Cell, Volume *134*

Supplemental Data

An RNAi Screen of Chromatin Proteins Identifies Tip60-p400 as a Regulator of Embryonic Stem Cell Identity

Thomas G. Fazzio, Jason T. Huff, and Barbara Panning

Supplemental Experimental Procedures

Antibodies

Antibodies used in this study were p400 (Bethyl Labs, BL1879); Dmap1 (Proteintech, 10411-1-AP); Oct-4 (H-134, Santa Cruz Biotechnology); Nanog (Cosmo Bio); Cyclophilin B (ab3565, Abcam); β-Actin (Santa Cruz Biotechnology); H2A.Z (Abcam, ab4174); H3K4me3 (Lake Placid Biologicals, AR-0169); H2AK5Ac Lake Placid Biologicals, AR-0201); H4K5Ac (Lake Placid Biologicals, AR-0119); H4K12Ac (Lake Placid Biologicals, AR-0106); H4K16Ac (Lake Placid Biologicals, AR-0107); tetra-acetylated H4 (Millipore, 06-866); p21 (NeoMarkers, Ab6); and γH2A.X (Upstate, 05-636).

Oligos

Oligonucleotides used for RT-PCR and ChIP:

RT-qPCR		
Dkk1	ACTCAAATGGCTTTGGTAATATGG	ATAATCTCTTCTGAATTCTGCCCA
Snai1	CTTGTGTCTGCACGACCTGTG	AGACTCTTGGTGCTTGTGGAG
Myc	AGCTGTTTGAAGGCTGGATTTC	GCAACATAGGATGGAGAGCAGA
Scamp1	CCTTGAGGTCTGTGGTATTGGA	TACACCCTTAGTGACCTCAGTGTC
Nodal	TCCTTCTTCTTCAAGCCTGTTG	CCAGATCCTCTTCTTGGCTCA
Akr1b8	TACTGTCACTCGAAGGGCATCT	ATCTCCTCGTCACTCAACTGGA
Nefl	AGCTAGAGGACAAGCAGAATGC	GCAAGCCACTGTAAGCAGAAC
Gata6	AGTCAAAAGCTTGCTCCGGTA	ACCTATGTAGAGGCCGTCTTGA
Gata4	CGGTTTTCTGGGAAACTGGA	AATTGGATTTGCGGTTGCTC
ChIP		
Rps9	TCTAAGGCGGATTTCTCTGAGAG	CTTTAGCTCCTGGTCGAGACG
Nodal	TACGCGTATGGCATATTTGTTAC	TCTTGATTCTTGGTGGTTTGAG
Zfp37	CCAAGGTATTGGGGTATTATGG	TACATCACAAAGCCCTCATTTC
Lars	GAACTCTGGTCTCAGCACATTG	GATTCTTGCTGTTCTTGAAACG
Col4a1	CTCAGAGTCCTACACTCATGTGC	ACTTTTAGGATGCGGAGAAGAG
Ahrgef12	ACAGGTGGGTATGAAGTTCCTC	CATGGTTTTCCTTCATTTCCTC
Gata6	TATTCACCAGCAGCGACTAGC	AATCTTCTAAAGCGAAGGACCG
Gata4	GGCAGGAGAAACTCACAAGTCC	GGATTCATGTTCCCTCCCAAAC
Speer2	CATGACCTTGCTCTGAGTCAAGT	CACCAATGTAGTGGCTAGACACTG

Generation of esiRNA library of chromatin structural and regulatory proteins

To obtain a comprehensive list of known or predicted chromatin structural and regulatory proteins, we searched the NCBI Gene database for an array of keywords mostly consisting of protein domains commonly found in chromatin proteins, such as “SANT”, “Set”, “PhD”, etc. In addition, we included “chromatin” as a search term to identify proteins that fell outside of the

domain search. The list of “hits” from the list was examined and non-chromatin genes that were included in the search were manually culled. Subsequent to completion of the screen, we found that 16 hits were reannotated as pseudogenes. Each of these 16 esiRNAs are predicted to target an authentic gene also identified as a hit in the screen. This serendipitously provided independent confirmation for the corresponding authentic genes. The final table (Table 1) includes only loci currently annotated as authentic genes. The primary screen data (Table S1) includes all hits from the screen, including these pseudogenes, with links to their current annotations.

To identify targets of each gene for RNAi, we used the web based program DEQOR (http://cluster-1.mpi-cbg.de/Deqor/deqor.html), which was designed for the purpose of identifying regions of genes for esiRNA design that maximally silence target genes and have minimal potential for cross-silencing(Henschel et al., 2004).

PCR, in vitro transcription, in vitro dicing and transfections were carried out in 96-well plates for the initial and secondary screens, and in 24-, 12-, or 6-well dishes for subsequent characterization. A two step PCR procedure similar to a previously described method(Foley and O'Farrell, 2004) was performed to generate templates for in vitro transcription. The initial PCR primers all had an eight base pair tag at their 5’ end: 5’GGGCGGGT, to allow annealing of a T7 anchor primer for the second round of PCR. Primary PCR reactions were performed using mouse cDNA prepared from several cell lines and mixed together as a template. A fraction of each PCR reaction was subjected to agarose gel electrophoresis to assess quality. PCR reactions that generated products of an incorrect size or with poor yield were individually optimized prior to the second PCR step. The second round of PCR utilized each primary PCR product, diluted 1/200, as a template and added the promoter for T7 RNA polymerase to each end. Cycling conditions for the second PCR were as follows:

1) 94^{o}C	2 min
2) 94 ^{o}C	30 sec
3) 42 ^{o}C	45 sec
4) 72 ^{o}C	1 min
5) Go to #2	4 times
6) 94 ^{o}C	30 sec
7) 60 ^{o}C	45 sec
8) 72 ^{o}C	1 min
9) Go to #6	29 times

10) 72 °C 5 min
11) 4 °C Hold

2 μl of each secondary PCR reaction was run on an agarose gel to assess quality and yield. In vitro transcription was performed using a RiboMAX Large Scale RNA Production System (Promega); 10 μl reactions were performed for each gene, using 4 μl unpurified second round PCR product as a template. In vitro transcription and annealing of the RNA strands were performed in one step in a thermal cycler as described(Kittler et al., 2005) and diluted with ultrapure water to 100 μl. 3-4 μl were run on an agarose gel to assess yield.

Purification of GST-RNaseIII was performed as described(Yang et al., 2002). Unpurified double stranded RNA (dsRNA) was diced in vitro using a modification of the procedure described in the ShortCut RNAi Kit (New England Biolabs). Reactions contained the following:

H_2O	120 ul
diluted dsRNA	20 ul
10X ShortCut Reaction Buffer	20 ul
200 mM $MnCl_2$	20 ul
GST-RNaseIII (~2.2 μg/μl)	20 ul

Dicing was carried out at 37°C for 90 minutes. esiRNAs were ethanol precipitated, washed with 75% ethanol, dried and re-suspended in 100μl of ultrapure water. 3μl of every fourth reaction was run on a 15% polyacrylamide 1X TBE Criterion gel (BioRad), stained with ethidium bromide and visualized to ensure efficient dicing and recovery. Plates of esiRNAs were stored at -20 °C.

For initial screening, we found that unpurified esiRNAs were highly effective in silencing of ESCs and immortal MEFs, as long as dicing was efficient. Subsequent experiments utilized esiRNAs purified as described(Myers and Ferrell, 2005), which allowed transfection of more precise quantities of esiRNA. Purified esiRNAs were transfected at a final concentration of 0.33 ng/μl.

ESCs were transfected in 96-well format as follows. Lipofectamine 2000 (Invitrogen) was diluted in OptiMEM medium with 16.5 μl OptiMEM and 0.33 μl Lipofectamine 2000 per well. This mix was incubated for fifteen minutes at room temperature. During the incubation, 2

μl esiRNA was diluted in 16.5 μl OptiMEM. 16.5 μl dilute Lipofectamine2000 was added to the diluted esiRNA, and the mixture was incubated an additional 15 minutes at room temperature. 67 μl of ESCs suspended in ESC medium without antibiotics at 300,000 cells per milliliter were then added to the lipid-esiRNA complexes and immediately plated on gelatin coated 96-well tissue culture plates, which were incubated at 37^{o}C with 5% CO_2. After 14-17 hours, 150 μl of fresh antibiotic-free ESC medium was added to each well. Phenotypes were recorded using an inverted light microscope after 72 and 96 hours. Images were collected on a Leica DM 1RB inverted microscope equipped with a Hamamatsu Orca AG camera. Transfection of MEFs and other somatic cells was performed identically, with three exceptions: 1. Cells were pre-plated 24 hours prior to transfection at 5000 cells per well, 2. Lipofectamine RNAiMAX (Invitrogen) was used in place of Lipofectamine 2000 and 3. Antibiotic-free medium appropriate for each cell type was substituted for ESC medium.

shRNAs

Annealed oligos encoding shRNAs targeting Dmap1 were cloned into pSICOR-mCherry as described (http://web.mit.edu/jacks-lab/protocols/pSico.html). The three shRNAs targeted regions of Dmap1 corresponding to the following 19-mer sequences:

Dmap1-A: GCAAGGCTCCCAAGAAGAA
Dmap1-B: GCATCAAGTTTCCAGATTT
Dmap1-C: GTGGTATGGTGTAAATAGA

Chromatin immunoprecipitation

Approximately 3 X 10^{7} ESCs were trypsinized and pipetted to single cell suspension in 10 ml ESC medium, to which was added 1/10 volume of Fix Solution(McConnell et al., 2004). Cells were rotated for 10 minutes and fixation was quenched by the addition of 0.5 ml 2.5M glycine, followed by 5 minutes of rotation. Cells were pelleted and resuspended in 1 ml PBS with 0.8 mM PMSF and 1 mM DTT. Cells were transferred to a microfuge tube, pelleted and resuspended in 1 ml Swelling Buffer (10mM Tris pH 8.0, 85mM KCl, 0.5% NP-40 with protease inhibitors), followed by rotation for 20 minutes at 4 degrees. Cells were pelleted and resuspended in 1 ml of room temperature Swelling Buffer with 3 mM $CaCl_2$, and 25 units of

micrococcal nuclease (Worthington) were added. Cells were rotated for five minutes and digestion was terminated with the addition of 20 μl 0.5M EGTA. Cells were pelleted and resuspended in 1 ml of Lysis Buffer(Chu et al., 2006) and sonicated with a Branson sonifier on setting 3 (5X 10 seconds each, resting on ice for 1 minute between pulses). Lysate was clarified twice by centrifugation at maximum speed in a microcentrifuge. Lysate concentration was determined by Bradford assay.

1/10 of the amount of the cross-linked lysate to be used in the immunoprecipitation was set aside as "Input" sample. Chromatin was immunoprecipitated with 1.5 μg p400, Nanog, H3K4me3 or H4ac antibody that had been pre-coupled overnight to 15 μl Protein A Dynabeads (Invitrogen) as described (http://www.fhcrc.org/science/labs/tsukiyama/protocols/chromatin_ip1.pdf). Immunoprecipitations were performed for 3 hours at 4 degrees with rotation. Beads were washed 5 times with 1 ml Lysis Buffer for 5 minutes each at 4 degrees with rotation. Chromatin was eluted 2 times with 2X Stop solution (20 mM Tris pH 8.0, 100 mM NaCl, 20 mM EDTA, 1% SDS) with shaking at 65 degrees. Cross-links were reversed overnight at 65 degrees in the presence of 20 μg proteinase K. DNA was phenol/chloroform extracted, ethanol precipitated and re-purified on Zymoclean columns (Zymo Research). For semi-quantitative PCR (Fig. 5a) Input DNA was diluted corresponding to 0.5%, 0.1% and 0.02% of the immunoprecipitated material, immunoprecipitated DNA was diluted 1:10, and 1 μl of each sample was added to a 20 μl PCR reaction. PCR was performed using 33 cycles.

Genome wide location analysis of p400

For genome-wide location analysis, 4 ng of input and immunoprecipitated DNA samples were amplified using a BioPrime Array CGH kit (Invitrogen), omitting fluorescent nucleotides and using 120 μM of each dNTP. Labeling, hybridization of amplified input and immunoprecipitated samples to Mouse MM8 promoter arrays and scanning were performed at NimbleGen Systems Inc. Raw Cy5 and Cy3 intensity values were extracted from the Nimblegen "pair" files for each of the two microarrays in the MM8 promoter set. Normalization of log2(ratio) probe values was performed using functions in the R (http://www.r-project.org) package limma in Bioconductor (Gentleman et al., 2004; Smyth, 2004). A variety of normalization methods were tested in comparison to the Nimblegen method. The final

normalized values reflect fitting of an intensity-dependent loess curve only to the control probes and then normalizing the chromosomal probes to the fit without background subtraction (using the “loess” method in the normalizeWithinArrays() function). This method yielded normalized log2(IP/input) values with a large dynamic range and empirical distributions most similar between the two arrays of the promoter set (data not shown).

Genes were split into 5 quantile bins containing equal numbers (quintiles) according to increasing expression (see below). The average enrichment surrounding annotated transcription start sites (Figure 5B) was calculated in R by fitting a loess line to the log2(IP/input) values in each quantile. A simple cutoff for bound versus unbound promoter regions was established as the inflection point of the bimodality of the distribution of mean p400 log2(IP/input) values, corresponding to a value of 0.1, as shown in Figure 5C, 5D, andS8. Alternatively, a more accurate estimate was obtained by fitting a kernel density estimate to the mean p400 log2(IP/input) values (Figure S6B). A normal distribution was then fit to the left mode as an estimate of the empirical distribution of unbound promoters, yielding approximately 45% unbound and 55% bound promoter regions in the total distribution. Finally, a completely different method was also used to assess the number of bound promoter regions: peaks were assigned with the moving average algorithm of TileMap (Ji and Wong, 2005) (Figure S6C). The log2(IP/input) values were loaded directly into TileMap and the program was run with moving windows of 5 probes and a maximal allowed gap of 1000 bp and unbalanced mixture subtraction with settings of “p” cutoff 0.5, “q” cutoff 0.2, and grid size 5000. The false discovery rate cutoff was varied and the resulting peaks were assigned to promoter regions, as plotted in Figure S6C. As a specificity control, the values of the probes were permuted and put through TileMap with identical settings, yielding very few peaks (Figure S6C). Similar results for TileMap were obtained using several different paramater settings (data not shown).

The normalized log2(IP/input) values for all probes in the dataset can be viewed in the UCSC Genome Browser (file available from GEO, http://www.ncbi.nlm.nih.gov/geo/), which was used to generate the right panel of Figure 5A. The enrichment within 1kb of each annotated TSS is provided in Table S7. These values were compared with published data for H3K4me3 and H3K27me3 location analysis at promoters in mouse E14 ESCs; data from

GEO Series GSE7815 (Maherali et al., 2007) were plotted against p400 enrichment using a density scatterplot and Polycomb targets (Boyer et al., 2006) were highlighted.

Expression Profiling

Sample preparation, labeling, and array hybridizations were performed according to standard protocols from the UCSF Shared Microarray Core Facilities and Agilent Technologies (http://www.arrays.ucsf.edu and http://www.agilent.com). Total RNA quality was assessed using a Pico Chip on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). RNA was amplified and labeled with Cy3-CTP or Cy5-CTP using the Agilent low RNA input fluorescent linear amplification kits following the manufacturers protocol (Agilent). Labeled cRNA was assessed using the Nandrop ND-100 (Nanodrop Technologies, Inc., Wilmington DE). Equal masses of sample were hybridized to Agilent whole mouse genome 4x44K Ink-jet arrays (Agilent). Two color (Cy5 and Cy3; competitive) hybridizations were performed to acquire the p400 knockdown data, whereas one color (Cy3; non-competitive) hybridizations were used to acquire the Tip60 knockdown data. For the p400 knockdown, hybridization samples were paired according to group designation and dye swaps were incorporated to correct for any dye bias. Hybridizations were performed for 14 hrs, according to the manufacturers protocol (Agilent). Arrays were scanned using the Agilent microarray scanner (Agilent) and raw signal intensities were extracted with Feature Extraction v8.9 software (Agilent).

A linear model was fit to the comparison to estimate the mean log2(fold change) in 4 biological replicates and to calculate a moderated t-statistic, B statistic, false discovery rate and p-value for each probe. Adjusted p-values were produced by the method proposed by (Holm, 1979). All procedures were carried out using functions in the R package limma in Bioconductor (Gentleman et al., 2004; Smyth, 2004).

We considered genes to be differentially expressed if at least one probe for the gene had a B statistic > 0 for the p400 or Tip60 knockdown, corresponding to odds greater than one-to-one of differential expression in the respective linear models. Enrichment of Gene Ontology terms and categories was performed with DAVID 2007 (http://david.abcc.ncifcrf.gov/).

Expression profiling data of Nanog or Oct4 knockdown were downloaded from Gene Expression Omnibus (GEO) DataSet GDS1824. Data of Oct4 overexpression or depletion were downloaded from GEO Series GSE5936. Data for Suz12 knockout were downloaded from the Young lab website (http://jura.wi.mit.edu/young_public/hES_PRC/Suz12_expression.zip). Data for Sox2 depletion were downloaded from GEO Series GSE5895. For each profiling experiment, a linear model was fit to the comparison against the appropriate ESC control to estimate the mean log2(fold change) using the R package limma in Bioconductor (Gentleman et al., 2004; Smyth, 2004). The median of the mean log2(fold change) estimates was used for genes with multiple probes.

Principal component analysis of the genes present in all of the datasets was performed using the correlation algorithm of the princomp() function in R and the complete results are in Table S6. The first two principal components presented in Figure 5D account for 63% of the total variance in the eight included datasets.

Expression levels were estimated for the quintiles in Figure 5B by back-calculating the intensity in the eGFP control knockdown cells from the log2(ratio) (M values) and average intensity (A values) provided from the limma package output for the p400 knockdown experiments.

Supplemental Data References

Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Chu, F., Nusinow, D. A., Chalkley, R. J., Plath, K., Panning, B., and Burlingame, A. L. (2006). Mapping post-translational modifications of the histone variant MacroH2A1 using tandem mass spectrometry. Mol Cell Proteomics *5*, 194-203.

Foley, E., and O'Farrell, P. H. (2004). Functional dissection of an innate immune response by a genome-wide RNAi screen. PLoS Biol *2*, E203.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J.*, et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol *5*, R80.

Henschel, A., Buchholz, F., and Habermann, B. (2004). DEQOR: a web-based tool for the design and quality control of siRNAs. Nucleic Acids Res *32*, W113-120.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. Scand J Stat *6*, 65-70.

Ji, H., and Wong, W. H. (2005). TileMap: create chromosomal map of tiling array hybridizations. Bioinformatics *21*, 3629-3636.

Kittler, R., Heninger, A. K., Franke, K., Habermann, B., and Buchholz, F. (2005). Production of endoribonuclease-prepared short interfering RNAs for gene silencing in mammalian cells. Nat Methods *2*, 779-784.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R.*, et al.* (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. Cell Stem Cell *1*, 55-70.

McConnell, A. D., Gelbart, M. E., and Tsukiyama, T. (2004). Histone fold protein Dls1p is required for Isw2-dependent chromatin remodeling in vivo. Mol Cell Biol *24*, 2605-2613.

Myers, J. W., and Ferrell, J. E. (2005). Silencing gene expression with Dicer-generated siRNA pools. In RNA Silencing, G. Carmichael, ed. (Totowa, Humana Press), pp. 352.

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol *3*, Article3.

Yang, D., Buchholz, F., Huang, Z., Goga, A., Chen, C. Y., Brodsky, F. M., and Bishop, J. M. (2002). Short RNA duplexes produced by hydrolysis with Escherichia coli RNase III mediate effective RNA interference in mammalian cells. Proc Natl Acad Sci U S A *99*, 9942-9947.

Table S2. References for mouse knockout phenotypes in Table 1

Gene	Reference
Rnf2	Voncken et al., *Proc. Natl. Acad. Sci. U S A*. **100**, 2468 (2003)
Sp1	Marin et al., *Cell*. **89**, 619 (1997)
Brca1	Hakem et al., *Cell*. **85**, 1009 (1996)
	Liu et al., *Genes Dev*. **10**,1835 (1996)
	Gowen et al., *Nat. Genet*. **12**, 191 (1996)
Mbd3	Hendrich et al., *Genes Dev*. **15**, 710 (2001)
Sin3A	Dannenberg et al. *Genes Dev*. **19**, 1531 (2005)
Arid4a	Wu et al., *Genes Dev*. **20**, 2589 (2006)
Brd4	Houzelstein et al., *Mol. Cell. Biol*. **22**, 3794 (2002)
Stat3	Takeda et al., *Proc. Natl. Acad. Sci. U S A*. **94**, 3801 (1997)
Trim28	Cammas et al., *Development*. **127**, 2955 (2000)
Ssrp1	Cao et al., *Mol. Cell. Biol.* **23**, 5301 (2003)
Htatip	Gorrini et al., *Nature*. **448**, 1063 (2007)
Trrap	Herceg et al., *Nat. Genet.* **29**, 206 (2001)
H2afz	Faast et al., *Curr. Biol*. **11**, 1183 (2001)
Smarcc1	Kim et al., *Mol. Cell. Biol*. **21**, 7787 (2001)
Smarca4	Bultman et al., *Mol. Cell*. **6**, 1287 (2000)
Gmmn	Hara et al., *Genes Cells*. **11**, 1281 (2006)
Mcm4	Shima et al., *Nat. Genet*. **39**, 93 (2007)
Sf3b1	Isono et al., *Genes Dev*. **19**, 536 (2005)

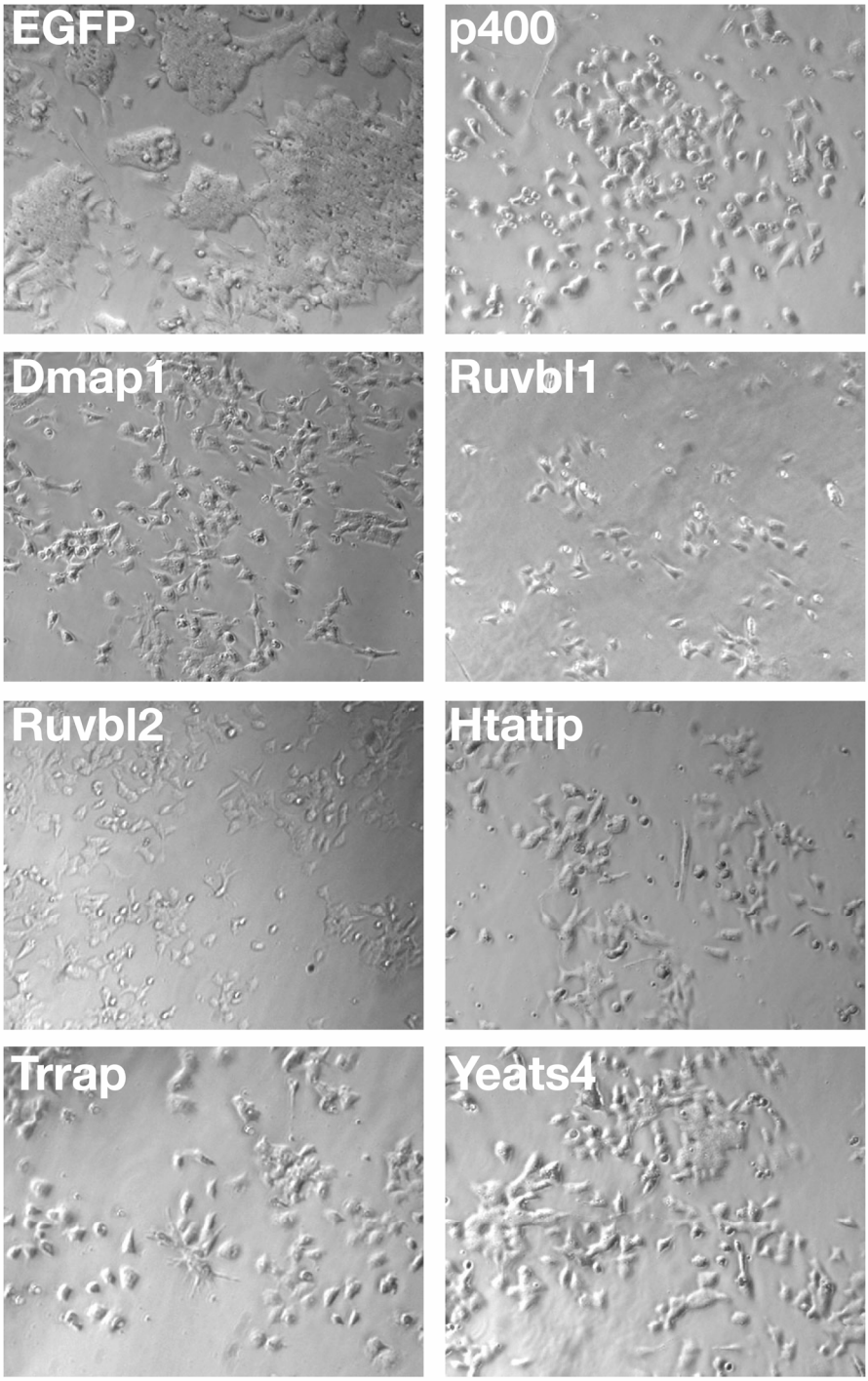

Figure S1. Validation of Tip60-p400 phenotypes using a second set of non-overlapping siRNAs. ESCs were transfected and imaged as in Figure 2A.

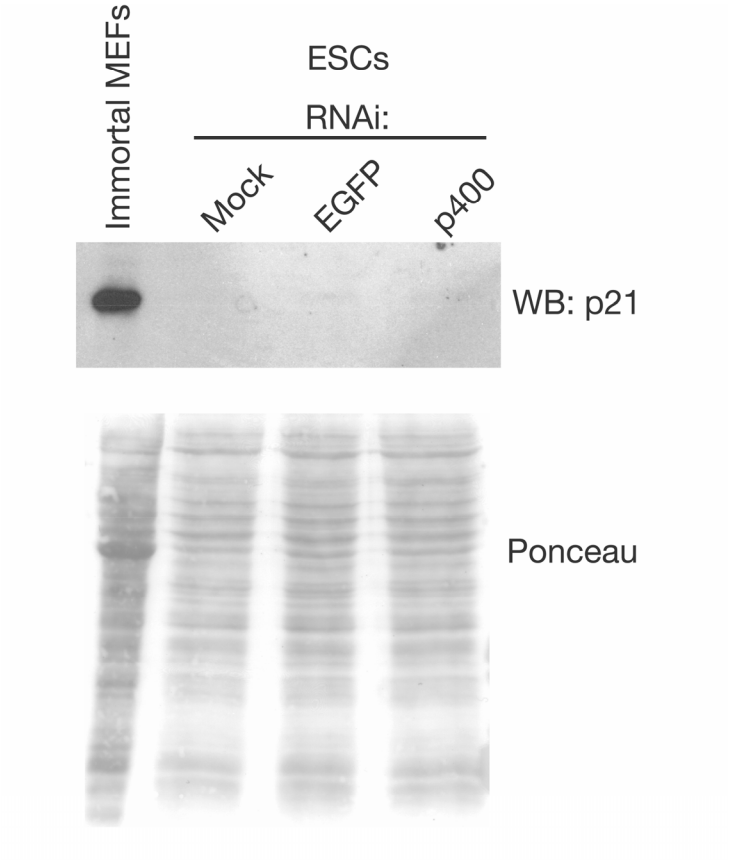

Figure S2. Depletion of p400 in ESCs does not induce p21 expression. ESCs were transfected with the indicated esiRNAs and harvested for Western blotting after 72 hours. Immortal mouse embryonic fibroblasts (MEFs) were run as a positive control for p21 expression.

A

Percent EBs recovered (n = 152 each)		
Mock	p400	Tip60
100	7.9	6.6

B

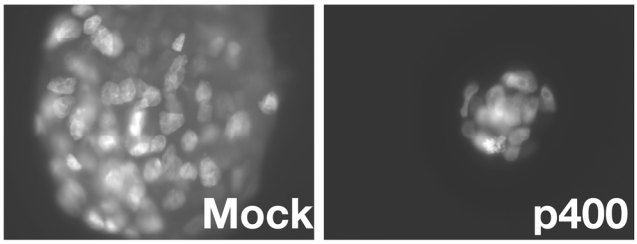

Figure S3. Tip60-p400 knockdown ESCs form few EBs after 4 days. (A) ESCs were transfected with the indicated siRNAs, suspended as in Figure 3B, and the number of EBs present on day 4 was noted. Data are presented as a percentage of the total number of suspension cultures in which an EB was recovered. (B) Images of day 4 EBs from mock- or p400-depleted ESCs.

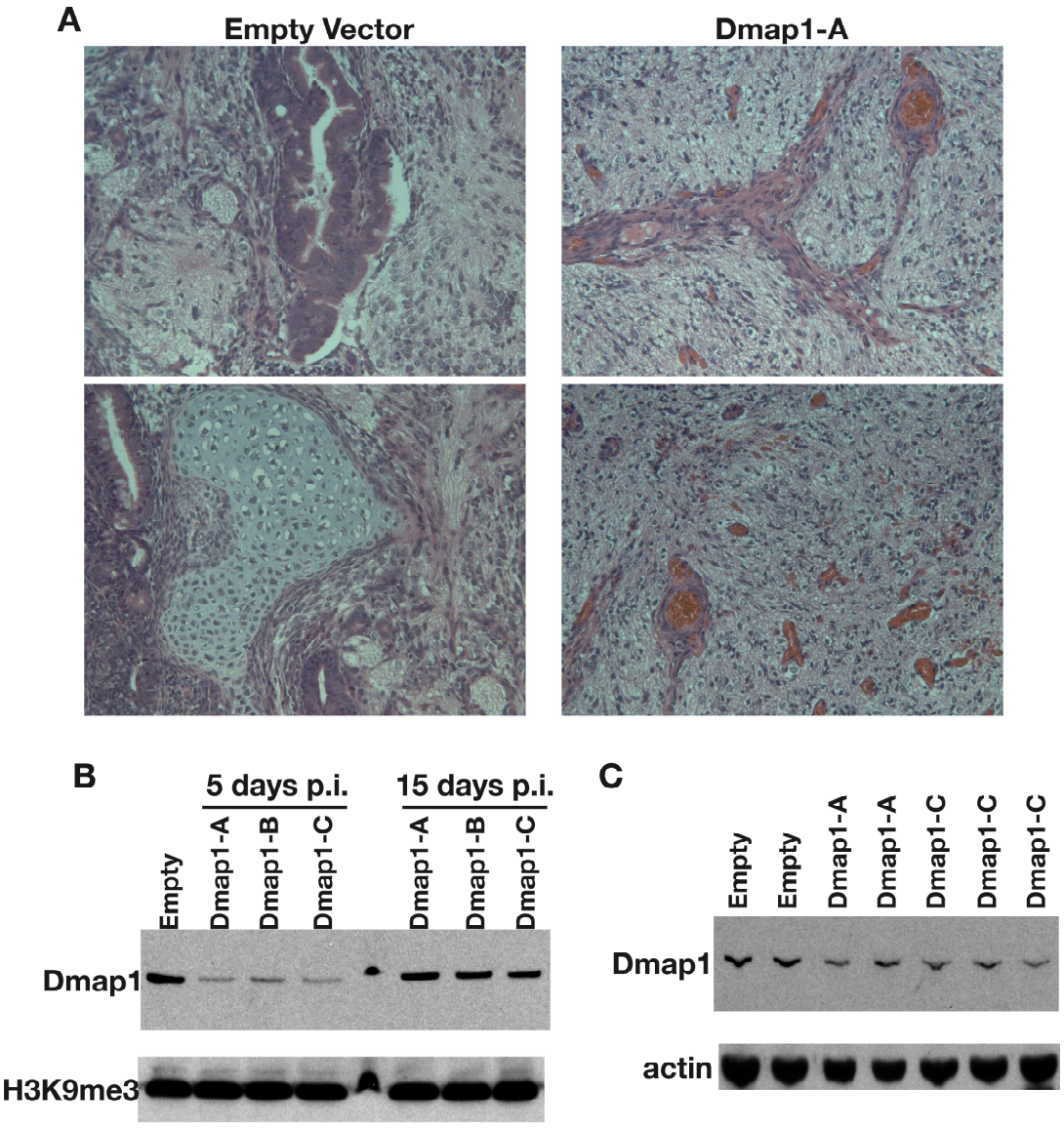

Figure S4. Dmap1 knockdown ESCs form smaller, less complex teratomas. (A) ESCs infected with empty or Dmap1 shRNA-expressing vectors were injected subcutaneously into nude mice and left for 21 days. Tumors were fixed in 4% paraformaldehyde, sectioned and stained with hematoxylin and eosin. (B) ESCs infected with the indicated shRNA-expressing viruses were sorted for the infected populations. Cells were harvested for western blotting at the time of injection (5 days post infection, p.i.) or passaged for an additional 10 days in culture. H3K9me3 is used as a loading control. (C) Samples of tumors derived from the indicated ESCs were blotted for Dmap1 and actin (as a loading control). As with the cultured ESCs, Dmap1 expression in the tumors appears to have partially recovered.

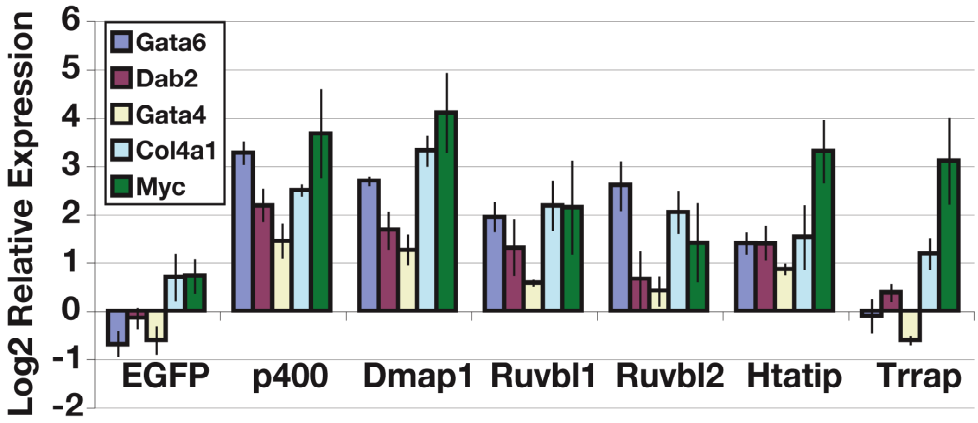

Figure S5. Validation of differentially expressed genes. RT-qPCR measuring mRNA levels of genes misregulated in p400-knockdown expression array data. mRNA levels of genes indicated to the left were measured in ESCs depleted of EGFP or Tip60-p400 subunits. mRNA levels were normalized to cyclophilin B. Averages and standard deviations for three biological replicates are shown.

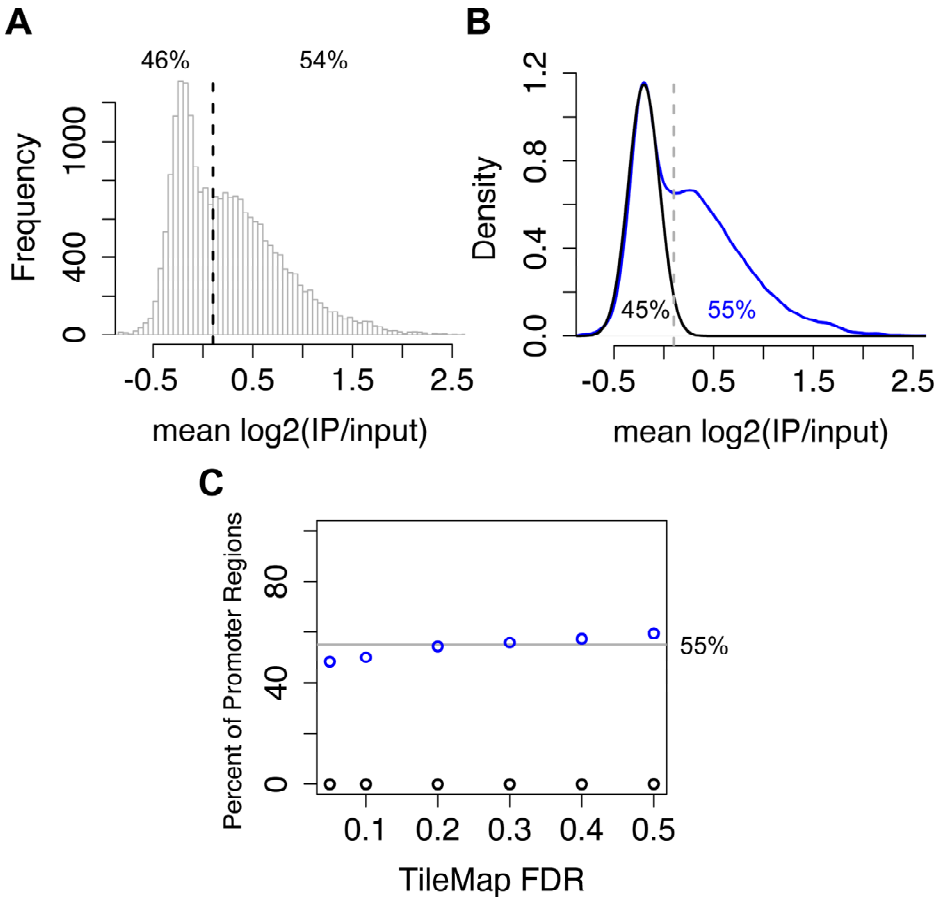

Figure S6. p400 binds at most promoter regions in ESCs. (a) A histogram of mean p400 log2(IP/input) at every promoter region of the two-microarray Nimblegen MM8 set. The approximate inflection point of the two modes is marked by a broken line corresponding to a hard cutoff with 46% of regions to the left (unbound) and 54% to the right (bound). (b) A second estimate of bound regions is provided where a curve is fit in place of a hard cutoff, yielding 45% unbound and 55% bound (see Materials and Methods for details). (c) Instead of averaging each promoter region, peaks were assigned by the moving average method of TileMap at various false discovery rate (FDR) cutoffs. The percent of promoter regions on the array containing at least one p400 peak is plotted (blue). For comparison a permuted set of the p400 data was put through the same procedure to indicate the specificty of peak-finding (black) (see Supplementary Methods for details).

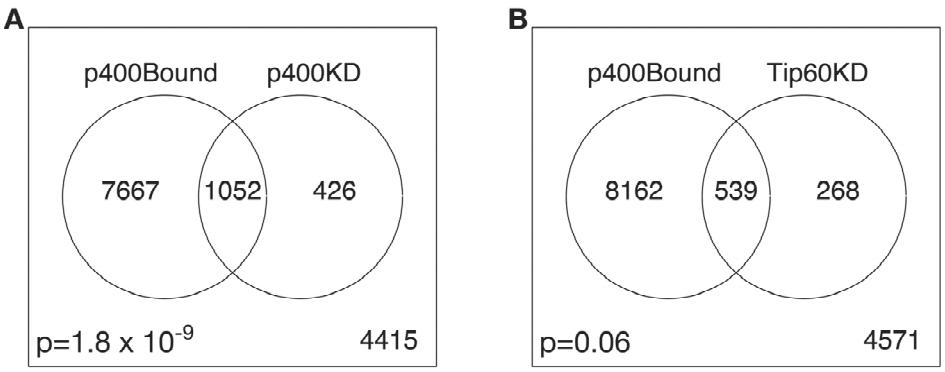

Figure S7. Overlap between p400 localization and genes misregulated upon Tip60 or p400 binding. Venn diagrams comparing overlap of genes bound by p400 at their promoters (left) with genes misregulated (right) upon knockdown of either p400 (A) or Tip60 (B). Genes that do not fall into either category are noted in the bottom right of each box. p-values assessing significance of each overlap are noted. In general, the Tip60 knockdown data was similar to p400, but more compressed, leading to weaker overlap.

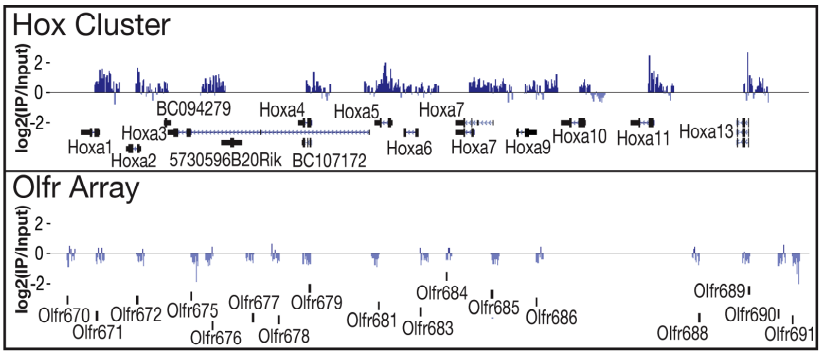

Figure S8. p400 binds to silent genes targeted by Polycomb complexes in ESCs. Distribution of p400 across a Hox gene cluster and an olfactory receptor (Olfr) gene array, graphed as in Fig. 5a. Hox genes are bound by Polycomb targets and contain the 'bivalent' chromatin signature, while Olfr genes are not bound by Polycomb and do not have the 'bivalent' mark. The name and location of each gene is indicated below the graph.

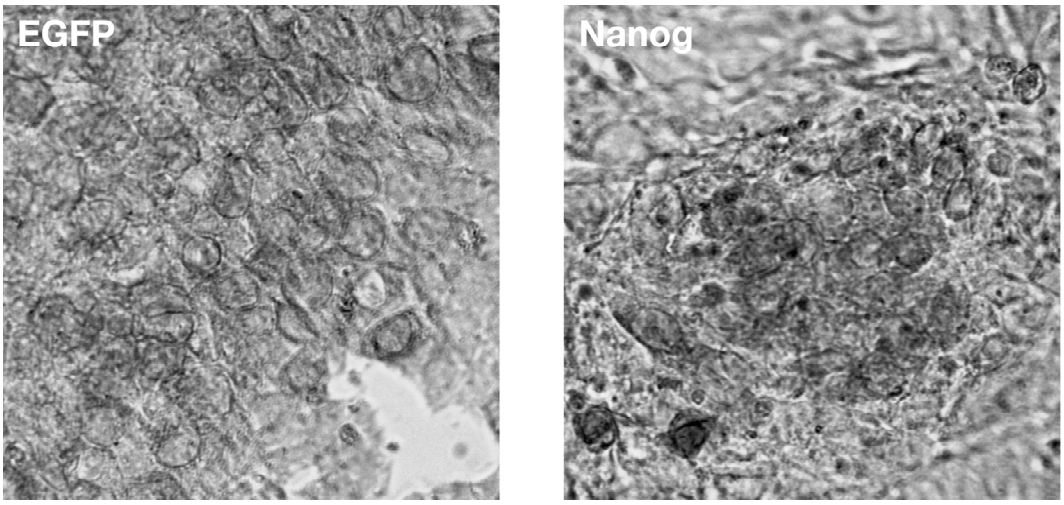

Figure S9. Nanog knockdown ESCs exhibit levels of alkaline phosphatase (AP) activity indistinguishable from controls. ESCs were transfected with the indicated esiRNAs for 64 hours and stained for AP activity as in Figure 3. Note that staining at this time point is weak compared with control knockdown ESCs in Figure 3, due to the fact that the ESC colonies shown above have not grown to the same density at this time point.

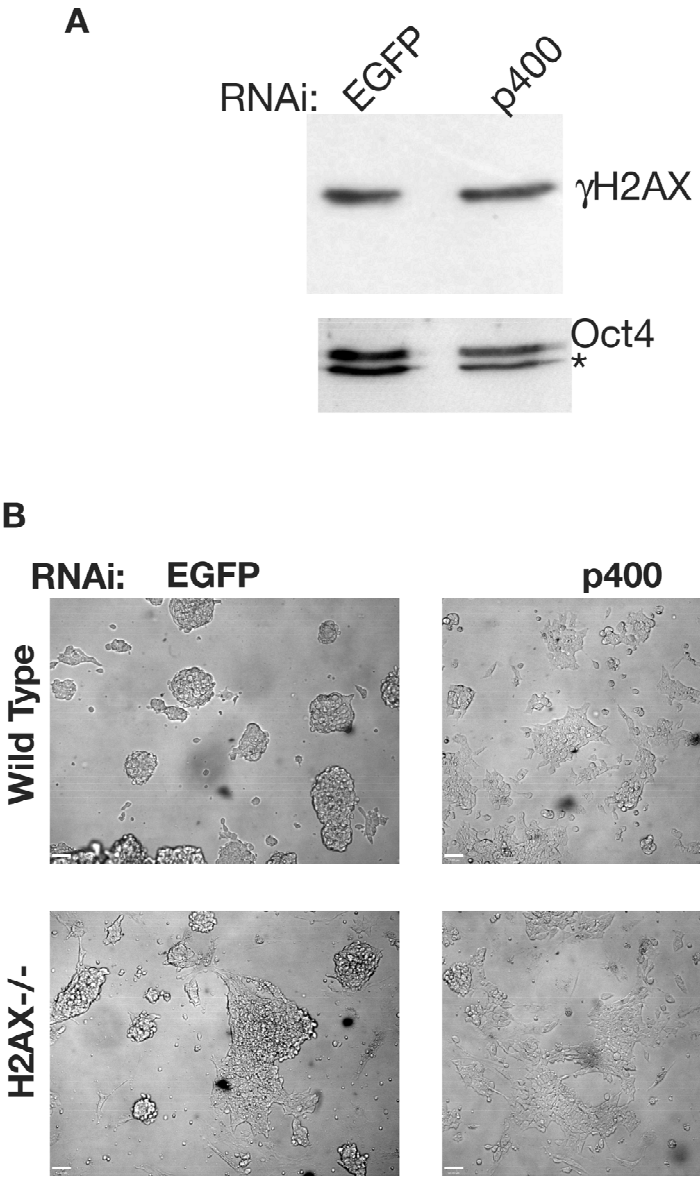

Figure S10. The self-renewal defect upon p400 depletion is independent of H2AX. (A) ESCs were transfected with EGFP or p400 esiRNAs for three days and harvested for Western blotting. Oct4 is shown as a loading control. The asterisk denotes a background band. (B) Wild type or H2AX-/- ESCs (Bassing et al, PNAS, 2002) were transfected with the indicated esiRNAs and imaged after three days.

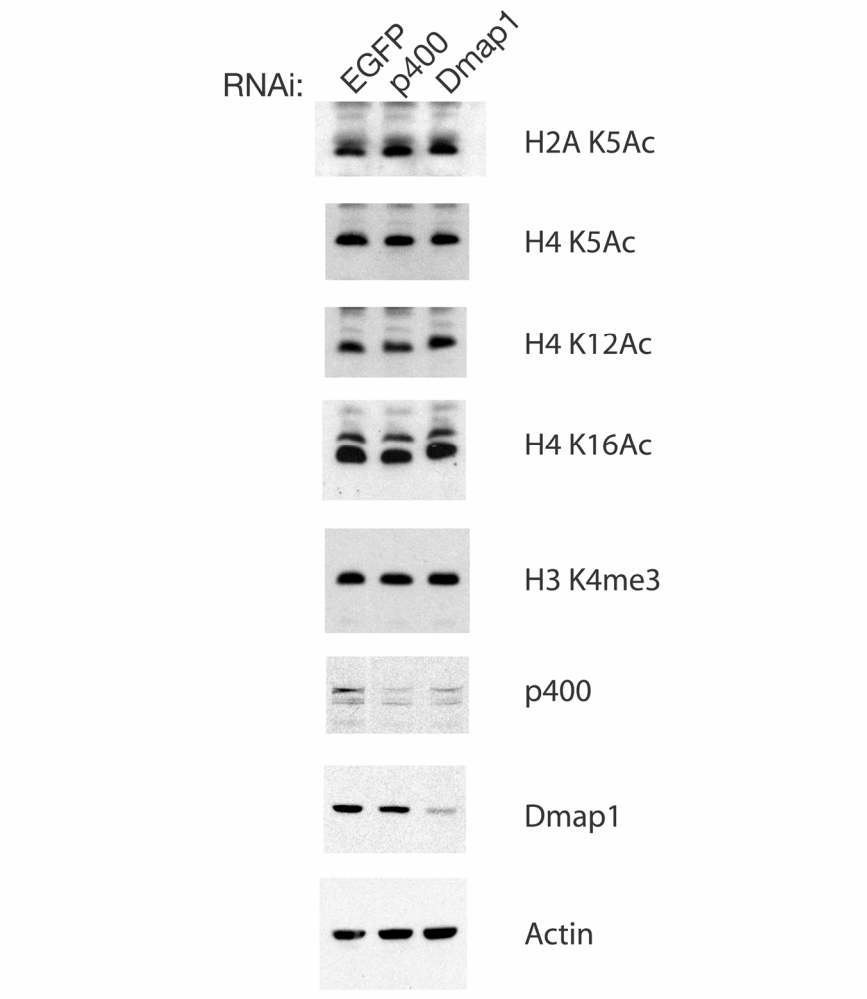

Figure S11. Bulk histone modifications are not affected in ESCs depleted of subunits of Tip60-p400 complex. ESCs were treated with esiRNA as indicated, harvested after 72 hours and subjected to western blotting.

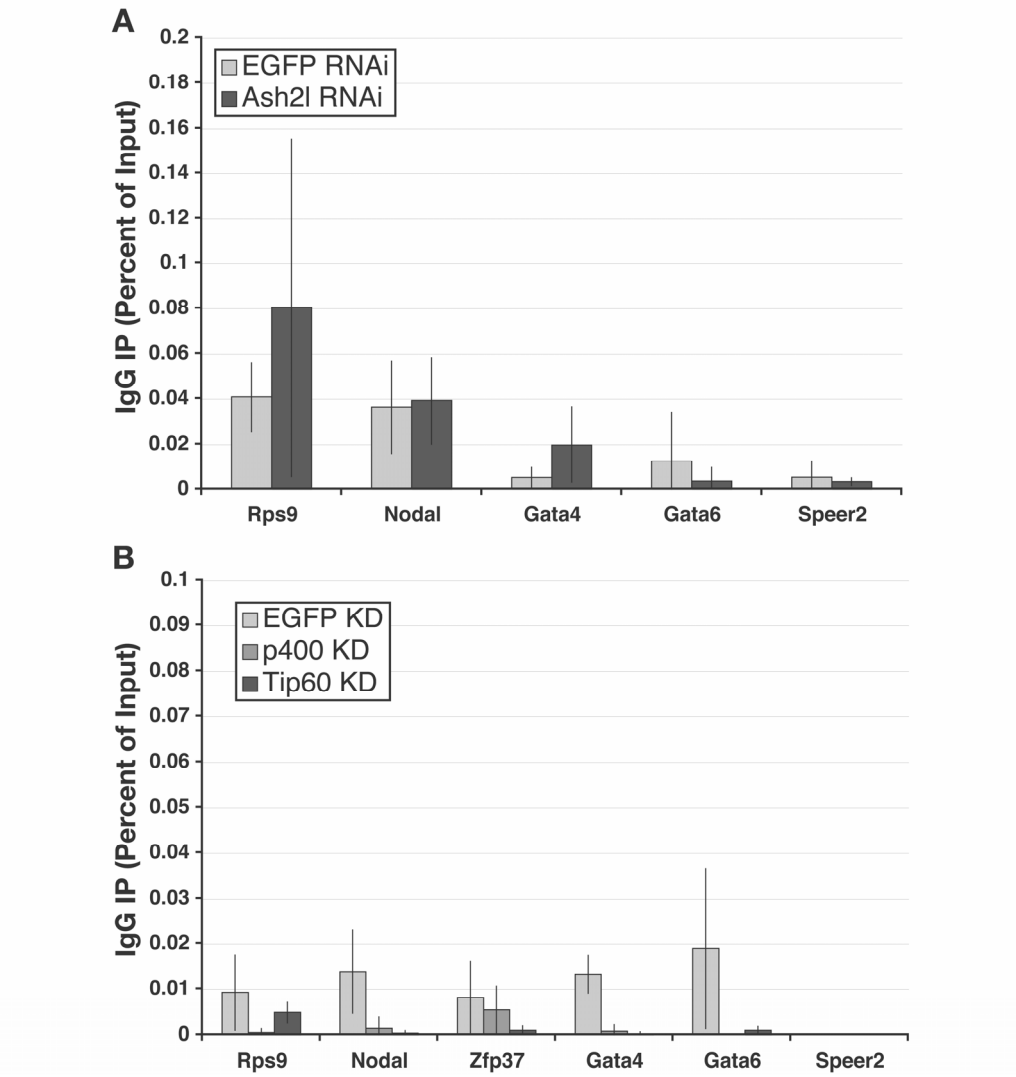

Figure S12. IgG ChIP controls. (A) qPCR as in Figure 5E for IgG ChIP controls. IgG IPs from the same knockdowns as in 5E were subjected to identical qPCRs. Note the minimal amount of IP as indicated on the y-axis. (B) As with (A), these are the IgG control ChIPs for Figure 6D.