites[43] and compared this overlap to that seen with a background set of all 1-kb tiles with an average methylation value of at least 0.4 in wild-type cells (**Fig. 4e**). Both *DNMT3A*$^{-/-}$ and *DNMT3B*$^{-/-}$ DMRs showed enrichment for the Polycomb repressive complex 2 (PRC2) components EZH2 and SUZ12 (**Fig. 4e**). The *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ DMRs showed strong enrichment for many sequence-specific transcription factors, including CTCF, whose binding is known to be affected by DNA methylation[44]. To determine whether these targets were associated with any specific biological processes, we took the set of 30,220 hypomethylated DMRs from the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells and used GREAT[45] to associate them with genes and calculate enrichments for functional ontologies. As expected, these regions were highly correlated with a large spectrum of developmental processes, including patterning, regionalization, growth factor stimulation and corresponding pathways (**Supplementary Fig. 4d** and **Supplementary Table 2**). The enriched categories suggest that the DMRs we identified are generally hypermethylated in wild-type pluripotent cells by DNMT3A and DNMT3B and normally lose methylation in a lineage- and tissue-specific manner as cells differentiate.

### *DNMT3A* deletion prevents gain of DNA methylation

To further investigate the knockout effects during differentiation, we derived early endodermal cells from the *DNMT3A*$^{-/-}$ cells (**Fig. 5a** and **Supplementary Fig. 5**). We then generated WGBS data and performed hierarchical clustering analysis, which included undifferentiated and sorted endoderm cells (dENs) of both wild-type and *DNMT3A*$^{-/-}$ cells (**Fig. 5b**). Undifferentiated cells in each case clustered more closely with their differentiated endoderm derivatives. We identified non-repetitive 1-kb tiles that had significantly different DNA methylation levels ($q$ value < 0.05, methylation difference ≥ 0.2) between wild-type ESCs and dEN cells to determine their comparable DNA methylation states in undifferentiated and differentiated *DNMT3A*$^{-/-}$ cells (**Fig. 5c**). We identified 575 DMRs that gained DNA methylation in wild-type dEN cells and were enriched for promoters and gene bodies (**Fig. 5d**). Strikingly, most of these DMRs remained hypomethylated in the dEN cells derived from *DNMT3A*$^{-/-}$ cells, suggesting that DNMT3A is required for their *de novo* methylation. For instance, we found that previously reported regions upstream of *FOXA2* and downstream of *DBX1* were no longer methylated in the knockout cells (**Fig. 5e**). We also differentiated the cells further toward the hepatoblast stage and stained the population with antibodies for FOXA2 and HNF4A. The loss of *DNMT3A* had no immediately detectable negative impact on the downstream differentiation in this lineage (**Fig. 5f**). In summary, we detect a unique molecular role for

© 2015 Nature America, Inc. All rights reserved.

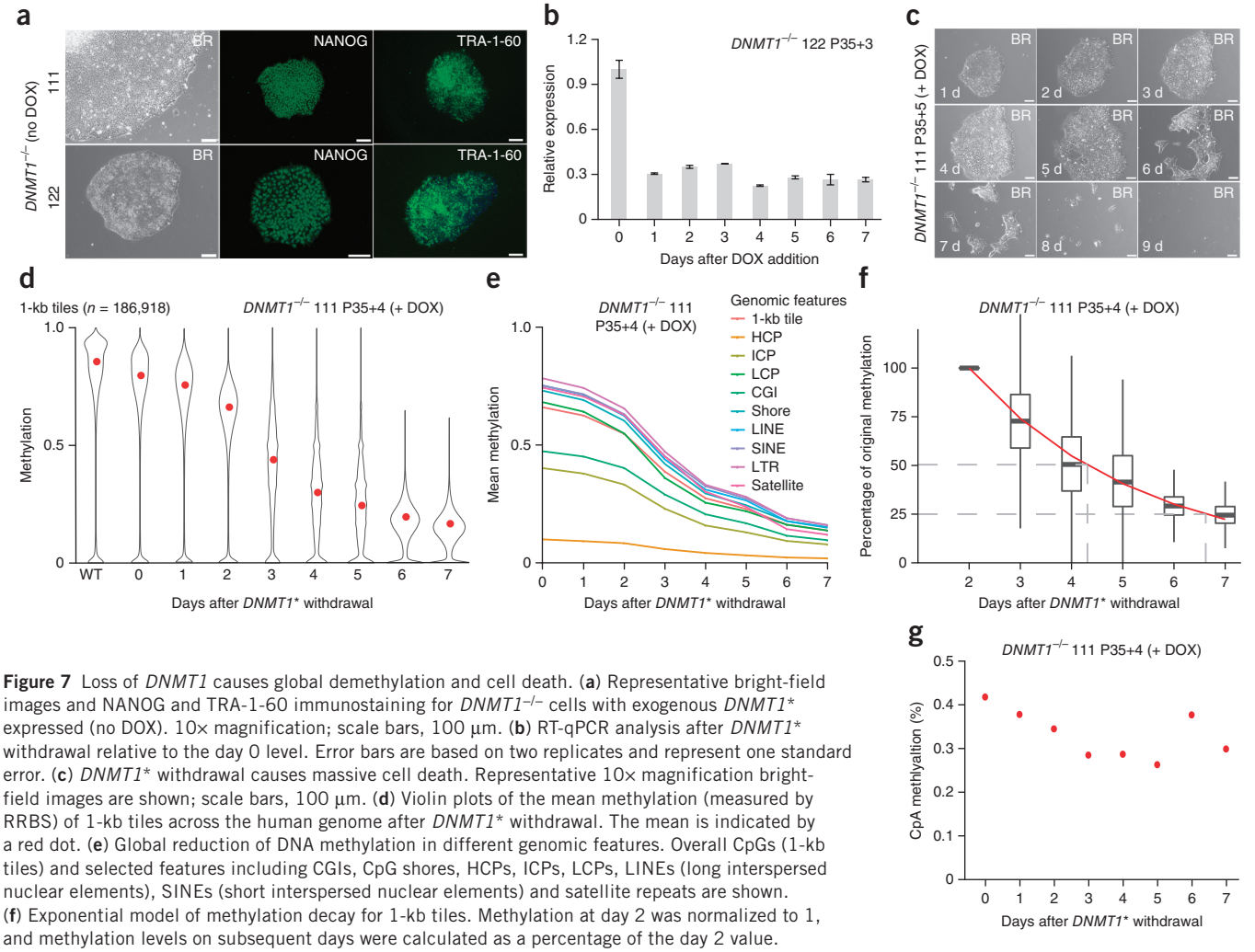



**Figure 7** Loss of *DNMT1* causes global demethylation and cell death. (**a**) Representative bright-field images and NANOG and TRA-1-60 immunostaining for *DNMT1*$^{-/-}$ cells with exogenous *DNMT1** expressed (no DOX). 10× magnification; scale bars, 100 µm. (**b**) RT-qPCR analysis after *DNMT1** withdrawal relative to the day 0 level. Error bars are based on two replicates and represent one standard error. (**c**) *DNMT1** withdrawal causes massive cell death. Representative 10× magnification bright-field images are shown; scale bars, 100 µm. (**d**) Violin plots of the mean methylation (measured by RRBS) of 1-kb tiles across the human genome after *DNMT1** withdrawal. The mean is indicated by a red dot. (**e**) Global reduction of DNA methylation in different genomic features. Overall CpGs (1-kb tiles) and selected features including CGIs, CpG shores, HCPs, ICPs, LCPs, LINEs (long interspersed nuclear elements), SINEs (short interspersed nuclear elements) and satellite repeats are shown. (**f**) Exponential model of methylation decay for 1-kb tiles. Methylation at day 2 was normalized to 1, and methylation levels on subsequent days were calculated as a percentage of the day 2 value. An exponential model was fitted starting at day 2 and is shown by the red line. The dashed gray lines mark the time points when 50% and 25% of the day 2 methylation levels remained. Days 0 and 1 were excluded from this model to avoid noise from any remaining DNMT1 protein. (**g**) Mean CpA methylation levels after *DNMT1** withdrawal.

DNMT3A during endoderm differentiation. Although we have not yet determined the functional relevance or downstream consequences of the targeted *de novo* methylation, our system provides a powerful new platform to dissect these effects in this and other lineages.

### *DNMT1* is essential for human ESC viability

As described above, we were unable to obtain homozygous-mutant lines for *DNMT1*, raising the possibility that human ESCs grown under standard fibroblast growth factor (FGF) conditions require DNA methylation for viability. This hypothesis was further supported by the failure to create straightforward homozygous mutants using an alternative TALEN-based strategy (**Supplementary Fig. 6a**) and by our inability to obtain any viable lines with efficient *DNMT1* knockdown, despite having tested a total of nine different short hairpin RNAs (shRNAs) (**Supplementary Fig. 6b**). Notably, the same TALEN-based strategy readily generated *DNMT3A* and *DNMT3B* homozygous-null lines, but no viable *DNMT1* homozygous-null lines emerged, consistent with the CRISPR experiments.

We therefore designed a rescue strategy that would enable the generation of homozygous-mutant lines without the immediate lethality of *DNMT1* loss (**Fig. 6a**). We first introduced two separate lentiviruses containing the sequence for tTA and TRE-*DNMT1** into the wild-type HUES64 line, creating a doxycycline-responsive (Tet-Off) *DNMT1** cell line that we could subsequently use for targeting (**Fig. 6b** and **Supplementary Fig. 6c**). We excluded the Kozak sequence (5′-GCCACC-3′) upstream of the start codon to prevent overexpression to levels that might cause adverse effects before targeting and created five point mutations within the target sequence and the protospacer-adjacent motif (PAM) sequence that had no effect on the protein sequence (**Fig. 6a** and **Supplementary Fig. 6d**) to prevent the gRNA from disrupting exogenous *DNMT1**.

Both qPCR and immunoblot analysis confirmed tight control of expression for exogenous *DNMT1**. In the absence of doxycycline, we observed an approximately threefold increase in overall DNMT1 levels that could be fully suppressed within 24 h of supplementation with doxycycline (**Fig. 6c**). We next used the CRISPR/Cas9 approach to knock out endogenous *DNMT1* while maintaining expression of exogenous *DNMT1**. We could now readily obtain homozygous-mutant lines in two independent rounds of targeting (**Supplementary Fig. 6e,f**), suggesting that our rescue strategy maintains a sufficient level of DNMT1 activity for survival. We selected two representative clones (111 and 122) for more detailed characterization and also confirmed by qPCR that *DNMT1* targeting did not affect the expression of *DNMT3A* and *DNMT3B* (**Supplementary Fig. 6g**).

© 2015 Nature America, Inc. All rights reserved.

### Loss of *DNMT1* causes global demethylation and cell death

*DNMT1* homozygous-knockout ESCs proliferated normally and could be maintained over several passages as long as exogenous *DNMT1** was expressed. The cells displayed a normal morphology and stained positive for NANOG and TRA-1-60 (**Fig. 7a**). Addition of doxycycline caused rapid downregulation of *DNMT1** (**Fig. 7b**), and cells began to die within a few days (**Fig. 7c** and **Supplementary Fig. 7a**). During this window, we observed increased DNA damage and arrest in G1 phase (**Supplementary Fig. 7b,c** and **Supplementary Table 3**), consistent with previous studies reporting that *DNMT1* depletion in human epidermal progenitors and primary human fibroblasts (IMR90) results in G1-phase arrest[20,46]. Using a catalytically inactive version of DNMT1 (Cys1226Trp)[47], which failed to rescue doxycycline-repressed lethality, we confirmed that this effect was due to the catalytic activity of DNMT1 (**Supplementary Fig. 7d–g**).

Despite increased arrest in G1 phase and onset of cell death, our inducible system enabled us to assess the effects on global DNA methylation upon acute loss of DNMT1 activity, which decreased substantially over a 1-week period (**Supplementary Fig. 7h**). We also used the catalytically inactive DNMT1 Cys1226Trp and showed that it could not rescue the cells and prevent the global loss of DNA methylation (**Supplementary Fig. 7i**). We then generated RRBS libraries to perform higher-resolution analysis (**Fig. 7d**). Consistent with global, replication-dependent depletion of methylation, we found that all genomic features were affected similarly (**Fig. 7e**). We fitted an exponential model of the methylation decay of individual 1-kb tiles and determined that it took 2.31 d for global DNA methylation to be reduced to half of the original levels (**Fig. 7f**). As expected, non-CpG methylation, which is the exclusive purview of *de novo* methyltransferases, was not affected by the acute loss of *DNMT1* (**Fig. 7g**).

## DISCUSSION

DNA methylation is essential for normal development. Over the past three decades, major advances in understanding its role have included the deletion of the responsible enzymes in mouse models, as well as the more recent emergence of comprehensive mapping tools. Here we have applied both to the human ESC model system to study the function of these enzymes, providing an unprecedented view of their respective shared and unique genomic targets.

All of the *DNMT3* knockouts, including the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells, were able to differentiate into the three germ layers. Overall CpG methylation was only mildly affected in the single knockouts, and even the *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells retained around 65% methylation after 17 passages (wild-type HUES64 cells, 83%, based on WGBS). This observation is consistent with previous mouse studies that suggest early differentiation may depend more on DNA methylation levels than on the continuous catalytic activity of these enzymes in the teratoma assay[39]. Our continued passaging suggests an average fidelity for DNMT1 of around 99.8%, which is higher than that estimated for mouse ESCs[2]. Further experimentation is required to explore the underlying mechanism; however, one can hypothesize that this difference simply reflects cell type–specific effects that have been noted for DNA methylation turnover rates[48]. It is also worth noting that mouse and human ESCs proliferate at very different rates (population doubling times of 13.5 ± 2.5 h and 36 ± 12 h, respectively)[49,50], which may have an impact on fidelity.

Our *DNMT3*-knockout human ESC lines provide a powerful model to study cell state–driven epigenetic effects during differentiation into specific cell types. It is likely that aberrant phenotypes arise only after differentiation, when *de novo* activity is required and/or compensation through DNMT3B is no longer available. *Dnmt3a*-knockout mice develop to term but die within a few weeks of birth[4], suggesting some essential although still poorly understood developmental function.

It is well established that mouse and human ESCs represent different pluripotent states, which may explain the sensitivity of human ESCs to loss of DNA methylation. All somatic cells require DNA methylation, and the more mature epiblast state modeled by current human ESCs may represent the first developmental period where maintenance becomes essential. In fact, although no data were shown, a similar effect has been suggested for mouse epiblast cells[51]. In the future, it will be interesting to test the effect of *DNMT1* deletion under recently described human 'naive' conditions to dissect the transition to DNMT1-dependent pluripotency[52–54].

Our study provides a comprehensive, high-resolution view of the largely conserved roles that the three catalytically active DNMTs have in human cells. The mutant lines represent powerful tools to study the function of these enzymes in various differentiation models and will therefore greatly accelerate the understanding of DNA methylation in human development and disease.

## METHODS

Methods and any associated references are available in the online version of the paper.

**Accession codes.** All data have been deposited in the Gene Expression Omnibus (GEO) under accession GSE63281.

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*


### ACKNOWLEDGMENTS

We thank all members of the Meissner laboratory, in particular C. Sindhu for helpful discussion and Z.D. Smith for critical feedback on the manuscript. We thank K. Musunuru (Harvard University) for providing the CRIPSR/Cas9 plasmids and thank Q. Ding from the Musunuru laboratory for technical support. J.L. was supported by a postdoctoral fellowship from the Human Frontiers Science Program. J.K.J. is supported by US NIH Director's Pioneer Award DP1 GM105378. A.M. is a New York Stem Cell Foundation Robertson Investigator. The work was funded by a US NIH (National Institute of General Medical Sciences) grant (P01GM099117) and the New York Stem Cell Foundation.


### AUTHOR CONTRIBUTIONS

J.L. and A.M. designed and conceived the study. J.L. generated all the cell lines and performed the experiments. R.K. performed the analysis. H.G. generated the WGBS libraries. A.G. and A.M. supervised the DNA methylation experiments. M.J.Z. and K.C. performed some analysis and assisted in the general data processing. A.M.T. and V.A. performed the experiments with the Scorecard assay. C.A.G. and J.D. assisted in endoderm and hepatocyte differentiation. C.G. performed the dot blot assay. R.P. supported the RNA profiling and FACS analysis. D.R., S.Q.T. and J.K.J. designed and generated the TALENs. W.M. and J.L.R. performed expression analysis. J.L., R.K. and A.M. interpreted the data and wrote the manuscript.

### COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.


1. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
2. Jackson, M. *et al.* Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol. Cell. Biol.* **24**, 8862–8871 (2004).
3. Li, E., Bestor, T.H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
4. Okano, M., Bell, D.W., Haber, D.A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* **99**, 247–257 (1999).


© 2015 Nature America, Inc. All rights reserved.

5. Jin, B. *et al.* DNA methyltransferase 3B (*DNMT3B*) mutations in ICF syndrome lead to altered epigenetic modifications and aberrant expression of genes regulating development, neurogenesis and immune function. *Hum. Mol. Genet.* **17**, 690–709 (2008).
6. Klein, C.J. *et al.* Mutations in *DNMT1* cause hereditary sensory neuropathy with dementia and hearing loss. *Nat. Genet.* **43**, 595–600 (2011).
7. Shah, M.Y. & Licht, J.D. *DNMT3A* mutations in acute myeloid leukemia. *Nat. Genet.* **43**, 289–290 (2011).
8. Winkelmann, J. *et al.* Mutations in *DNMT1* cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum. Mol. Genet.* **21**, 2205–2210 (2012).
9. Gifford, C.A. *et al.* Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**, 1149–1163 (2013).
10. Habibi, E. *et al.* Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **13**, 360–369 (2013).
11. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
12. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
13. Smith, Z.D. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
14. Ziller, M.J. *et al.* Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477–481 (2013).
15. Dodge, J.E., Ramsahoye, B.H., Wo, Z.G., Okano, M. & Li, E. *De novo* methylation of MMLV provirus in embryonic stem cells: CpG versus non-CpG methylation. *Gene* **289**, 41–48 (2002).
16. Meissner, A. *et al.* Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877 (2005).
17. Ziller, M.J. *et al.* Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet.* **7**, e1002389 (2011).
18. Fan, G. *et al.* DNA hypomethylation perturbs the function and survival of CNS neurons in postnatal animals. *J. Neurosci.* **21**, 788–797 (2001).
19. Jackson-Grusby, L. *et al.* Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat. Genet.* **27**, 31–39 (2001).
20. Sen, G.L., Reuter, J.A., Webster, D.E., Zhu, L. & Khavari, P.A. DNMT1 maintains progenitor function in self-renewing somatic tissue. *Nature* **463**, 563–567 (2010).
21. Trowbridge, J.J., Snow, J.W., Kim, J. & Orkin, S.H. DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell Stem Cell* **5**, 442–449 (2009).
22. Tsumura, A. *et al.* Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805–814 (2006).
23. Smith, Z.D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
24. Chen, T. *et al.* Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. *Nat. Genet.* **39**, 391–396 (2007).
25. Egger, G. *et al.* Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proc. Natl. Acad. Sci. USA* **103**, 14080–14085 (2006).
26. Rhee, I. *et al.* DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* **416**, 552–556 (2002).
27. Rhee, I. *et al.* CpG methylation is maintained in human cancer cells lacking *DNMT1*. *Nature* **404**, 1003–1007 (2000).
28. Martins-Taylor, K., Schroeder, D.I., LaSalle, J.M., Lalande, M. & Xu, R.H. Role of DNMT3B in the regulation of early neural and neural crest specifiers. *Epigenetics* **7**, 71–82 (2012).
29. Horii, T., Tamura, D., Morita, S., Kimura, M. & Hatada, I. Generation of an ICF syndrome model by efficient genome editing of human induced pluripotent stem cells using the CRISPR system. *Int. J. Mol. Sci.* **14**, 19774–19781 (2013).
30. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
31. Hockemeyer, D. *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* **29**, 731–734 (2011).
32. Jinek, M. *et al.* RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013).
33. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
34. Miller, J.C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
35. Ding, Q. *et al.* Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* **12**, 393–394 (2013).
36. Gordon, C.A., Hartono, S.R. & Chedin, F. Inactive DNMT3B splice variants modulate *de novo* DNA methylation. *PLoS ONE* **8**, e69486 (2013).
37. Tsankov, A.M. *et al.* Transcription factor binding dynamics during human ES cell differentiation. *Nature* **518**, 344–349 (2015).
38. Bock, C. *et al.* Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439–452 (2011).
39. Chen, T., Ueda, Y., Dodge, J.E., Wang, Z. & Li, E. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* **23**, 5594–5605 (2003).
40. Smith, Z.D. *et al.* DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611–615 (2014).
41. Illingworth, R. *et al.* A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* **6**, e22 (2008).
42. Landau, D.A. *et al.* Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell* **26**, 813–825 (2014).
43. Gerstein, M.B. *et al.* Architecture of the human regulatory network derived from ENCODE data. *Nature* **489**, 91–100 (2012).
44. Stadler, M.B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).
45. McLean, C.Y. *et al.* GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
46. Barra, V., Schillaci, T., Lentini, L., Costa, G. & Di Leonardo, A. Bypass of cell cycle arrest induced by transient *DNMT1* post-transcriptional silencing triggers aneuploidy in human cells. *Cell Div.* **7**, 2 (2012).
47. Clements, E.G. *et al.* DNMT1 modulates gene expression without its catalytic activity partially through its interactions with histone-modifying enzymes. *Nucleic Acids Res.* **40**, 4334–4346 (2012).
48. Shipony, Z. *et al.* Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells. *Nature* **513**, 115–119 (2014).
49. Orford, K.W. & Scadden, D.T. Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation. *Nat. Rev. Genet.* **9**, 115–128 (2008).
50. Li, L. *et al.* Individual cell movement, asymmetric colony expansion, Rho-associated kinase, and E-cadherin impact the clonogenicity of human embryonic stem cells. *Biophys. J.* **98**, 2442–2451 (2010).
51. Hutnick, L.K., Huang, X., Loo, T.C., Ma, Z. & Fan, G. Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation–independent mechanism. *J. Biol. Chem.* **285**, 21082–21091 (2010).
52. Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–286 (2013).
53. Chan, Y.S. *et al.* Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**, 663–675 (2013).
54. Theunissen, T.W. *et al.* Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487 (2014).
55. Di Giaimo, R. *et al.* The expression of *de novo* DNA methylase DNMT3b, of the methyl-CpG binding protein MBD2b and of 5-MCDG glycosylase shows two waves of induction during CaCO-2 cell differentiation. *Gene* **351**, 73–81 (2005).


© 2015 Nature America, Inc. All rights reserved.

## ONLINE METHODS

**Cell culture.** Human ESCs were expanded on CF1 mouse embryonic fibroblasts (Global Stem) in KO-DMEM (Life Technologies) containing 20% Knockout Serum Replacement (KOSR; Life Technologies) and FGF2 (10 ng/ml; Millipore). Cells were passaged by enzymatic dissociation using collagenase IV (1 mg/ml; Life Technologies). For the feeder-free culture condition, plates were coated with Geltrex (Life Technologies), and the culture medium was mTeSR1 (Stemcell Technologies). HUES64 cells were derived in the Melton and Eggan laboratories[56], have been maintained in our laboratory since then using the above conditions and are routinely tested to be free of mycoplasma.

**Electroporation.** Human ESCs were first collected using 1 mg/ml collagenase type IV (Invitrogen) and then dissociated into single cells using TrypLE (Invitrogen). Cells ($1 \times 10^7$) resuspended in buffered saline (PBS) were electroporated with 15 μg of Cas9 and 15 μg of gRNA plasmids (Gene Pulser Xcell System, Bio-Rad: 250 V, 500 μF, 0.4-cm cuvettes). The Cas9 plasmid and gRNA backbone have been described previously[35]. FACS was carried out 48 h after transfection based on fluorescent marker expression. Sorted cells were replated as single cells at a very low density on irradiated CF1 feeders in hESC medium supplemented with Rho kinase (ROCK) inhibitor (Stemgent, Y-27632) for the first 24 h. Individual colonies were picked and expanded under the feeder-free/mTeSR condition. Genomic DNA was then purified. Mutations were validated by PCR and DNA sequencing.

**Constructs.** GIPZ *DNMT1* shRNA was obtained from Thermo Scientific (RHS4531-NM_001130823). TRIPZ doxycycline-inducible *DNMT1* shRNA was obtained from Thermo Scientific (RHS4740-EG1786). To create plenti-ef1a-tTA-2a-mCherry, the tTA-2a-mCherry construct was inserted into the lentivirus backbone (Addgene, 17448), the sequence encoding tTA was cloned from Addgene plasmid 26803 (ref. 57) and the sequence encoding mCherry was cloned from Addgene plasmid 23243 (ref. 58). To create plenti-ef1a-puro$^r$-TRE-DNMT1*, *DNMT1* was cloned from Addgene plasmid 36939 (ref. 59); after mutations were added, *DNMT1** was cloned into the plenti-ef1a-puro$^r$-TRE backbone[60], which was a gift from the laboratory of L. Xiao (Zhejiang University). To create plenti-ef1α-DNMT1$^{WT}$-IRES-nlsEGFP, *DNMT1* was cloned from Addgene plasmid 36939 (ref. 59). To create plenti-ef1α-DNMT1$^{C1226W}$-IRES-nlsEGFP, a point mutation in *DNMT1* was introduced so that the encoded protein harbored a single amino acid replacement at position 1,226 from cysteine to tryptophan.

**Primers and targeting strategy.** Primer sequences are listed in **Supplementary Table 4**. Sequence information for the TALEN targeting strategy is provided in **Supplementary Table 5**.

**Virus production and infection of human ESCs.** Virus was produced by cotransfecting HEK293 cells with lentivirus plasmid together with delta 8.91 and VSVG using FuGENE HD (Promega). Twenty-four hours after transfection, the medium was changed to mTeSR1 for virus collection. Supernatant containing virus was collected twice, 48 h and 72 h after transfection, separately. Virus was stored at 4 °C. Before infection, virus was filtered through a 0.45-μm filter.

For virus infection of human ESCs, cells were plated as small clumps on plates precoated with Geltrex. Twenty-four hours after plating, cells were infected with virus supernatant for 3 h in an incubator. The medium was changed back to mTeSR1. The 3-h infection was repeated the next day. In this case, cells were maintained in the pluripotent state.

**Isolation of genomic DNA from 96-well plates.** The cells in each well were ~90% confluent. Cells were washed twice with PBS and then lysed with 50 μl/well of lysis buffer (10 mM Tris, pH 7.5, 10 mM EDTA, 10 mM NaCl, 0.5% sarcosyl or SDS, 1 mg/ml proteinase K) at 37 °C overnight. Genomic DNA was precipitated by adding 100 μl/well of cold NaCl/ethanol (1.5% volumes of 5 M NaCl in 100% cold ethanol) at room temperature for 2–4 h. Genomic DNA precipitate was washed three times with 70% ethanol, each time discarding the alcohol by inversion. DNA was air dried for 15–20 min at room temperature. DNA was resuspended in 30–100 μl of DNase-free, RNase-free double-distilled water.

**Antibodies.** The following antibodies were used for immunostaining: OCT4 (BD Transduction Laboratories, 611202); NANOG (R&D, AF1997); TRA-1-60 (Stemgent, 090068); Annexin V (Abcam, ab54775); FOXA2 (R&D, AF2400); HNF4A (Abcam, ab41898); and γ-H2A.X phosphorylated at Ser139 (Abcam, ab2893). The following antibodies were used for immunoblotting: DNMT3A (Cell Signaling Technology, 2160); DNMT3B (Santa Cruz Biotechnology, sc-20704); DNMT1 (Activemotif, 39204); and actin (Abcam, ab8226).

**Immunostaining.** Human ESCs were passaged on Geltrex-coated plates, and cells were cultured in mTeSR1 medium. Cells were fixed in 4% paraformaldehyde for 15 min at room temperature. Permeabilization and blocking were performed in 5% BSA and 1% Triton X-100 in PBS for 30 min. Cells were stained with primary antibody at 4 °C overnight. Secondary antibody was applied for 2 h at room temperature. Nuclei were stained with Hoechst 33342 (Life Technologies). Images were acquired using an Olympus IX71 microscope and MetaMorph Advanced software.

**RNA extraction.** RNA was isolated from human ESCs using TRIzol (Invitrogen, 15596-026). Further purification was carried out with RNeasy columns (Qiagen, 74104) and DNase treatment.

**Embryoid body formation.** Human ESCs were expanded in 10-cm$^2$ dishes on irradiated CF1 mouse embryonic fibroblasts (Global Stem). Cells were collected using 1 mg/ml collagenase IV (Life Technologies) once they were confluent. After three washes with 1× PBS, cell clumps were plated into six-well low-attachment dishes (Corning) in KO-DMEM supplemented with 20% KOSR, GlutaMAX (200 mM; Life Technologies) and MEM Non-Essential Amino Acids (Life Technologies). Medium was changed every 2–3 d.

**Directed differentiation into the three germ layers.** Human ESCs were plated as clumps in six-well plates coated with Geltrex in mTeSR1 culture medium. On day 3, directed differentiation was induced, and the medium was changed every day.

Ectoderm differentiation medium contained 2 μM A83-01 TGF-β inhibitor (Tocris), 2 μM PNU-74654 WNT3A inhibitor (Tocris) and 2 μM Dorsomorphin BMP inhibitor (Tocris) with 15% KOSR, MEM Non-Essential Amino Acids and 55 μM 2-mercaptoethanol.

Mesoderm differentiation medium contained 100 ng/ml Activin A (Life Technologies), 10 ng/ml bFGF (Millipore), 100 ng/ml BMP4 (Life Technologies) and 100 ng/ml VEGF (Life Technologies) with 0.5% FBS (Hyclone), MEM Non-Essential Amino Acids, 55 μM 2-mercaptoethanol and GlutaMAX in DMEM/F12 medium (Life Technologies) for the first 24 h of differentiation. For the remaining 4 d, the medium contained 10 ng/ml bFGF, 100 ng/ml BMP4 and 100 ng/ml VEGF with 0.5% FBS, MEM Non-Essential Amino Acids, 55 μM 2-mercaptoethanol and GlutaMAX in DMEM/F12 medium.

Endoderm differentiation medium contained 100 ng/ml Activin A and 2 μM/ml lithium chloride (Sigma) with 0.5% FBS, 200 mM GlutaMAX, MEM Non-Essential Amino Acids and 55 μM 2-mercaptoethanol in RPMI medium (Life Technologies).

**Teratoma formation.** Teratoma formation assays were performed as a service by the Harvard Genome Modification Facilities. Briefly, human ESCs were injected into the kidney capsules of 3 immunosuppressed mice as 1 million cells per animal to form teratomas. The formed teratomas were fixed in 4% paraformaldehyde pH 7.4, embedded in paraffin, sectioned at 10- to 12-μm thickness, and stained with hematoxylin and eosin (H&E) for examination. Histology examination was performed by the histology core facility at the Harvard Stem Cell Institute.

**Hepatoblast differentiation.** Differentiation was performed as described previously[61] with some modifications. Briefly, cells were dissociated with Accutase and seeded onto Geltrex-coated dishes with ROCK inhibitor at high confluence in mTeSR1. Within 24 h of plating, definitive endoderm differentiation was

© 2015 Nature America, Inc. All rights reserved.

started, and cells were cultured in RPMI/2% B27 (without insulin), 100 ng/ml Activin A, 10 ng/ml BMP4 and 20 ng/ml FGF2 for 48 h. For days 3–5 of differentiation, cells were cultured in RPMI/2% B27 (without insulin) and 100 ng/ml Activin A. For days 6–10 of differentiation, cells were cultured in RPMI/2% B27 (with insulin) plus 20 ng/ml BMP4 and 10 ng/ml FGF2. Medium was changed every 24 h.

**Cell cycle assays.** Live cells were incubated with Hoechst 33342 in the incubator for 15 min. Cells were then trypsinized into single cells, washed with PBS once and analyzed by FACS directly.

**Dot-blot assays.** Isolated genomic DNA was denatured in 0.4 M NaCl and 10 mM EDTA for 10 min at 99 °C. Denatured DNA was spotted in a two-fold serial dilution on a Zeta Membrane (Bio-Rad) using a Bio-Dot apparatus (Bio-Rad). The blotted membrane was rinsed in 2× SSC, air dried and UV cross-linked at 120,000 μJ/cm$^2$. The membrane was blocked in 1× TBS containing 5% nonfat dry milk for 30 min. Monoclonal antibody to 5-methylcytosine (Diagenode, 33D3) was diluted 1:250 in TBST containing 0.3% Tween-20 and 5% nonfat dry milk, and was incubated overnight at 4 °C. The membrane was washed three times for 5 min each in TBS and was incubated for 1 h at room temperature with horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG (Jackson Lab, 115-025-174). The membrane was washed three times in TBS and visualized by chemiluminescence using the Gel Doc XR (Bio-Rad). The same membrane was incubated 30 min with SYBR Gold Nucleic Acid stain (Invitrogen, S-11494) in 1× TAE buffer. Total DNA loading was visualized by UV transillumination.

**Scorecard assays and data analysis.** Experiments were carried out using the TaqMan hPSC Scorecard kit (Life Technologies, A15872). Data were analyzed with TaqMan hPSC Scorecard Analysis Software on the Life Technologies website.

**RRBS and WGBS library construction.** RRBS library construction was performed as described[62]. WGBS library construction was performed as described[9].

**RRBS and WGBS data processing and analysis.** WGBS raw sequencing reads were aligned using MAQ in bisulfite mode against human genome version hg19/GRCh37, discarding duplicate reads. DNA methylation calling was performed on the basis of an extended custom software pipeline published previously for RRBS[63]. Briefly, the methylation level at a CpG or CpA site was the number of reads with that site methylated divided by the total number of reads covering the site.

To ensure the comparability of regional DNA methylation levels across all samples, only CpGs covered by ≥5× in all samples qualified for the computation of regional DNA methylation levels. We defined the average methylation for a genomic region as the coverage-weighted mean of the methylation levels of the individual CpGs within the region. Subsequently, we averaged a region's DNA methylation level over replicates. DMRs were identified using a two-sample weighted *t* test using the methylation values of the CpGs within each region as individual sampling events (the minimum number of CpGs in each region for each sample was two). We performed multiple-testing correction as described[64] using the R qvalue package.

We used two sets of CpG island annotations. The first used a bioinformatics definition and designated a region as a CpG island if it had a GC content higher than 0.5 and a ratio greater than 0.6 of observed CpG dinucleotides to what would be expected on the basis of the GC content over a length of 700 bp. The second was a set of experimentally verified CpG islands[41].

We defined promoters as the regions 2,000 bp upstream and 500 bp downstream of the RefSeq TSS set. These were divided into HCPs, ICPs and LCPs as published previously[65]. Intergenic regions, exons and introns were defined on the basis of the RefSeq gene models downloaded from the UCSC browser.

We used GREAT (Genomic Region Enrichment of Annotation Tool)[45] to associate DMRs with genes. We limited the associations to genes within 20 kb of DMRs and used a false discovery rate (FDR) threshold of $1 \times 10^{-6}$ and a fold enrichment threshold of 2. The complete set of enriched terms across all ontologies tested is in **Supplementary Table 2**.

The fidelity of DNMT1 was calculated by fitting an exponential model to the mean DNA methylation of different genomic features using the nls function in R. The exponential model has the form $y = a^x$, where $a$ is the fidelity and $x$ is the estimated number of population doublings based on passage number, days per passage (eight) and doubling time (28.8 h).

The PDR was calculated for each CpG by examining the individual reads covering that CpG and classifying them as concordant (all CpGs methylated or all CpGs unmethylated), filtering out reads with fewer than three CpGs covered.

We downloaded the conservative set of uniform peaks called by the ENCODE Project[43] and computed overlaps of these peaks with DMRs identified in *DNMT3A*$^{-/-}$, *DNMT3B*$^{-/-}$ and *DNMT3A*$^{-/-}$; *DNMT3B*$^{-/-}$ cells. We calculated the significance of the enrichment with Fisher's exact test, using the DMR set as the foreground and the set of non-repetitive 1-kb tiles with a methylation level of at least 0.4 as the background. Human imprinted control regions were taken from ref. 66. Mouse *Dnmt3a*$^{-/-}$, *Dnmt3b*$^{-/-}$ and *Dnmt3a*$^{-/-}$; *Dnmt3b*$^{-/-}$, *Dnmt1*$^{-/-}$ and *Dnmt3a*$^{-/-}$; *Dnmt3b*$^{-/-}$ with *Dnmt1* knockdown[16] data were aligned to mm9 and otherwise processed similarly to the human RRBS data.


56. Chen, A.E. *et al.* Optimal timing of inner cell mass isolation increases the efficiency of human embryonic stem cell derivation and allows generation of sibling cell lines. *Cell Stem Cell* **4**, 103–106 (2009).
57. Hsiao, E.C. *et al.* Constitutive $G_s$ activation using a single-construct tetracycline-inducible expression system in embryonic stem cells and mice. *Stem Cell Res. Ther.* **2**, 11 (2011).
58. Papapetrou, E.P. *et al.* Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc. Natl. Acad. Sci. USA* **106**, 12759–12764 (2009).
59. Li, H. *et al.* The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. *J. Biol. Chem.* **281**, 19489–19500 (2006).
60. Rao, L. *et al.* Highly efficient derivation of skeletal myotubes from human embryonic stem cells. *Stem Cell Rev.* **8**, 1109–1119 (2012).
61. Si-Tayeb, K. *et al.* Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. *Hepatology* **51**, 297–305 (2010).
62. Boyle, P. *et al.* Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biol.* **13**, R92 (2012).
63. Bock, C. *et al.* Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat. Biotechnol.* **28**, 1106–1114 (2010).
64. Storey, J.D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* **100**, 9440–9445 (2003).
65. Saxonov, S., Berg, P. & Brutlag, D.L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. USA* **103**, 1412–1417 (2006).
66. Woodfine, K., Huddleston, J.E. & Murrell, A. Quantitative analysis of DNA methylation at all human imprinted regions reveals preservation of epigenetic stability in adult somatic tissue. *Epigenetics Chromatin* **4**, 1 (2011).


© 2015 Nature America, Inc. All rights reserved.


doi:10.1038/ng.3258