Stem Cell Reports
Article

Alternative SET/TAFI Promoters Regulate Embryonic Stem Cell Differentiation

Raghu Ram Edupuganti,[1,5,6] Arigela Harikumar,[1,5] Yair Aaronson,[1] Alva Biran,[1] Badi Sri Sailaja,[1] Malka Nissim-Rafinia,[1] Gajendra Kumar Azad,[1,4] Malkiel A. Cohen,[2] Jung Eun Park,[3] Chikdu S. Shivalila,[2] Styliani Markoulaki,[2] Siu Kwan Sze,[3] Rudolf Jaenisch,[2] and Eran Meshorer[1,4,*]

[1]The Department of Genetics, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Jerusalem 91904, Israel
[2]Whitehead Institute for Biomedical Research, Cambridge, MA, USA
[3]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore
[4]The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
[5]Co-first author
[6]Present address: Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Radboud University Nijmegen, 6525 Nijmegen, the Netherlands
*Correspondence: meshorer@huji.ac.il
http://dx.doi.org/10.1016/j.stemcr.2017.08.021

SUMMARY

Embryonic stem cells (ESCs) are regulated by pluripotency-related transcription factors in concert with chromatin regulators. To identify additional stem cell regulators, we screened a library of endogenously labeled fluorescent fusion proteins in mouse ESCs for fluorescence loss during differentiation. We identified SET, which displayed a rapid isoform shift during early differentiation from the predominant isoform in ESCs, SETα, to the primary isoform in differentiated cells, SETβ, through alternative promoters. SETα is selectively bound and regulated by pluripotency factors. SET depletion causes proliferation slowdown and perturbed neuronal differentiation *in vitro* and developmental arrest *in vivo*, and photobleaching methods demonstrate SET's role in maintaining a dynamic chromatin state in ESCs. This work identifies an important regulator of pluripotency and early differentiation, which is controlled by alternative promoter usage.

INTRODUCTION

Embryonic stem cells (ESCs), derived from the inner cell mass of preimplantation embryo, have the capability to give rise to all cell types of an adult organism (Evans and Kaufman, 1981). Although considerable attention has been devoted to the biology of ESCs, we are still far from understanding the complete underlying molecular mechanisms that govern pluripotency and lineage flexibility. To date, a core set of transcription factors (TFs) in concert with chromatin regulators has been identified, maintaining the "stem cell state" (Chambers and Tomlinson, 2009; Lessard and Crabtree, 2010; Loh et al., 2011).

Chromatin has been at the focal point in stem cell biology due to a variety of roles it plays in conferring and maintaining pluripotency (Fazzio et al., 2008; Gaspar-Maia et al., 2009; Lessard and Crabtree, 2010). Extensive modifications and rearrangements both at the global and local levels take place in chromatin structure during differentiation of ESCs (Meshorer and Misteli, 2006; Yamazaki et al., 2007), from a more dynamic, permissive structure in the pluripotent state to a restricted conformation following differentiation (Efroni et al., 2008). Some of the factors responsible for this hyperdynamic plasticity have recently begun to emerge, and include histone acetylation and methylation (Melcer et al., 2012), several histone modifiers and chromatin remodeling proteins (Cervoni et al.,

2002; Gaspar-Maia et al., 2009; Ho et al., 2009), as well as the nuclear lamina protein Lamin A (Melcer et al., 2012). Despite this extensive research, additional regulators of pluripotency are still being identified (Betschinger et al., 2013; Cheloufi et al., 2015; Ho et al., 2015; Respuela et al., 2016), and it is clear that additional factors await discovery.

To identify additional stem cell regulators, we generated a library of endogenously labeled fluorescent fusion proteins in R1 mouse ESCs (Harikumar et al., 2017 [this issue of *Stem Cell Reports*]). Here, we report that SET nuclear oncogene (SET, also known as TAF-I), a multifunctional linker histone chaperone, undergoes an isoform switch during early ESC differentiation via alternative promoters, and is involved in regulating pluripotency, proliferation, and differentiation of mouse ESCs.

RESULTS

Screening for Downregulated Proteins during ESC Differentiation

To identify potential regulators of pluripotency and ESC differentiation, we screened for proteins that are downregulated early upon retinoic acid (RA)-induced differentiation. Fluorescence levels, representing endogenous protein levels, were monitored using live time-lapse microscopy

over 4 days. Using this screen, we identified SET nuclear oncogene (SET), which was rapidly downregulated after the induction of differentiation (Figures 1A–1C and Movie S1). As expected from previous reports (Kato et al., 2011), SET was mainly found in the nucleus (Figures 1A and S1A–S1C). Interestingly, SET protein levels appear to reflect the differentiation state of ESCs in culture, with undifferentiated colonies showing high fluorescence and early differentiating colonies or colony edges showing reduced fluorescence (Figure S1B). These results suggest that SET is predominantly expressed in ESCs (Figures S1B–S1D) and is reduced during ESC differentiation, and that our N-terminal YFP tag of SET (Figure 1B) did not alter its nuclear localization.

Alternative Promoters Give Rise to Two Different SET Isoforms

The *Set* gene is located on chromosomes 2 and 9 in mouse and human, respectively. The *Set* gene is alternatively spliced, with four transcripts predicted to give rise to protein products of varying sizes (Figure S1E). SET is well conserved across species, with mouse and human SET proteins sharing 94% similarity. SET has two prominent isoforms, SETα and SETβ (Matsumoto et al., 1993; Nagata et al., 1995) (Figures 1E–1G and S1E–S1H). SET isoforms share most of the coding sequence except the first exons (Figure 1B), thus giving rise to almost identical proteins, which differ only at their N terminus (Figure 1D). SETα has a 36-amino-acid (aa) α-specific region and SETβ has a 24-aa β-specific region (Nagata et al., 1995). A dimerization domain and a highly acidic C-terminal domain, which is important for binding acetylated proteins such as p53 (Wang et al., 2016), follow these unique N-terminal isoform-specific regions (Figure 1D). SETβ is the most widely expressed isoform in differentiated cells, whereas SETα is expressed in a limited number of differentiated cell types and is usually expressed at considerably lower to non-existent levels compared with the β isoform (Nagata et al., 1998). In our YFP-SET clone, the YFP was integrated in intron 1, after the SETα-specific 5′ exon (Figure 1B), and therefore only SETα, and not SETβ, is tagged by YFP, conveniently allowing us to distinguish between the two isoforms. To measure the levels of the SETα- and SETβ-specific isoforms in ESCs, we performed qPCR analysis using primers specific to the unique 5′ exons. The expression level of SETα was considerably higher than that of SETβ in ESCs both at RNA and protein level (Figures S1H and 1G). Interestingly, SETα mRNA decreased rapidly during differentiation with a gradual concomitant rise in SETβ levels (Figures 1E and 1F), demonstrating an isoform switch at the transcriptional level. Western blots using anti-SET antibodies that recognize both isoforms show that the SETα decreased rapidly and SETβ levels increased moder-

ately during differentiation (Figure 1G). Finally, RNA sequencing (RNA-seq) tracks from ESCs and mouse embryonic fibroblasts (MEFs) from different sources (Shen et al., 2012; Yue et al., 2014) provided further support for the SETα/SETβ isoform switch between ESCs and MEFs at the RNA level (Figure 1H). Since we readily detected both YFP-tagged SET and native SET in the YFP-SET clone (Figure S1F), we conclude that SET expression is biallelic.

Core Pluripotency Factors Bind and Regulate SETα Expression

The unique elevated levels of SETα in ESCs and its abrupt decrease during differentiation called for testing its expression regulation. To this end, we used publicly available datasets (Chen et al., 2008; Marson et al., 2008) of epigenetic modifications and TF binding maps, as well as the BindDB webtool, recently developed by our group (Aaronson et al., 2016; Livyatan et al., 2015), enabling *in silico* reverse-chromatin immunoprecipitation (ChIP) analysis, to search for potential emerging features of the SET promoter(s). Our *in silico* analysis revealed that the upstream regions of both SETα and SETβ 5′ exons are enriched for H3K4me3 (Figure S2A, bottom), a mark of active transcription. This nicely depicts the existence and location of the alternative SET promoters. As expected from an active gene, we did not find any enrichment for H3K27me3 in these promoter regions (Figure S2A). Analyzing the binding of TFs to SET promoters, we found that at least nine TFs, many of which are ESC specific, bind the SETα, but not the SETβ, promoter (Figure 2A). BindDB analysis revealed that none of the pluripotency factors bind the SETβ promoter, which was instead bound by factors such as TOP-OIIa, TET1, OGT, HDAC2, FBXL10, CAPG, and CTR9, suggesting a more poised state (Figure 2A). The binding of SETα promoter by OCT4, SOX2, KLF4, and NANOG was confirmed by ChIP-qPCR (Figure 2B). These data suggest that the two isoforms are distinctly regulated and that the SETα promoter alone is bound by the pluripotency network TFs.

OCT4 Regulates SETα Expression in ESCs

Next, we wished to test the functional significance of the binding of the pluripotency factors to the SETα promoter. To determine the effect of OCT4 on SETα expression, we took advantage of the Zhbtc4 ESC system, in which endogenous *Oct4* gene is under the control of doxycycline (Dox) (Niwa et al., 2000). Addition of Dox completely abolishes OCT4 expression (Figures 2C and 2F). While OCT4-depleted cells began to differentiate after 48 hr, 24 hr after Dox addition the Zhbtc4 ESC colonies remained undifferentiated and appeared similar to the control colonies in morphology (Figure S2B) and Nanog expression
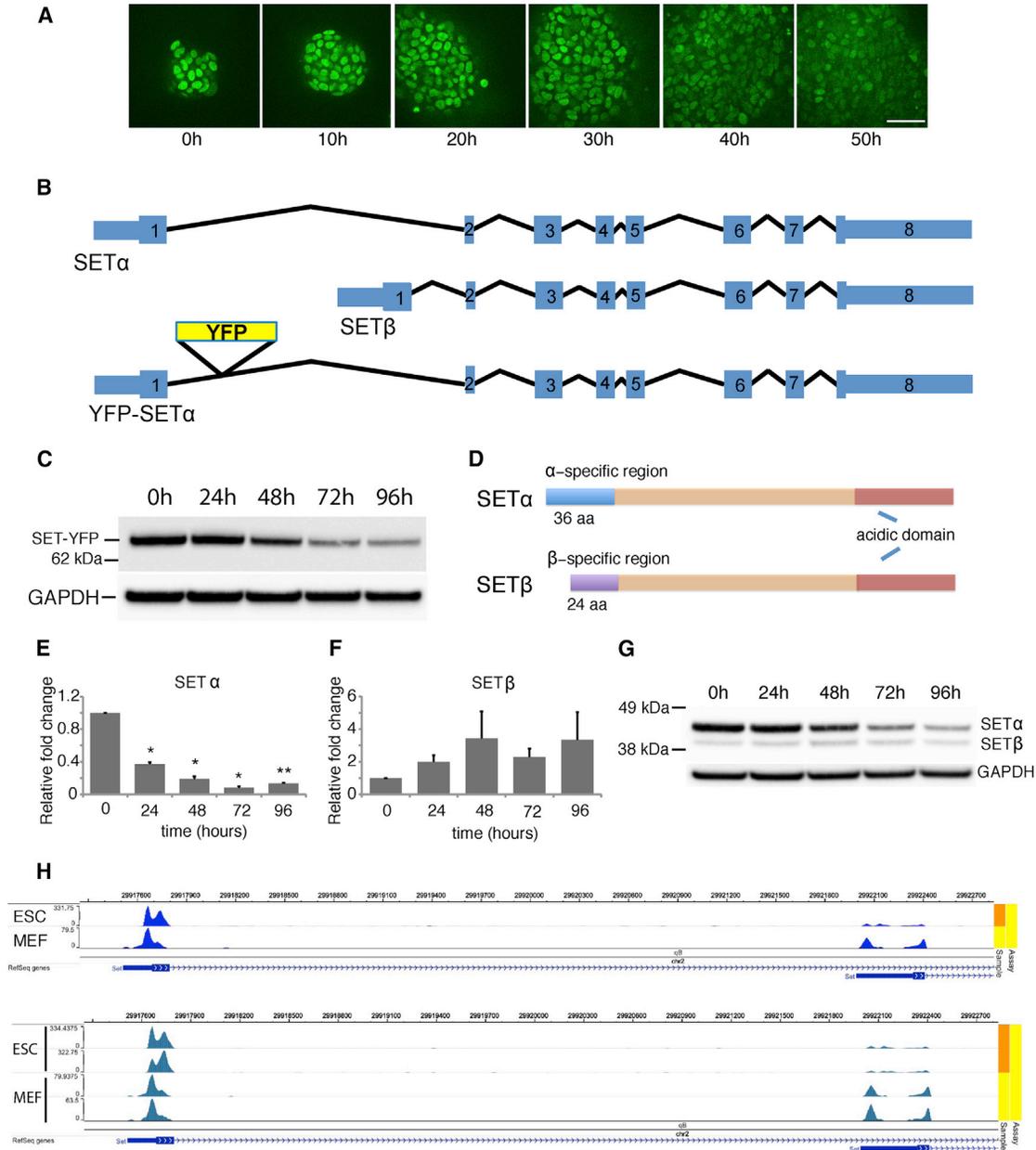
Figure 1. A SET Isoform Switch during Early ESC Differentiation

(A) Time-lapse images of SETα-YFP cells during the first 50 hr of RA-induced ESC differentiation. Scale bar, 50 μm.

(B) Schematic showing the gene structure of the two SET isoforms and YFP-SETα.

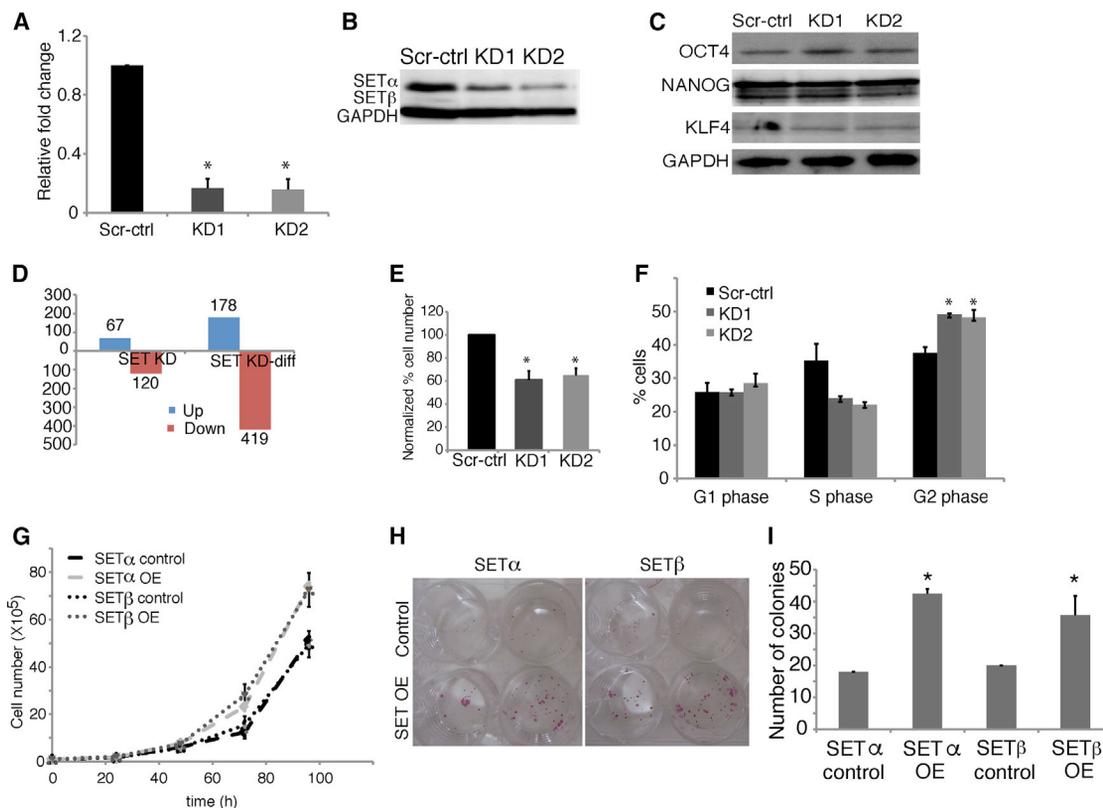(C) Anti-GFP western blots of SET-YFP during ESC differentiation. GAPDH was used as control.

(D) Protein domain model depicting the SETα and SETβ isoforms. Shown are SETα-specific region (blue), SETβ-specific region (violet), dimerization domain (orange), and acidic domain (red).

(E and F) qRT-PCR analysis of SETα (E) (*p < 0.001, **p ≤ 1.6 × 10^{-5}, 2-tailed Student's t test) and SETβ (F) mRNA level during ESC differentiation (Data are shown as mean ± SD; n = 3 independent experiments).

(G) Anti-SET western blots showing SETα and SETβ isoforms during ESC differentiation. GAPDH was used as loading control. Differentiation was carried out in ESC medium with RA (1 μM) and without LIF on gelatin-coated plates.

(H) RNA-seq tracks showing evidence for SETα and SETβ expression in ESCs and MEFs.

Figure 2. SETα Expression Is Regulated by Pluripotency Factors in ESCs

(A) *In silico* reverse-ChIP analysis for SET. For every query, our BindDB webtool and pipeline returns the TFs and chromatin modifications that are enriched at the promoter region (proximal or distal). Dark green, binding; light green, no binding. Green, blue, and red boxes depict TFs or histone marks associated with active chromatin, pluripotency, or poised chromatin, respectively.

(B) ChIP analysis of pluripotency factors at the SETα promoter (left). SETα gene body corresponding to a region in intron 7 was used as a negative control (right) (data are shown as mean ± SD; n = 3 independent experiments).

(C) qRT-PCR analysis of *Oct4* and *Nanog* mRNA levels in Zhbtc4 cells before and after Dox treatment. Expression levels were normalized to GAPDH (data are shown as mean ± SD; n = 3 independent experiments).

(D and E) qRT-PCR analysis of SETα (D) and SETβ (E) mRNA levels in Zhbtc4 cells before and after Dox treatment. Expression levels were normalized to GAPDH expression (data are shown as mean ± SD; n = 3 independent experiments).

(F) Western blots of SETα, SETβ, and OCT4 in Zhbtc4 cells before and after Dox treatment. GAPDH was used as a loading control.

(G) Binding profiles of RNAPII at SETα and SETβ promoters before (0 hr) and after (24 hr) Dox addition. Data are from Rahl et al. (2010).

(Figure 2C). Importantly, SETα, but not SETβ, was selectively decreased (by ∼50%) in the Dox-treated cells after 24 hr (Figures 2D and 2E), suggesting that SETα, but not SETβ, is controlled by OCT4. These results were confirmed at the protein level (Figure 2F), although due to the higher stability of the protein the effect is more subtle. Next, we examined the previously published RNA polymerase II (RNAPII) occupancy data in the Zhbtc4 cells after OCT4 depletion (Rahl et al., 2010), and found a selective decrease at the SETα promoter 24 hr after Dox addition (Figure 2G).

Figure 3. SET Regulates ESC Proliferation

(A) qRT-PCR analysis of SET knockdown efficiency (n = 3 independent biological experiments; *p < 0.005, 2-tailed t test). Expression levels were normalized to GAPDH. Error bars represent ±SD.

(B) Western blots of SET in control (Scr-ctrl) and SET-KD cells. GAPDH was used as control.

(C) Western blots of pluripotency factors in SET-KD clones. GAPDH was used as control.

(D) The number of upregulated (blue) and downregulated (red) genes in undifferentiated SET-KD ESCs (left) and in RA-induced SET-KD cells (right).

(E) Reduced proliferation in SET-KD ESCs (*p < 0.05, 2-tailed t test). Cell number of SET-depleted ESCs was normalized to Scr-ctrl cell number. Error bars represent ±SD (n = 3 independent biological experiments).

(F) Cell-cycle analysis of SET-KD cells. Cell number in each phase was normalized to that of Scr-ctrl (n = 3 independent experiments; error bars represent ±SD; *p < 0.05, 2-tailed t test).

(G) KH2-ESC proliferation assay following overexpression (OE) of SETα or SETβ (n = 3 independent experiments; error bars represent ±SD).

(H) Representative images of clone formation assay of control (top), SETα-OE (bottom left), and SETβ-OE (bottom right) wells. Colonies were detected with AP staining (pink).

(I) Quantification of (H) (n = 3 independent experiments; error bars represent ±SD; *p < 0.05, 2-tailed t test).

Taken together, these data demonstrate a highly selective promoter-specific regulation of SETα, but not SETβ, by OCT4 and likely other pluripotency factors (Figure S2C).

SET Regulates Cell Proliferation and Survival of ESCs

To explore the role of SET in ESCs, we generated stable knockdown (KD) clones of total SET, and used the CRISPR/Cas9 system to selectively disrupt either SETα, SETβ, or both (SET-DKO) in mouse ESCs (Figures S3A–S3C, 3A, and 3B). All KD and knockout (KO) clones were verified by qPCR and western blotting, and KO clones were further validated by sequencing. Interestingly, SETα KO ESCs showed increased levels of SETβ transcripts (Figure S3C). The SET-DKO clones grew slowly and formed small colonies. We therefore reverted to analyzing the stable SET-KD clones, where most, but not all, of the protein is depleted (Figures 3A and 3B). The morphology of the SET-KD clones remained unaltered when grown on MEFs (Figure S3D), as well as the expression level of OCT4, NANOG, and KLF4 (Figures S3E and 3C). However, there was a slight increase in the expression level of differentiation markers of all three germ lineages (Figure S3F).
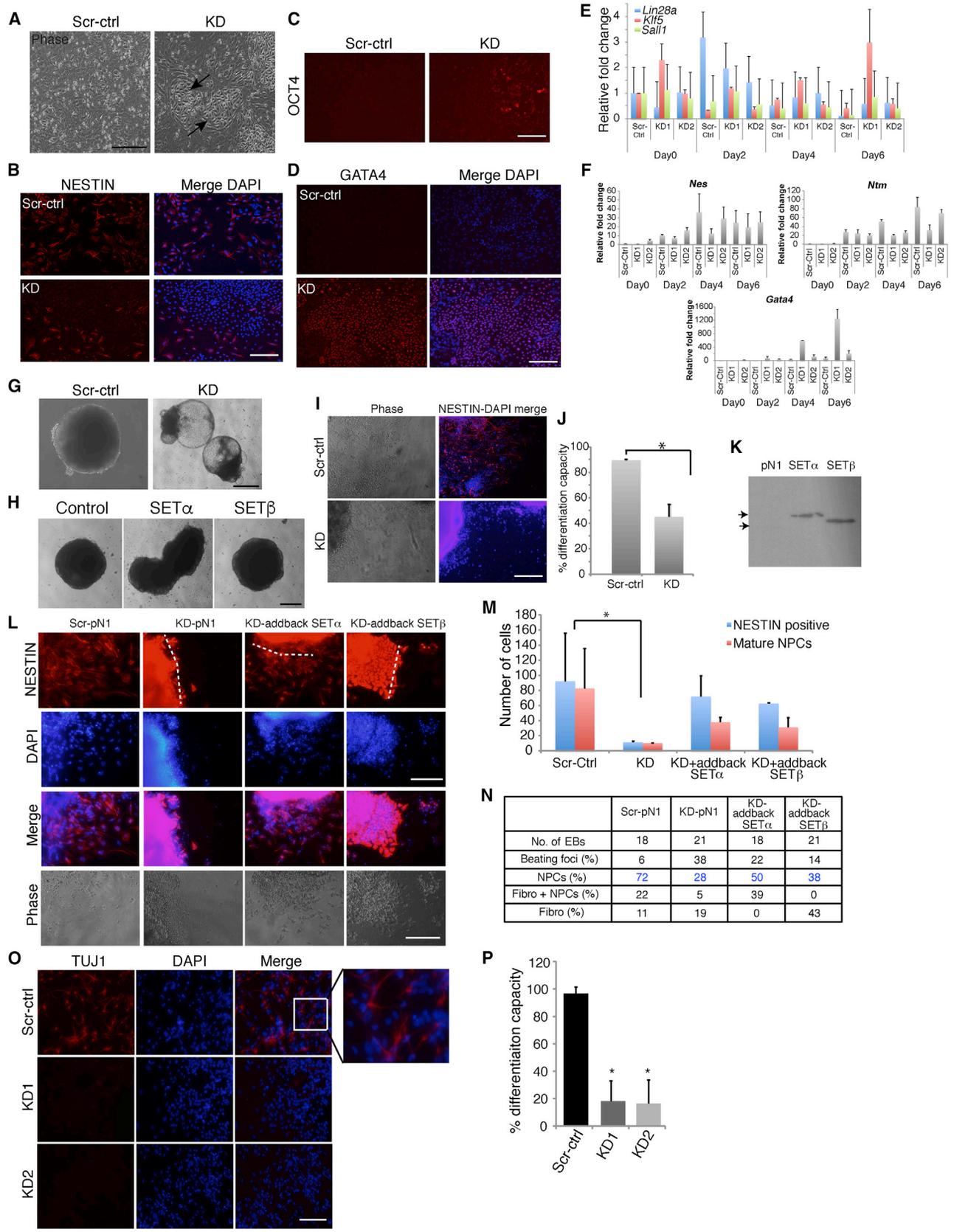
(legend on next page)

Expression microarrays confirmed a relatively minor effect on gene expression with 67 and 120 genes reproducibly upregulated and downregulated (1.5-fold cutoff), respectively, in the undifferentiated SET-KD clones (Figures 3D and S3G–H; Table S1). A closer look at the downregulated genes revealed a class of genes that are specifically involved in the G_2/M phase of the cell cycle, of which *Cdc16* and *Anapc4* govern the exit from mitosis. To test the role of SET on cell-cycle kinetics, we performed proliferation assays in SET-KD and SET-overexpressing (OE) cells. After 96 hr, we observed a ~40% decrease in cell numbers in the SET-KD clones compared with controls (Figure 3E), without any apparent influence on cell death. Fluorescence-activated cell sorting-based cell-cycle analysis using propidium iodide revealed a shift from S to G_2/M phase in the SET-KD population (Figure 3F), suggesting that SET depletion slows down cells in the G_2/M phase. To test the effects of SET overexpression on proliferation, we used KH2 ESCs (Supplemental Experimental Procedures) to individually overexpress either YFP-SETα, hemagglutinin (HA)-SETα, or HA-SETβ in ESCs (Figures S3I–S3K). We found that SET overexpression increased ESC proliferation rates by ~1.4-fold (Figure 3G).

We next tested the effect of the SET isoforms on imparting survival advantage to ESCs. We performed a clone-formation assay of SETα- or SETβ-OE KH2 cells with and without Dox, and stained the cells for alkaline phosphatase

(AP) 6 days later. SETα and SETβ overexpression increased colony numbers by ~2.3-fold and ~1.7-fold, respectively (Figures 3H and 3I), indicating that SET can impart survival advantage, consistent with its effect on proliferation. The fact that both SETα and SETβ had similar effects suggests that the N-terminal region of SET is not involved in these processes. These results are also in line with observations in cancer cell lines, where SET isoforms are greatly overexpressed, increasing the cancer cells' survival rate and proliferative capacity (Carlson et al., 1998; Fukukawa et al., 2000).

SET Regulates Neuroectodermal Differentiation of ESCs

To investigate the effect of SET-KD on ESC differentiation, we induced control (Scr-ctrl) and SET-KD cells to differentiate with RA (Figure 4A). By day 4, while Scr-ctrl cells differentiated normally into NESTIN-positive cells as expected (Figure 4B, top), the SET-KD clones formed Nestin-negative circular colonies of small round cells, surrounded by NESTIN-positive cells (Figure 4B, bottom). The cells within the circular colonies were negative for the pluripotency marker OCT4 (Figure 4C), suggesting that they did not remain undifferentiated. Testing different lineage markers we found that the OCT4-negative circular colonies expressed the early endodermal marker GATA4 (Figure 4D). The Scr-ctrl cells were all negative for GATA4, as expected

Figure 4. SET Regulates ESC Neuronal Differentiation
(A) Phase contrast images of Scr-ctrl cells (left) and SET-KD cells (right). SET depletion results in circular colonies and aberrant neuronal differentiation (arrows). Scale bars, 200 μm.
(B) Immunofluorescence (IF) for NESTIN (red) and RA-differentiated Scr-ctrl (top) and SET-KD cells (bottom). Nuclei were counterstained with DAPI (blue, right). Note the characteristic circular structures in the KD clones. Scale bar, 200 μm.
(C) IF for OCT4 (red) in RA-differentiated Scr-ctrl (left) and SET-KD clones (right). Scale bar, 200 μm.
(D) IF for GATA4 (red) in RA-differentiated Scr-ctrl (top) and SET-KD ESCs (bottom). Nuclei were counterstained with DAPI (blue, right). Scale bar, 200 μm.
(E and F) Time-course qRT-PCR analysis of pluripotency (E) and differentiation (F) factor levels in Scr-Ctrl and KD clones during RA-induced differentiation. Expression levels were normalized to GAPDH (n = 3 independent biological experiments; error bars represent ±SEM).
(G) EBs derived from Scr-ctrl cells (left) and SET-KD clones (right). Note formation of cystic EBs in the SET-KD clones. Scale bar, 500 μm.
(H) Normal EB formation in SETα- and SETβ-overexpressing clones. Scale bar, 500 μm.
(I) IF for NESTIN (red) in NPCs derived from Scr-ctrl (top) and SET-KD clones (bottom) without addback vectors. Phase-contrast images are shown on the left. Scale bar, 200 μm.
(J) Quantification of (I) (n = 3 independent experiments; error bars represent ±SD; *p < 0.02, 2-tailed Student's t test).
(K) Western blots using anti-HA antibodies showing HA-SETα and HA-SETβ overexpression in addback clones. Arrowheads indicate SET isoform positions.
(L) NESTIN immunostaining (red) in NPCs derived from Scr-ctrl cells (Scr-pN1, left), SET-KD cells (second from left), and in SET-KD-addback SETα (second from right) or SETβ (right) clones. Empty vectors (pN1) were used as controls. Nuclei were counterstained with DAPI (blue). Phase contrast is shown at the bottom. Dotted white lines represent the border between EBs and differentiating NPCs. Scale bars, 100 μm.
(M) Quantification of the addback experiments (n = 3 independent experiments; error bars represent ±SD; *p < 0.05, 2-tailed t test).
(N) The relative abundance (in percentage) of the different cell types observed during NPC differentiation of Scr-ctrl, SET-KD and the various addback clones.
(O) IF for TUJ1 (red) in differentiated neurons derived from Scr-ctrl (top) and SET-KD ESCs (middle and bottom). Nuclei were counterstained with DAPI (blue). Scale bar, 100 μm.
(P) Quantification of neuronal differentiation capacity in Scr-ctrl and KD clones (n = 3 independent experiments; error bars represent ±SD; *p < 0.05, 2-tailed Student's t test).

(Figure 4D, top). To rule out the possibility that the formation of the circular colonies was due to the more slowly differentiating SET-KD ESCs, we maintained these cultures in RA-containing differentiation medium for 8 days. The cells did not show any progression in differentiation and slowly died (data not shown). These data indicate that the SET-KD clones are defective in ectodermal differentiation. To rule out the possibility that the expression changes we observed were due to different readout of cells at different stages of differentiation, we performed a time-course differentiation experiment. The SET-KD cells were differentiated for 6 days and the expression of pluripotency and neuronal differentiation markers was quantified by qPCR. Here also we observed a similar general trend in marker expression in the KD clones, i.e., increased pluripotency factor expression (Figure 4E) and increased endodermal gene expression (Figure 4F, bottom). These results suggest that the effects of SET on differentiation are not due to different readouts of cells at different stages.

We next used non-directed differentiation into embryoid bodies (EBs). By day 5, the SET-KD EBs gave rise to small fluid-filled EBs, gradually growing into large EBs by day 8 (Figure 4G). Interestingly, many of the fluid filled EBs formed from the SET-KD clones gave rise to spontaneously beating cellular aggregates in suspension (Movie S2), suggesting cardiomyocyte differentiation. Control EBs appeared normal and did not produce any beating EBs. Overexpression of either SETα or SETβ did not alter EB formation and differentiation (Figure 4H), suggesting that SET is important, but not limiting, during this process.

Since RA differentiation of SET-KD cells yielded considerably fewer NESTIN-positive cells, we further tested the potential of the SET-KD cells to generate neuronal progenitor cells (NPCs) (Lee et al., 2000). The SET-KD cells gave rise to a significantly lower number of NPCs (by >50%, p < 0.05) compared with controls (Figures 4I and 4J). To rule out unspecific clonal or knockdown artifacts, we stably reintroduced SETα or SETβ or a control pN1 vector into the SET-KD clones (Figure 4K) and repeated the neuronal differentiation. All clones formed normal EBs, but when differentiated into NPCs, SETα addback only partially rescued the differentiation phenotype (~45%) (Figures 4L and 4M), while SETβ addback gave rise to fibroblast-like cells which, curiously, stained positive for NESTIN (~40%). We also observed, once again, increased tendency to spontaneously form beating foci only in the SET-KD clones (Figure 4N). As expected, pN1 addback was similar to that of SET-KD clones, failing to generate NPCs (~10%), while the Scr-ctrl differentiated normally. These data indicate that SET plays an important function in neurogenesis.

We also tested whether the few emerging SET-KD NPCs could give rise to mature neurons (Efroni et al., 2008; Lee et al., 2000). The Scr-ctrl NPCs differentiated into Tuj1-positive neurons efficiently and formed networks (Figure 4O, top). In contrast, SET-KD NPCs formed very few Tuj1-positive neurons and grew into fibroblast-like cells (Figures 4O and 4P). Unexpectedly, the SET-KD clones gave rise to a large number of single beating cells of varying shapes and sizes (Movie S3). These findings once again indicate that SET depletion propels ESCs toward an endodermal lineage in the expense of neuroectoderm.

To gain molecular insight, we performed gene expression microarray analyses of RA-induced (4 days) SET-KD clones versus Scr-ctrl. RA-induced SET-KD cells had 178 upregulated and 419 downregulated genes compared with controls (Figures 3D and S4A; Tables S2 and S3). Interestingly, some of the upregulated genes are involved in the maintenance of pluripotency (Figure S4A). Changes in selected genes were reconfirmed by qRT-PCR (Figure S4B). Additional prominent examples of altered genes include several Hox cluster genes (*Hoxb13*, *Hoxd1*, *Hoxd4*, *Hoxd8*, and *Hoxd13*), which failed to be induced during differentiation. *Hoxd4*, *Hoxd8*, and *Hoxd13* were all shown to play important roles during neuronal differentiation (Zha et al., 2012), possibly explaining the aberrant neuronal induction phenotypes described above. In addition, SET-KD cells failed to upregulate many genes involved in neurogenesis (Table S4) confirmed by gene ontology analysis (Figure S4C). This suggests that SET acts as an upstream activator of a battery of transcription factors that are required for neuronal differentiation. We cannot exclude the possibility that SET is acting on a few neuronal master regulators but in either case, SET, and particularly SETβ, appears to control neuronal gene expression during differentiation.

We next set out to test the effects of SETα on directed differentiation to NPCs. For this we used the KH2 system, which allowed us to knock in, using the FLPe recombinase, SETα or SETβ under a Dox-controlled promoter (Beard et al., 2006). First, we assessed the effect of SET isoform overexpression on pluripotency marker levels in self-renewing ESCs. qPCR analysis revealed no apparent effect on pluripotency marker expression on SET-OE isoforms for 2 days (Figure S4D). We next sought to understand the effect of SET overexpression on ESC differentiation. Individual SET isoforms were overexpressed in ESCs differentiated into EBs for 4 days. We found that while SETβ overexpression had no discernible effect, SETα isoform overexpression suppressed mesodermal (Brachyury) and endodermal (Gata4) marker expression (Figure 5A), consistent with the SET-depletion experiments, where endodermal lineage markers were upregulated in the SET-depleted cells (Figure S3F). To test the effects of SETα OE on neuronal differentiation, we used both KH2 cells and KH2-HA-SETα cells, differentiated in either the continued presence of 0.5 μg/mL Dox, constantly driving SETα overexpression, or with transient Dox (1 μg/mL) induction for
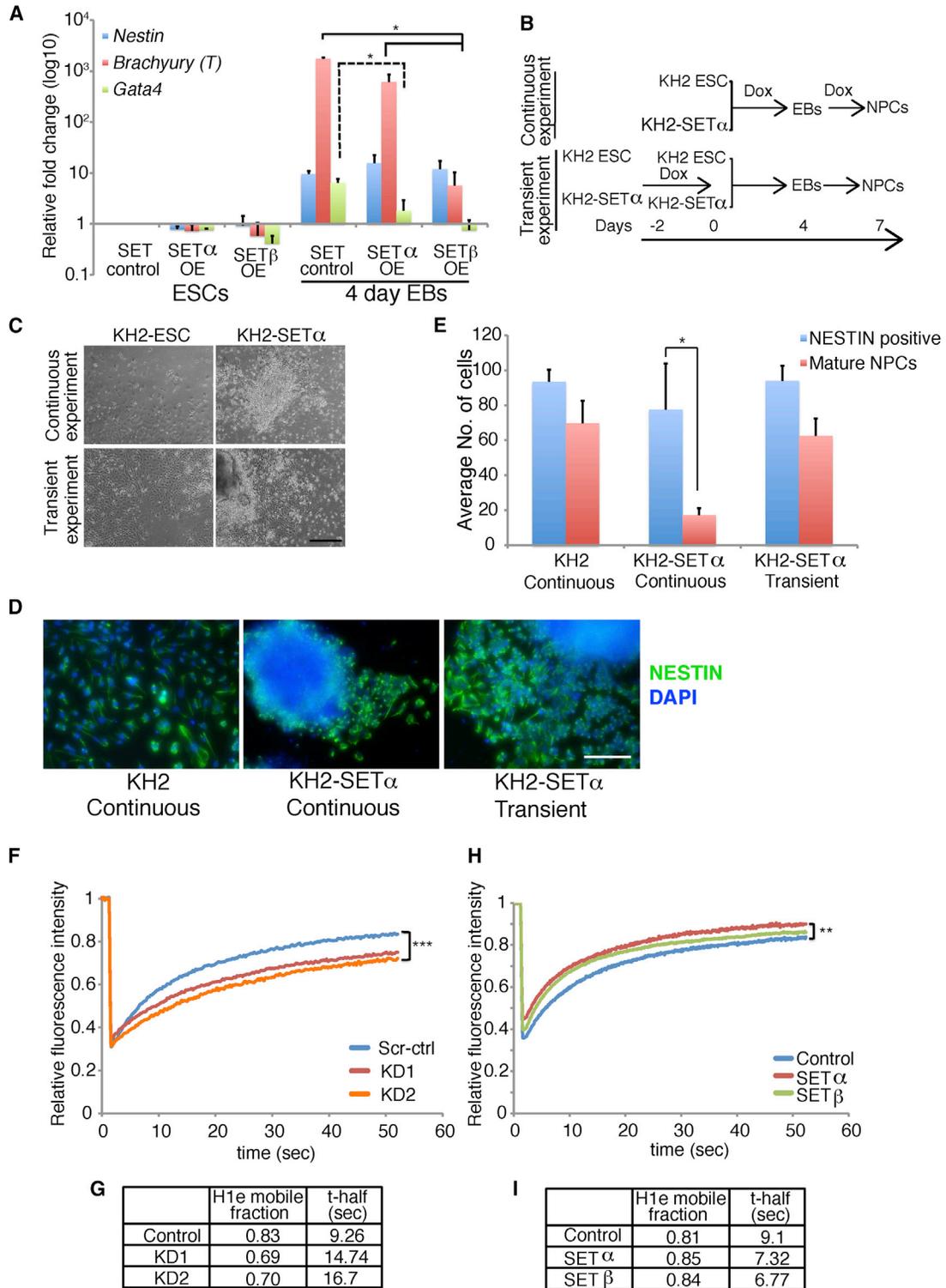
Figure 5. SETα Is Essential for Maintaining the ESC State

(A) qRT-PCR analysis of the lineage-specific factors Nestin (blue), Brachyury (red), and Gata4 (green) in SET-control, SETα-OE, and SETβ-OE ESCs and in 4-day differentiated EBs (n = 3 independent experiments; error bars represent ±SD; *p < 0.05; 2-tailed t test). Expression levels were normalized to GAPDH.

(B) Experimental layout of NPC differentiation from SET-OE ESCs.

2 days only (Figure 5B). Continued SETα expression resulted in fewer and aberrant NPCs (∼75% less), while transient SETα expression resulted in normal NESTIN-positive NPCs, similarly to control KH2-ESCs (Figures 5C–5E). Moreover, continuous SETβ overexpression had no discernible effect on NPC generation. These data show that high levels of SETα are detrimental for early neuronal differentiation, and that SETα levels must decline to ensure proper neurogenesis. Based on all combined data, we propose that SETβ is a lineage-choice factor that promotes neurogenesis during early differentiation and that SETα antagonizes this process.

Since SET overexpression had dramatic effects on *in vitro* differentiation, we next wished to examine the effects of SETα and SETβ on differentiation *in vivo*. We injected SET-KD, SET-OE, and control cells subcutaneously in severe-combined-immunodeficient (SCID) mice. Control ESCs, uninduced KH2-HA-SETα ESCs, uninduced KH2-HA-SETβ ESCs, as well as SET-KD cells all formed teratomas, with cells of all three germ layers present, although the SET-KD teratomas appeared smaller. Unexpectedly, when SETα was induced by addition of Dox to the mice's drinking water, no teratomas were produced (Figure S5A) and we could not detect any undifferentiated cell mass at the site of injection. This suggests that high levels of SETα expression are incompatible with differentiation both *in vitro* and *in vivo*, although only *in vitro* SETα promoted proliferation. To further test this phenomenon, we injected wild-type, SETα-KO, SETβ-KO, and SET-DKO ESCs into SCID mice. Remarkably, while both the SETα-KO and the SET-DKO formed teratomas, albeit the latter's were smaller than controls, the SETβ-KO ESCs failed to make teratomas. Since the DKO were able to differentiate *in vivo*, we attribute the failure of the SETβ-KO ESCs to generate teratomas to the elevated levels SETα in these cells (Figure S3B). This is in line with our SETα-OE experiment whereby increased levels of SETα resulted in loss of teratoma-forming capacity (Figure S5A). Finally, to test whether SET is essential for embryonic development, we co-injected SET guide RNA (gRNA) with Cas9 RNA into fertilized zygotes and found develop-

mental arrest (in ∼25% of the progeny) at around embryonic day 6.5 (Figures S5B and S5C).

SET Maintains H1 Dynamics on Chromatin in ESCs

SET has been recently reported to act as a histone H1 chaperone in HeLa cells (Kato et al., 2011). Since in ESCs, chromatin protein dynamics, including H1, is elevated (Christophorou et al., 2014; Melcer et al., 2012; Meshorer et al., 2006), it was tempting to speculate that SETα might contribute to this hyperdynamic plasticity. To test this, we performed fluorescence recovery after photobleaching (FRAP) analysis in SET-KD and control cells expressing histone H1e-Cherry as conducted previously (Melcer et al., 2012), and found that H1 recovery was significantly reduced in the SET-KD cells compared with controls (p < 0.001) (Figures 5F and 5G). This was mostly due to differences in the bleach depth, which represent the highly mobile fraction of the protein (Nissim-Rafinia and Meshorer, 2011). Since SETα is the predominant form in undifferentiated ESCs, it is likely responsible for this effect, although SETβ is also present at low levels. To unequivocally distinguish between the contribution of the two SET isoforms, we repeated the FRAP experiments in KH2 ESCs expressing SETα or SETβ separately, and found significantly elevated H1 dynamics (p < 0.02) when SETα was overexpressed in KH2 cells (Figures 5H and 5I). SETβ showed a similar but statistically insignificant trend. These data demonstrate that SETα is chiefly responsible for histone H1 dynamics in ESCs, but also that both variants possess H1 chaperone activity to some extent. To confirm this, we also tested the effect of SET depletion on histone H1e mobility in differentiated cells (MEFs), and found comparable effects (Figure S5D). SETα itself is also highly dynamic in both ESCs and differentiated cells (Figure S5E), suggesting weak and transient binding to chromatin. Repeated SET ChIP and SET ChIP-sequencing attempts were unsuccessful with only a small fraction of successfully precipitated DNA, supporting the dynamic nature of SET in ESCs. Taken together, these results suggest that SET is a regulator of linker histone dynamics in both ESCs and differentiated cells, and imply that SETα is more efficient in maintaining

(C) Phase-contrast images of continued (top) or transient (bottom) SETα expression during NPC differentiation. Scale bar, 200 μm.

(D) Immunostaining for NESTIN in NPCs derived from KH2 cells (left), KH2 cells continually expressing SETα (middle), and KH2 cells transiently expressing SETα (right). Nuclei were counterstained with DAPI (blue). Scale bar, 100 μm.

(E) Quantification of the transient and continuous addback experiments (n = 3 independent experiments; error bars represent ±SD; *p < 0.05, 2-tailed Student's t test).

(F) FRAP curves of histone H1e-Cherry in control (blue) and SET-KD (red, orange) cells (n = 3 independent experiments; ***p < 10^{-4}; 2-tailed t test).

(G) Kinetic parameters of SET-KD FRAP experiments

(H) FRAP curves of H1e-Cherry in control (blue), SETα-OE (red), and SETβ-OE (green) cells (n = 3 independent experiments; **p < 0.02, 2-tailed Student's t test).

(I) Kinetic parameters of SET-OE FRAP experiments.

a dynamic linker histone state, potentially contributing to the hyperdynamics phenotype observed in ESCs (Melcer et al., 2012; Meshorer et al., 2006).

DISCUSSION

In this study, using our clone library (Harikumar et al., 2017) we identified SET, a nuclear protein previously not implicated in ESC biology or pluripotency, to play a role in maintaining ESCs as well as in lineage choice decisions during differentiation. Notably, according to the human embryo resource (HumER: https://intranet.cmrb.eu/Human_embryos/) (Vassena et al., 2011), SET is one of the predominantly expressed proteins in early human development. Based on computational analysis and conditional OCT4 depletion, we show here that SET isoform expression is regulated by two alternative promoters and that in ESCs, SETα is controlled by multiple TFs, including OCT4. The two SET isoforms differ only in their N-terminal portion, while most of the protein (∼90%) is shared. We propose that SETα is expressed in ESCs to allow SET's beneficial effects on proliferation and linker chaperone activities, without interfering with differentiation, while SETβ is essential for proper differentiation. Supporting this, we found that co-injections of SET gRNA with Cas9 RNA into fertilized zygotes results in halted development around embryonic day 6.5. Therefore, since SETβ's specific roles, such as activation of *Hox* cluster genes, apoptosis, etc., are not compatible with the undifferentiated ESC state, whereas SETα has a superior linker chaperone activity, and both share similar properties in maintaining proliferation, the switch between SETα and SETβ is crucial for proper differentiation. Immunoprecipitation followed by mass spectrometry analysis of SETα and SETβ separately revealed that although most associated proteins are shared, each isoform has a handful of specific partners that likely act in concert to confer specificity (Data not shown).

Histone chaperones such as SET affect various processes of histone metabolism ranging from synthesis to deposition on chromatin (Avvakumov et al., 2011), and SET itself has been shown to decondense sperm chromatin (Matsumoto et al., 1999). It is therefore tempting to speculate that high levels of SET would act in a similar fashion in ESCs, contributing to a decondensed chromatin conformation. Supporting this hypothesis, our FRAP analysis indicates that SET is involved in histone H1 release from chromatin in ESCs. Hyperdynamic association of chromatin proteins is a hallmark of pluripotency (Meshorer et al., 2006), and was found to be controlled by histone acetylation and methylation (Melcer et al., 2012). In this regard, SET seems to have opposing actions on chromatin plasticity. On the one hand it increases linker histone dynamics, likely reflecting its role as an H1 chaperone (Kato et al., 2011), on the other hand, SET restricts histone H4 acetylation, likely reflecting its association with the INHAT complex (Seo et al., 2001). Increased histone acetylation was shown to enhance chromatin protein dynamics in ESCs (Melcer et al., 2012). How can these seemingly opposing actions of SET be resolved? One option is that its effect on chromatin protein dynamics is restricted to linker histones and that core histones would not be similarly affected. An additional option is that its function as a linker histone chaperone is more important than its function in restricting histone acetylation. The latter is somewhat more likely since increased histone acetylation in ESCs, although it supports the pluripotent state, has very little effect on gene expression (Boudadi et al., 2013; Hezroni et al., 2011). Taken together, our results suggest that in addition to the other mechanisms previously described that support a dynamic chromatin state in ESCs (Melcer et al., 2012), high levels of SETα also act in a similar direction to keep chromatin in its characteristic hyperdynamic ESC state.

In summary, the identification of SET, with its alternative isoforms, as a regulator of proliferation and differentiation in ESCs adds to the list of key factors that aid in maintaining the stem cell state. SET is not a typical TF or chromatin remodeler that directly regulates specific events in stem cells. We envisage that SET acts as a potentiating and balancing factor of several key processes in ESCs, rather than a bona fide TF. We were able to identify different functions of the two SET isoforms in ESCs and during differentiation. Based on its intriguing expression pattern during ESC differentiation, its role in embryonic development should now be explored.

EXPERIMENTAL PROCEDURES

Detailed descriptions are provided in Supplemental Experimental Procedures.

All mice were obtained from the Jackson Laboratory and maintained in the Whitehead Institute animal facility. All experiments were approved by the Committee on Animal Care (CAC) at the Massachusetts Institute of Technology, and animal procedures were performed following the NIH guidelines.

All short hairpin RNAs (shRNAs) were selected from the TRC (The RNAi Consortium, Broad Institute) database. The synthesized shRNA oligos were cloned into the lentiviral vector pLVTHM. For proliferation measurements, ESCs were plated in gelatin-coated (0.2%) 6-well plates at a density of 10^5 per well. Dox was added at a final concentration of 1 μg/mL and medium was replaced daily. For neuronal differentiation, 2×10^6 ESCs were seeded on bacterial culture dishes for EB formation. EBs were grown for 4 days in ESC medium with 10% fetal bovine serum without leukemia inhibitory factor (LIF). At day 4, EBs were plated on poly-L-ornithine/fibronectin (PLO/FN)-coated plates for NPC differentiation in DMEM

Nutrient Mixture-F12 (HAM) medium containing ITS (insulin/transferrin/selenium) and FN. NPCs were grown for 3 days in DMEM-F12 ITS/FN medium and fixed with 4% paraformaldehyde. For dopaminergic neuronal differentiation, NPCs were trypsinized and plated on PLO/FN-coated plates. Cells were fed with DMEM-F12 medium containing N2 plus medium/basic fibroblast growth factor/ascorbic acid every day for 4 days. Differentiation of dopaminergic neurons was induced by removing growth factors. Neuronal differentiation was continued for 10 days with continued medium changes every 2 days.

ACCESSION NUMBERS

The NCBI GEO accession number associated with this paper is GEO: GSE49361.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, 5 figures, 10 tables, and 3 movies and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2017.08.021.

REFERENCES

Aaronson, Y., Livyatan, I., Gokhman, D., and Meshorer, E. (2016). Systematic identification of gene family regulators in mouse and human embryonic stem cells. Nucleic Acids Res. *44*, 4080–4089.

Avvakumov, N., Nourani, A., and Cote, J. (2011). Histone chaperones: modulators of chromatin marks. Mol. Cell *41*, 502–514.

Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis *44*, 23–28.

Betschinger, J., Nichols, J., Dietmann, S., Corrin, P.D., Paddison, P.J., and Smith, A. (2013). Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. Cell *153*, 335–347.

Boudadi, E., Stower, H., Halsall, J.A., Rutledge, C.E., Leeb, M., Wutz, A., O'Neill, L.P., Nightingale, K.P., and Turner, B.M. (2013). The histone deacetylase inhibitor sodium valproate causes limited transcriptional change in mouse embryonic stem cells but selectively overrides Polycomb-mediated Hoxb silencing. Epigenetics Chromatin *6*, 11.

Carlson, S.G., Eng, E., Kim, E.G., Perlman, E.J., Copeland, T.D., and Ballermann, B.J. (1998). Expression of SET, an inhibitor of protein phosphatase 2A, in renal development and Wilms' tumor. J. Am. Soc. Nephrol. *9*, 1873–1880.

Cervoni, N., Detich, N., Seo, S.B., Chakravarti, D., and Szyf, M. (2002). The oncoprotein Set/TAF-1beta, an inhibitor of histone acetyltransferase, inhibits active demethylation of DNA, integrating DNA methylation and transcriptional silencing. J. Biol. Chem. *277*, 25026–25031.

Chambers, I., and Tomlinson, S.R. (2009). The transcriptional foundation of pluripotency. Development *136*, 2311–2322.

Cheloufi, S., Elling, U., Hopfgartner, B., Jung, Y.L., Murn, J., Ninova, M., Hubmann, M., Badeaux, A.I., Euong Ang, C., Tenen, D., et al. (2015). The histone chaperone CAF-1 safeguards somatic cell identity. Nature *528*, 218–224.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.

Christophorou, M.A., Castelo-Branco, G., Halley-Stott, R.P., Oliveira, C.S., Loos, R., Radzisheuskaya, A., Mowen, K.A., Bertone, P., Silva, J.C., Zernicka-Goetz, M., et al. (2014). Citrullination regulates pluripotency and histone H1 binding to chromatin. Nature *507*, 104–108.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D., Buetow, K.H., et al. (2008). Global transcription in pluripotent embryonic stem cells. Cell Stem Cell *2*, 437–447.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Fazzio, T.G., Huff, J.T., and Panning, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell *134*, 162–174.

Fukukawa, C., Shima, H., Tanuma, N., Ogawa, K., and Kikuchi, K. (2000). Up-regulation of I-2(PP2A)/SET gene expression in rat primary hepatomas and regenerating livers. Cancer Lett. *161*, 89–95.

Gaspar-Maia, A., Alajem, A., Polesso, F., Sridharan, R., Mason, M.J., Heidersbach, A., Ramalho-Santos, J., McManus, M.T., Plath, K., Meshorer, E., et al. (2009). Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature *460*, 863–868.

Harikumar, A., Edupuganti, R.R., Sorek, M., Azad, G.K., Markoulaki, S., Sehnalová, P., Legartová, S., Bártová, E., Farkash-Amar, S., Jaenisch, R., et al. (2017). An endogenously tagged fluorescent

fusion protein library in mouse embryonic stem cells. Stem Cell Reports 9, this issue, 1304–1314.

Hezroni, H., Tzchori, I., Davidi, A., Mattout, A., Biran, A., Nissim-Rafinia, M., Westphal, H., and Meshorer, E. (2011). H3K9 histone acetylation predicts pluripotency and reprogramming capacity of ES cells. Nucleus 2, 300–311.

Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A.I., Ranish, J., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proc. Natl. Acad. Sci. USA 106, 5181–5186.

Ho, L., Tan, S.Y., Wee, S., Wu, Y., Tan, S.J., Ramakrishna, N.B., Chng, S.C., Nama, S., Szczerbinska, I., Chan, Y.S., et al. (2015). ELABELA is an endogenous growth factor that sustains hESC self-renewal via the PI3K/AKT pathway. Cell Stem Cell 17, 435–447.

Kato, K., Okuwaki, M., and Nagata, K. (2011). Role of Template Activating Factor-I as a chaperone in linker histone dynamics. J. Cell Sci. 124, 3254–3265.

Lee, S.H., Lumelsky, N., Studer, L., Auerbach, J.M., and McKay, R.D. (2000). Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells. Nat. Biotechnol. 18, 675–679.

Lessard, J.A., and Crabtree, G.R. (2010). Chromatin regulatory mechanisms in pluripotency. Annu. Rev. Cell Dev. Biol. 26, 503–532.

Livyatan, I., Aaronson, Y., Gokhman, D., Ashkenazi, R., and Meshorer, E. (2015). BindDB: an integrated database and webtool platform for "reverse-ChIP" epigenomic analysis. Cell Stem Cell 17, 647–648.

Loh, Y.H., Yang, L., Yang, J.C., Li, H., Collins, J.J., and Daley, G.Q. (2011). Genomic approaches to deconstruct pluripotency. Annu. Rev. Genomics Hum. Genet. 12, 165–185.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell 134, 521–533.

Matsumoto, K., Nagata, K., Ui, M., and Hanaoka, F. (1993). Template activating factor I, a novel host factor required to stimulate the adenovirus core DNA replication. J. Biol. Chem. 268, 10582–10587.

Matsumoto, K., Nagata, K., Okuwaki, M., and Tsujimoto, M. (1999). Histone- and chromatin-binding activity of template activating factor-I. FEBS Lett. 463, 285–288.

Melcer, S., Hezroni, H., Rand, E., Nissim-Rafinia, M., Skoultchi, A., Stewart, C.L., Bustin, M., and Meshorer, E. (2012). Histone modifications and lamin A regulate chromatin protein dynamics in early embryonic stem cell differentiation. Nat. Commun. 3, 910.

Meshorer, E., and Misteli, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. Nat. Rev. Mol. Cell Biol. 7, 540–546.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev. Cell 10, 105–116.

Nagata, K., Kawase, H., Handa, H., Yano, K., Yamasaki, M., Ishimi, Y., Okuda, A., Kikuchi, A., and Matsumoto, K. (1995). Replication factor encoded by a putative oncogene, set, associated with myeloid leukemogenesis. Proc. Natl. Acad. Sci. USA 92, 4279–4283.

Nagata, K., Saito, S., Okuwaki, M., Kawase, H., Furuya, A., Kusano, A., Hanai, N., Okuda, A., and Kikuchi, A. (1998). Cellular localization and expression of template-activating factor I in different cell types. Exp. Cell Res. 240, 274–281.

Nissim-Rafinia, M., and Meshorer, E. (2011). Photobleaching assays (FRAP & FLIP) to measure chromatin protein dynamics in living embryonic stem cells. J. Vis. Exp. http://dx.doi.org/10.3791/2696.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat. Genet. 24, 372–376.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell 141, 432–445.

Respuela, P., Nikolic, M., Tan, M., Frommolt, P., Zhao, Y., Wysocka, J., and Rada-Iglesias, A. (2016). Foxd3 promotes exit from naive pluripotency through enhancer decommissioning and inhibits germline specification. Cell Stem Cell 18, 118–133.

Seo, S.B., McNamara, P., Heo, S., Turner, A., Lane, W.S., and Chakravarti, D. (2001). Regulation of histone acetylation and transcription by INHAT, a human cellular complex containing the set oncoprotein. Cell 104, 119–130.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature 488, 116–120.

Vassena, R., Boue, S., Gonzalez-Roca, E., Aran, B., Auer, H., Veiga, A., and Izpisua Belmonte, J.C. (2011). Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. Development 138, 3699–3709.

Wang, D., Kon, N., Lasso, G., Jiang, L., Leng, W., Zhu, W.G., Qin, J., Honig, B., and Gu, W. (2016). Acetylation-regulated interaction between p53 and SET reveals a widespread regulatory mode. Nature 538, 118–122.

Yamazaki, T., Kobayakawa, S., Yamagata, K., Abe, K., and Baba, T. (2007). Molecular dynamics of heterochromatin protein 1beta, HP1beta, during mouse preimplantation development. J. Reprod. Dev. 53, 1035–1041.

Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., Ma, Z., Davis, C., Pope, B.D., et al. (2014). A comparative encyclopedia of DNA elements in the mouse genome. Nature 515, 355–364.

Zha, Y., Ding, E., Yang, L., Mao, L., Wang, X., McCarthy, B.A., Huang, S., and Ding, H.F. (2012). Functional dissection of HOXD cluster genes in regulation of neuroblastoma cell proliferation and differentiation. PLoS One 7, e40728.