Figure S1 - Characterization of DMRs

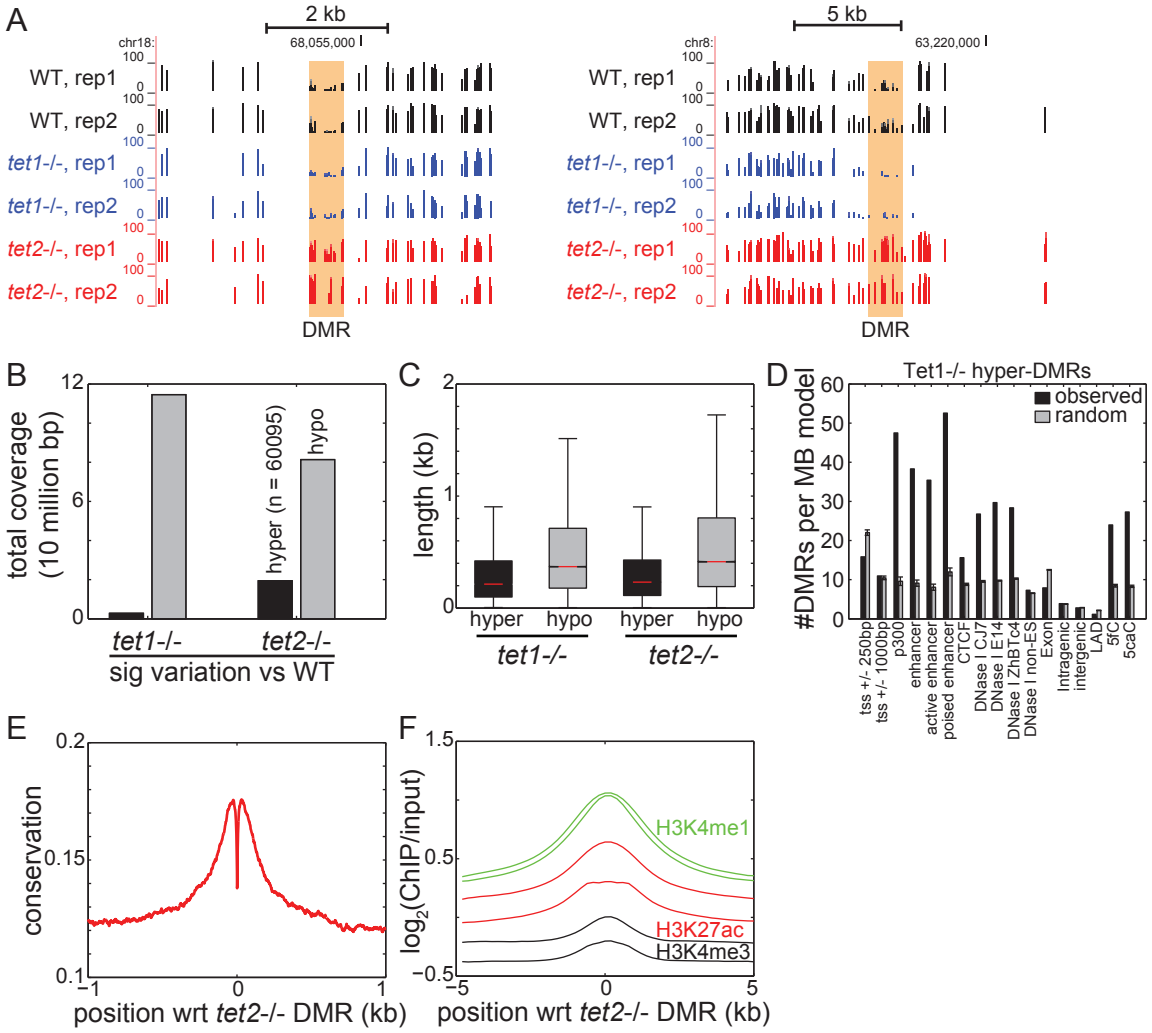


Figure S2 - Snapshots of hypermethylated enhancers

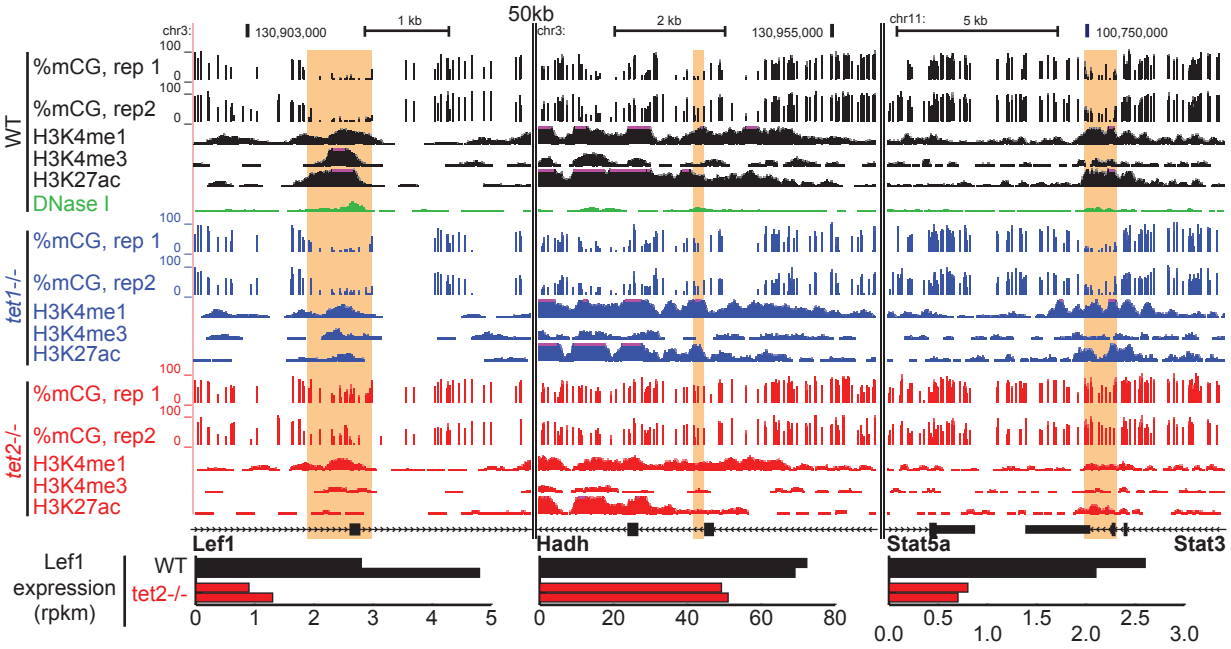


Figure S3 - Enhancer enrichment at differentially expressed genes

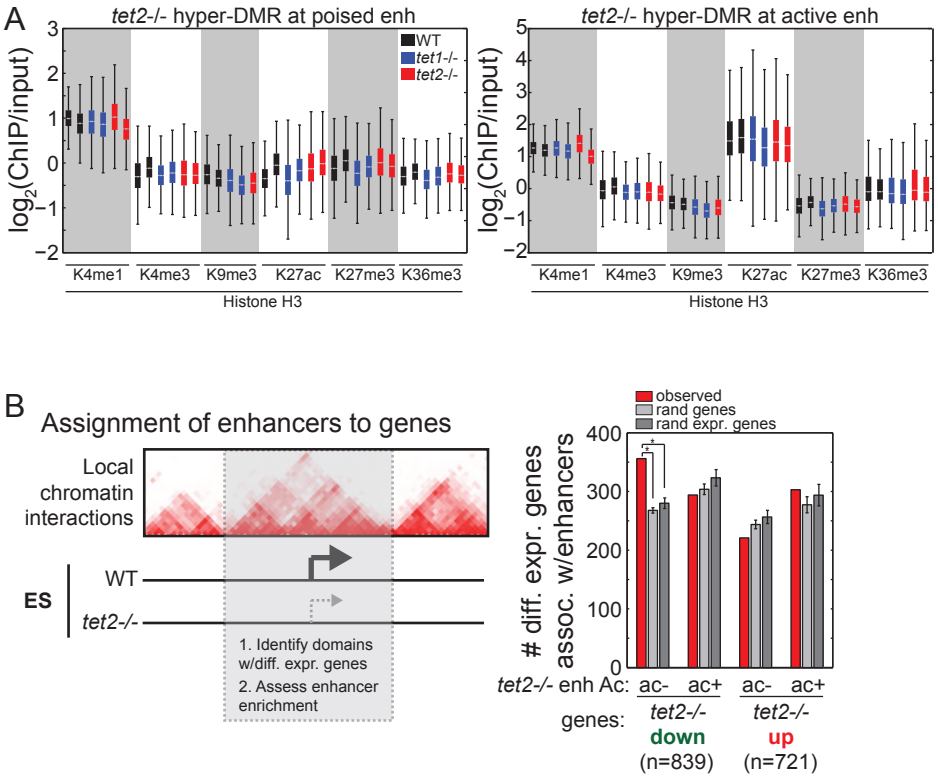
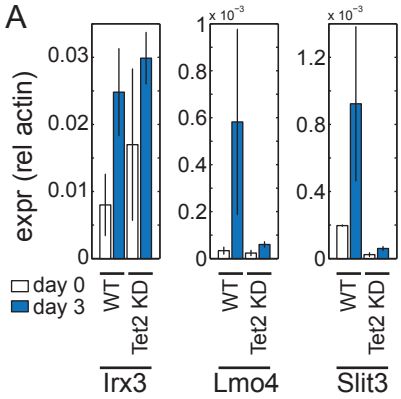
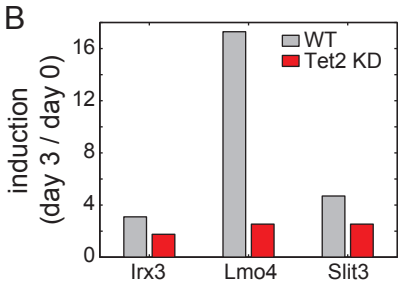
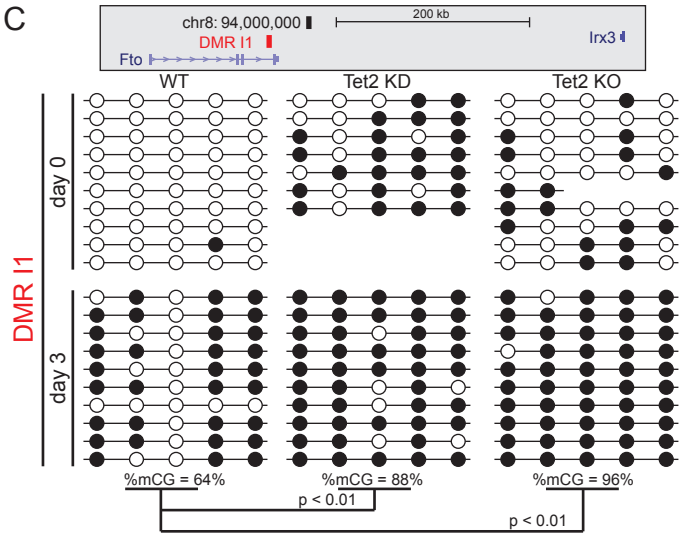
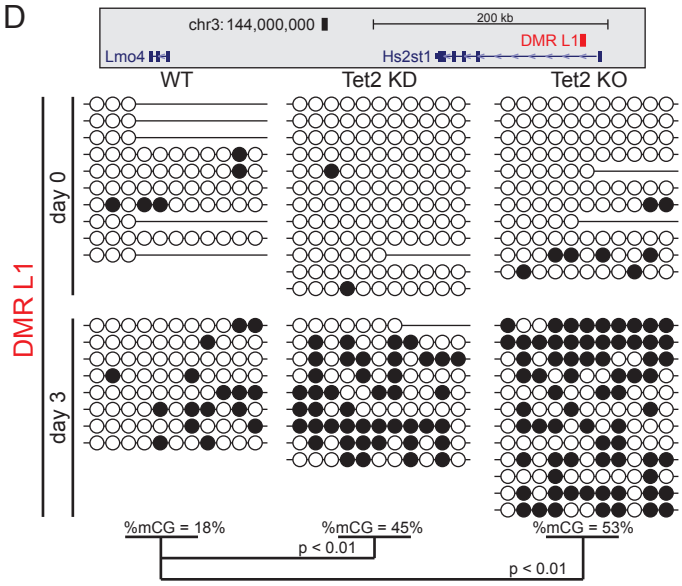


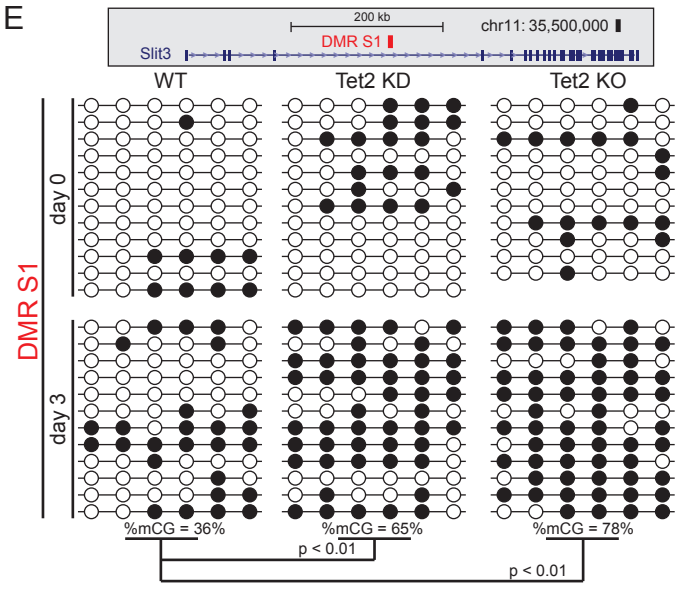
Figure S4 - Comparison to Tet2 knockdown cells











Supplemental Text

Supplemental Figure Legends

Figure S1, related to Figure 2: Characterization of DMRs.

(A) UCSC Genome Browser snapshots of DNA methylation at Tet2$^{-/-}$ hyper-DMRs. (B) Genomic coverage of DMRs in Tet2$^{-/-}$ and Tet2$^{-/-}$ cells. (C) Length distribution of DMRs. Boxplot edges indicate the 25th and 75th percentiles, and whiskers indicate non-outlier extremes. (D) Relative enrichment of Tet1$^{-/-}$ hyper-DMRs at various genomic features. Error bars indicate standard deviation. (E) Average PhastCons conservation score centered at Tet2$^{-/-}$ DMRs. (F) Enrichment of histone modifications centered at Tet2$^{-/-}$ DMRs.

Figure S2, related to Figure 3: Snapshots of hypermethylated enhancers.

UCSC Genome Browser snapshots illustrating the enrichment of DNA methylation and histone modifications at WT enhancers that lose H3K27ac in Tet2$^{-/-}$ cells. The expression of genes near these enhancers is indicated in the bar chart below.

Figure S3, related to Figure 4: Enhancer enrichment at differentially expressed genes.

(A) For enhancers bearing (left) poised and (right) active chromatin signatures in WT cells that overlap with Tet2$^{-/-}$ hyper-DMRs, shown is the enrichment of various histone modifications. Boxplot edges indicate the 25th and 75th percentiles, and whiskers indicate non-outlier extremes. (B) (left) Schematic illustrating the assignment of differentially expressed genes to enhancers by using local chromatin interactions from Hi-C experiments (Dixon et al., 2012). (right) Enrichment of domains containing *Tet2*$^{-/-}$ differentially expressed genes with enhancers that lose or retain H3K27ac. Comparisons are made to randomly selected genes (light gray) and randomly selected genes expressed in ESCs (dark gray). Error bars indicate standard deviation. (C-D) Enrichment of hypermethylated hypo-Ac enhancers identified in this study with genes repressed in (C) Tet2 knockdown mESCs (Koh et al., 2011) and (D) independently derived Tet2 knock-out mESCs (Dawlaty et al., 2013).

Figure S4, relatedp to Figure 5: Comparison to Tet2 knockdown cells.

(A) Wild-type and Tet2 knockdown mESCs (Huang et al., 2014) were differentiated towards NPCs, and the expression of three delayed induction genes were monitored by qRT-PCR at day 0 and day 3 after differentiation. Error bars indicate standard deviation. (B) Shown is the relative induction ratio (day 3 / day 0) for delayed induction genes Irx3, Lmo4, and Slit3 in wild-type and Tet2 knockdown cells. (C-E) Locus-specific bisulfite sequencing for DMRs near (C) Irx3, (D) Lmo4, and (E) Slit3. Shown are data for wild-type (left), Tet2 knockdown (middle), and Tet2 knockout cells (right) at DMRs identified in Tet2$^{-/-}$ cells before (top) and after (middle/bottom)

differentiation. White circles indicate unmethylated cytosine and black circles indicate methylated cytosine for individual clones. P-values indicated are from applying Fisher's Exact Test.

Supplemental Table Legends

Table S1, related to Figure 2: DMRs identified in Tet1$^{-/-}$ and Tet2$^{-/-}$ compared to wild-type cells.

Table S2, related to Figure 4: Enhancers hypermethylated and hypo-acetylated in Tet2$^{-/-}$ compared to wild-type cells.

Supplemental Experimental Procedures

Assessing 5hmC by mass spectrometry

4 µg of genomic DNA was digested by nuclease P1 (Sigma), venom phosphodiesterase I (Type VI) (Sigma), and alkaline phosphatase (Sigma) according to published protocols (Crain, 1990). After a brief desalting with ammonium-equilibrated cation exchange resin (Poly-Prep Columns AG 50W-X8, Bio-Rad) and filter, 10 µL (out of 40 µL) recovered solution was injected into LC-MS/MS. The nucleosides were separated by reverse phase ultra-performance liquid chromatography on a C18 column, with online mass spectrometry detection using Agilent 6410 QQQ triple-quadrupole LC mass spectrometer set to multiple reaction monitoring (MRM) in positive electrospray ionization mode. The nucleosides were quantified using the nucleoside to base ion mass transitions of 258 to 142 (5hmC) and 228 to 112 (C). Quantification and detection limits were determined by comparison with the standard curves obtained from nucleoside standards running at the same volume and time.

Statistics for TAB-Seq

For TAB-Seq, we estimated the non-conversion rate of unmodified cytosine using a 10-kb unmethylated fragment of lambda phage genome as: $n_{C\text{-}WT,rep1} = 0.0058$, $n_{C\text{-}WT,rep2} = 0.0058$, $n_{C\text{-}Tet1\ KO,rep1} = 0.0051$, $n_{C\text{-}Tet1\ KO,rep2} = 0.0053$, $n_{C\text{-}Tet2\ KO,rep1} = 0.0046$, and $n_{C\text{-}Tet2\ KO,rep2} = 0.0042$. We then estimated the non-conversion rate of 5mC from fully methylated cytosines for each cell line as: $n_{5mC\text{-}WT,rep1} = 0.0394$, $n_{5mC\text{-}WT,rep2} = 0.0272$, $n_{5mC\text{-}Tet1\ KO,rep1} = 0.0217$, $n_{5mC\text{-}Tet1\ KO,rep2} = 0.0237$, $n_{5mC\text{-}Tet2\ KO,rep1} = 0.0222$, and $n_{5mC\text{-}Tet2\ KO,rep2} = 0.0187$.

Estimating the abundance of 5mC and 5hmC

To measure the absolute abundance of 5hmC using TAB-Seq, let the non-conversion rate of unmethylated cytosine be n_C and let the non-conversion rate of 5mC be n_{mC}, as measured from bases in the corresponding cell-type having 0% and 100% methylation from methylC-Seq, respectively. In TAB-Seq, the error rate of bases with 0% 5mC are dominated by n_C, whereas bases with 100% 5mC are dominated by n_{mC}, and bases inbetween are a linear combination of the two. To estimate 5hmC at a genomic locus with C_{TAB} cytosine and T_{TAB} thymine base-calls from TAB-Seq and C_{trad} cytosine and T_{trad} thymine base-calls from methylC-Seq, estimate the fraction of methylated bases as $f_{h/mC} = C_{trad} / (C_{trad} + T_{trad})$ and the fraction of unmethylated bases as $f_C = 1 - f_{mC}$. Then the TAB-Seq error rate at this base is estimated as $e = (f_C n_C) + (f_{h/mC} n_{mC})$. Thus, the fraction of 5hmC bases, denoted as "corrected 5hmC" is estimated as $f_{hmC} = C_{TAB} / (C_{TAB} + T_{TAB}) - e$. Finally, 5mC is the estimated as $f_{mC} = f_{h/mC} - f_{hmC}$.

Primer sequences for qRT-PCR

We extracted total RNA by Trizol reagent (Life Technologies) and synthesized cDNA using SuperScript III (Life Technologies), per manufacturer instructions. Finally, qRT-PCR (duplicate biological replicates, each consisting of triplicate technical replicates) was performed on a Roche Light Cycler 480 using the following primers:

Actin, forward: ctaaggccaaccgtgaaaag
Actin, reverse: accagaggcatacagggaca
Slit3, forward: gccacctcagtgagaacctc
Slit3, reverse: tgtccctcaaagcccaga
Lmo4, forward: ttgcaatataggggagaagca
Lmo4, reverse: tccatggcatagagcagaaa
Irx3, forward: aaaagttactcaagacagctttcca
Irx3, reverse: cgatttaaaaatggttgaaaagttaag

Locus-specific bisulfite sequencing

After bisulfite conversion of genomic DNA, semi-nested PCR was performed with Taq polymerase (Qiagen) with [$MgCl_2$] = 2.0 mM using the following conditions: 1) 95C for 4 minutes, 2) 30 cycles of (95C for 45 seconds, 57.5C for 1.5 minutes, 72C for 1.5 minutes), 3) 72C for 5 minutes. The PCR primer sequences used are:

DMR I1:

Forward primer 1: AATTTTTGTGTTATTTGAGAGATTG
Forward primer 2: GGTTTTTGGATTTGGTATAGTTTT
Reverse primer 1 and 2: ACAACCAATTCATATCACTCATTTA

DMR L1:

Forward primer 1: TTGTTTAAAGAAGTTTTATGAGGGTT
Forward primer 2: TTTTTTATTGGGTTGGGTTATAAG
Reverse primer 1 and 2: CAATATCTTTATTACCTCCTCAAAATTC

DMR S1:

Forward primer 1: ATATTTTAAGGTATTTTAAAAATAAAAGGT
Forward primer 2: TTAATGAAGATGAATTTTAGTTTATAAATT
Reverse primer 1 and 2: CAATAAATAAAAAATATTTTCATTATAACC

PCR products were cloned into pGEM-T Easy vectors (Promega), transformed into Stellar Competent cells (Clontech), and plated on LB/ampicillin/X-gal media. Plasmids were extracted from white colonies and sequenced (Eton Bioscience Inc).

Supplemental References

Crain, P.F. (1990). Preparation and enzymatic hydrolysis of DNA and RNA for mass spectrometry. Methods Enzymol *193*, 782-790.

Dawlaty, M.M., Breiling, A., Le, T., Raddatz, G., Barrasa, M.I., Cheng, A.W., Gao, Q., Powell, B.E., Li, Z., Xu, M.*, et al.* (2013). Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev Cell *24*, 310-323.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Huang, Y., Chavez, L., Chang, X., Wang, X., Pastor, W.A., Kang, J., Zepeda-Martinez, J.A., Pape, U.J., Jacobsen, S.E., Peters, B.*, et al.* (2014). Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. Proc Natl Acad Sci U S A *111*, 1361-1366.

Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C.A., Mostoslavsky, G.*, et al.* (2011). Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell *8*, 200-213.