nature cell biology

ARTICLES

The histone deacetylase SIRT6 controls embryonic stem cell fate via TET-mediated production of 5-hydroxymethylcytosine

Jean-Pierre Etchegaray[1,2,10], Lukas Chavez[3,8,9], Yun Huang[3,8,9], Kenneth N. Ross[1,2], Jiho Choi[1,2], Barbara Martinez-Pastor[1,2], Ryan M. Walsh[1,2], Cesar A. Sommer[4], Matthias Lienhard[3], Adrianne Gladden[5], Sita Kugel[1,2], Dafne M. Silberman[6], Sridhar Ramaswamy[1,2], Gustavo Mostoslavsky[4], Konrad Hochedlinger[1,2,7], Alon Goren[5,10], Anjana Rao[3] and Raul Mostoslavsky[1,2,10]

How embryonic stem cells (ESCs) commit to specific cell lineages and yield all cell types of a fully formed organism remains a major question. ESC differentiation is accompanied by large-scale histone and DNA modifications, but the relations between these epigenetic categories are not understood. Here we demonstrate the interplay between the histone deacetylase sirtuin 6 (SIRT6) and the ten-eleven translocation enzymes (TETs). SIRT6 targets acetylated histone H3 at Lys 9 and 56 (H3K9ac and H3K56ac), while TETs convert 5-methylcytosine into 5-hydroxymethylcytosine (5hmC). ESCs derived from *Sirt6* knockout (S6KO) mice are skewed towards neuroectoderm development. This phenotype involves derepression of OCT4, SOX2 and NANOG, which causes an upregulation of TET-dependent production of 5hmC. Genome-wide analysis revealed neural genes marked with 5hmC in S6KO ESCs, thereby implicating TET enzymes in the neuroectoderm-skewed differentiation phenotype. We demonstrate that SIRT6 functions as a chromatin regulator safeguarding the balance between pluripotency and differentiation through Tet-mediated production of 5hmC.

During early stages of development, ESCs proliferate and differentiate into all somatic cell types. ESC differentiation requires global changes of chromatin architecture to elicit specific epigenetic programs of gene expression associated with each somatic cell type. Chromatin alterations including changes in histone modifications and DNA methylation patterns play a critical role during the commitment, establishment and maintenance of a particular cell lineage during early embryogenesis[1]. Notably, the interplay between these chromatin alterations, and how they execute epigenetic programs of gene expression during ESC differentiation remain largely unknown.

DNA methylation is usually linked to chromatin compaction and gene inactivation, which constitutes a critical process to establish cell lineage specification during ESC differentiation[2]. DNA methylation is a reversible process catalysed by the Fe^{2+}- and α-ketoglutarate-dependent dioxygenases, TET enzymes[3,4]. There are three TET orthologues in the mouse, TET1, TET2 and TET3. These enzymes revert the methylation status of DNA by successive oxidation of 5-methylcytosine (5mC) into 5hmC, 5-carboxycytosine (5caC) and 5-formylcytosine (5fC), which are intermediates of an active DNA demethylation mechanism[5,6]. Increased levels of 5hmC are tightly associated with the maintenance of the pluripotency state of ESCs (refs 7–9). The expression of TET1 and TET2, maintained at high levels in ESCs, diminishes during differentiation, which correlates with repression of pluripotent genes and activation of developmental genes[4,10–13]. The TET-dependent production of 5hmC has been implicated in cell lineage specification of ESCs (ref. 14). However, upstream regulatory mechanisms underlying the participation of TET enzymes and the potential role of 5hmC as a direct epigenetic

[1]The Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, Massachusetts 02114, USA. [2]The MGH Center for Regenerative Medicine, Harvard Medical School, Boston, Massachusetts 02114, USA. [3]La Jolla Institute for Allergy and Immunology, Sanford Consortium for Regenerative Medicine, UCSD Department of Pharmacology, UCSD Moores Cancer Center, La Jolla, California 92037, USA. [4]The Center for Regenerative Medicine (CReM), Boston Medical Center, Boston University School of Medicine, Boston, Massachusetts 02118, USA. [5]Broad Technology Labs (BTL), The Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA. [6]Department of Human Biochemistry, Medical School, CEFyBO-UBA-CONICET, Buenos Aires, CP1121, Argentina. [7]Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA. [8]Present addresses: Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, Heidelberg 69120, Germany (L.C.); Institute of Biosciences & Technology, Texas A&M University Health Science Center, Houston, Texas 77030, USA (Y.H.). [9]These authors contributed equally to this work.
[10]Correspondence should be addressed to J-P.E., A.Goren or R.M.
(e-mail: etchegaray.jean-pierre@mgh.harvard.edu or agoren@broadinstitute.org or rmostoslavsky@mgh.harvard.edu)

Received 20 May 2014; accepted 3 March 2015; published online 27 April 2015; DOI: 10.1038/ncb3147

component regulating specific genes during ESC differentiation remain undetermined.

One of the histone modifications involved in ESC function is acetylation of Lys 56 in histone H3 (H3K56ac), which has been linked to the pluripotent transcriptional network in human ESCs (ref. 15). More specifically, H3K56ac levels correlate with the transcriptional activation of pluripotent genes, its levels diminishing significantly on those genes during ESC differentiation[15]. How this mark is regulated during ESC differentiation remained unclear. The NAD-dependent histone deacetylase Sirt6 was shown to target H3K56ac in mouse ESCs (refs 16,17) and is one of seven mammalian members of the sirtuin protein network, with roles in genome stability, glucose metabolism and tumour suppression[16–21]. Owing to its ability to specifically target H3K56ac, we investigated the potential participation of SIRT6 in ESC differentiation. Our results demonstrate that SIRT6 directly regulates the expression of the core pluripotent genes *Oct4*, *Sox2* and *Nanog*, via deacetylation of H3K56ac, which in turn controls ESC differentiation through Tet-mediated oxidation of 5mC into 5hmC.

RESULTS

SIRT6 deletion skews ESC differentiation potential towards neuroectoderm

ESCs derived from S6KO mouse embryos showed skewed differentiation compared with those derived from their wild-type WT littermates. When cultured to form embryoid bodies (EBs), S6KO EBs from three different ESC lines were significantly smaller in size compared with their WT counterparts (Fig. 1a). Immunofluorescence analysis showed expression of the endoderm marker GATA4 to be downregulated in S6KO EBs, while expression of the neuroectoderm marker GFAP was upregulated (Fig. 1b). GFAP is also upregulated in EBs derived from S6KO induced pluripotent stem cells (iPSCs; Supplementary Fig. 1A). The skewing of S6KO EBs towards neuroectoderm and away from endoderm, mesoderm and trophectoderm was confirmed by examining the expression of additional markers (Fig. 1c and Supplementary Fig. 1B). Moreover, when we subjected ESCs to an *in vitro* neurogenesis protocol, we found a striking increase in the number of nestin- and β-III tubulin-expressing neurons in S6KO versus WT controls (Fig. 1d and Supplementary Fig. 1C,D). The expression of nestin was upregulated in S6KO EBs even under normal culturing conditions (Supplementary Fig. 1E). Notably, even before differentiation, S6KO ESCs exhibited a downregulation of genes associated with endoderm, mesoderm and trophectoderm, while neuroectoderm-related genes were upregulated, consistent with a primed differentiation state in the absence of SIRT6 (Fig. 1e and Supplementary Fig. 1F). These results point towards a previously unidentified role for Sirt6 in regulating cell lineage specification during ESC differentiation.

Pluripotent genes are not repressed during differentiation of ESCs lacking SIRT6

Persistent expression of Oct4, Sox2 and Nanog is critical to maintain the pluripotency state, but it needs to be silenced on ESC differentiation[22]. However, during early stages of ESC differentiation, OCT4 and SOX2 were shown to orchestrate germ layer fate decisions. OCT4 was found to suppress neuroectoderm differentiation while promoting development of the mesoderm. Divergently, SOX2 inhibits mesodermal differentiation, but promotes the development of neuroectoderm. An overall downregulation of Nanog is a causal event to impel the differentiation state[23]. Thus, we assessed whether lack of Sirt6 could alter expression of these pluripotent genes following differentiation. Notably, Sirt6 deficiency was characterized by persistent expression of the core pluripotency genes *Oct4*, *Sox2* and *Nanog* and their protein products in both ESCs and EBs (Fig. 1f–i). Using an alternative differentiation protocol (retinoic acid, RA) we find that even though OCT4 was normally repressed, the expression of SOX2 and NANOG persisted on forced differentiation (Supplementary Fig. 1G). These data suggest that Sirt6 negatively regulates the expression of these core pluripotent genes to achieve proper ESC differentiation.

SIRT6 regulates levels of H3K56ac and H3K9ac at the promoters of *Oct4*, *Sox2* and *Nanog*

To determine the mechanism by which SIRT6 regulates expression of the core pluripotent genes, we subjected ESCs from both S6KO and WT to chromatin immunoprecipitation (ChIP) before and after differentiation. We first assessed binding of SIRT6 to the pluripotent genes. Notably, SIRT6 was found at the promoter regions of *Oct4*, *Sox2* and *Nanog* both before and after differentiation (Fig. 2a). We then tested for the presence of H3K56Ac, one of the SIRT6 substrates previously linked to pluripotent gene expression[15]. Consistently, the levels of H3K56ac at these promoters were increased in S6KO compared with WT ESCs and EBs (Fig. 2b). Additionally, another SIRT6 substrate, H3K9ac, was also increased at the promoter regions of these core pluripotent genes in S6KO versus WT ESCs (Fig. 2c). Furthermore, the recruitment of SIRT6 was extended inside the *Oct4* locus, showing maximum binding at the promoter (primers A and B at −195 and −13 positions, respectively) and towards the 3′ end of exon 1 (primer C at position +719) before and after differentiation (Fig. 2d,e), which is paralleled by an increase in H3K56ac in S6KO versus WT EBs (Fig. 2f). The recruitment of Sirt6 was not apparent at position +4220 (primer D, between exons 4 and 5) towards the end of *Oct4* paralleling the unchanged levels of H3K56ac between WT and S6KO EBs. Moreover, Sirt6-dependent deacetylation of both H3K9ac and H3K56ac on the *Oct4, Sox2* and *Nanog* loci was confirmed by genome-wide analyses (Fig. 2g,h). Notably, an increase in these histone modifications is retained after RA-dependent differentiation, particularly at the Sox2 locus (Supplementary Fig. 1H). These results support the idea that Sirt6 negatively regulates expression of the core pluripotency genes *Oct4*, *Sox2* and *Nanog*, thereby emphasizing a critical role for Sirt6 in stem cell function. The inability to suppress these core pluripotent genes might in part account for the skewed differentiation towards the neural lineage in S6KO compared with WT ESCs and iPSCs.

Global increase of TET expression and 5hmC in S6KO versus WT ESCs

Tet1 and *Tet2* genes are postulated targets of OCT4 and SOX2 (refs 14,24). Strikingly, we observed significant upregulation of Tet1 and Tet2 expression in S6KO compared with WT ESCs, at both messenger RNA and protein levels (Fig. 3a,b). Slot blot analysis showed a striking global increase of 5hmC in S6KO compared with WT ESCs, without a global alteration in the levels of 5mC (Fig. 3c,d).

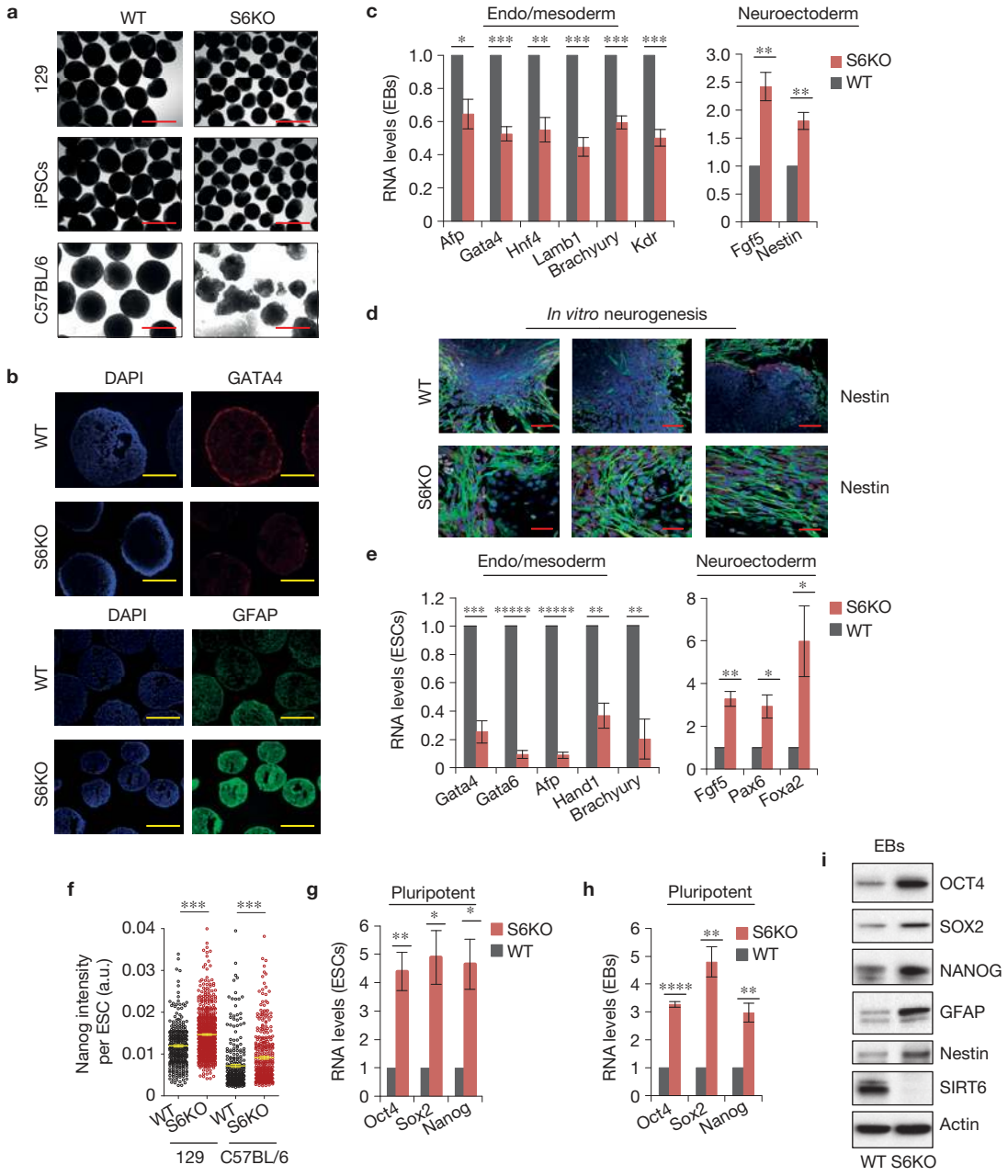


Figure 1 SIRT6 deficiency skews ESC differentiation towards neuroectoderm and promotes expression of *Oct4*, *Sox2* and *Nanog*. (**a**) EBs derived from WT and S6KO ESCs (129, iPSCs from 129 and C57BL/6 mouse strain). Scale bars, 250 μm. (**b**) Immunofluorescence of EBs from WT and S6KO (129 mouse strain) for *Gata4* (scale bars, 250 μm) and *Gfap* (scale bars, 500 μm). (**c**) Expression of endoderm, mesoderm and neuroectoderm genes in WT versus S6KO EBs. Quantitative PCR with reverse transcription (qRT–PCR) data are expressed relative to WT EBs. (**d**) Immunofluorescence of *in vitro*-generated neurons from WT and S6KO EBs for nestin (green). Nuclei were stained with DAPI. Scale bars, 100 μm. (**e**) Expression of endoderm, mesoderm and neuroectoderm genes in WT versus S6KO ESCs. qRT–PCR data are expressed relative to WT EBs. (**f**) Quantification of *Nanog* levels (mean intensity, a.u.) per cell by immunostaining in WT and S6KO ESCs from 129 and C57BL/6 genetic backgrounds. Yellow bars represent mean ± s.e.m. $^{***}P<0.001$ by one-way ANOVA followed by Tukey test analysis (a.u., arbitrary units). (**g**) Core pluripotent gene expression in WT versus S6KO ESCs assessed by qRT–PCR analysis. (**h**) Core pluripotent gene expression in WT versus S6KO EBs assessed by qRT–PCR analysis. (**i**) Western blot analysis for the core pluripotent factors on WT versus S6KO EBs. These are representatives of at least $n=3$ experimental replicates. The data in **c**,**e**–**h** are at least $n=3$ experimental replicates, values are mean ± s.e.m. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.0001$, $^{*****}P<0.00001$, by *t*-test analysis. qRT–PCR analysis in **c**,**e**,**g**,**h** is from independent RNA preparations. Uncropped images of blots are shown in Supplementary Fig. 9.

Concordantly to the upregulation of *Tet* genes in S6KO, we found an increased recruitment of OCT4 and SOX2 to both *Tet1* and *Tet2* genes in S6KO versus WT ESCs by ChIP analysis, using primers for previously identified OCT4:SOX2 consensus binding sites[14] (Fig. 3e,f). Furthermore, levels of OCT4 and SOX2 binding to their own genes were enhanced in S6KO versus WT ESCs (Fig. 3g). These data

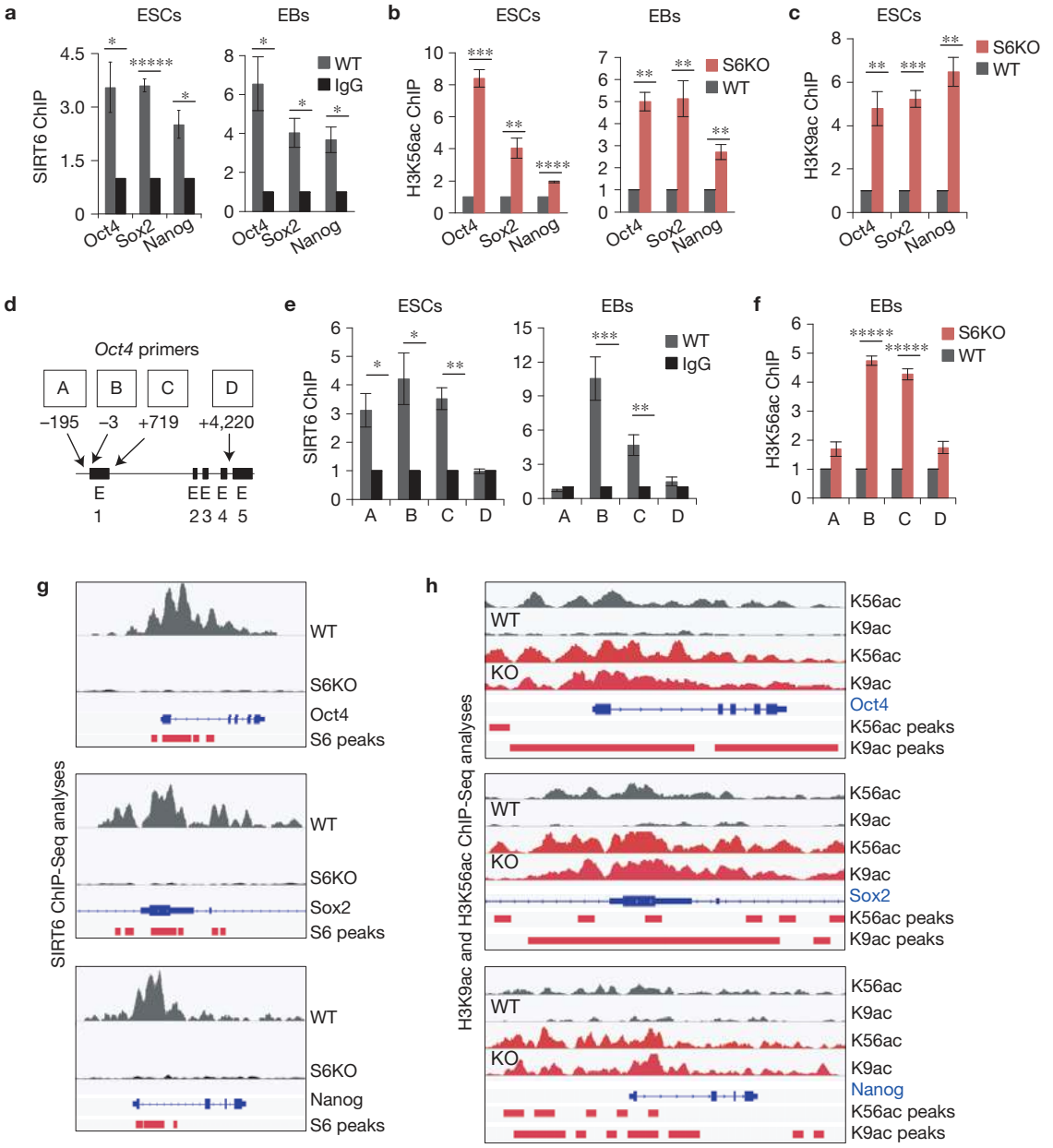


Figure 2 SIRT6-dependent regulation of core pluripotent genes. (**a**) ChIP analysis for Sirt6 on core pluripotent gene promoters in WT ESCs and EBs. Data are expressed relative to IgG ChIP control. (**b**) ChIP analysis for H3K56ac on core pluripotent gene promoters in both ESCs and EBs from WT and S6KO. Data are expressed relative to WT values. (**c**) ChIP analysis for H3K9ac on core pluripotent gene promoters in EBs from WT and S6KO. Data are expressed relative to WT values. (**d**) Schematic diagram of the *Oct4* locus depicting primers used for ChIP assays in **e**,**f**. (**e**) ChIP analysis for SIRT6 on the *Oct4* locus in WT ESCs and EBs. Data are expressed relative to IgG ChIP control. (**f**) ChIP analysis for H3K56ac on the *Oct4* locus in EBs from WT and S6KO. Data are expressed relative to WT values. (**g**) SIRT6 ChIP-Seq binding profiles on *Oct4*, *Sox2* and *Nanog* genes in WT and S6KO ESCs. Images were created with the Integrative Genomic Viewer[51]. Data are normalized to total counts, and the scale range is 0.0–7.0. (**h**) ChIP-Seq binding profiles of histone marks H3K56ac and H3K9ac on *Oct4*, *Sox2* and *Nanog* genes in WT and S6KO ESCs. Images were created with the Integrative Genomic Viewer[51]. Data are normalized to total counts, and the scale range is 0.0–2.0. The red bars under each plot in **g**,**h** represent statistically significant peaks for each ChIP-Seq analysis. The data in **a**–**c**,**e**,**f** are $n=3$ experimental replicates (independent chromatin samples from different cell preparations); values are mean ± s.e.m. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.0001$, $^{*****}P<0.00001$, by *t*-test analysis.

support a positive role for OCT4:SOX2-dependent regulation of Tet expression, which is enhanced in S6KO compared with WT ESCs, and thereby suggest an important function for SIRT6, as a potential modulator of a Tet-dependent mechanism associated with ESC differentiation.

Depletion of TET1 or TET2 rescues the differentiation phenotype in S6KO ESCs

The upregulation of TETs and 5hmC production in S6KO led us to examine the genetic interaction between SIRT6 and TETs in ESCs. Strikingly, short hairpin RNA (shRNA)-mediated depletion of

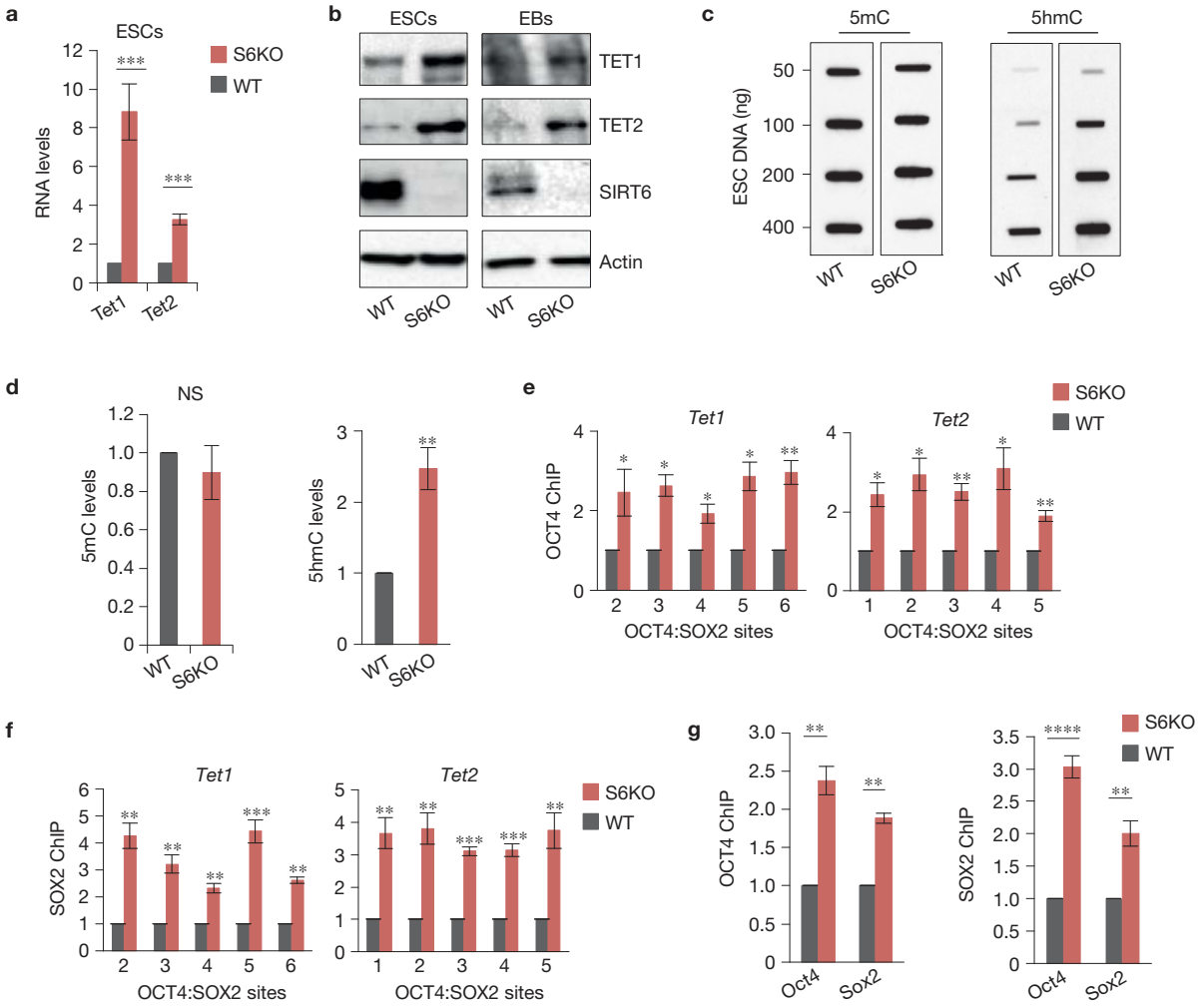


Figure 3 OCT4:SOX2-dependent upregulation of TETs in S6KO versus WT ESCs and EBs. (**a**) *Tet1* and *Tet2* gene expression in WT versus S6KO ESCs. qRT–PCR data are expressed relative to WT ESCs. Data are from $n=3$ experimental replicates (independent RNA preparations) and values are mean ± s.e.m. $^{***}P<0.001$, by *t*-test analysis. (**b**) Western blot analysis for TET1 and TET2 in both ESCs and EBs. A representative of $n=3$ biological replicates is shown. (**c**) Global 5mC and 5hmC levels assayed by slot blot analysis in WT versus S6KO ESCs. (**d**) Graphs showing fold change of 5mC and 5hmC from **h**. (**e**) ChIP analysis for OCT4 on both *Tet1* and *Tet2* at OCT4:SOX2 predicted binding sites[14] in WT versus S6KO ESCs. Data are expressed relative to WT values. (**f**) ChIP analysis for SOX2 on both *Tet1* and *Tet2* at OCT4:SOX2 predicted binding sites[14] in WT versus S6KO ESCs. Data are expressed relative to WT values. The data are $n=3$ experimental replicates. Values are mean ± s.e.m. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, by *t*-test analysis. (**g**) ChIP analysis for OCT4 or SOX2 on both *Oct4* and *Sox2* genes in WT versus S6KO ESCs. Data are expressed relative to WT values. The data in **d**–**g** are $n=3$ experimental replicates (independent chromatin samples from different cell preparations). Values are mean ± s.e.m. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.0001$, by *t*-test analysis. NS, not significant. Uncropped images of blots are shown in Supplementary Fig. 9.

either *Tet1* or *Tet2* (ref. 14) fully rescued not only the abnormal morphology of EBs derived from S6KO ESCs (Fig. 4a), but also the skewed differentiation, as demonstrated by the normalized expression of the neuroectoderm marker GFAP (Fig. 4b,c). Furthermore, the expression of additional germ layer markers—the neuroectoderm markers *Fgf5* and nestin and the endoderm markers *Gata4* and *Gata6*—was restored following *Tet* knockdown (Fig. 4d). As expected, the levels of *Tet1* and *Tet2* transcripts, as well as *Oct4* and *Nanog* transcripts, were brought back to near WT levels on *Tet* knockdown in S6KO EBs (Fig. 4e), as were the elevated levels of 5hmC in S6KO ESCs (Fig. 4f,g). Interestingly, knockdown of either *Tet1* or *Tet2* caused downregulation of both *Tet* genes (Fig. 4e), thereby implicating both TETs in the S6KO differentiation phenotype. These results further support a role for SIRT6 in suppressing expression of the core pluripotent genes in both pluripotent and differentiating ESCs, thus indirectly controlling the levels of TET enzymes to facilitate proper differentiation.

SIRT6 controls levels of 5hmC of genes associated with neuroectoderm

To elucidate the role of TET proteins in WT and S6KO ESCs, we analysed the genome-wide distribution of 5hmC in WT and S6KO ESCs by cytosine 5-methylenesulphonate immunoprecipitation[25]. We then compared 5hmC levels in adjacent, non-overlapping 500-base-pair (bp) windows genome-wide[26]. S6KO ESCs showed significant ($P<1\times10^{-3}$) gain of 5hmC at more windows (2,218) than lost 5hmC

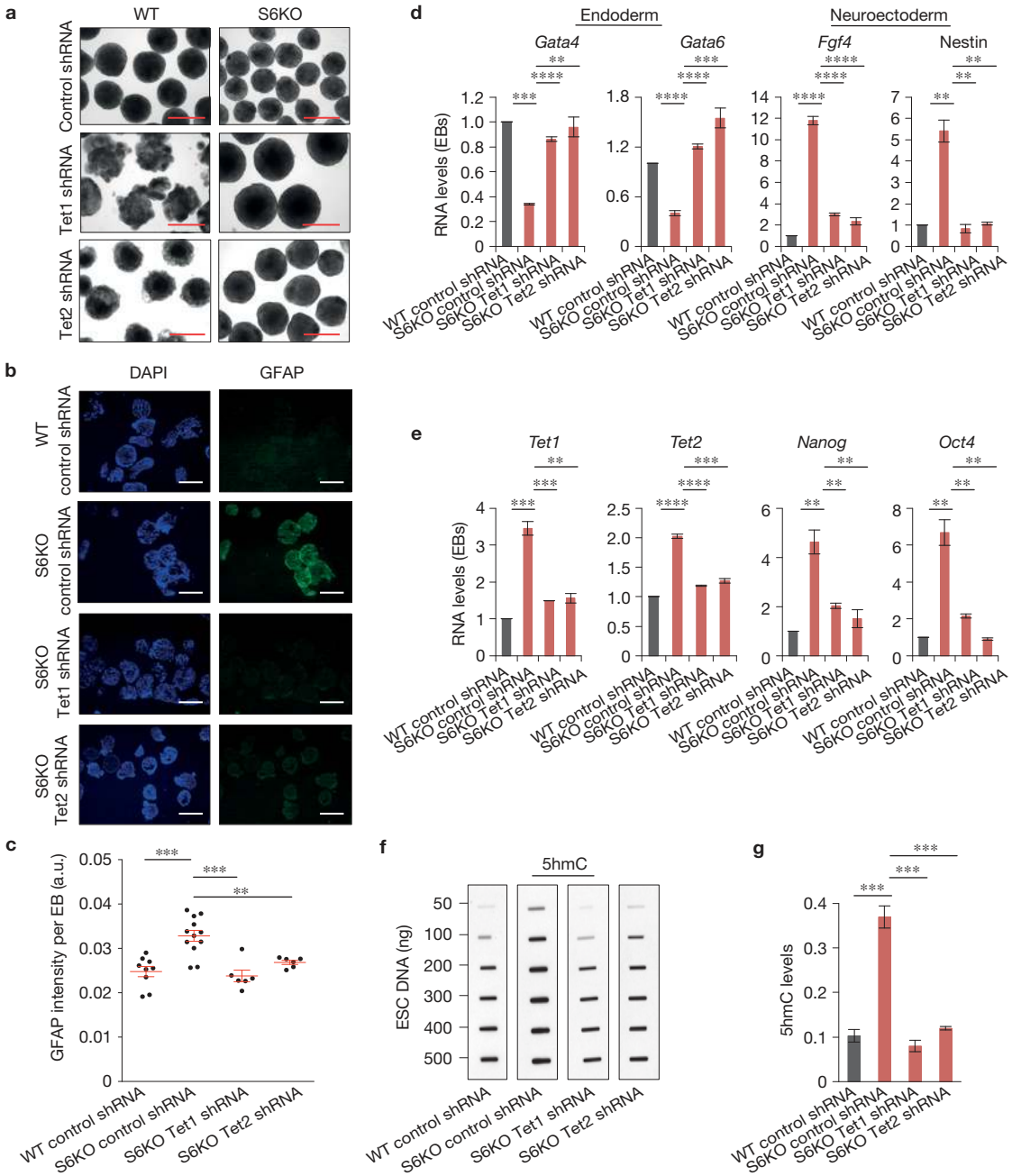


Figure 4 Tet knockdown rescues the differentiation phenotype of S6KO ESCs and global levels of 5hmC. (**a**) EBs derived from WT and S6KO ESCs stably infected with shRNA targeting *Tet1* or *Tet2*. Scale bars, 500 μm. (**b**) Immunofluorescence of Tet knockdown EBs for GFAP. Scale bars, 500 μm. (**c**) Graph showing quantification of GFAP intensity (mean intensity, a.u.) per EB in **b**. Data are represented as detection of GFAP per EB ($n=3$ EB preparations from each genotype). Red bars represent mean ± s.e.m. $^{**}P<0.01$, $^{***}P<0.001$ by one-way ANOVA followed by Tukey test analysis. (a.u., arbitrary units). (**d**) Gene expression of endoderm, and neuroectoderm genes in Tet knockdown EBs. qRT–PCR data are expressed relative to WT EBs stably transfected with shRNA control. (**e**) Expression of *Tet* and core pluripotent genes in *Tet* knockdown EBs analysed as described above. (**f**) Global 5hmC levels assayed by slot blot analysis in *Tet* knockdown ESCs. (**g**) Graphs show fold change of 5hmC from **f**. The data are represented as $n=3$ experimental replicates (independent DNA preparations from different cell cultures). Values are mean ± s.e.m. $^{***}P<0.001$, by *t*-test analysis. The data in **d**,**e** and **g** are $n=3$ experimental replicates (qRT–PCR analysis from independent RNA preparations). Values are mean ± s.e.m. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.0001$, by *t*-test analysis. qRT–PCR analysis on **d**,**e** are from independent RNA preparations. Uncropped images of blots are shown in Supplementary Fig. 9.

(1,562 windows; Supplementary Table 1), confirming at a genome-wide level the overall gain of 5hmC observed by dot blot analysis (Fig. 3c,d). Differentially hydroxymethylated regions (DHMRs) with gain of 5hmC were enriched over DHMRs with loss of 5hmC at promoters ($P<1.3\times10^{-252}$ and $P<6.3\times10^{-16}$, respectively) as well as CpG islands ($P<1.4\times10^{-269}$ and $P<1.2\times10^{-48}$, respectively;

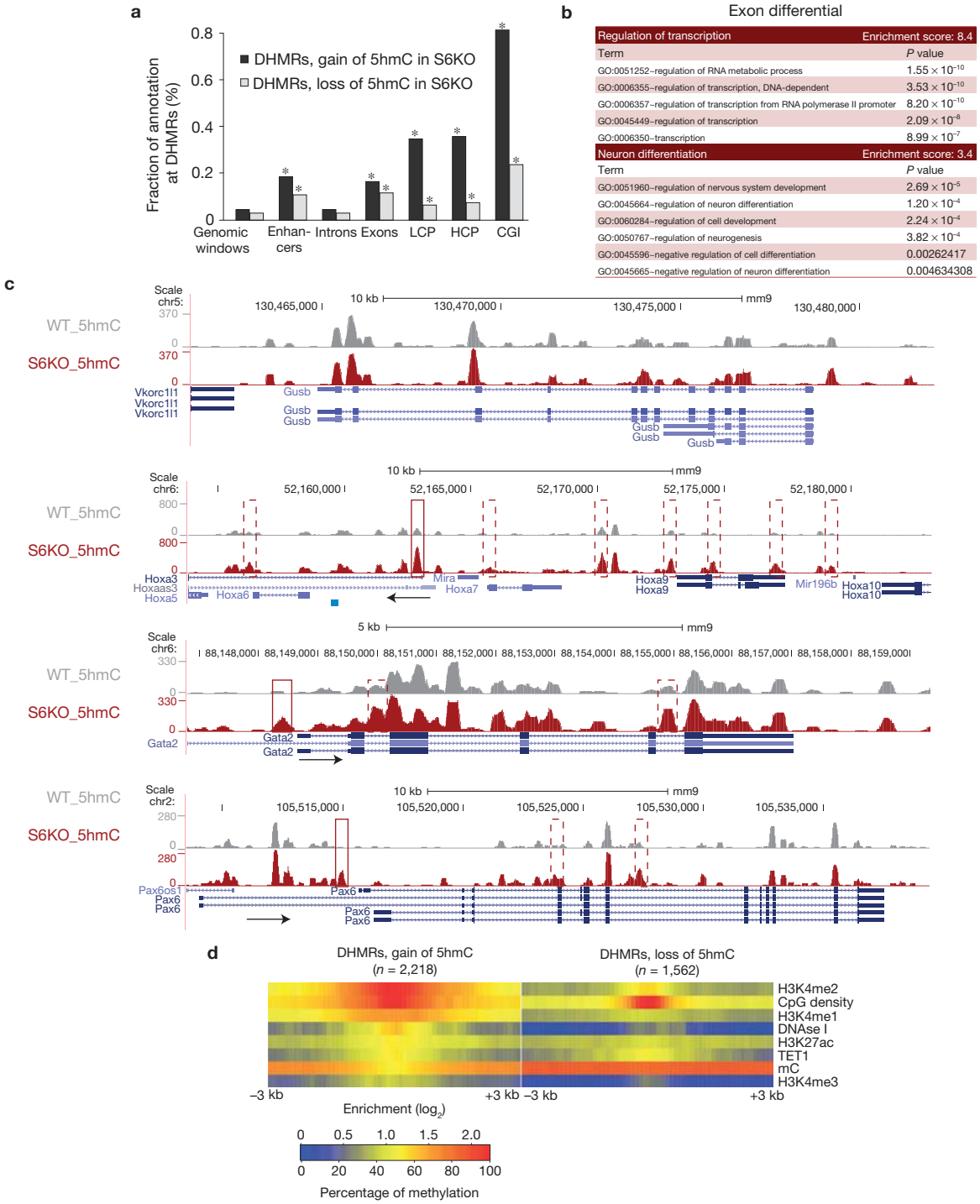


Figure 5 Characterization of genomic regions with change of 5hmC in S6KO compared with WT ESCs. (**a**) Total DHMRs with gain ($n=2,218$ genomic windows) or loss ($n=1,562$ genomic windows) of 5hmC in S6KO represent 0.04% or 0.03% of the genome, respectively (genomic windows). Both classes of DHMRs are significantly enriched at enhancers, exons, LCPs HCPs, and CpG islands (*P value $\leq 1\times10^{-3}$, Fisher's exact test), where DHMRs with gain or loss of 5hmC are similarly enriched at exons (P value $\leq 6.3\times10^{-155}$ and P value $\leq 1.5\times10^{-110}$, respectively), whereas DHMRs with gain of 5hmC are much stronger enriched than DHMRs with loss of 5hmC at promoters (LCPs and HCPs, P value $\leq 1.3\times10^{-252}$ and P value $\leq 6.3\times10^{-16}$, respectively), and at CpG islands (P value $\leq 1.4\times10^{-269}$ and P value $\leq 1.2\times10^{-48}$, respectively). (**b**) Functional annotation[45] of genes with change of 5hmC (gain or loss) at exons reveals significantly enriched gene ontology clusters associated with regulation of transcription and neuron differentiation. (**c**) UCSC browser visualization of gain of 5hmC in S6KO (red) versus WT (grey) at the promoter of *Hoxa3*, *Gata2* and *Pax6* genes, and at multiple other genomic regions in the vicinity. No changes in 5hmC levels on the *Gusb* gene, is shown as an analytical control. (**d**) Enrichment analysis of histone H3 modifications strongly connects H3K4me2 and low-methylated regions to DHMRs with gain of 5hmC in S6KO. Both, gain and loss of 5hmC occurs at regions with elevated levels of CpG density and TET1 binding. The marks are sorted by high (top) to low (bottom) enrichment at the centre of DHMRs with gain of 5hmC in S6KO.

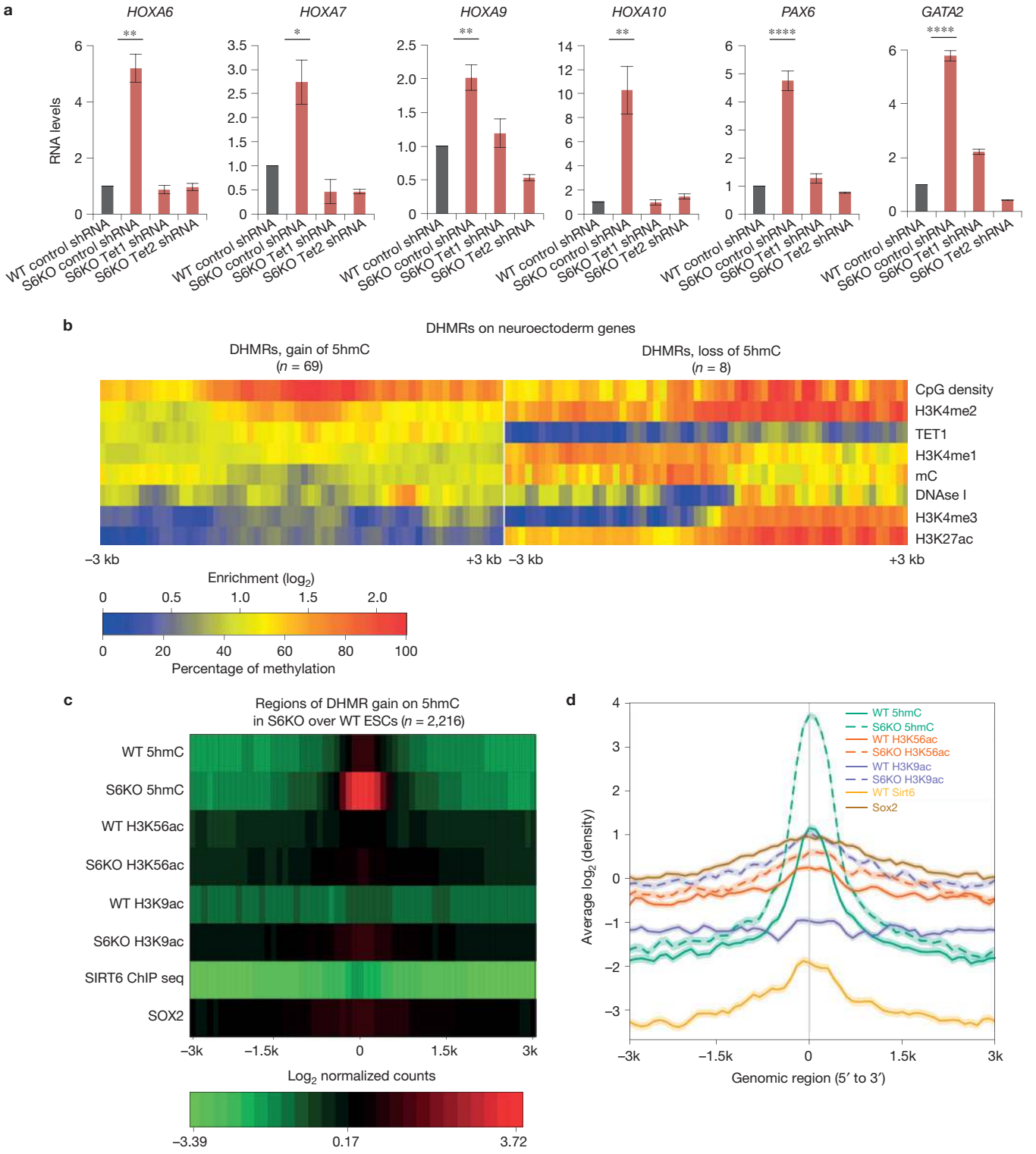


Figure 6 Association of 5hmC with H3K4me2, but not H3K9ac and/or H3K56ac, in neuroectoderm genes upregulated in S6KO compared with WT ESCs. (**a**) Expression of neuroectoderm genes with gain of 5hmC in *Tet* knockdown ESCs. qRT–PCR data are expressed relative to WT control shRNA. Data are $n = 3$ experimental replicates (independent RNA preparations), and values are mean ± s.e.m. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{****}P < 0.0001$, by *t*-test analyses. (**b**) Enrichment analysis of histone H3 modifications connect H3K4me2 to upregulated neuroectoderm genes with gain of 5hmC in S6KO. The marks are sorted by high (top) to low (bottom) enrichment at the centre of DHMRs with gain of 5hmC in S6KO. (**c**) Heat map plot of regions of DHMR gain on 5hmC in S6KO over WT in ESCs showing average profile for ±3 k centred around the 2,216 regions with DHMR gains for the factors 5hmC, H3K56ac, H3K9ac, SIRT6 and SOX2 in WT and S6KO ESCs. Each row of the heat map represents mean values of enrichment *z*-scores in 10-bp windows in the ±3 kb region. (**d**) Enrichment line plot of average profile for regions of DHMR gain on 5hmC in S6KO over WT mouse ESCs ($n = 2,216$) for data in **c**. The semi-transparent band behind the line shows the s.e.m. for each average profile.

Fig. 5a). By grouping gene promoters according to their CpG densities (±1 kilobase (kb) around their transcription start sites), we found DHMRs to be equally enriched in low- (LCP) as well as high- (HCP) CpG-density promoters (Fig. 5a). However, gain of 5hmC within exons was preferentially observed at genes important for regulation of transcription and neuronal differentiation (Fig. 5b and Supplementary Table 2 and Supplementary Figs 2 and 3). Notably, the presence of 5hmC at exons was recently shown to positively correlate with gene expression[27]. In particular, the *Hoxa* gene cluster, which is implicated in neural crest development, exhibits a significant enrichment of 5hmC, along with *Gata2* and *Pax6*, also implicated in neurogenesis (Fig. 5c). However, the housekeeping gene β-glucuronidase (*Gusb*) shows no difference in 5hmC levels between WT and S6KO ESCs (Fig. 5c), which ensures specificity in this analysis. By correlating gain or loss of 5hmC with publicly available data on histone marks, bisulphite-derived methylation, and transcription factor binding data (Supplementary Table 3), we found that regions that gained 5hmC in S6KO compared with WT ESCs were also enriched in H3K4me2, an epigenetic mark associated with both promoters and enhancers that is involved in transcriptional activation and the binding of TET1 (Fig. 5d). Interestingly, the gain of 5hmC in S6KO, compared with WT ESCs, occurs at regions with low 5mC in WT ESCs (Fig. 5d), confirming an important role for SIRT6 in regulating 5hmC at low-methylated regions, which have been previously associated with distal regulatory elements[28] (Fig. 5d). Concomitant to the enrichment of 5hmC and H3K4me2, we found the expression of the *Hoxa* gene cluster together with *Gata2* and *Pax6*, along with other genes of the neural lineage, to be upregulated in S6KO versus WT ESCs (Fig. 6a and Supplementary Fig. 4A). Importantly, the upregulation of these genes was rescued on *Tet1* or *Tet2* knockdown (Fig. 6a and Supplementary Fig. 4A). The association between 5hmC and H3K4me2 was further evaluated by correlation with public data as described above, on a set of genes from the neural lineage whose expression is upregulated in S6KO and rescued on *Tet* knockdown (Fig. 6b). This analysis shows a strong association between 5hmC with H3K4me2 within neuroectoderm genes whose expression is increased in S6KO versus WT ESCs (Fig. 6a,b). We also address any potential interplay of 5hmC with H3K9ac and/or H3K56ac by genome-wide analyses. Consistent with our previous work, we found a global elevation of both H3K56ac and H3K9ac at various genomic regions in ESCs, which is maintained after RA-dependent differentiation (Supplementary Fig. 7). We found SIRT6 and its histone targets, H3K56ac and H3K9ac, to mark genes involved in transcription, metabolism, RNA processing, cell cycle, chromatin organization, and DNA repair as well as ESC marker genes associated with the maintenance of pluripotency (Supplementary Tables 4–8). Notably, we found no significant correlation on the enrichments of 5hmC, H3K56ac, H3K9ac, or SIRT6 binding in S6KO versus WT ESCs (Fig. 6c,d and Supplementary Fig. 8A–D). We also evaluated the role of SOX2, which was shown to promote expression of the neural lineage[23], by comparing the genomic regions enriched for 5hmC with the SOX2 ChIP-Seq data set from ref. 29, and found no significant correlation with the gain of 5hmC in S6KO versus WT ESCs (Fig. 6c,d). We also found no correlation of H3K56ac and H3K9ac with SOX2 targets (Supplementary Fig. 8 and Supplementary Table 9). Overall, these data indicate that enrichment of 5hmC in S6KO is not directly interlocked with H3K56ac and/or H3K9ac, as well as SIRT6- or SOX2-targeted genomic regions. Therefore, SIRT6 plays a hierarchical role in regulating ESC differentiation by modulating TET-dependent production of 5hmC through direct repression of the core pluripotent gene network via deacetylation of H3K56ac and H3K9ac. Furthermore, these data strongly suggest that 5hmC may function as a positive transcriptional determinant to control the expression of genes associated with neuroectoderm development.

SIRT6 regulates developmental programs *in vivo*

To further evaluate the role of Sirt6 in ESC differentiation *in vivo*, we injected S6KO and WT ESCs subcutaneously into immunodeficient mice and followed teratoma formation. Teratomas derived from S6KO were significantly smaller compared with their WT counterparts (Fig. 7a). The neuroectoderm marker β-III tubulin and OCT4 were expressed at elevated levels in S6KO-derived teratomas (Fig. 7b). Remarkably, knocking down either *Tet1* or *Tet2* rescued the smaller size of teratomas derived from S6KO ESCs (Fig. 7c). Furthermore, the upregulation of OCT4, β-III tubulin and GFAP expression was also rescued on either *Tet1* or *Tet2* knockdown (Fig. 7d,e). To explore this phenotype further, we integrated a GFP-encoding gene within the *Rosa* locus in both WT and S6KO iPSCs. These cells were then injected into mouse blastocysts and chimaerism was determined in mid-gestation (E12.5) embryos by immunohistochemistry (IHC) using an anti-GFP antibody. As expected, WT GFP–iPSCs gave rise to most if not all tissues of the GFP chimaeras (Fig. 8a and Supplementary Table 8B). However, mice originated from blastocysts injected with S6KO iPSCs exhibited a weak GFP staining and in some cases almost exclusively in tissues of the central nervous system, such as diencephalon and partial regions of the neural cord (marked by yellow arrows; Fig. 8a and Supplementary Table 8B).

Depletion of SIRT6 in human ESCs leads to a differentiation defect similarly to mouse S6KO ESCs

To assess whether this SIRT6-dependent developmental mechanism is evolutionarily conserved, we explored the differentiation capacity of human ESCs (hESCs) on shRNA-mediated depletion of SIRT6, and found that SIRT6-depleted human EBs (hEBs) were significantly smaller compared with their controls (Fig. 8c). Furthermore, the expression of TET1, TET2 and OCT4 was elevated in SIRT6-depleted hESCs (Fig. 8d). Paralleling the role of SIRT6 in mouse ESCs, we found from genome-wide analysis in ref. 30 that human SIRT6 is recruited at human *Oct4* and *Sox2* genes (Supplementary Fig. 4B), supporting an evolutionarily conserved role for SIRT6 in regulating the core pluripotent genes. Consistently, we found minimal or no binding of SIRT6 to the *Tet* genes (Supplementary Fig. 4C), again suggesting that SIRT6 modulates TET-dependent production of 5hmC, indirectly. Furthermore, the neuroectoderm marker nestin was upregulated in SIRT6-deficient hESCs (Fig. 8d), supporting the predisposition towards the neural cell lineage in the absence of SIRT6. Hence, we propose a model whereby SIRT6 controls ESC differentiation by repressing the expression of *Oct4* and *Sox2*, consequently diminishing the OCT4:SOX2-activated expression of Tet enzymes, and thereby limiting the levels of 5hmC at specific genomic regions to allow balanced transcription of developmentally regulated genes (Fig. 8e).

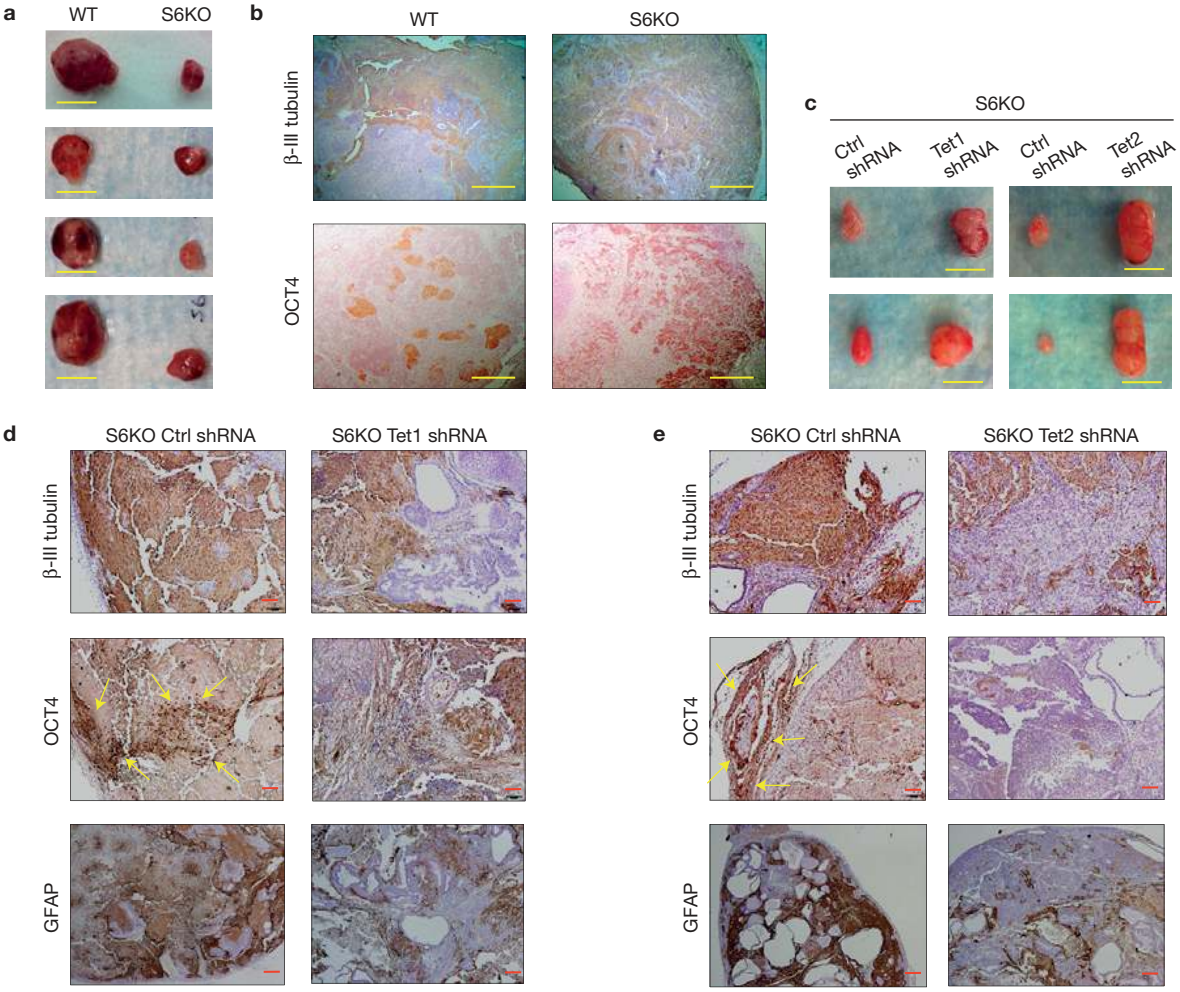
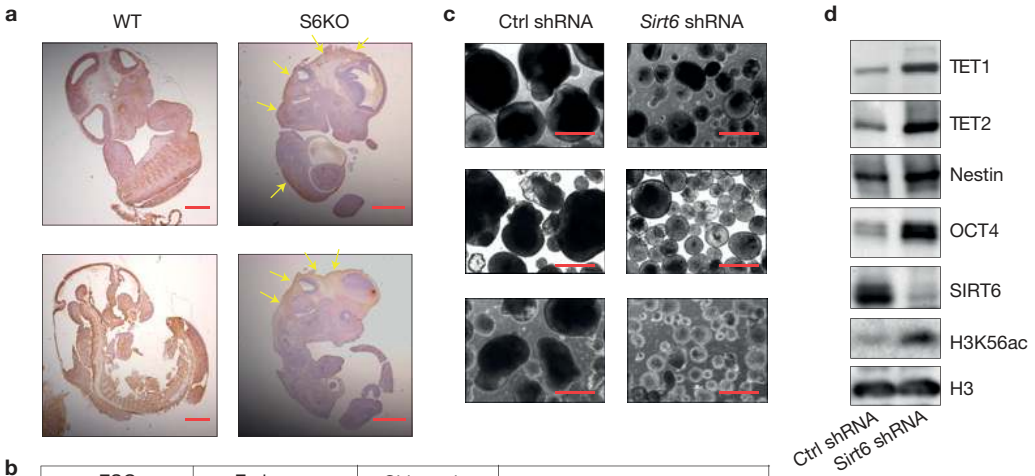


Figure 7 SIRT6 deficiency triggers an *in vivo* differentiation defect in mouse and in human EBs, which is rescued by Tet knockdown on mouse teratomas. (**a**) Teratomas from C57BL/6 WT and S6KO ESCs. Data are shown as $n = 4$ biological replicates (teratomas derived from 4 distinct mice). Scale bars, 1 cm. (**b**) IHC analysis for β-III tubulin and OCT4 on WT versus S6KO teratomas. Pictures are taken at ×5. IHC of one representative from $n = 3$ experimental replicates is shown (teratomas derived from 3 distinct mice). Scale bars, 250 μm. (**c**) Teratomas derived from stably infected S6KO ESCs with shRNA targeting either *Tet1* or *Tet2* versus control shRNA. Scale bars, 1 cm. (**d**,**e**) IHC staining for β-III tubulin, OCT4 and GFAP of teratomas from **c**. One representative from $n = 3$ experimental replicates is shown. Areas of OCT4-positive nuclear staining are demarked by the yellow arrows. Scale bars, 100 μm.

DISCUSSION

A critical step during ESC differentiation involves the silencing of the pluripotent gene network to allow expression of cell-lineage-specific genes. The core pluripotent genes *Oct4* and *Nanog* undergo transcriptional silencing through DNA methylation at their regulatory regions, which is maintained in differentiated somatic cells[31,32]. Our work demonstrates an additional mechanism to repress expression of the core pluripotent genes during ESC differentiation, which involves the histone deacetylase SIRT6. Previous studies showed that SIRT6 is a critical modulator of glycolytic metabolism, DNA repair and cancer[18–21]. Our work determines an important function for SIRT6 as a regulator of ESC differentiation, by repressing the expression of *Oct4*, *Sox2* and *Nanog*. Importantly, derepression of these core pluripotent genes in S6KO during ESC differentiation suggests that lack of SIRT6 could potentially increase the efficiency of reprogramming. Indeed, we found a ~10-fold increase in iPSC formation during reprogramming of mouse neural progenitor cells derived from S6KO mice compared with WT controls (Supplementary Fig. 5F–H).

Somatic cells switch from an oxidative metabolic state to a glycolytic state during reprogramming to iPSCs (refs 33–37). Furthermore, glycolysis is critical for the maintenance of pluripotency[38]. Both MEFs and ESCs lacking SIRT6 exhibit a higher rate of glycolysis compared with WT (ref. 19). However, the differentiation phenotype of S6KO EBs was not rescued following glycolysis inhibition, by knocking down *Pdk1* (Supplementary Fig. 5A). These data suggest a predominant role for SIRT6 in regulating the expression of pluripotent genes and 5hmC levels, independent of its function in metabolism, during ESC differentiation. Importantly, ectopic expression of human SIRT6 rescues the differentiation phenotype of S6KO EBs, thereby establishing specificity for the role of SIRT6 in the S6KO differentiation phenotype as well as its conserved role between mice and humans (Supplementary Fig. 6A–C).



b

ESC genotype	Embryos developed	Chimaeric embryos	Degree of chimaerism
S6KO iPSCs (1)	21	6	3 weak, 3 neural bias
S6KO iPSCs (2)	16	7	2 very weak, 5 weak
WT iPSCs	9	6	5 strong, 1 weak

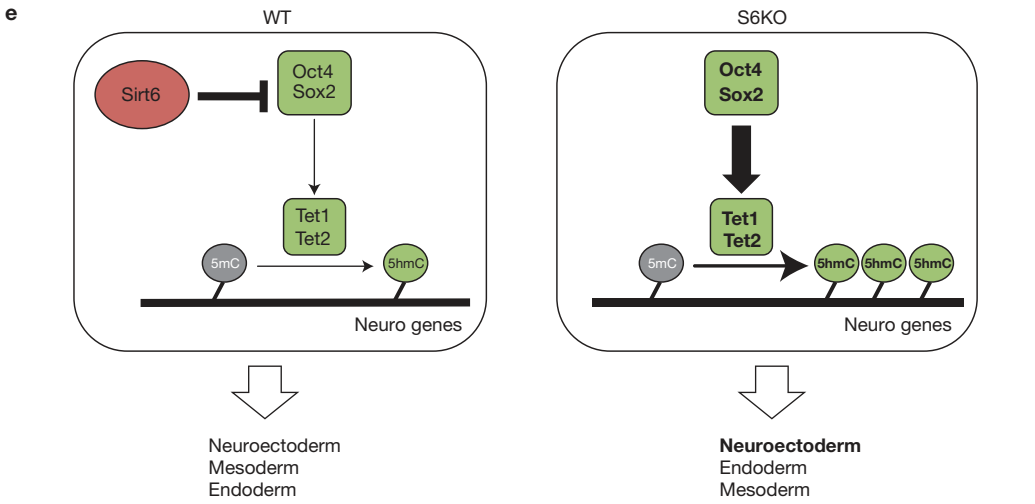


Figure 8 Sirt6 deficiency triggers an *in vivo* differentiation defect in mouse and in human EBs. (**a**) IHC analysis for GFP on chimaeric mice (E12.5) derived from WT or S6KO iPSCs. Note that the contribution of S6KO iPSCs is restricted to some neural tissues (yellow arrows). Scale bars, 1 mm. (**b**) Table showing the numbers of embryos developed and the degree of chimaerism based on IHC with anti-GFP antibody. Note that S6KO iPSCs exhibit a weak degree of chimaerism or are restricted to the neural tissue in three embryos. (**c**) hEBs stably infected with an shRNA control or shRNA targeting Sirt6. Scale bars, 500 μm. A representative of $n=5$ biological replicates is shown. (**d**) Western blot analysis for Tet enzymes, Oct4 and the neuroectoderm marker nestin on hEBs stably infected with an shRNA control or shRNA targeting Sirt6. A representative from $n=3$ experimental replicates is shown. (**e**) Schematic representation depicting the role of Sirt6 as a regulator of ESC differentiation via repression of Oct4 and Sox2 gene expression that in turn controls Tet-dependent oxidation of 5mC into 5hmC, which is needed to achieve proper development of the germ layers. Sirt6 depletion causes a derepression of Oct4 and Sox2 triggering an upregulation of Tet-dependent 5hmC production that results in skewed development towards neuroectoderm. Uncropped images of blots are shown in Supplementary Fig. 9.

The biological significance of TET-mediated 5mC oxidation in epigenetic regulation remains poorly understood, especially with regard to its relevance during ESC differentiation, whereby pluripotent cells commit to specific cell lineages. In addition to being a DNA demethylation intermediate, 5hmC is recognized by chromatin regulatory proteins and therefore postulated to function as an epigenetic mark[39]. However, specific target genes, whose expression is associated with 5hmC as an epigenetic determinant of cell lineage specification during ESC differentiation, are unknown. We identified genes associated with the neural lineage whose expression is directly correlated with an enrichment of 5hmC at promoters and exons in S6KO ESCs (Fig. 5a–c). Furthermore, the enrichment of 5hmC in these genes occurs near H3K4me2 marks, an epigenetic signature involved in transcription activation (Figs 5d and 6b). The upregulated expression of TET enzymes in S6KO ESCs suggests a potential increase in further oxidized forms of 5mC. Indeed, we detected elevated levels of 5caC in S6KO versus WT ESCs by slot blot analysis, which are rescued on *Tet1* or *Tet2* knockdown (Supplementary Fig. 5B–E).

Therefore, S6KO ESCs represent a relevant biological system to further analyse the importance of TET-mediated DNA oxidation during cell fate choices.

Recent studies have shown that *Tet1*/*Tet2* double knockout ESCs are depleted of 5hmC, which correlates with developmental defects in teratomas and chimaeric embryos[40]. *Tet1*/*Tet2*-deficient mice were obtained at lower frequency indicating a phenotype of partial lethality and thereby supporting the critical role of TET enzymes in embryonic development[40]. In this context, S6KO mice from specific genetic backgrounds (C57BL/6) are born at lower Mendelian ratios (see Supplementary Table 11), possibly due to embryonic lethality, thereby supporting that imbalanced levels of TET proteins impair embryonic development.

Changes in overall chromatin architecture are required during the transition from pluripotency to differentiated states. However, the interplay between chromatin regulators and epigenetic determinants associated with the establishment of transcriptional programs during cell fate choices remain poorly understood. Our work identifies the chromatin enzyme SIRT6 as a key regulator of ESC differentiation, acting through sequential regulation of the core pluripotent genes and TET-mediated production of 5hmC to control expression of genes involved in neural cell fate. Thus, in the absence of SIRT6, neural-related genes are marked with elevated levels of 5hmC, suggesting that this modification might function as either an epigenetic determinant or a facilitator of local DNA demethylation that channels ESC commitment to the neural cell lineage. Interestingly, TET2-dependent hydroxylation of 5mC was also found to be required for the transcriptional activation of the *Hoxa* cluster, which is critical for cell lineage specification in NT2 cells, an embryonic carcinoma cell line that can be differentiated with RA into neural cells[41]. Together, our data argue that the differentiation defect we observe in S6KO EBs is linked to Tet function, resulting in a predisposition towards a neuroectoderm developmental pathway.

Collectively, our studies unravel a molecular mechanism implicating SIRT6 as a critical regulator of ESC differentiation that involves the core pluripotent genes and TET-dependent production of 5hmC. Future studies will determine, at a gene-by-gene level, whether 5hmC and its further oxidized forms 5fC and 5caC participate in gene expression and ESC differentiation as molecular intermediates in the process of DNA demethylation, as epigenetic marks that recruit chromatin and transcriptional regulators to gene regulatory regions, or whether both mechanisms apply. □

METHODS

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

Note: Supplementary Information is available in the online version of the paper

ACKNOWLEDGEMENTS

This work was supported in part by NIH grants GM093072-01, DK088190-01A1 (R.M.), 5R01HD058013-05 (to K.H.) and HD065812, CA151535 (to A.R.). R.M. is the Kristine and Bob Higgins MGH Research Scholar, the Warshaw Institute Fellow, and a Howard Goodman Awardee. L.C. was the recipient of a Feodor Lynen Research Fellowship from the Alexander von Humboldt Foundation. Y.H. was supported by a postdoctoral fellowship from the Leukemia and Lymphoma Society. C.A.S. is the recipient of the Evans Center Fellow Award. A.Gladden and A.Goren were the recipient of the Broad Institute SPARC (Scientific Projects to Accelerate Research and Collaboration) program. We thank O. Bar-Nur and S. Cheloufi for advice on the reprogramming of neural progenitor cells, and E. Kelliher for setting up the automation of the ChIP-Seq experiments.

AUTHOR CONTRIBUTIONS

J-P.E. conceived and designed most of the experiments, collected and analysed data, and wrote the manuscript; L.C., M.L. and K.N.R. performed computational analyses; Y.H. performed ChIP-Seq experiments; J.C. performed human ESC experiments; C.A.S. and G.M. performed iPSC experiments; B.M-P. performed and analysed immunofluorescence experiments; R.M.W. performed neurogenesis and immunofluorescence experiments; A.Gladden performed ChIP-Seq experiments; S.K. performed cloning of human SIRT6 expression system; D.M.S. performed *in vitro* experiments; S.R. supervised computational analyses; K.H. supervised the human ESC experiments; A.Goren conceived, supervised and analysed all the ChIP-Seq experiments and their computational analysis; A.R. supervised computational analyses, 5hMeC Chip-Seq experiments and edited the manuscript; R.M. conceived and supervised the study, analysed the data and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://dx.doi.org/10.1038/ncb3147
Reprints and permissions information is available online at www.nature.com/reprints

1. Chen, T. & Dent, S. Y. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat. Rev. Genet.* **15,** 93–106 (2014).
2. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14,** 204–220 (2013).
3. Iyer, L. M., Tahiliani, M., Rao, A. & Aravind, L. Prediction of novel family of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* **11,** 1698–1710 (2009).
4. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324,** 930–935 (2009).
5. He, Y. F. *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333,** 1303–1306 (2011).
6. Ito, S. *et al.* Tet proteins can convert 5-methycytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333,** 1300–1303 (2011).
7. Ficz, G. *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473,** 398–402 (2011).
8. Williams, K. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473,** 343–348 (2011).
9. Wu, H. *et al.* Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.* **25,** 679–684 (2011).
10. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324,** 929–930 (2009).
11. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466,** 1129–1133 (2010).
12. Ko, M. *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468,** 839–843 (2010).
13. Szwagierczak, A., Bultmann, S., Schmidt, C. S., Spada, F. & Leonhardt, H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res.* **38,** e181 (2010).
14. Koh, K. P. *et al.* Tet1 and tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8,** 200–213 (2011).
15. Xie, W. *et al.* Histone H3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. *Mol. Cell* **33,** 417–427 (2009).
16. Yang, B., Zwaans, B. M. M., Eckersdoroff, M. & Lombard, D. B. The sirtuin Sirt6 deacetylates H3K56Ac *in vivo* to promote genomic stability. *Cell Cycle* **16,** 2662–2663 (2009).
17. Michishita, E. *et al.* Cell cycle-dependent deacetylation of telomeric histone H3 lysine K56 by human Sirt6. *Cell Cycle* **16,** 2664–2666 (2009).
18. Mostoslavsky, R. *et al.* Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. *Cell* **124,** 315–329 (2006).
19. Zhong, L. *et al.* The histone deacetylase SIRT6 regulates glucose homeostasis via Hif1α. *Cell* **140,** 280–293 (2010).
20. Sebastián, C. *et al.* The histone deacetylase SIRT6 is a tumor suppressor that controls cancer metabolism. *Cell* **151,** 1185–1199 (2012).
21. Toiber, D. *et al.* SIRT6 recruits SNF2H to DNA break sites, preventing instability through chromatin remodeling. *Mol. Cell* **51,** 454–468 (2013).
22. Young, R. A. Control of the embryonic stem cell state. *Cell* **144,** 940–954 (2011).
23. Thomson, M. *et al.* Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145,** 875–889 (2011).
24. Wu, Y. *et al.* Oct4 and the small molecule inhibitor, SC1, regulates Tet2 expression in mouse embryonic stem cells. *Mol. Biol. Rep.* **40,** 2897–2906 (2013).
25. Pastor, W. A. *et al.* Genome-wide mapping of 5-hydroximethylcytosine in embryonic stem cells. *Nature* **473,** 394–397 (2011).

26. Lienhard, M., Grimm, C., Morkel, M., Herwig, R. & Chavez, L. MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. *Bioinformatics* **30,** 284–286 (2014).
27. Huang, Y. *et al.* Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA* **111,** 1361–1366 (2014).
28. Stadler, M. B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480,** 490–495 (2011).
29. Lodato, M. A. *et al.* SOX2 co-occupies distal enhancer elements with distinct POU factors in ESCs and NPCs to specific cell state. *PLoS Genet.* **9,** e1003288 (2013).
30. Ram, O. *et al.* Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell* **147,** 1628–1639 (2011).
31. Epsztejn-Litman, S. *et al.* De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* **15,** 1176–1183 (2008).
32. Li, J. Y. *et al.* Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. *Mol. Cell. Biol.* **27,** 8748–8759 (2007).
33. Folmes, C. D. *et al.* Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metab.* **14,** 264–271 (2011).
34. Panopoulos, A. D. *et al.* The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. *Cell Res.* **22,** 168–177 (2012).
35. Prigione, A. & Adjaye, J. Modulation of mitochondrial biogenesis and bioenergetic metabolism upon *in vitro* and *in vivo* differentiation of human ES and iPS cells. *Int. J. Dev. Biol.* **54,** 1729–1741 (2010).
36. Varum, S. *et al.* Energy metabolism in human pluripotent stem cells and their differentiated counterparts. *PLoS ONE* **6,** e20914 (2011).
37. Zhou, W. *et al.* HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transitions. *EMBO J.* **31,** 2103–2116 (2012).
38. Zhang, J., Nuebel, E., Daley, G. Q., Koehler, C. M. & Teitell, M. A. Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal. *Cell Stem Cell* **11,** 589–595 (2012).
39. Spruijt, C. G. *et al.* Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* **152,** 1146–1159 (2013).
40. Dawlaty, M. M. *et al.* Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Cell* **24,** 310–323 (2013).
41. Bocker, M. T. *et al.* Hydroxylation of 5-methylcytosine by TET2 maintains the active state of the mammalian HOXA cluster. *Nat. Commun.* **3,** 818 (2012).

DOI: 10.1038/ncb3147

METHODS

Mouse ESC cultures, formation and immunostaining of embryoid bodies. Mouse ESCs (mESCs) derived from Sirt6 KO and WT 129 mouse strain[18] were maintained on γ-irradiated mouse embryonic fibroblasts (MEFs) in knockout DMEM medium (GIBCO) containing 15% ES-qualified FBS, 0.1 mM each of non-essential amino acids, 2 mM L-glutamine, 0.1 mM β-mercaptoethanol, 50 units ml^{-1} penicillin/streptomycin (Invitrogen) and supplemented with leukaemia inhibiting factor (LIF). For all experiments described, cells were trypsinized and plated for 30 min on standard tissue culture dishes to remove feeder cells before floating ESCs were collected and re-plated on gelatin-coated dishes or wells before differentiation towards embryoid bodies (EBs). ESCs were dissociated with trypsin (day 0) and cultured at a density of 5,000 cells ml^{-1} in Iscove's modified Dulbecco's medium (IMDM) with 15% FCS (Atlanta Biologicals), 10% protein-free hybridoma medium (PFHM-II; Gibco), 2 mM L-glutamine (Gibco), 200 μg ml^{-1} transferrin (Roche), 0.5 mM ascorbic acid (Sigma) and 4.5°—10^{-4} M monothioglycerol (MTG; Sigma). Differentiation was carried out using the standard hanging drop method and EBs were collected at day 10. For immunofluorescence analysis, EBs were fixed and cryopreserved according to the procedure described in ref. 42. Gfap immunofluorescence on EBs was performed according to the procedure described in ref. 43. Image acquisition was done in a Leica fluorescence microscope under non-saturating exposure conditions. The quantification of Gfap intensity per EB was performed with CellProfiler software. Briefly, images were segmented using DAPI staining to generate masks matching EBs. This mask was applied in the Gfap image to calculate the mean intensity of Gfap staining in each EB.

Generation of Tet1 and Tet2 knockdown ESC lines. *Tet1* and *Tet2* shRNA constructs were used as previously described[14]. Briefly, ESCs were transfected by electroporation at 320 V, 250 μF with Tet1, Tet2 or control pSuper.puro shRNA constructs and selected on puromycin-resistant γ-irradiated feeders. Clones were picked and screened by qRT–PCR for knockdown of Tet1 and Tet2 expression. Established stable clones were subsequently maintained in the absence of puromycin without loss of knockdown assessed after up to 10 serial passages.

Hydroxymethylated DNA immunoprecipitation assay. Short reads derived from S6KO and WT ESC cytosine 5-methylenesulphonate immunoprecipitation (CMS IP) and input assays were mapped against mm9 as previously described[25]. After discarding all reads that map to multiple positions in the genome, there are 11,679,660 reads for S6KO CMS IP, 13,916,523 reads for WT CMS IP, 24,157,471 reads for S6KO input, and 16,743,260 reads for WT input.

To identify genome-wide differentially hydroxymethylated regions (DHMRs) in S6KO compared with WT, we calculated differential coverage at genome-wide 500-bp windows by employing the Bioconductor package MEDIPS (ref. 26) v1.12.0 (extend = 200, uniq = T, window_size = 500, BSgenome = BSgenome.Mmusculus.UCSC.mm9, adj = F, diff.method = edgeR, P value = 1×10^{-3}), resulting in 1,562 genomic windows with significant loss of 5hmC in S6KO, and 2,218 genomic windows with significant gain of 5hmC (Supplementary Table 1).

To assign DHMRs to genes, we downloaded the mm9 table refGene (track RefSeqGenes) via the Table Browser at the UCSC Genome Bioinformatics Site (20 December 2013). To assign DHMRs to CpG islands, we downloaded the mm9 table cpgIslandExt (track CpG Island) via the Table Browser at the UCSC Genome Bioinformatics Site (20 December 2013). To assign DHMRs to enhancers, we downloaded mm9 enhancers provided at http://chromosome.sdsc.edu/mouse/download/mESC.zip (ref. 44). We assigned DHMRs to promoters if they overlap with the ±1 kb region around a transcription start site. We have calculated the CpG observed/expected ratio for all gene promoters by considering the number of Cs, Gs and CpGs in the ±1 kb region around their transcription start sites. By fitting a Gamma Normal Mixture to the resulting promoter CpG observed/expected ratios, we have identified the intersection of the Gamma and the Gaussian distributions at CpG obs/exp = 0.399 as a threshold to distinguish gene promoters into low-CpG-density and high-CpG-density promoters. We assigned DHMRs to exons, CpG islands or enhancers, respectively, if their genomic coordinates overlap. Enrichment of GO terms associated with the previously identified genes containing DHMRs in their promoters or exons has been calculated using DAVID (ref. 45; Supplementary Table 2).

We tested the enrichment of DNA methylation, transcription factor binding sites, and several histone modifications within and surrounding our previously defined DHMRs (Figs 4d and 5c). For this, we accessed publicly available data sets as listed in Supplementary Table 3. External data sets were downloaded from NCBI GEO/SRA as sra-lite packages according to the given access numbers. Fastq files were extracted using the sra toolkit[46] sratoolkit.2.1.9-centos_linux64/fastq-dump). Fastq files of the individual lanes were mapped against mm9 using bowtie[47] (*bowtie-0.12.7/bowtie-q-n 2-k 1–best–strata-m 1–maxbts 10000 mm9-y–chunkmbs 1028–sam*). For each data set, the mapping results of the individual lanes were sorted, pooled, and exported as one bam file using samtools[48] (SAMtools-0.1.18/samtools).

Alignments (bam files) and our previously defined DHMRs (annotations) were imported into the R/Bioconductor environment (http://www.r-project.org, http://www.bioconductor.org). Alignments were shifted 60 bases towards their 3′-end to account for an estimated fragment length of 120 base pairs. For each tested annotation and for each alignment, coverage was calculated as follows: the midpoint of each region associated with an annotation was calculated and the range of −3 kb to +3 kb around this midpoint was divided into 60 bins of constant length (bins have a length of 50 bp). Alignment coverage was calculated at these bins. To avoid division by zero, we added one to each bin of the immunoprecipitate and of the control sample. Subsequently, immunoprecipitate signals at the tested bins were divided by a library size correction value (see below). Enrichment is defined by the $\log_2$ of corrected immunoprecipitate signals divided by the control signals, where the IP data sets in Supplementary Table 3 are considered as immunoprecipitate and two separate ChIP-Seq input data sets are considered as control as indicated. For the average enrichment profiles, we calculated the $\log_2$ of the mean over all corrected immunoprecipitate over control ratios at every tested position of stacked annotations. For each immunoprecipitate sample, the correction value for the library size was estimated by calculating the mean over the ratios of immunoprecipitate over control sample short-read coverage at genome-wide 100-bp windows.

For DNA methylation data, we have accessed the processed single CpG table given in ref. 28 and for each tested bin we calculated the mean mCpG signal (that is, the mean of nMeth/nTot*100 values of all CpGs that fall into the bin). For the ChIP-chip data set[28] we accessed the available processed $\log_2$ ratios, calculated the mean for each probe over both replicates, and for each tested bin, we calculated the mean of included probes.

Sirt6 ChIP-Seq data analysis. To establish the most suitable conditions for Sirt6 ChIP-Seq analysis, we first tested four different antibodies by western blotting (Aviva System Biology, OAAB16579; Abcam, ab62739; Cell Signaling Technology, 12486; Sigma-Aldrich, S4322). Next, as western blot and ChIP-Seq conditions may differ, we systematically determined the antibody and concentrations that produce the highest signal-to-noise ratio. To reduce experimental variability, we used our automated ChIP-Seq methodology[49]. For each of the two most promising antibodies—Aviva System Biology, OAAB16579 and Cell Signaling Technology, 12486—we split the sheared chromatin between 3 ratios of antibody/chromatin (0.5 ml, 1 ml and 5 ml of each antibody/3 × 10^6 cells), and performed ChIP-Seq. In addition, as a positive control, we used the same chromatin preparation for ChIP-Seq with an antibody targeting H3K27ac (Cell Signaling Technologies D5E4, optimized at 1 ml/20 × 10^6 cells). Our results indicated that CST in a ratio of 1 ml/20 × 10^6 cells outperformed the rest, and we used the data for further analysis.

Reads from Sirt6 ChIP-Seq for Sirt6 wild-type (S6WT) and knockout (S6KO) and WCE control for mouse ES cells were aligned to mouse genome mm9 using bwa and duplicate reads were marked with Picard tools (http://picard.sourceforge.net). Peaks were called using MACS2 (ref. 50) with the false discovery rate $q = 0.01$. Called peaks were checked visually with the IGV browser[51]. Peaks were assigned to genes using the TSS and NDG methods of PeakAnnotator in the PeakAnalyzer package where genes with peaks in the promoters were defined to be those genes with a peak within a distance of ±1 kb from the TSS and genes with a peak in an exon were defined as genes with the 'Overlap-Center' falling in the gene's exon.

Sox2 ChIP-Seq analysis. Data for Sox2 ChIP-Seq from mouse ESCs was downloaded from GEO (GSM1050291 or SRR630006.sra for Sox2 ChIP-Seq in mESC, GSM1050292 or SRR630007.sra for mESC WCE rep1, and GSM1050293 or SRR630008.sra for mESC WCE rep2; ref. 29) in the form of SRA files. SRA files were converted to fastq files using fastq-dump. Trimming of low-quality reads and clipping of sequencing adaptors was done using the program Trimmomatic[52]. Trimmed reads from Sox2 ChIP-Seq in mouse ESCs and WCE control were aligned to mouse genome mm9 using bowtie[47] (bowtie -q -p 11–phred33-quals -v 2 -e 70 -m 3 –best–strata–sam) and samtools[48] was used to remove duplicate reads, sort and index bam files, and merge the two WCE controls. Peaks were called using MACS2 (ref. 50) with the false discovery rate $q = 0.01$. Called peaks were checked visually with the IGV browser[51]. Peaks were assigned to genes using the TSS and NDG methods of PeakAnnotator in the PeakAnalyzer package where genes with peaks in the promoters were defined to be those genes with a peak within a distance of ±1 kb from the TSS and genes with a peak in an exon were defined as genes with the 'Overlap-Center' falling in the gene's exon.

H3K56ac and H3K9ac ChIP-Seq data analysis. Reads from H3K56ac and H3K9ac ChIP-Seq for Sirt6 wild-type (S6WT) and knockout (S6KO) and WCE control for mouse ESCs (both untreated and treated with differentiating agent retinoic acid (RA)) were aligned to mouse genome mm9 using bwa and duplicate reads were marked with Picard tools (http://picard.sourceforge.net). R Bioconductor tool MEDIPS was used to find differential ChIP-Seq binding profiles for the histone marks between S6WT and S6KO in both untreated and RA-treated

samples (using MEDIPS parameters: extend = 200, uniq = T, window_size = 500, BSgenome = BSgenome.Mmusculus.UCSC.mm9, adj = F, diff.method = edgeR, P value $= 1 \times 10^{-3}$). Adjacent blocks of gain or loss were merged using the MEDIPS function mergeFrames. This resulted in genomic windows with significant loss and gain of histone marks H3K56ac and H3K9ac as shown in Table 1. as follows:

Table 1 Genomic windows with gain and loss of H3K56ac and H3K9ac.

	Number of merged S6KO gain blocks	Number of merged S6KO loss blocks
H3K56ac	18,722	22,126
H3K9ac	39,096	28,450
H3K56ac with RA	14,887	11,536
H3K9ac with RA	23,794	24,695

Differential H3K56ac and H3K9ac blocks were checked visually with the IGV browser[51]. The merged blocks of differential gain or loss from MEDIPS were assigned to genes using the TSS and NDG methods of PeakAnnotator in the PeakAnalyzer package where genes with blocks with gains or losses in the promoters were defined to be those genes where a merged differential block fell within a distance of ±1 kb from the TSS and genes with gains or losses in an exon were defined as genes where the 'Overlap-Center' fell in the gene's exon.

Summary ChIP-Seq heat map images. Average binding profiles for areas of DHMR gain on 5hmC in S6KO (Supplementary Fig. 12A,B) and histone mark gain (Supplementary Fig. 12C–F) were produced by calculating the mean coverage values of 100-bp windows in the ±3 kb region centred in each of the areas of S6KO gain. The heat maps and line plots represent the mean enrichment by bin across all areas of S6KO gain after correcting for the direction of the DNA strand. The semi-transparent bands behind the lines of Supplementary Fig. 12B,D,F represent the standard error of the mean across the ±3 kb region.

Quantification of Nanog in mESCs. ESCs were grown in 96-well clear-bottom plates (BD Falcon) that were pretreated with 0.2% gelatin. Cells were fixed in PBS 4% (w/v) paraformaldehyde. Next, cells were permeabilized with PBS 0.5% (v/v) Triton X-100, washed with PBS 0.2% (v/v) Tween20 and blocked with PBS 0.2% (v/v) Tween20, 5% goat serum. Cells were incubated with a 1:100 dilution of anti-Nanog antibody (Abcam) for 1 h at room temperature in blocking solution. Cells were washed 3 times with PBS 0.2% (v/v) Tween20 and subsequently incubated with secondary antibody conjugated with Alexa 488 for 1 h at room temperature in blocking solution in the dark. After 2 washes, cell were stained with 300 nM DAPI for 2 min, washed once with PBS and stored in PBS at 4 °C until image acquisition. For quantification of Nanog intensity in ESCs, 16 pictures of each well were automatically acquired in a BD-Pathway 435 high-throughput microscope (Becton Dickinson) with AttoVision 1.5 software using a dry ×40 magnification lens. Images were analysed with CellProfiler software. Briefly, images were segmented using the DAPI staining to generate masks matching cell nuclei. This mask was applied in the Nanog image to calculate the mean intensity of Nanog staining in each ESC. All of the images for quantitative analyses were acquired under non-saturating exposure conditions.

Human embryonic stem cell culture. Human ESC lines were cultured with MEFs (Globalstem) pre-plated at 12–15,000 cells cm^{-2}. Medium containing DMEM/F12, 20% knockout serum replacement, 1 mM L-glutamine, 100 μM MEM non-essential amino acids, and 0.1 mM β-mercaptoethanol was used. FGF-2 at 10 ng ml^{-1} was added after sterile filtration and cells were fed daily and passaged every week using StemPro EZPassage (Invitrogen).

Human EB formation. Human ESC lines were mechanically dissociated using StemPro EZPassage (Invitrogen) to make clumps of consistent size. Human EBs were generated in suspension culture over a low-speed shaker for 10 days. Culturing conditions included medium containing DMEM/F12, 20% knockout serum replacement, 1 mM L-glutamine, 100 μM M MEM non-essential amino acids and 0.1 mM β-mercaptoethanol, in the absence of FGF-2.

Generation of ESCs from C57BL/6. Four- to five-week-old Sirt6± female mice (C57BL/6) were induced to superovulation (7.5 I.U. PMSG administration followed, 48 h after, by 7.5 I.U. hCG administration via intraperitoneal injection) and inbred. Blastocysts were collected at day 3.5 after vaginal plug check, and then cultured in knockout DMEM medium (GIBCO) as ESCs. Briefly, pre-implanted E3.5 blastocysts were retrieved from the uterine horn by flushing and collected under a dissection microscope. Each blastocyst was plated onto MEF feeder (Globalstem) containing NDiff 227 medium (StemCells) supplemented with LIF and two inhibitors[53]. The cells were cultured for 4 days without changing medium. After that, the cells were cultured for an additional 3–4 days and fed daily with fresh medium until picking. Outgrowth from the culture was picked, trypsinized, and re-plated onto MEF feeders containing KO DMEM (Invitrogen) supplemented with 15% FBS (Hyclone), 1 mM L-glutamine, 100 μM M MEM non-essential amino acids, and 0.1 mM β-mercaptoethanol, and LIF.

Generation of iPSCs. MEFs were collected from E13.5 WT and S6KO embryos using standard procedures. iPSC lines were generated with the STEMCCA lentiviral reprogramming system essentially as described previously[54]. Briefly, MEFs were co-infected at passage three or four with the Dox-inducible pHAGE2-Tet-STEMCCA vector and a lentivirus expressing the reverse tetracycline transcriptional activator (rtTA) and cultured on mitomycin C-treated fibroblasts in ESC medium (DMEM supplemented with 15% FBS, L-glutamine, penicillin/streptomycin, non-essential amino acids, β-mercaptoethanol and 1,000 U ml^{-1} LIF (ESGRO; Chemicon; Millipore). Doxycycline was added at a final concentration of 1 μg ml^{-1} and removed at day 10 post-infection. iPSC colonies were mechanically isolated five days following Dox removal, expanded in ESC medium and characterized by immunofluorescence and alkaline phosphatase staining.

Neurogenesis assay. ESCs were dissociated using trypsin (0.05%) and purified by attachment to tissue culture dishes for 1 h. EBs were allowed 3 to 4 days to form after plating of ESCs in bacterial dishes in DMEM media containing 10% defined FBS (Sigma–Aldrich), 2 mM L-glutamine (Invitrogen), 1×NEAA (Invitrogen), 10 mM Hepes (Invitrogen), 1 mM 2-mercaptoethanol, 100 units ml^{-1} penicillin, and 100 g ml^{-1} streptomycin (Invitrogen) (EB media). EBs were allowed one day to attach to tissue culture dishes, and neuronal precursors were then selected for by incubation in DMEM/F-12 media containing apotransferrin (50 g ml^{-1}; Sigma–Aldrich), insulin (5 g ml^{-1}; Sigma–Aldrich), sodium selenite (30 nM; Sigma–Aldrich), fibronectin (250 ng ml^{-1}; Sigma–Aldrich), 100 units ml^{-1} penicillin, and 100 g ml^{-1} streptomycin (Invitrogen) (ITSFn media) for 7–10 days. Cells were subsequently dissociated by trypsin (0.05%), and neuronal precursors were expanded and patterned for 4 days after plating onto fibronectin-coated/polyornithine-coated plates at a density of 75,000 cells per square centimetre in DMEM/F-12 media containing apotransferrin (100 g ml^{-1}), insulin (5 g ml^{-1}), sodium selenite (30 nM), progesterone (20 nM), putrescine (100 nM), penicillin (100 units ml^{-1}), streptomycin (100 g ml^{-1}), laminin (1 g ml^{-1}), basic fibroblast growth factor (FGF2; 10 ng ml^{-1}) (R & D Systems), Shh (500 ng ml^{-1}) (R & D Systems), and FGF8 (100 ng ml^{-1}) (R & D Systems) (N3 media). The cells were subsequently differentiated in N3 media containing 200 M ascorbic acid for 3–14 days (stage 5). Cells used for immunofluorescent staining were fixed in 4% formaldehyde for 20 min and rinsed three times with PBS followed by an overnight incubation at 4 °C with anti-TUJ1 (β-III tubulin) antibody. Samples were then incubated with Alexa Fluor 590 secondary antibody for 2 h and images were captured on a Leica Microscope.

Western blot analysis. Western analysis was carried out as previously described[19]. The antibodies used are as follows: anti-Sirt6 (Abcam, ab62739), anti-Oct4 (Santa Cruz, sc5279), anti-Sox2 (Millipore, AB5603), anti-Nanog (Abcam, ab80892), anti-Tet1 (Millipore, 09-872), anti-Tet2 (Abcam, ab94580), anti-nestin (Abcam, ab6142), anti-H3K56ac (Abcam, abab76307), anti-H3 (Abcam, ab1791) and anti-actin (Sigma-Aldrich, A2066).

Slot blot analysis. Genomic DNA samples were prepared with twofold serial dilutions in TE buffer and then denatured in 0.4 M NaOH/10 mM EDTA at 95 °C for 10 min and followed by adding an equal volume of cold 2 M ammonium acetate (pH 7.0). Denatured DNA samples were spotted on a nitrocellulose membrane in an assembled Bio-Dot apparatus (Bio-Rad) according to the manufacturer's instructions. The membrane was washed with 2×SSC buffer and ultraviolet-crosslinked for 10 min. Then the membrane was blocked with 5% non-fat milk for 1 h and incubated with anti-5hmC (Active Motif) or anti-5mC (Active Motif) for HRP-conjugated secondary antibodies and enhanced chemiluminescence detection. The membrane was subsequently stained with methylene blue to confirm corresponding amounts of DNA for each sample. Quantification was performed by ImageJ software analysis.

Chromatin immunoprecipitation assays. Chromatin immunoprecipitation (ChIP) assays were performed as previously described[55]. Briefly, ESCs were crosslinked with 1% formaldehyde/PBS for 15 min at room temperature. Crosslinking was quenched by addition of 0.125 M glycine. ESCs were washes twice with ice-cold PBS, and then collected in RIPA buffer as described previously[55]. Samples were then sonicated to generate DNA fragments of approximately 0.5 kb. Approximately, 1 μg of pre-cleared protein extract was used for immunoprecipitation with anti-Sirt6 (CST), anti-H3K56ac (Epitomics) or anti-IgG (Abcam) using protein A Sepharose beads (GE Lifesciences) followed by overnight incubation

DOI: 10.1038/ncb3147

at 4 °C. Bead-containing samples were then washed as previously described[55]. Immunocomplexes were eluted by incubation at 65 °C for 10 min in the presence of 1% SDS, and crosslinking was reversed by 5 h incubation at 65 °C in the presence of 200 mM NaCl. DNA was purified by the QIAquick spin kit (QIAGEN) and further assessed by qPCR using the LightCycler 480 system from Roche. Data were normalized to input and expressed relative to the nonspecific IgG ChIP control (Fig. 1f) or to the ChIPs from WT samples (Fig. 1g). The primer sequences can be obtained on request.

Teratoma formation assays. Teratomas were obtained by subcutaneous injection of 1×10^6 WT versus S6KO ESCs (C57BL/6) near each quadriceps flank of SCID mice (Charles River Laboratories). Teratomas were collected when the size exceeded 2.0 cm in diameter and were fixed overnight in 4% paraformaldehyde. Teratomas derived from S6KO were significantly smaller than WT and they never reached more than 0.5 cm in diameter. Paraffin sections and staining were performed by the Specialized Rodent Histopathology Core Facility at Harvard Medical School. Animal handling and maintenance were performed in accordance with institutional guidelines. Immunohistochemistry analysis on teratomas was performed by the Histopathology Core Facility at Massachusetts General Hospital.

Chimaera formation. ESCs derived from WT or S6KO (129 mouse strain) were gene targeted with ROSA26–EGFP as described previously[56] to facilitate tracking of ESC-derived tissues after blastocyst injection. Immunohistochemistry analysis using anti-GFP (Clontech) was performed by the Histopathology Core Facility at Massachusetts General Hospital.

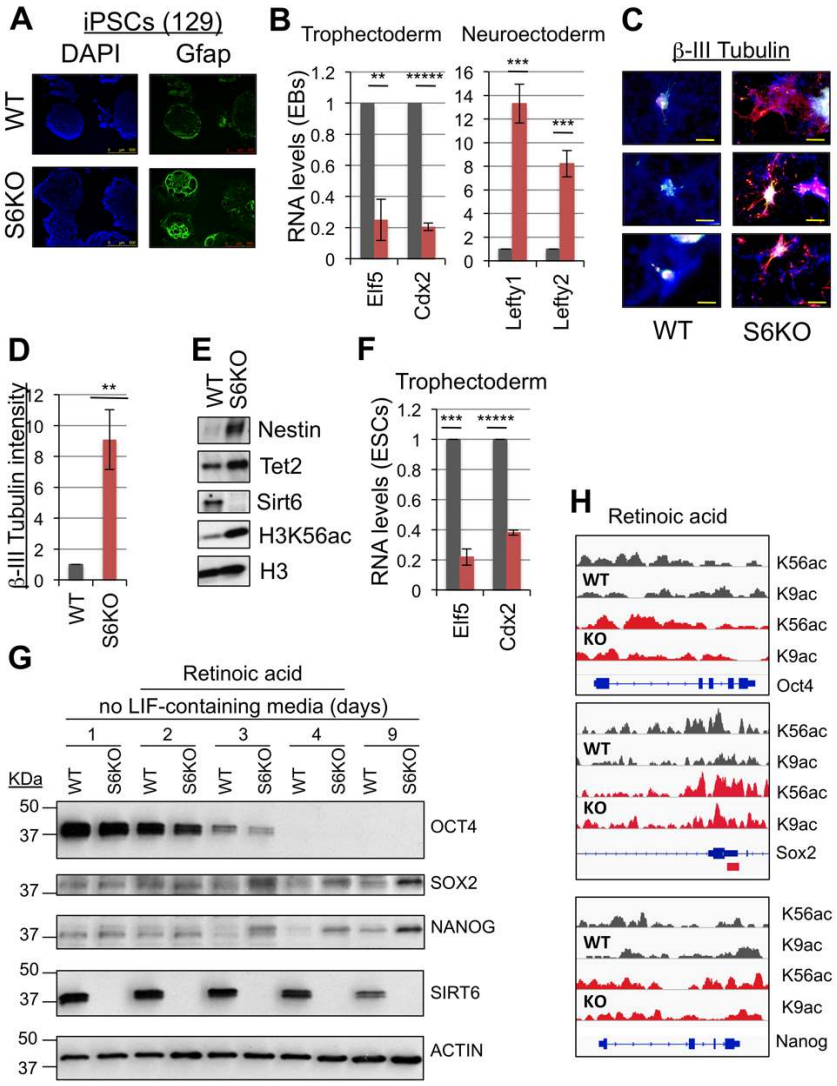
RNA extraction and quantitative PCR with reverse transcription. Total RNA was extracted with the TriPure Isolation Reagent (Roche) as described by the manufacturer. For cDNA synthesis, 1 mg of total RNA was retro-transcribed by using the QuantiTect Reverse Transcription Kit (Qiagen). Real-time PCR was performed using the SYBR green master mix (Roche), following the manufacturer's instructions, with the exception that the final volume was 12.5 ml of SYBR green reaction mix. Real-time monitoring of PCR amplification was performed using the LightCycler 480 detection system (Roche). Data were expressed as relative mRNA levels normalized to the β-actin expression level in each sample. The primer sequences can be obtained on request.

GEO accession number. The GEO number for the data sets in this manuscript is GSE65836. The GEO numbers for the data sets regarding Sox2 ChIP-Seq are GSM1050291, GSM1050292 and GSM1050293. The GEO number for the human Sirt6 ChIP-Seq data set is GSE32509.

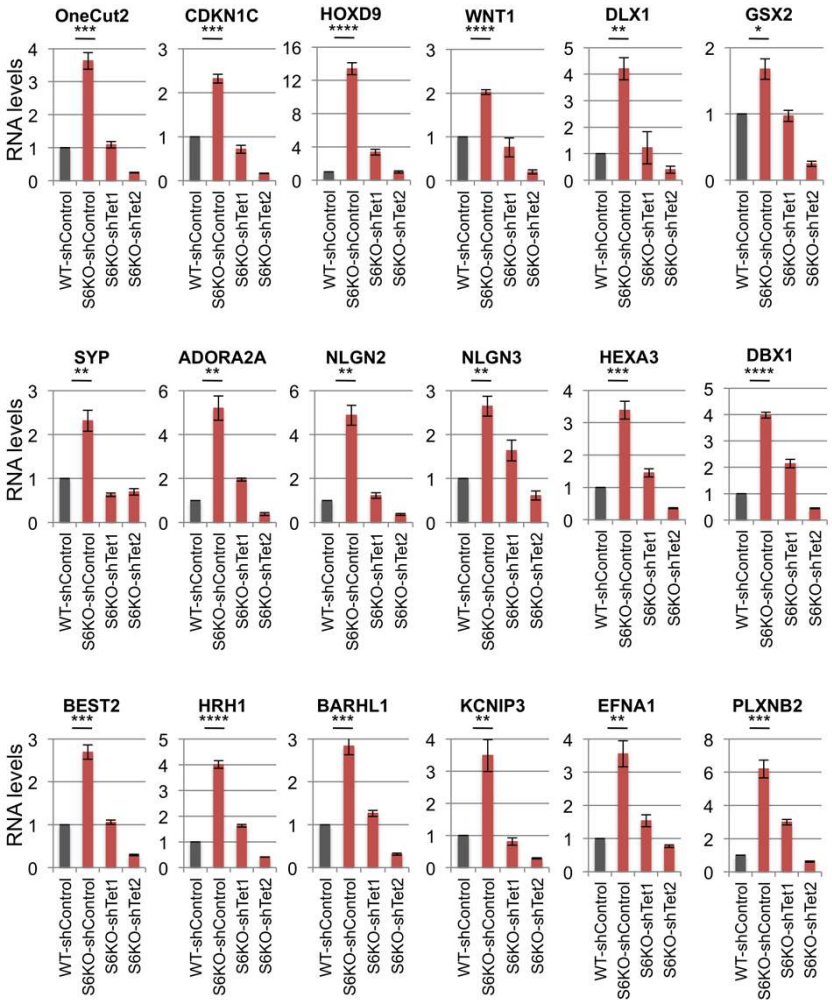
42. Gomes, I. C., Acquarone, M., Maciel Rde, M., Erlich, R. B. & Rehen, S. K. Analysis of pluripotent stem cells by using cryosections of embryoid bodies. *J. Vis. Exp.* **46,** http://dx.doi.org/10.3791/2344 (2010).
43. Gaspar-Maia, A. *et al.* Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* **460,** 863–868 (2013).
44. Shen, Y. *et al.* A map of the cis-regulatory sequence in the mouse genome. *Nature* **488,** 116–120 (2012).
45. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4,** 44–57 (2009).
46. Leinonen, R., Sugawara, H. & Shumway, M. The sequence read archive. *Nucleic Acids Res.* **39,** 19–21 (2011).
47. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memoryefficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
48. Li, H. *et al.* The sequence alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
49. Garber, M. *et al.* A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol. Cell* **47,** 810–822 (2012).
50. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).
51. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29,** 24–26 (2011).
52. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* **30,** 2114–2120 (2014).
53. Ying, Q-L., Stavridis, M., Griffiths, D., Li, M. & Smith, A. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotechnol.* **21,** 183–186 (2003).
54. Sommer, C. A. *et al.* Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* **27,** 543–549 (2009).
55. Gomes, P. N. *et al.* Gene-specific requirement for P-TEFb activity and RNA polymerase II phosphorylation within the p53 transcriptional program. *Genes Dev.* **20,** 601–612 (2006).
56. Stadtfeld, M. *et al.* Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **465,** 175–181 (2010).

DOI: 10.1038/ncb3147

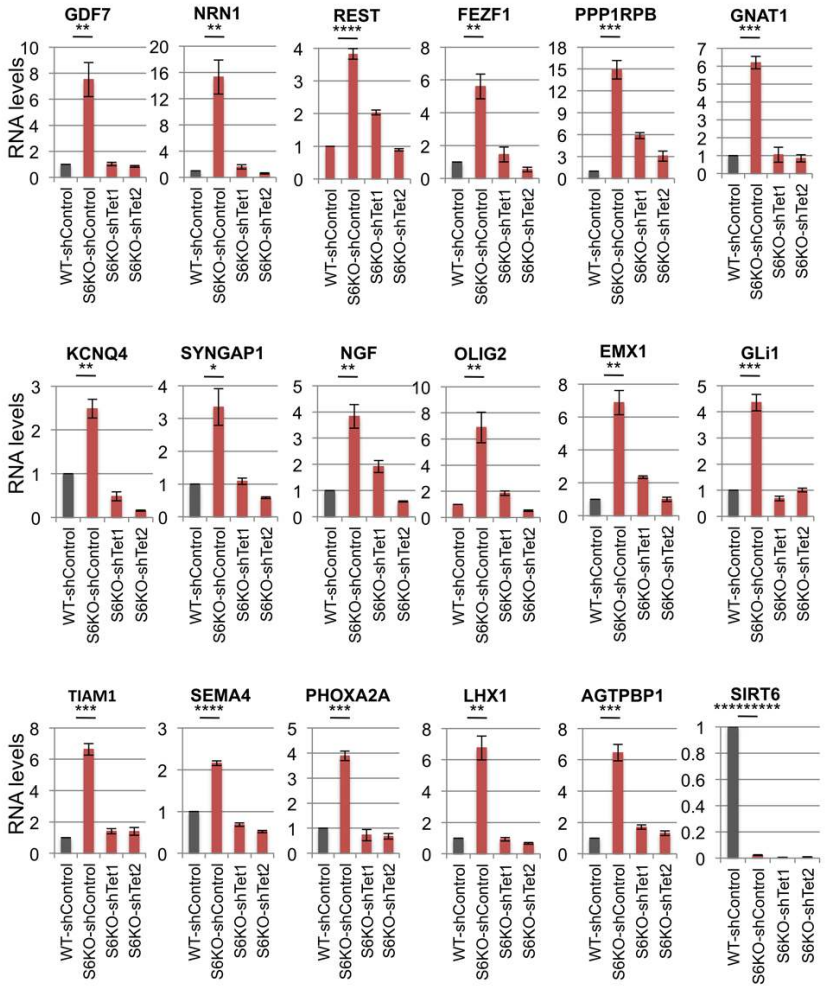
nature cell biology



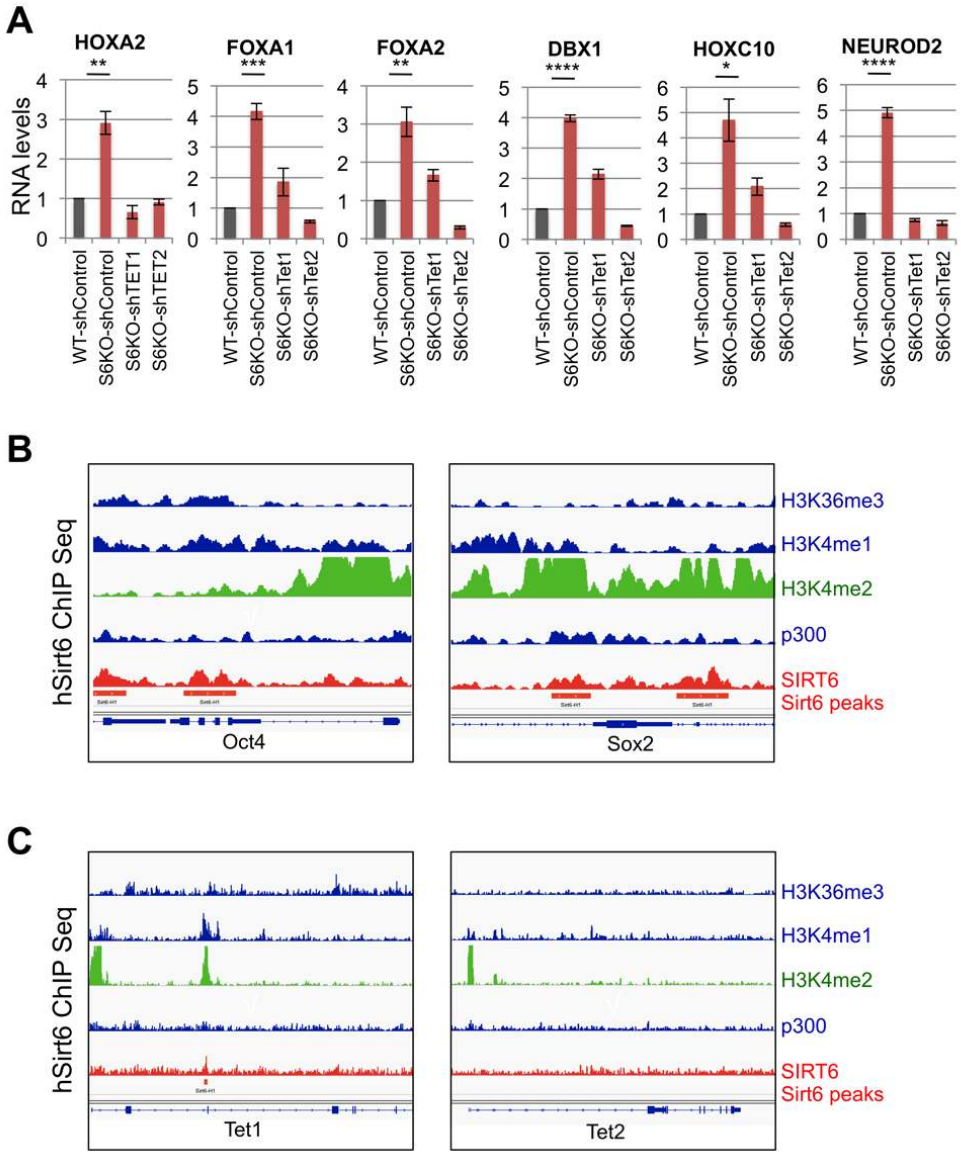
Supplementary Figure 1 Gene expression linked to the skewed phenotype towards neuroectoderm in S6KO EBs, Related to Figure 1. (A) Immunofluorescence of EBs derived from WT and S6KO iPSCs (129 mouse strain) for Gfap (green). Scale bar, 500 μm. (B) Gene expression of trophectoderm and neuroectoderm genes in WT *versus* S6KO EBs. qRT-PCR data is expressed relative to WT EBs. Data is expressed relative to WT values. The data are n = 3 experimental replicates (from independent RNA preparations). Values are mean +/- s.e.m. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{*****}P < 0.00001$, by *t*-test analysis. (C) Immunofluorescence of *in vitro* generated neurons from WT and S6KO EBs for β-III Tubulin (red). Undifferentiated neurons are visualized with dapi staining (blue). This is a representative of n = 3 experimental replicates (independent experiments). Scale bar, 100 mm.(D) Intensity of immunofluorescence of images from n = 3 experimental replicates (independent cell preparations) was analyzed by Image J. Data was normalized to WT values and represented as mean +/- s.e.m. $^{**}P < 0.01$, by *t*-test analysis. (E) Western blots showing expression of Nestin and Tet2 in WT *versus* S6KO EBs cultured under regular EB-medium. Expression of Sirt6 and its target H3K56ac are also shown. (F) Gene expression of trophectoderm genes in WT *versus* S6KO ESCs. qRT-PCR data is expressed relative to WT EBs. Data is expressed relative to WT values. The data are n = 3 experimental replicates (from independent RNA preparations). Values are mean +/- s.e.m. $^{***}P < 0.001$, $^{*****}P < 0.00001$, by *t*-test analysis. (G) Inability to silence expression of Sox2 and Nanog upon retinoic acid-induced differentiation in S6KO ESCs. Western blot analysis showing the expression of core pluripotent proteins in WT *versus* S6KO ESCs during a time-course retinoic acid-induced differentiation assay. Molecular weight markers are indicated. (H) Levels of H3K9ac and H3K56ac in S6KO *versus* WT ESCs after retinoic acid (RA)-mediated differentiation. ChIP-Seq binding profiles of the histone marks H3K56ac and H3K9ac on *Oct4*, *Sox2* and *Nanog* genes in WT *versus* S6KO ESCs after retinoic acid-mediated differentiation. Images were created with the Integrative Genomic Viewer (IGV)[53]. Data are normalized to total counts, and the scale range is 0.0 – 1.0. The data on panels (B) and (F) are n = 3 experimental replicates, values are mean +/- s.e.m. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{*****}P < 0.00001$ by *t*-test throughout the figure.



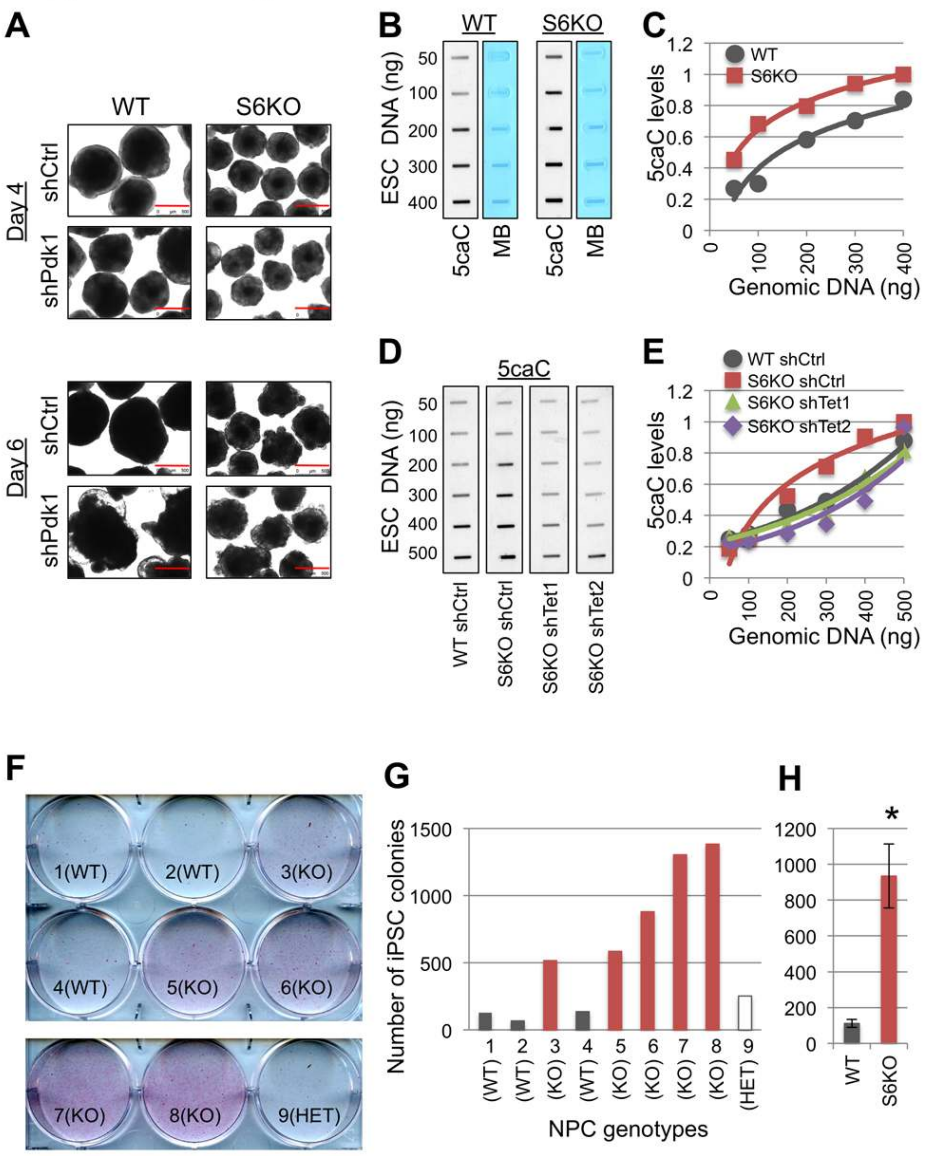
Supplementary Figures 2 and 3 The upregulated expression of neuronal differentiation related-genes exhibiting a 5hmC gain in S6KO over WT ESCs is rescued upon Tet knockdown, Related to Figures 5 and 6. Gene expression analysis by qRT-PCR is expressed relative to WT ESCs. Data are n = 3 experimental replicates (independent RNA preparations), values are mean +/- s.e.m. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$, $*********P < 0.00000001$, by *t*-test throughout the figure.



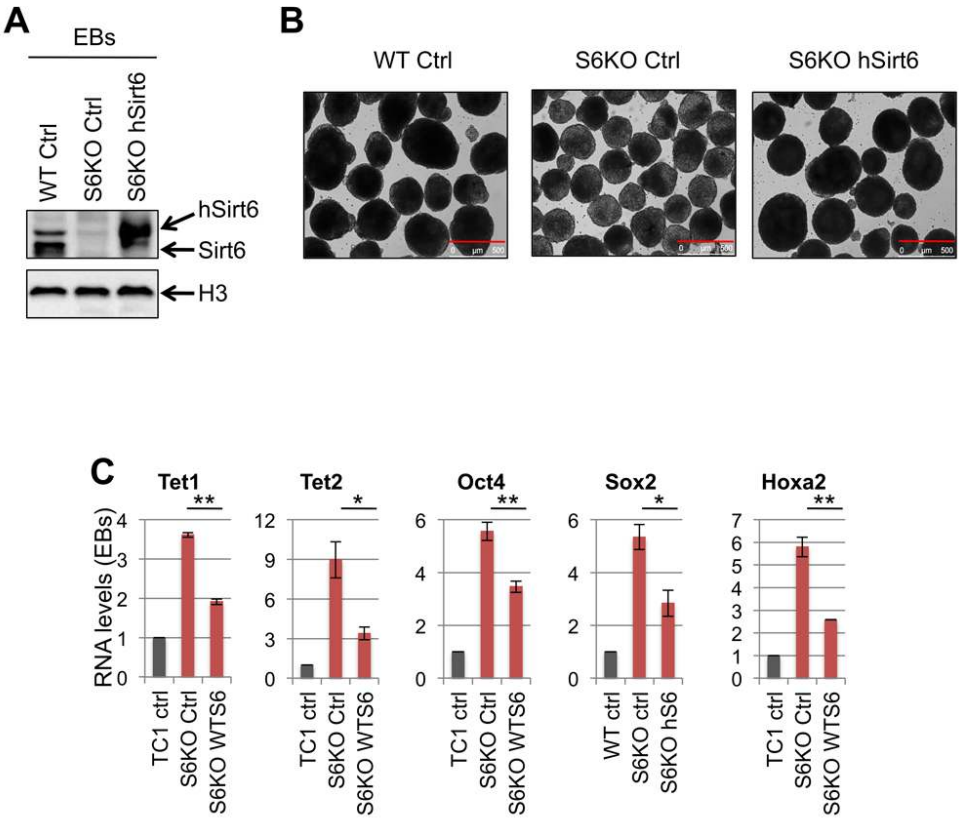
Supplementary Figures 3 The upregulated expression of neuronal differentiation related-genes exhibiting a 5hmC gain in S6KO over WT ESCs is rescued upon Tet knockdown, Related to Figures 5 and 6. Gene expression analysis by qRT-PCR is expressed relative to WT ESCs. Data are n = 3 experimental replicates (independent RNA preparations), values are mean +/- s.e.m. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$, $*********P < 0.00000001$, by *t*-test throughout the figure.



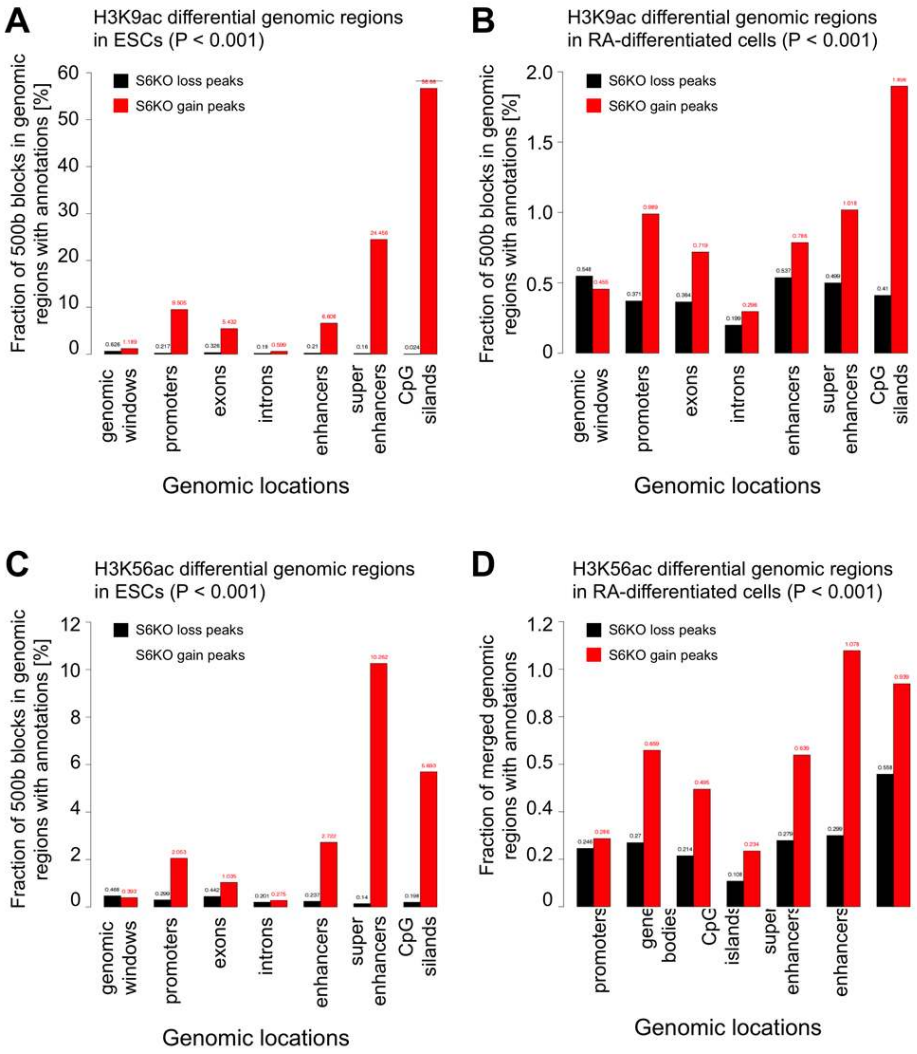
Supplementary Figure 4 (A) The upregulated expression of neuronal differentiation related-genes exhibiting a 5hmC gain in S6KO over WT ESCs is rescued upon Tet knockdown, Related to Figures 5 and 6. Gene expression analysis by qRT-PCR is expressed relative to WT ESCs. Data are n = 3 experimental replicates (independent RNA preparations), values are mean +/- s.e.m. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$, by *t*-test throughout the figure. (B) Sirt6 is recruited to the core pluripotent factors in human ESCs, Related to Figure 8. Genome wide Sirt6 ChIP-Seq data originated by Ram and colleagues[30] was used to determine recruitment of human Sirt6 (hSirt6) to *Oct4* and *Sox2*. (C) A small but significant peak for hSirt6 binding is apparent in *Tet1*, but not *Tet2* gene. Statistically enriched hSirt6 peaks are shown as red boxes. Histone marks (H3K4me1, H3K4me2 and H3K36me3) and p300 signals are also shown for reference. Supplementary Table S10 shows all the genomic regions targeted by hSirt6 analyzed using the same approach as in Ram and colleagues[30]. The accession number for these data can be found at: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE32509.



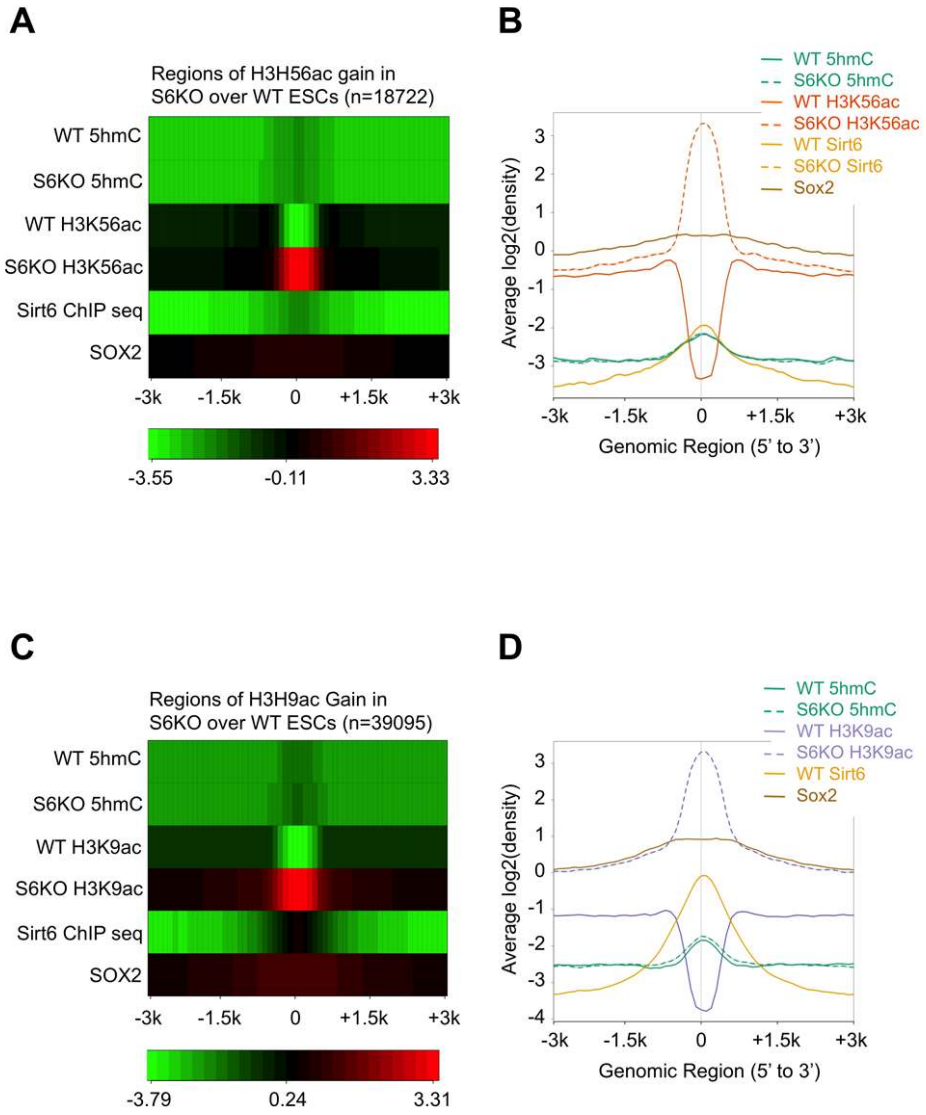
Supplementary Figure 5 (A) Pdk1 knockdown does not rescue the differentiation phenotype in S6KO EBs, Related to Figure 3. EBs derived from WT and S6KO ESCs (129 mouse strain). Scale bar, 500 μm. Pictures were taken at days 4 and 6 during EB formation. Data are representative of n = 3 experimental replicates. (B) Elevated levels of 5caC in S6KO *versus* WT ESCs. Related to Figure 4. Global 5caC levels assayed by slot blot analysis in ESCs. (C) Graph showing densitometric quantification of 5caC levels at each concentration of genomic DNA from panel (B). Data are representative of n = 2 experimental replicates. (D) Tet downregulation rescues high levels of 5caC in S6KO *versus* WT ESCs. Global 5caC levels assayed by slot blot analysis in Tet knockdown ESCs. (E) Graph showing densitometric quantification of 5caC levels at each concentration of genomic DNA from panel (D). Data are representative of n = 2 experimental replicates. (F) Increase efficiency of somatic cell reprogramming in S6KO *versus* WT NPCs. iPSC colony formation assay measured with alkaline phosphatase. (G) Graph showing iPSC colonies from each genotype. (H) Graph showing the average of WT (n = 3) *versus* S6KO (n = 5) NPCs reprogrammed into iPSCs. Colonies were quantified by image J, and values are mean +/- s.e.m. $*P < 0.05$, by *t*-test analysis.



Supplementary Figure 6 Ectopic expression of human Sirt6 rescues the differentiation phenotype of S6KO EBs. (A) Western blot analysis showing expression of endogenous Sirt6 and ectopic hSirt6. Total histone H3 is shown as loading control. (B) Embryoid bodies grown till day 10. Genotypes are indicated. Scale bar, 500 mm. (C) Gene expression analysis by qRT-PCR is expressed relative to WT ESCs. Data are n = 3 experimental replicates (independent RNA preparations), values are mean +/- s.e.m. $*P < 0.05$, $**P < 0.01$, by *t*-test analysis.

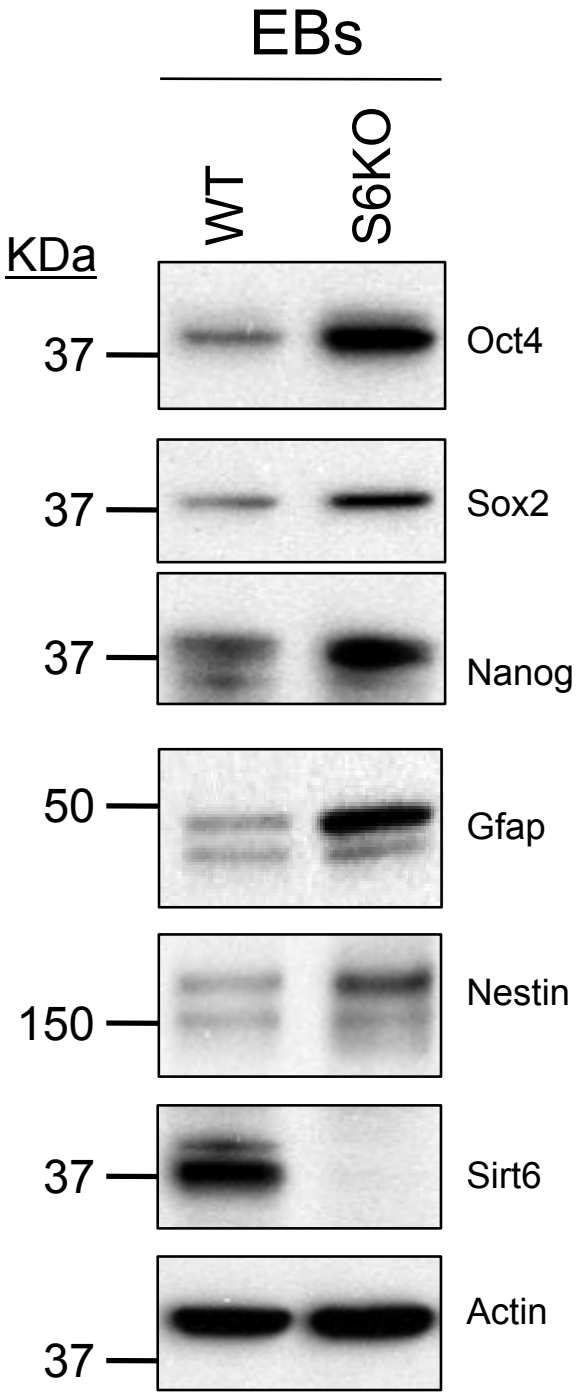


Supplementary Figure 7 Characterization of genomic regions with changes of H3K9ac and H3K56ac in S6KO *versus* WT ESCs, before and after retinoic acid (RA)-mediated differentiation. MEDIPS software was used to find regions with differential histones H3K9ac and H3K56ac in 500 base windows with p<0.001 and these regions with gains and losses upon Sirt6 KO were mapped to promoters, gene bodies, CpG islands, enhancers, and super-enhancers.(A) Genomic positions of H3K9ac gains and losses in mouse ESCs. (B) Genomic positions of H3K9ac gains and losses in RA-differentiated mouse ESCs. (C) Genomic positions of H3K56ac gains and losses in mouse ESCs. (D) Genomic positions of H3K56ac gains and losses in RA-differentiated mouse ESCs.



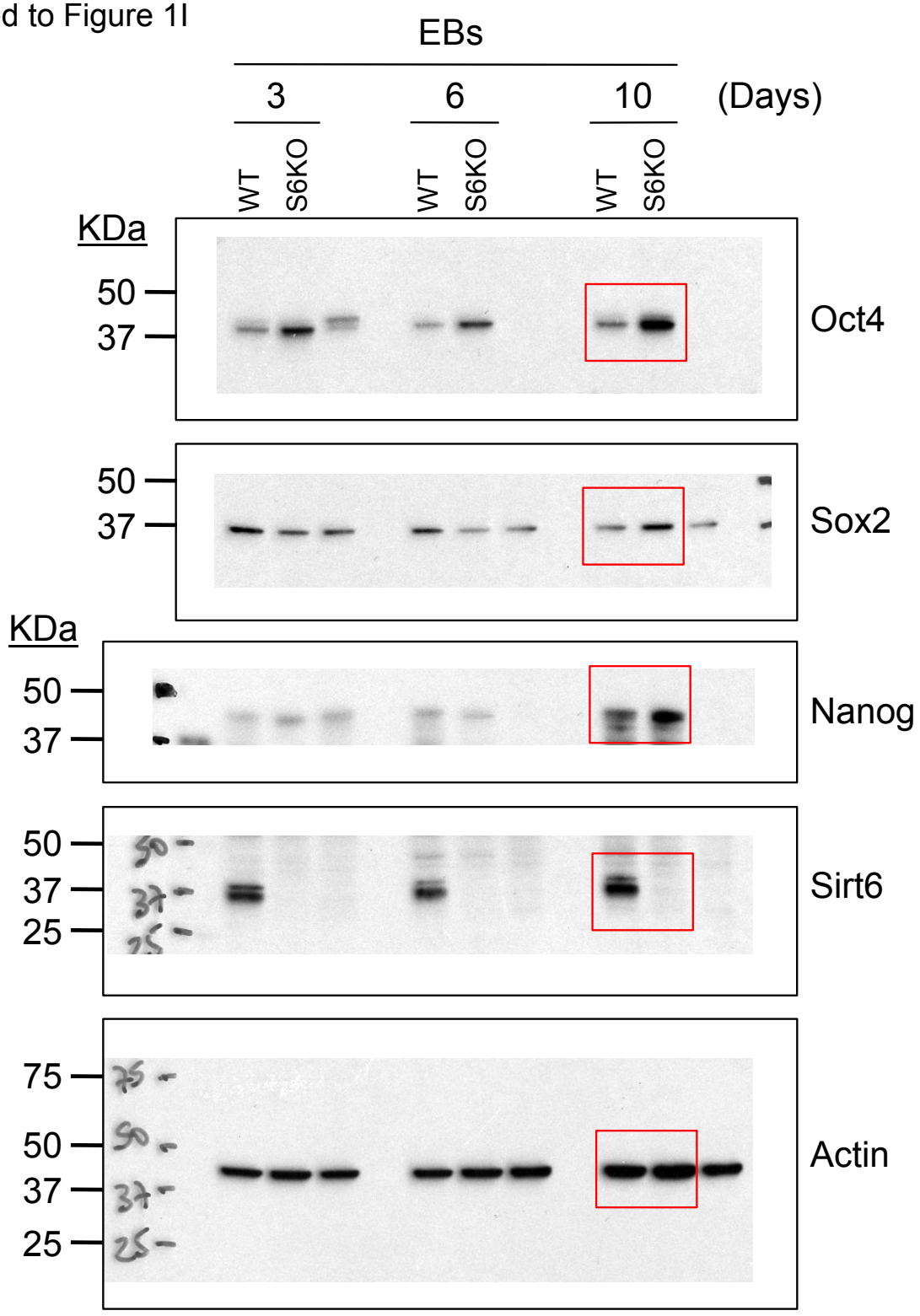
Supplementary Figure 8 Enrichment of 5hmC does not correlate with H3K9ac and/or H3K56ac in S6KO *versus* WT ESCs. (A) Heat map plot of regions of H3K56ac gains found with MEDIPS software in S6KO over WT in mouse ESCs show average profile for a +/-3kb band centered around the 18722 regions with H3K56ac gain ($p<0.001$) for the factors 5hmC, H3K56ac, H3K9ac, Sirt6, and Sox2 in WT and S6KO ESCs. (B) Enrichment line plot of average profile for regions of H3K56ac gain in S6KO over WT in mouse ESCs (n=18722) for the data on panel (A). Semi-transparent band behind line shows standard error of the mean for each average profile. (C) Heat map plot of regions of H3K9ac gains found with MEDIPS software in S6KO over WT mouse ESCs show average profile for a +/-3kb band centered around the 39095 regions with H3K9ac gain ($p<0.001$) for the factors 5hmC, H3K56ac, H3K9ac, Sirt6, and Sox2 in WT and S6KO ESCs. (D) Enrichment line plot of average profile for regions of H3K9ac gain in S6KO over WT in mouse ES cells (n=39095) for the data on panel (C). Semi-transparent band behind line shows standard error of the mean for each average profile.

Related to Figure 1I



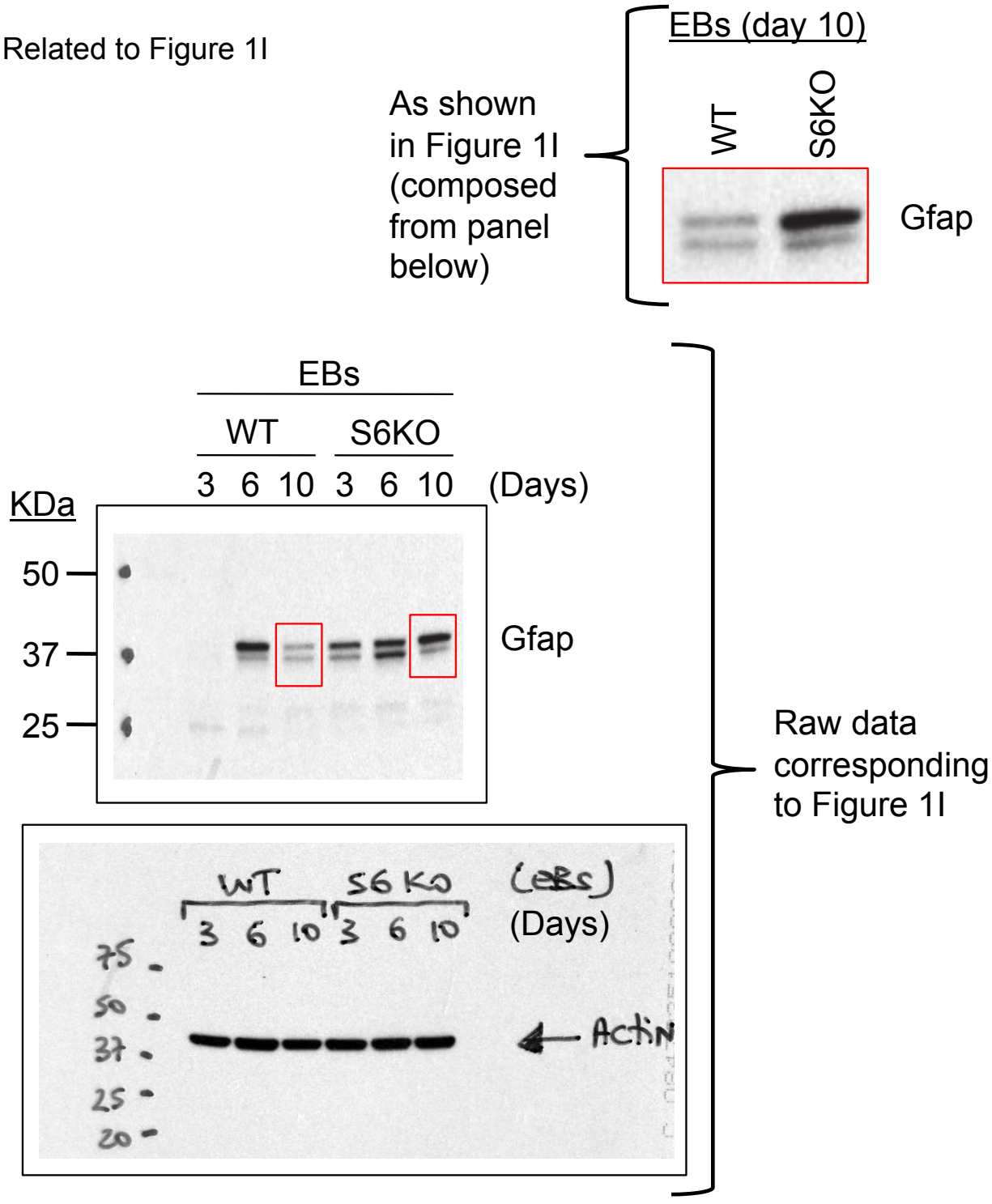
Supplementary Figure 9 This is a copy of Figure 1I, which is based on Supplementary Figures (B), (C) and (D). These are unprocessed scanned Western blot panels showing a time course (days 3, 6 and 10) differentiation of EB formation. Molecular weight (KDa) ladders are shown for each Western blot panel.

Related to Figure 1l



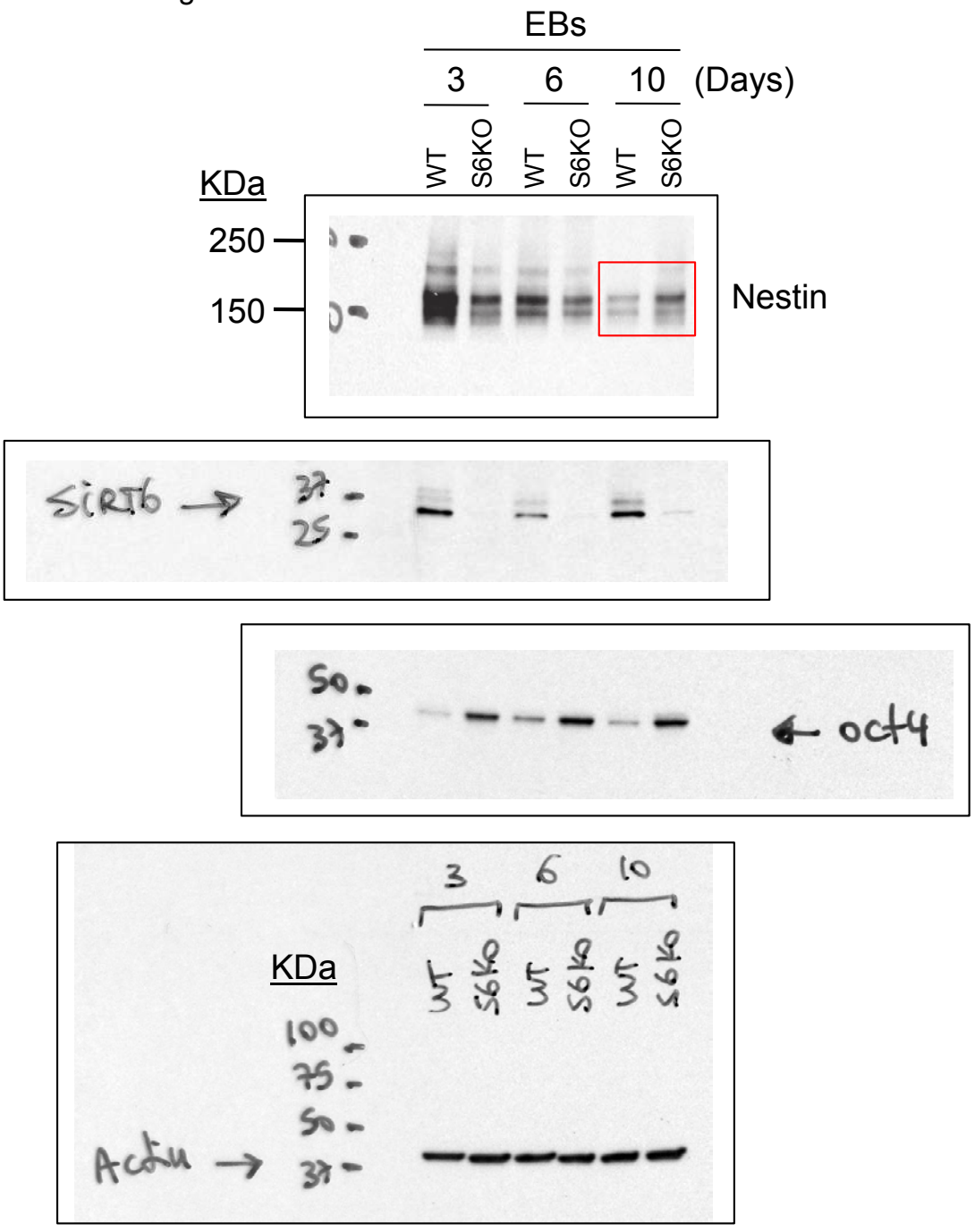
Supplementary Figure 9 continued The red boxes highlight the expression of Oct4, Sox2, Nanog, Sirt6 and Actin (as loading control), as shown in Figure 1l.

Related to Figure 1l

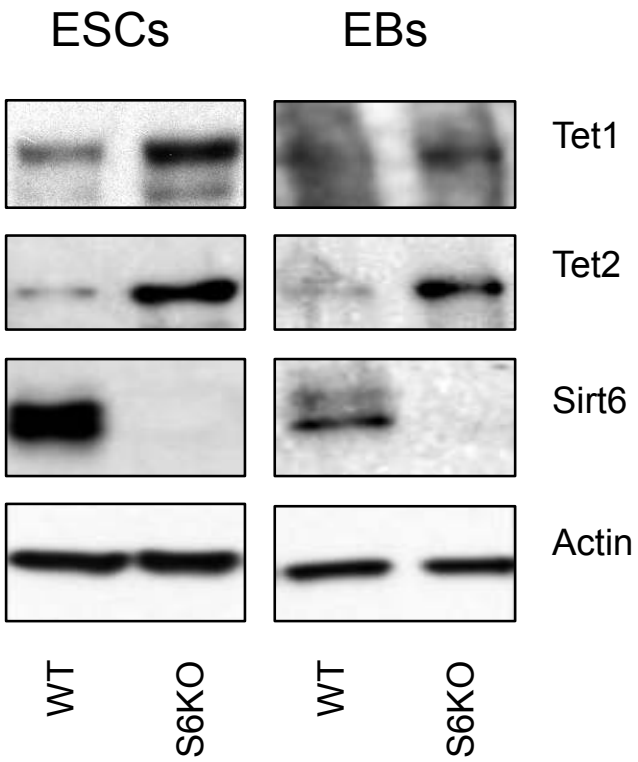


Supplementary Figure 9 continued The red boxes highlight the expression of Gfap as shown in Figure 1l. Expression of Actin is included as loading control.

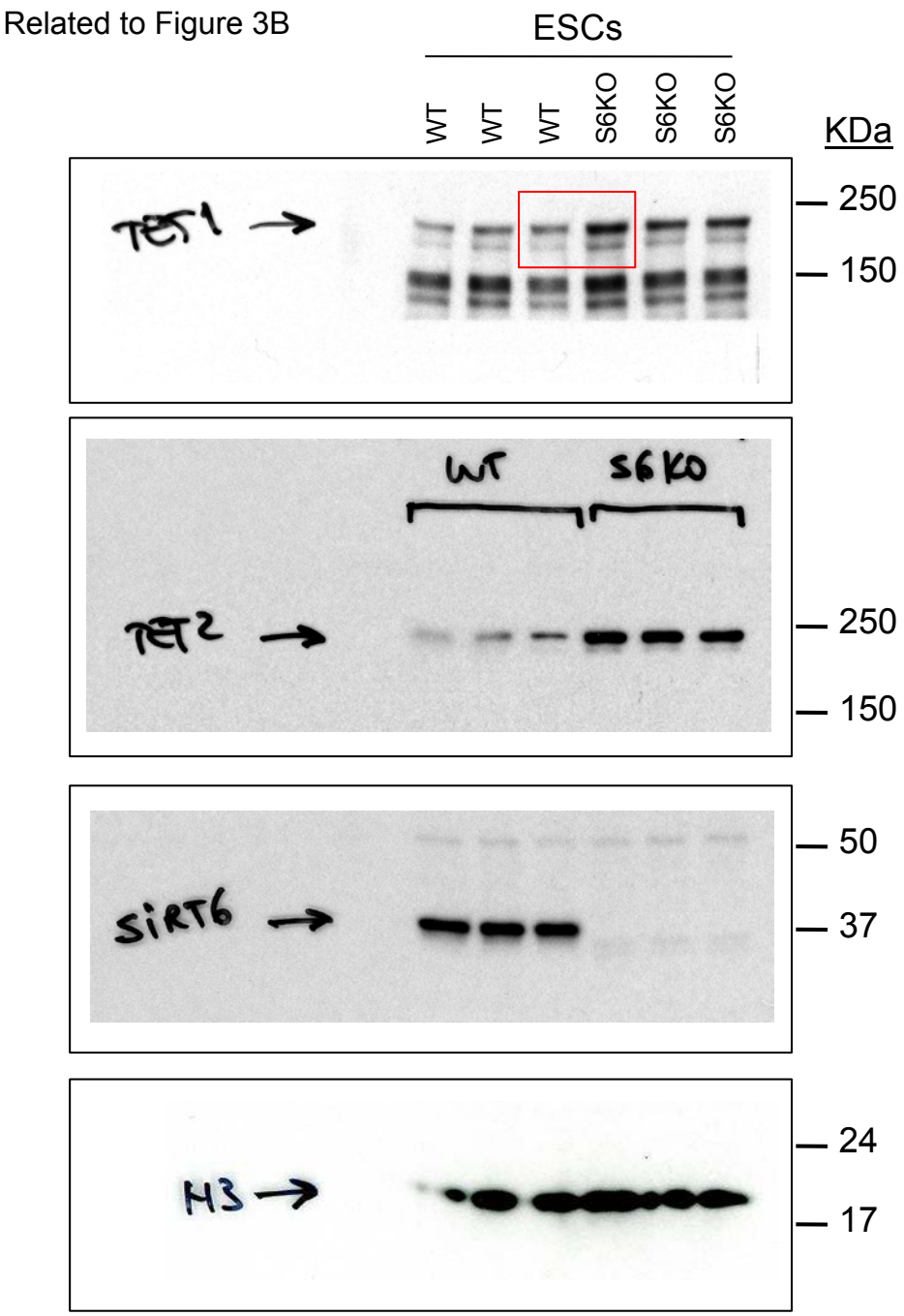
Related to Figure 1I



Supplementary Figure 9 continued The red box highlights the expression of Nestin as shown in Figure 1I. Expression of Sirt6, Oct4 and Actin (as loading control) are included, since these are distinct Western blots from panel (B).

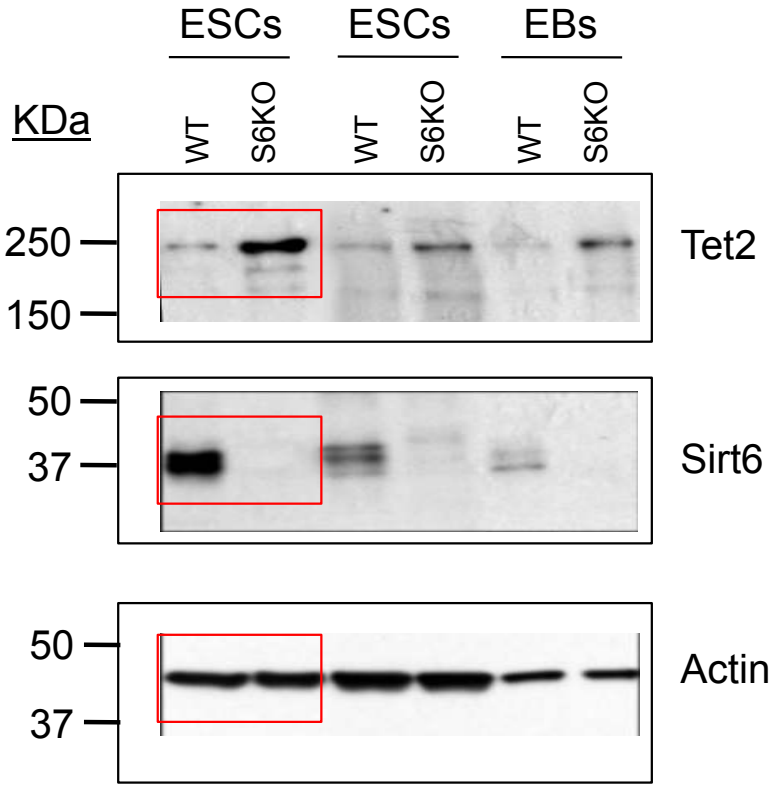


Supplementary Figure 9 continued This is a copy of Figure 3B, which is based on Supplementary Figures (F), (G) and (H). These are unprocessed scanned Western blot panels showing protein expression from ESCs and EBs along with molecular weight (KDa) ladders.

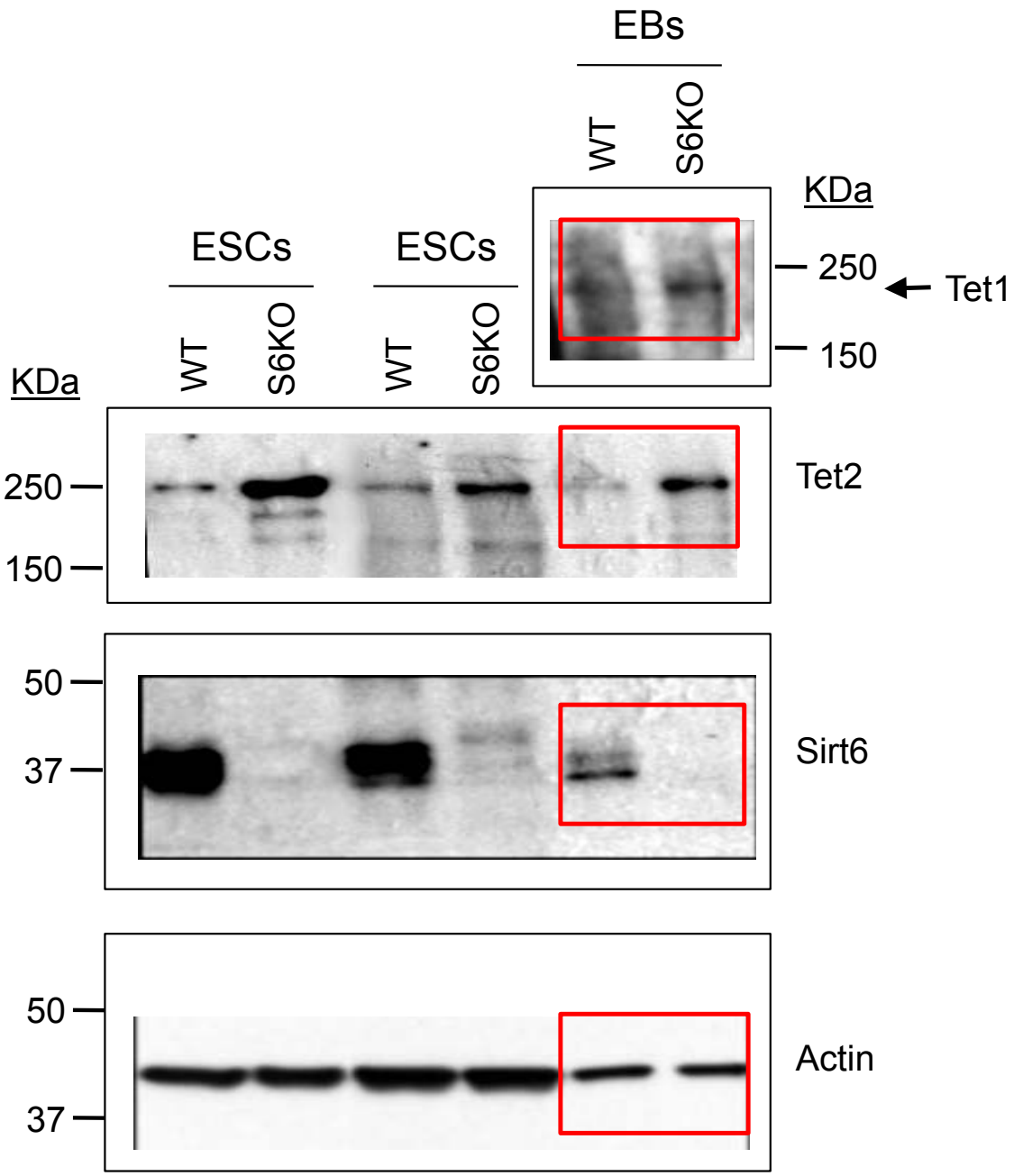


Supplementary Figure 9 continued The red box in the upper panel shows expression of Tet1 as shown in Figure 3B. Expression of Tet2 and Sirt6 are also included. Histone H3 was used as loading control.

Related to Figure 3B



Supplementary Figure 9 continued The red boxes highlight the expression of Tet2, Sirt6 and Actin in ESCs, as shown in Figure 3B.



Supplementary Figure 9 continued The red boxes highlight the expression of Tet1, Tet2, Sirt6 and Actin in EBs, as shown in Figure 3B.

Supplementary Table Legends

Supplementary Table 1 Identification of differentially hydroxymethylated regions (DHMRs).

Supplementary Table 2 Gene ontology annotation of DHMRs.

Supplementary Table 3 Enrichment of histone modifications and transcription factor binding.

Supplementary Table 4 Lists of genes with Sirt6 peaks associated with promoters (TSS +/- 1kb) and exons in mouse ESCs. List of mouse ESC gene markers with Sirt6 peaks associated with transcription start sites (TSS), exons and introns in mouse ESCs.

Supplementary Table 5 Gene lists with differential H3K56ac (gain or loss of H3K56ac in S6KO compared to WT) at promoters (TSS +/- 1kb) or exons in mouse ESCs.

Supplementary Table 6 Genes lists with differential H3K9ac (gain or loss of H3K9ac in S6KO compared to WT) at promoters (TSS +/- 1kb) or exons in untreated mouse ESCs. A second list of promoter gain with p<1e-7 and log fold change > 4 was included because DAVID analysis is limited to lists of 3000 or fewer genes.

Supplementary Table 7 Genes lists with differential H3K56ac (gain or loss of H3K56ac in S6KO compared to WT) at promoters (TSS +/- 1kb) or exons in mouse ESCs treated with differentiating agent, retinoic acid (RA).

Supplementary Table 8 Genes lists with differential H3K9ac (gain or loss of H3K9ac in S6KO compared to WT) at promoters (TSS +/- 1kb) or exons in mouse ESCs treated with the differentiating agent, retinoic acid (RA).

Supplementary Table 9 Lists of genes with Sox2 peaks associated with promoters (TSS +/- 1kb) and exons in mouse ESCs. Data is from Lodato and colleagues[29] using GEO sample GSM1050291 or SR1050291 along with corresponding whole cell extract (WCE) controls in standard ChIP-Seq pipeline with MAC2 for peak detection.

Supplementary Table 10 List of 2,604 genomic regions ("peaks") called for Sirt6 using the same approach employed in Ram and colleagues[30].

Supplementary Table S11 Mendelian frequencies calculated from n = 35 mouse litters (total of 223 born pups).