Figure S1.

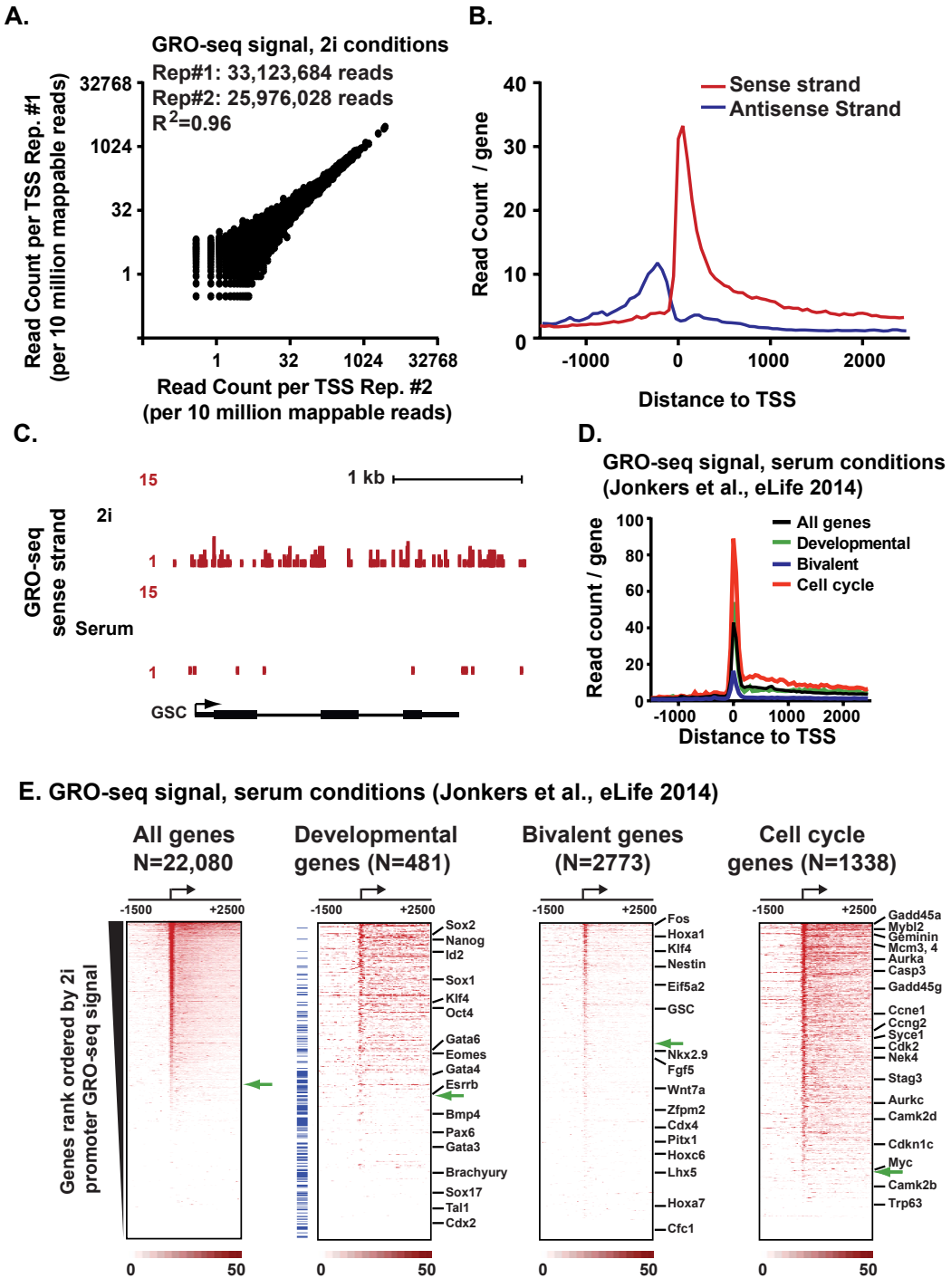


Figure S1. GRO-seq Analysis of Engaged Pol II in ESCs Grown in 2i and Serum-containing Conditions are in Good Agreement. Related to Figure 1.

(A) Independent replicates of GRO-seq performed in 2i (this work) are in good agreement. The number of reads that mapped near each of 22,080 unique mouse TSSs

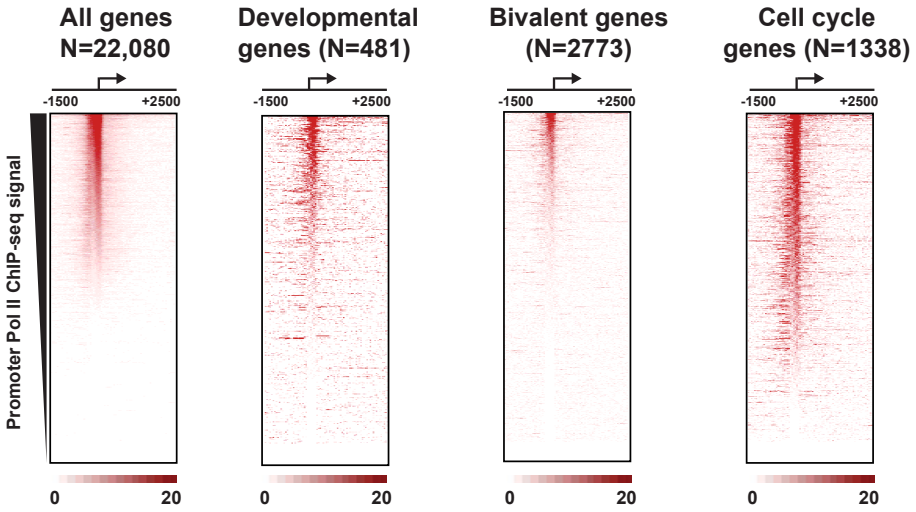
is shown for each replicate. Read counts were density normalized per 10 million mappable reads for each sample (total number of reads=59,099,712).
(B) Composite metagene analysis of sense and antisense strand GRO-seq reads exhibits the expected distribution of engaged Pol II (N=22,080 TSSs). Sense-strand GRO-seq signal (red) is apparent in the downstream direction at protein-coding genes, with the peak indicative of paused Pol II present ~50 nt from the TSS. Anti-sense reads (blue) arise from divergent promoters firing in the opposite direction.
(C) Representative example of GRO-seq signal at bivalent genes in 2i and serum-containing media. Goosecoid (GSC) is expressed at a higher level in 2i than in serum (Marks et al., 2012), and accordingly shows more GRO-seq signal across the gene body. However, GSC does not show evidence of pausing in either media condition.
(D) Average sense-strand GRO-seq signal around all genes (black) as compared to the signal derived from developmental genes (green), bivalent genes (blue) and cell cycle regulators (red).
(E) Heatmap depiction of GRO-seq reads from ESCs cultured in serum, around mouse RefSeq TSSs. Genes are shown as in Figure 1A, and are ranked by decreasing promoter GRO-seq signal (±150nt) observed in cells grown in 2i, emphasizing the strong correspondence between the two GRO-seq data sets. Gene groups shown are identical to those in Figure 1A.

Figure S2.

A.

	GRO-seq 2i	GRO-seq serum	Pol II ChIP-seq serum (Total)	Pol II ChIP-seq serum (1, 8WG16)	Pol II ChIP-seq 2i (8WG16)	Pol II ChIP-seq 2i (S5P)	Pol II ChIP-seq serum (2, 8WG16)
GRO-seq, 2i	1						
GRO-seq, serum	0.88	1					
Pol II ChIP-seq, serum (Total)	0.85	0.86	1				
Pol II ChIP-seq, serum (1, 8WG16)	0.82	0.83	0.94	1			
Pol II ChIP-seq, 2i (8WG16)	0.82	0.79	0.87	0.85	1		
Pol II ChIP-seq, 2i (S5P)	0.82	0.80	0.88	0.85	0.82	1	
Pol II ChIP-seq, serum (2, 8WG16)	0.79	0.77	0.82	0.80	0.80	0.79	1

B. Pol II ChIP-seq signal (2i conditions, 8WG16 antibody, Marks et al. Cell 2012)



C.

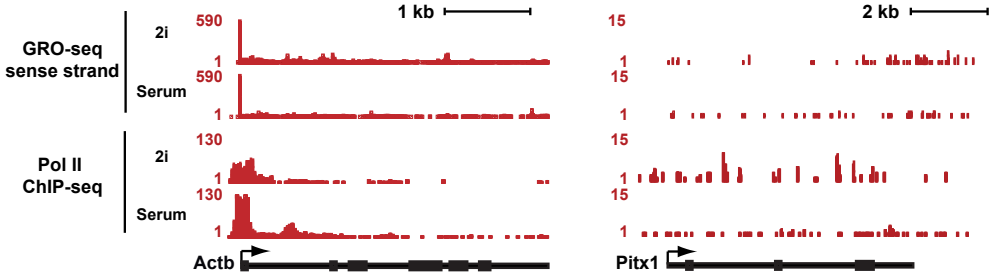


Figure S2. Comparison of GRO-seq and Pol II ChIP-seq Datasets from ESCs. Related to Figure 2.

(A) The level of promoter-proximal Pol II signal was determined around RefSeq TSSs for multiple GRO-seq and Pol II ChIP-seq data sets representing different cell culture conditions and antibodies. Signals were: ±150nt sense-strand signal for GRO-seq and

±150 bp (strand-independent) signal for Pol II ChIP-seq. The correlation between data sets is shown as Spearman's rho. Clustering of data reveals that the two GRO-seq data sets are highly similar, with both showing strong agreement to total Pol II ChIP-seq from ESCs grown in serum.
Data sets in the order listed in the table are:
1) GRO-seq 2i, this work
2) GRO-seq serum (Jonkers et al., 2014)
3) Pol II ChIP-seq, serum (Total Pol II antibody) (Rahl et al., 2010)
4) Pol II ChIP-seq, serum (1, 8WG16 antibody preferring hypophosphorylated Pol II CTD), (Shen et al., 2012)
5) Pol II ChIP-seq, 2i (8WG16 antibody) (Marks et al., 2012)
6) Pol II ChIP-seq, 2i (H5 antibody preferring Pol II CTD phosphorylated at Serine 5) (Tee et al., 2014)
7) Pol II ChIP-seq, serum (2, 8WG16 antibody), (Marks et al., 2012)
(B) Heatmap depiction of Pol II ChIP-seq reads from ESCs cultured in 2i (Marks et al., 2012), centered around RefSeq TSSs as shown in Figure 1A. Genes are ranked by decreasing promoter Pol II ChIP-seq signal (±150nt).
(C) Examples of genes that display pausing across data sets, Actb (left), or lack significant Pol II, Pitx1 (right). Data is shown as in Figure 2C.

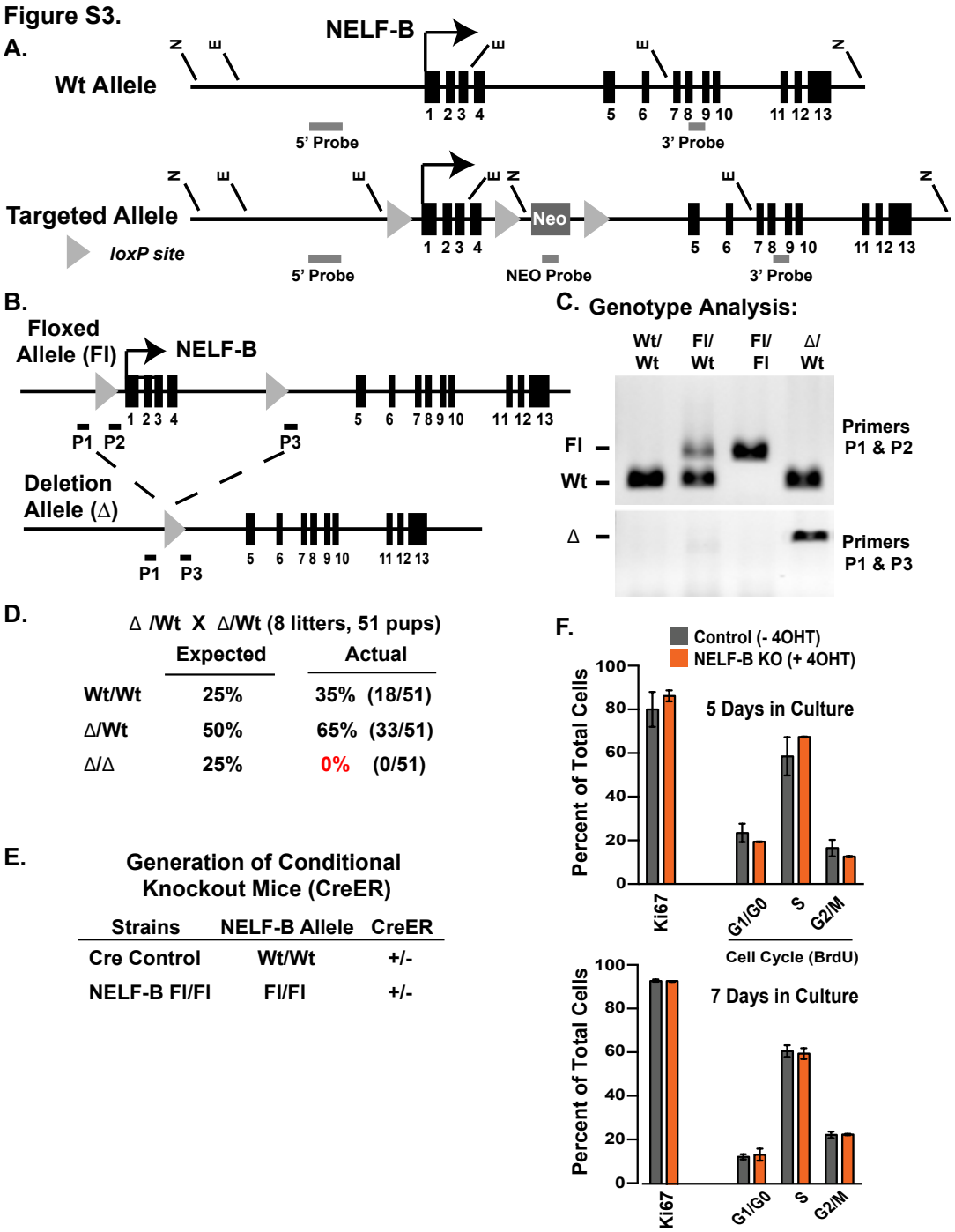


Figure S3. Generation of NELF-B Conditional KO mice. Related to Figure 3.

(A) Schematic of the murine NELF-B gene (mm9; chr2:25,055,232 to 25,067,009) with wild type and targeted alleles. The transcription start site is shown as an arrow and exons indicated as numbered black boxes. For the targeted allele, the positions of the loxP sites (grey triangles) are shown, the selectable neomycin cassette (Neo) is represented as a grey rectangle, and the locations of the 5', 3', and NEO Southern blot

probes are given. Restriction sites used for Southern analysis are noted (E = EcoRI, N = NheI).
(B) Generating conditional NELF-B knockout mice. LoxP sites (grey triangles) flank the promoter and first four exons of NELF-B (Floxed allele, Fl, top). Activation of Cre recombinase results in recombination between LoxP sites and deletion of this region of NELF-B (Deletion allele, Δ, bottom). The locations of genotyping primers (P1, P2, & P3) are shown.
(C) Example of genotype analysis from mouse-tail sections. Primer combinations (indicated at right) are able to distinguish between wild type (Wt), Floxed (Fl), and deletion (Δ) NELF alleles.
(D) Genotype analysis of eight independent litters from NELF-$B^{\Delta/Wt}$ x NELF-$B^{\Delta/Wt}$ intercrosses revealed no viable offspring carrying a homozygous deletion of NELF-B.
(E) Genotypes of Cre control and conditional NELF-B knockout mice used for derivation of murine ESCs are listed. Cre control mice carry one copy of the Cre-ER driver (referred to as Cre control), to account for potential off-target effects of Cre recombinase activity. The experimental mice, called conditional NELF-B KO throughout are homozygous for the NELF-B Floxed allele, and possess a single copy of Cre-ER (NELF-$B^{Fl/Fl}$, Cre-ER +/-). The cells derived from these mice are therefore capable of recombination and NELF-B KO upon treatment with 4OHT.
(F) Control (NELF-$B^{Fl/Fl}$, -4OHT) and NELF-B KO (NELF-$B^{Fl/Fl}$, +4OHT) ESCs were labeled *in vivo* with BrDU to assess cell cycle profiles and immunostained with an antibody that detects Ki67 to examine proliferation. Quantification by flow cytometry revealed no significant differences in cell cycle distribution or proliferation between Control (grey) and NELF-B KO ESCs (orange) after 5 or 7 days of culture. A minimum of 25,000 events was recorded for each population of cells. Data represent N=3 independent clones ±SEM.

Figure S4.

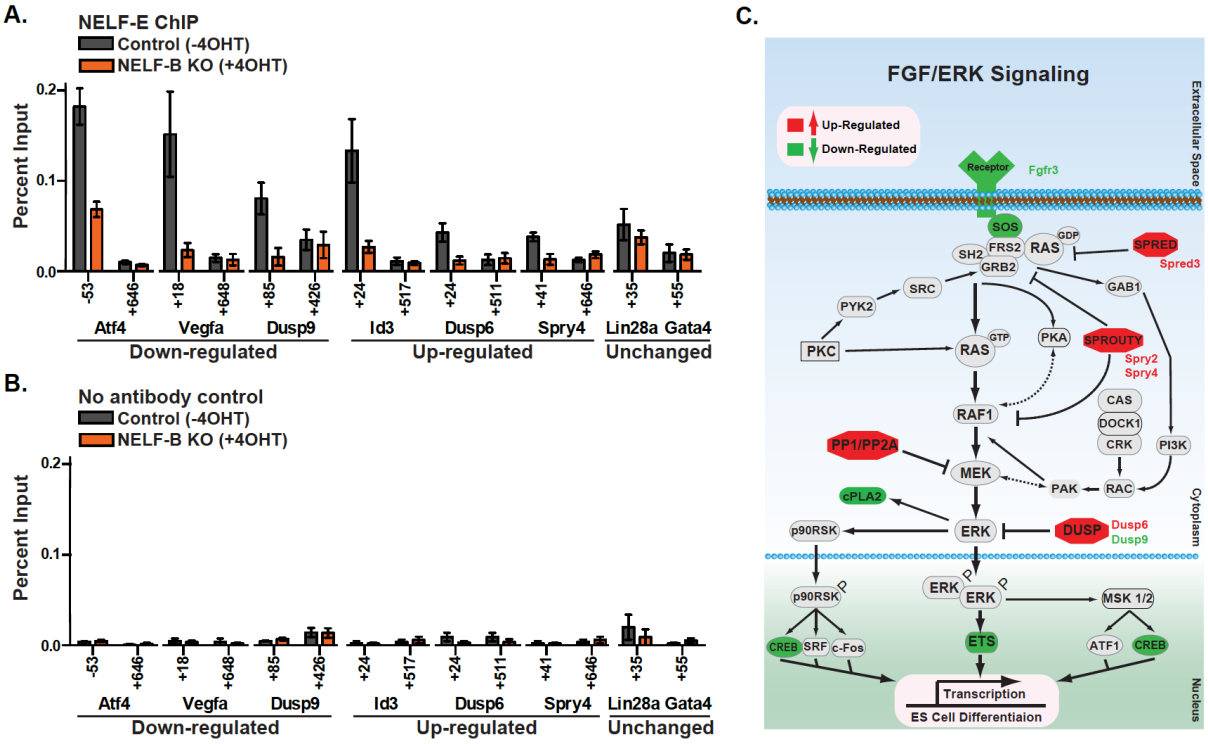


Figure S4. NELF-B KO Disrupts NELF Complex and Perturbs Expression of FGF/ERK Signaling Machineries. Related to Figure 4.

(A and B) NELF-B$^{Fl/Fl}$ ESCs were cultured in 2i media with vehicle or 4OHT treatment for 4 days, and ChIP material isolated at day 5. The percentage of input obtained in Control and NELF-B KO ESCs is plotted for the genes designated. The positions indicated represent the center of each primer pair with respect to the transcription start site for that gene. Data represent N=3 independent clones ±SEM. Immunoprecipitations were performed with an antibody against (A) NELF-E and (B) No antibody as a control. (C) Shown is a schematic representation of proteins involved in the FGF/ERK signaling cascade (adapted from Ingenuity Systems). Proteins encoded by genes that are significantly dysregulated in NELF-B KO ESCs are color-coded based on whether NELF-B loss results in the up-regulation (red) or down-regulation (green) of the indicated gene. Proteins shown in grey are not significantly altered upon NELF-B KO. Several repressors of ERK signaling (represented with an octagonal 'stop-sign' shape) are significantly up-regulated in NELF-B KO ESCs. Conversely, several receptors and activators of FGF/ERK signaling are down-regulated upon loss of NELF-B. 'Ps' represent the phosphorylated form of proteins.

Figure S5.

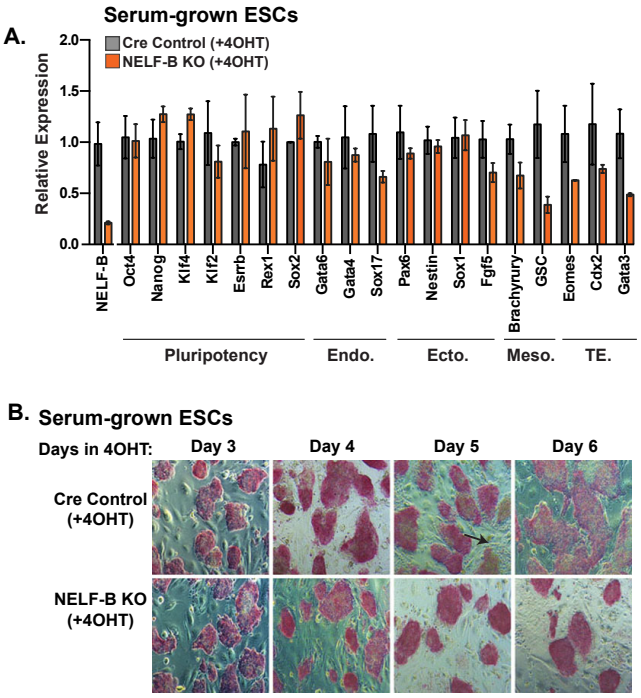


Figure S5. NELF-B KO Reduces Spontaneous Differentiation of ESCs Grown in Serum as Compared to Cre control ESCs treated in parallel with 4OHT. Related to Figure 5.

(A) NELF-B KO neither alters expression of pluripotency markers nor increases expression of lineage-specifying genes. Cre Control (NELF-B $^{Wt/Wt}$, CreER $^{+/-}$) and NELF-B KO (NELF-B $^{Fl/Fl}$, CreER $^{+/-}$) ESCs were maintained undifferentiated on feeders in M15 media plus serum and LIF. All ESCs were grown in 4OHT for 2 days before passage onto new feeders, with an additional 24 hour pulse of 4OHT, and then grown in the absence of 4OHT for 2 additional days. At day 5 after the initial treatment with 4OHT, total RNA was harvested and reverse transcribed to examine expression of NELF-B and select markers of pluripotency and differentiation into different lineages: endoderm (Endo.), ectoderm (Ecto.), mesoderm (Meso.) and trophectoderm (TE) were examined by qPCR. Data represent N=3 independent clones ±SEM.

(B) Representative images show the morphology, size, and alkaline phosphatase activity of Cre Control (NELF-B $^{Wt/Wt}$, CreER $^{+/-}$) and NELF-B KO (NELF-B $^{Fl/Fl}$, CreER $^{+/-}$) ESC colonies. ESCs were plated on feeders in M15 media plus Lif and 4OHT. Cells were fed daily and maintained without passage to allow for spontaneous differentiation. On the days indicated, fixed cells were stained for alkaline phosphatase (AP) and imaged. Note the increase in size of and change in morphology, as well as the appearance of dispersed, alkaline phosphatase negative refractory endodermal cells (arrow) surrounding Cre Control but not NELF-B KO ESC colonies.

Figure S6.

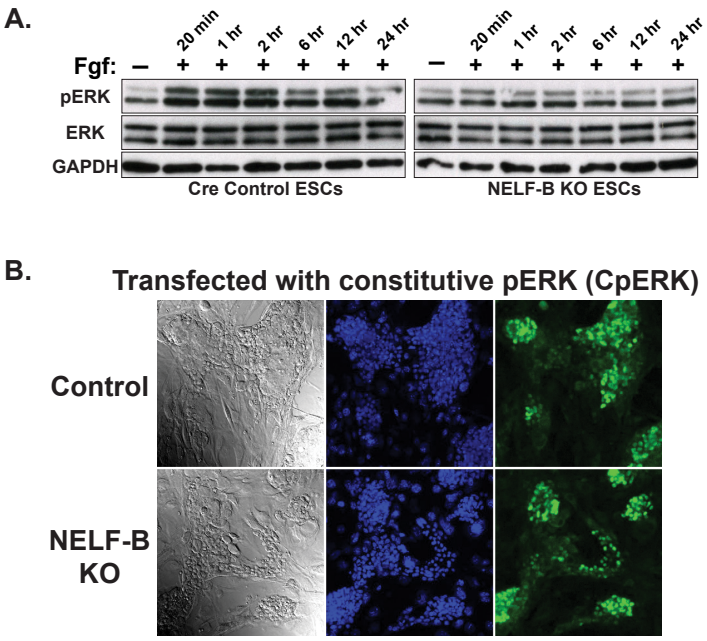
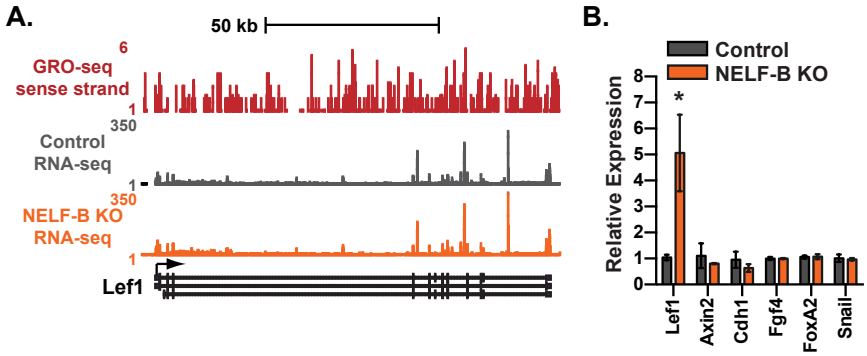


Figure S6. Characterization of Additional ESC clones Following Fgf4 Stimulation or Transfection with Constitutive pERK. Related to Figure 6.

(A) Western analysis reveals a robust activation of phosphorylated ERK (pERK) following Fgf stimulation in Cre control but not NELF-B KO ESCs. Distinct ESC clones of each genotype were cultured and analyzed in an identical manner to Figure 6A. Levels of total ERK are also shown, along with GAPDH as a loading control.

(B) Independent Cre Control and NELF-B KO ESC clones were cultured and analyzed in an identical manner to Figure 6B. Cre control and NELF-$B^{Fl/Fl}$ ESCs were treated with 4OHT for 1 day prior to transfection with a CpERK-expressing plasmid. ESCs plated in serum were then analyzed by immunofluorescence for DAPI and Gata4 on Day 5 following 4OHT treatment. Differentiation is rescued in NELF-B KO ESCs following exogenous ERK activation using a constitutively active ERK protein. Immunofluorescence images for DAPI and Gata4 are shown as merged z-stack sections.

Figure S7.



C.

Wnt target gene expression upon NELF-B knock-down in ESCs (Amleh et al., PLOS One, 2009)

Unchanged								Changed
Atoh1	Cdx1	Fgf20	Gja1	L1cam	NeuroD1	Ptgs2	Tiam1	Cyr61
Axin2	Cdx4	Fgf28	Grem1	LBH	Neurog1	Pttg1	Tnfrsf11a	Dll1
Bglap	Cldn1	Fgf4	Hnf1a	Lgr5	Nkx2-2	Runx2	Tnfrsf11b	Id2
Birc5	Ctla4	Fgf9	Igf1	Met	Nos2	Sall4	Tnfsf9	Lef1
BMP4	Cxcl8	FN1	Igf2	MMP2	NRCAM	Sfrp2	Twist	Stra6
Brachyury	Dkk1	Fosl1	Il6	MMP3	Oct4	Six3	Vcan	Tnfrsf19
c-Jun	Eda	Foxn1	Irx3	MMP7	Ovol1	Snai1	Vegfc	
c-Myc	Edn1	Fst	Isl1	MMP9	Pitx2	Sox17	WISP1	
Cacna1g	Efnb1	Fzd7	ISLR	Msl1	Plaur	Sox9	WISP2	
CCND1	Egfr	Gast	Jag1	Mycbp	Postn	SP5	WISP3	
CD44	Emp1	Gbx2	Klf5	n-Myc	PPARD	Tcf4	Wnt3a	
Cdh1	Enpp2	Gcg	Krt14	Nanog	Prl2c3	Tert	Wrch-1	
Cdkn2a	EphB1							

Figure S7. Lack of Effects of NELF-B Deletion on Lef1/ Wnt Target Genes. Related to Figure 7.

(A) Lef1 does not appear to be a target of NELF-mediated pausing in ESCs grown in 2i conditions. GRO-seq signal is shown, revealing uniform, low-level distribution of Pol II signal across the Lef1 gene body and no evidence of pausing. RNA-seq analysis in 2i shows no sign of up-regulation of Lef1 in cells lacking NELF-B.

(B) Lef1 is up-regulated upon ablation of NELF-B when ESCs are grown in serum conditions. However, this alone does not appear to impact canonical Wnt signaling, since traditional targets of Lef1 are not altered.

(C) Evaluation of the microarray data from (Amleh et al., 2009) similarly indicates that up-regulation of Lef1 in these cells fails to change expression of the vast majority of canonical Lef1/Wnt target genes. 104 Lef1/Wnt target genes in ESCs were obtained from the Nusse Lab, Stanford University (http://www.stanford.edu/group/nusselab/cgi-bin/wnt/target_genes). 98 of 104 Lef1/Wnt targets were unaffected in the previous microarray experiment, suggesting an absence of bona fide activation of Wnt signaling in these cells.

Supplemental Table Legend

Table S1. Genes Significantly Altered in Expression Upon Loss of NELF-B in ESCs Grown in 2i. Related to Figure 3.

Total RNA preparations were obtained from conditional NELF-B KO ESCs on Day 5 after treatment +/- 4OHT. Ribosomal RNA was removed prior to library construction by hybridizing to ribo-depletion beads that contain biotinylated capture probes (Ribo-Zero, Epicentre). RNA was then fragmented and libraries were prepared according to the TruSeq Stranded Total RNA Gold Kit (Illumina) using random hexamer priming. Reads were mapped to the mouse mm9 assembly, and 771 unique genes (Entrez IDs) displayed significantly altered expression upon NELF-B KO (P-value <0.01), as described in Experimental Procedures. This table includes the Entrez Gene IDs, Common Gene Names and RefSeq IDs, as well as P-value and Bejamini-Hochberg adjusted P-value (used for determination of significance).
(See attached excel spreadsheet, Table S1).

Supplemental Experimental Procedures

ESC growth, creation of NELF-B KO mouse and sample preparation

Cell Culture Conditions

All ESC culture was conducted at 37°C in 5% CO_2. Except where noted otherwise, ESCs were passaged every 2 days. For 2i conditions, ESCs were maintained without feeders in knockout DMEM (KO-DMEM), 15% knockout serum replacement (KOSR, Invitrogen), 1 mM NaPyruvate (Millipore), 1% NEAA, 1% BME, 1% Pen/Strep, 1% Glutamax, 1000 U/ml ESGRO, 1 μM MEK inhibitor (PD0325901, Stemgent), and 3 μM GSK3 inhibitor (CHIR99021, Stemgent). To activate Cre recombinase and recombine out the floxed NELF-B allele, NELF-$B^{Fl/Fl}$, $CreER^{+/-}$ ESCs were treated with 100 nM 4OHT (Sigma) for 4 days (unless otherwise indicated). ESCs were then grown without 4OHT for the completion of the experiment.

Non-restrictive media was M15 (15% ES-grade FBS, DMEM, 1% nucleosides, 1% non-essential amino acids (NEAA), 1% BME, 1% Pen/Strep, 1% Glutamax, and 1000 U/mL ESGRO) with ESCs plated on mitomycin-C (Sigma) inactivated feeder cells (STO) cells. ESCs grown in M15 were treated with vehicle or 4OHT for 3 days to trigger recombination. Cre control cells were treated similarly with 4OHT and gene expression analyzed to account for off-target effects of 4OHT or Cre activity in our assays (e.g. Figure S5).

Generation of conditional NELF-B KO mice

The targeting vector for NELF-B KO consisted of a 4.6-kb NotI/SfiI fragment located 5' of the NELF-B promoter, as the 5' homology arm (mm9 chr2: 25,072,048-25,067,537), and a loxP site-flanked neomycin selectable marker. The 3' homology arm (mm9 chr2: 25,064,655-25,059,393) consisted of a 5.3-kb AscI/NruI fragment located 3' of exon 4 to immediately 3' of exon 7. The parent vector, 3loxp3NwCD, consists of the negative selection marker DTA (Diphtheria Toxin). Homologous recombination in C57BL/6 ES cells resulted in the replacement of exons 1-4 of NELF-B with the selectable marker at sequences located from chr2: 25,067,537 to 25,064,655 (Figure S3A). 5', 3', and neomycin probes were used to screen ESC clones for positive integration by Southern blot. 5' and 3' external probes were generated by PCR and cloned into the pCR4.0 backbone. Clones were confirmed for homologous recombination and single neomycin integration. Cre electroporation was used to remove the *neomycin* cassette, and blastocyst injection of clone G3 generated chimeras. From breeding of G3 chimeras with C57BL/6N Tac wildtype mice, heterozygous mice were obtained. All of the resulting Floxed NELF-B mice were maintained on the C57BL/6 background.

To generate mice wherein recombination is driven by 4OHT, NELF-B$^{Fl/Fl}$ mice were crossed with CreER$^{+/+}$ mice (Jackson Laboratory stock #004847; B6;129-Gt(ROSA)26Sor<tm1(cre/ERT)Nat>/J). This CreER knock-in allele constitutively expresses a cytoplasmic Cre protein fused to a mutant version of the estrogen receptor ligand-binding domain. Upon treatment with 4OHT the CreER protein localizes to the nucleus, resulting in nuclear loxP site recombination. CreER$^{+/+}$ mice were also crossed to wild type C57BL/6 mice for derivation of Cre control ESCs (Figure S3E).

Genotype analysis

DNA from tail sections and/or cells was lysed in 25 mM NaOH and 0.2 mM EDTA at 98^{o}C for a minimum of 2 hours. Samples were neutralized with an equal volume of 40 mM Tris-HCl pH 8.0, and 2 uL was used for subsequent genotyping. Taq DNA polymerase (Invitrogen) PCR amplified products were able to resolve NELF-B Wt, Fl, and Deletion (Δ) alleles (Figure S3C), in addition to screening for the presence of Cre. Primers available on request.

Embryo harvest for establishment of ESCs

For ESC derivation, mice were mated to generate litters of Cre control (NELF-B$^{Wt/Wt}$, CreER $^{+/-}$) and Conditional NELF-B KO (NELF-B$^{Fl/Fl}$, CreER $^{+/-}$) embryos. At E3.5, pregnant dams were euthanized, the uterus removed, and blastocysts were flushed out. Blastocysts were transferred to individual wells of a 96-well plate containing mitotically inactivated feeder cells in warmed M15 media. Blastocysts were allowed to attach for 2 days before changing media. The media was subsequently changed every 2 days until the ICM expanded enough to passage. ESCs were trypsinized for 3-5 minutes at 37^{o}C, an equal volume of M15 media containing serum was then added to inactivate the trypsin and colonies were pipetted up and down to dissociate cells. ESCs were transferred to one new well of a 96-well plate containing feeders – without centrifugation – indicating P1. Additional M15 media was added to each well to dilute the trypsin. Media was changed daily and when ESCs reached ~75% confluency, they were passaged from 96-well to 24-well and then to 6-well plates. ESC clones were genotyped and characterized, with aliquots of each frozen for future use.

GRO-seq assay

GRO-seq experiments were performed essentially as described in (Core et al., 2008) in the C2 ESC line, of C57BL/6 genetic background (Gertsenstein et al., 2010), grown in 2i media. Briefly: Nuclei were harvested by swelling trypsinized cells for 10 minutes in 50 mL cold lysis buffer (10 mM Tris pH 7.5, 10 mM NaCl, 3 mM $CaCl_2$, 2 mM $MgCl_2$, 0.5% NP40, 5 mM DTT, 300 mM sucrose, 1 mM PMSF, 2 U/ml SUPERase-IN RNAse Inhibitor (Ambion), 5 tablets Roche protease inhibitors), then gently dounce homogenized, centrifuged and washed once with 50 mL lysis buffer. Nuclei were

resuspended in freezing buffer as in (Core et al., 2008) at the concentration of $5x10^6$ nuclei per 100 uL, and stored at -80°C until use.

To avoid a bias against representation of C-rich regions in the data set, the final concentration of CTP (Roche) was increased in the run-on ten-fold above that used in (Core et al., 2008) to 10 uM. To permit a quick, reproducible stop to the run-on reaction, the reaction was terminated after 5 minutes by the addition of Trizol (Invitrogen). Following base hydrolysis, primary Br-U immunoprecipitation and end repair, the libraries were made using the Illumina small RNA TruSeq kit except the 3' adapter was ligated using T4 RNA Ligase II, truncated KQ enzyme (NEB). A secondary BrU IP was performed before ligation of the 5' adapter (Illumina) followed by a tertiary BrU IP to produce triple-selected run-on RNA. RNA was reverse transcribed and amplified with 15 cycles of PCR (Illumina). Libraries were PAGE purified, selecting for products larger than 140 bp.

RNA-seq of NELF-B KO ESCs

Two distinct conditional NELF-B KO ESC clones were grown +/- 4OHT in 2i media. On day 5 after initiating recombination, total RNA was prepared using Trizol (Invitrogen) and manufacturer recommended methods. DNA contamination was removed using the RNeasy Mini Kit (Qiagen) per manufacturer's instructions, and RNA quality was assessed using a Bioanalyser Nano ChIP (Agilent). Ribosomal RNA was removed prior to library construction by hybridizing to ribo-depletion beads that contain biotinylated capture probes (Ribo-Zero, Epicentre). RNA was then fragmented and libraries were prepared according to the TruSeq Stranded Total RNA Gold Kit (Illumina) using random hexamer priming. Illumina Spike-ins were used for normalization.

Start-seq assay

For Start-RNA-seq, three distinct clones of Control and NELF-B KO ESCs were grown as described for RNA-seq. Start-RNAs were prepared from mouse ESCs essentially as described in (Nechaev et al., 2010). In brief, approximately $2x10^7$ ESCs were trypsinized and collected by centrifugation. After washing with ice-cold 1x PBS, cells were swelled in 10 ml of Swelling Buffer (10 mM Tris [pH 7.5], 10 mM NaCl, 2 mM $MgCl_2$, 3 mM CaCl, 0.3 M sucrose, 0.5% Igepal, 5 mM dithiothreitol, 1 mM PMSF, protease inhibitors and SUPERase-IN RNAse inhibitor (Ambion)) by incubating for 15 minutes on ice followed by 14 strokes with a loose pestle. The dounced cells were spun for 5 minutes at 500x g, the supernatant (cytoplasm) was discarded, the pellet resuspended in 30 ml of Swelling Buffer and spun as above. The supernatant was discarded and the nuclei pellet was resuspended in 1 ml of Swelling Buffer, aliquoted and stored at -80C. Libraries were prepared according to the TruSeq Small RNA Kit

(Illumina). To normalize samples, 15 synthetic capped RNAs were spiked into the Trizol preparation at a specific quantity per 10^7 cells, as in (Henriques et al., 2013).

ChIP

Control and NELF-B KO ESCs grown in 2i as described for RNA- and Start-seq were collected for ChIP analyses. Pol II (Rpb3 subunit) and NELF-E ChIP were performed as previously described (Gilchrist et al., 2010) with several modifications. ESCs were grown to 70% confluence and crosslinked in the tissue culture dish with 1% formaldehyde for 20 minutes at room temperature. Crosslinked ESCs were harvested by scraping, resuspended at 1ml/10^8 cells in sonication buffer containing 0.5% SDS, and lysed for 10 minutes on ice. Sonication was performed in a Biorupter (Diagenode) for 12 rounds of 3X 30-second bursts. ChIP was performed using 36 uL of an antibody targeting the Rpb3 subunit of Pol II, 21 uL of an antibody targeting NELF-E (Sigma #HPA007187), or no antibody as a control using 7.5 x 10^6 ESCs per immunoprecipitation.

ESC Colony morphology and immunostaining

To examine differentiation potential after NELF deletion (Figure 5D), conditional NELF-B KO ESCs treated +/- 4OHT were seeded on feeders at day 2 at a high density of 500,000 cells/well of 6-well TC dish. After plating, cells were cultured for three days without passage in order to allow for spontaneous differentiation. Brightfield images were taken using AxioVision software (Carl Zeiss) and an Axiovert 40 CFL Zeiss microscope.

For immunostaining, ESCs were fixed for 10 minutes at room temperature with 4% paraformaldehyde in PBS, permeabilized using 0.5% Triton X-100 for 15 minutes and blocked for 30 minutes at room temperature using 10% donkey serum (Jackson Laboratories) in PBS. Following this incubation, primary antibody diluted in blocking solution was added to the cells for 1 hour at room temperature in the dark. Cells were then washed 5 times with blocking solution and the secondary antibodies diluted in blocking solution were added, and incubated for 20 minutes at room temperature in the dark. Cells were washed an additional 5 times with 1X PBS, and the slides were mounted with glass coverslips using ProLong Gold Antifade Reagent with DAPI (Invitrogen), and allowed to set overnight. Slides were then imaged on the Zeiss LSM710 confocal microscope and analyzed using ZEN 2011 software.
Antibodies used were: Gata4 (Santa Cruz Biotechnology, #sc25310, 1:50) and Nanog (Cosmo Bio #RCAB0002P-F, 1:1000). All images are representative of experiments performed in several distinct clones from each genotype. ESC colonies of similar size were evaluated, as shown in the bright field images, to facilitate direct comparisons.

To further evaluate ESC responsiveness to differentiation cues and to eliminate potential off-target effects of 4OHT or Cre toxicity (Figure S5B), Cre control and NELF-B KO ESCs were seeded on feeders at day 0 and grown for 6 days without passage in non-restrictive M15 media plus 100 nM 4OHT (total time in 4OHT = 6 days). On each day, colony size and morphology were examined for indications of increased confluency and spontaneous differentiation, such as ruffling activity and the appearance of dispersed, refractory endodermal cells. On days 3, 4, 5, and 6 cells were stained for alkaline phosphatase.

Alkaline Phosphatase Staining

Staining was performed on fixed cells using the alkaline phosphatase detection kit (Stemgent) according to manufacturer recommendations. Briefly, ESCs were fixed with fixation solution for 5 minutes at room temperature. Fixative was removed and cells were rinsed with 2 mls of 1X PBS plus 0.05% Tween-20. Alkaline phosphatase stain solution was added to the cells and plates were incubated at room temperature in the dark for 10 minutes. The staining reaction was quenched by rinsing the cells twice with 1X PBS. Cells were then viewed and imaged using a Carl Zeiss Axiovert 40 CFL microscope fitted with an AxioCam Mrc camera and Axiovision Rel 4.8 software.

RNA isolation and qRT-PCR in non-restrictive media

For expression analysis in NELF-B KO ESCs grown in non-restrictive media, Cre control and conditional NELF-B KO ESCs were grown for five days following initiation of recombination with 4OHT or vehicle treatment (total time in 4OHT=72 hours). Total RNA was prepared using Trizol and RNeasy kits (Qiagen) according to the manufacturer recommendations, including on-column DNase treatment. A total of 250 ng total RNA was used to synthesize cDNA using the iScript cDNA synthesis kit (Bio-Rad). qPCR was performed using Power SYBR Green Master Mix (Applied Biosystems) and then analyzed on a C1000 Touch Thermal Cycler with CFX384 Real-Time System (Bio-Rad). Values were normalized to beta-actin and represent the average ± range for 2 distinct ESC lines with respect to the control. Primer sequences are available upon request.

Fgf4 treatment of ESCs

Cre control and NELF-B KO ESCs were plated on feeders in M15 media with a 24 hour pulse of 4OHT (After 24 hours, cells were washed 2x with PBS to clear cells of 4OHT and fresh M15 media was added). 48 hours after initial plating, ESCs were trypsinized and passaged onto new feeders in M15 with a second 24 hour pulse of 4OHT (Again, after the 24 hour pulse, 4OHT was washed from the cells and fresh M15 added). 96 hours after initial plating, 10 ng/mL Fgf4 (R & D Systems) was used to stimulate the FGF/ERK pathway in the presence of 1 ug/mL heparin (Sigma). ESCs treated for the times indicated were harvested using trypsin and the MEFs were rapidly depleted by

plating ESCs on gelatinized tissue culture plates in M15 media +/- Fgf4/heparin. The ESCs in suspension were pelleted at 1000xG for 5 minutes and western analysis performed as described below.

Analysis of ESCs expressing constitutively active ERK

A previously described expression vector (Robinson et al., 1998) consisting of a fusion between ERK2 and MEK1, with mutations in the nuclear export signal of MEK1, produces a form of ERK2 that is able to shuttle to the nucleus, become phosphorylated and activate target genes. The vector termed ERK2-MEK1-LA, which we shortened to CpERK, was expressed in Cre control and NELF-B KO ESCs.

ESCs were plated in media with 4OHT for 24 hours. The next day, ESCs were trypsinized, re-plated without 4OHT and transfected with CpERK and a GFP expressing plasmid (pEGFP-N1: to assist in the isolation of transfected cells). The next day, ESCs were trypsinized and re-plated on feeders (in M15 media without 4OHT) in either 8-well chamber slides (Lab-TEK #154941) for Immunofluorescence analysis, or in 6-well plates for FACS and quantitative RT-PCR analysis. Initial 4OHT treatment and transfection was performed in 2i media without feeders, so that transfected feeders did not interfere with subsequent FACS isolation step.

Five days after the initial treatment with 4OHT, ESCs in 8-well chamber slides were fixed for 10 minutes at room temperature with 4% paraformaldehyde and washed 3x with 1X PBS. ESCs were permeabilized in 0.5% Triton X-100 for 15 minutes, washed with 1X PBS, and blocked for 30 minutes at room temperature using 10% donkey serum (Jackson Laboratories) in 1X PBS. Following this incubation the block was removed and primary antibody diluted in blocking solution was added to the cells for 1 hour at room temperature in the dark. Cells were washed 5 times with blocking solution and the secondary antibodies diluted in blocking solution were added, and incubated for 20 minutes at room temperature in the dark. Cells were washed an additional 5 times with 1X PBS, and the slides were mounted with glass coverslips using ProLong Gold Antifade Reagent with DAPI (Invitrogen), and allowed to set overnight. Imaging and immunofluorescence was performed as described above.

For gene expression analysis, transfected ESCs plated in 6-well plates were harvested using trypsin and sorted into populations of GFP positive and negative cells with Fluorescence Activated Cell Sorting (FACS), using the BD Biosciences FACSAria II. Total RNA was harvested from sorted cells, reverse transcribed, and analyzed by qPCR as described above.

Western Blotting

At the indicated days of culture, ESCs were lysed directly in the tissue culture plate with RIPA buffer (25 mM Tris-HCl, 150 mM NaCl, 15 NP-40, 1% NaDeoxycholate, 0.1% SDS) containing 1X Halt Protease & Phosphatase Inhibitor Cocktail (Thermo) on ice for 5 minutes. Cell lysates were scraped from dish and transferred to a 1.5 mL tube on ice. Samples were sonicated for one round of 3x 30-second bursts using a Bioruptor. Protein concentration was determined using the BCA Protein Assay Kit (Pierce) and 20 ug of protein was probed.

Antibodies were: NELF-B (custom antibody, described below, 1:1000 dilution), NELF-E (Sigma HPA007187, 1:2000), Actin (Chemicon MAB1501, 1:40,000). A polyclonal rabbit antibody was raised against a purified NELF-B protein comprising residues 142-384 of human NELF-B expressed in *E. coli* with a GST tag. Following purification, the GST tag was removed using a TEV cleavage site between NELF-B and the GST tag. The antibody recognizes the full length NELF-B protein as the major reactive band in a western, and this band is lost upon NELF-B KO (Figs. 3A). For Fig. 6A, antibodies were: NELF-B (1:1000), ERK1/2 (p44/42, Cell Signaling #9102, 1:1000), Phosphorylated ERK1/2 (phos-p44/42, Cell Signaling #4376, 1:1000), and GAPDH (Abcam #ab8254, 1:5000). Westerns shown in Figure 6A and Figure S6A are representative of analysis performed in multiple ESC clones from each genotype.

ESC proliferation assay

Control and NELF-B KO ESCs were passaged in 2i media every two days and seeded at identical densities at each passage. To document ESC colony morphology, bright field images were taken at 10x using an Axiovert 40 CFL Zeiss microscope and processed using Axiovision software. For Figure 3B, adherent cells were released with trypsin on the indicated days in culture and viable, trypan blue negative cells were counted. Cell numbers were used to calculate total fold expansion for the entire time course.

Cell cycle analysis

To examine cell cycle distribution and proliferation by flow cytometry (Figure S3F), Control (NELF-B$^{Fl/Fl}$, -4OHT) and NELF-B KO (NELF-B$^{Fl/Fl}$, +4OHT) ESCs were prepared according to instructions in the FITC BrdU Flow Kit (BD Biosciences #559619, 1:50). Briefly, ESCs were pulse labeled with 10 uM BrdU diluted in 2i media for 45 minutes at 37^{o}C in 5% CO_2. ESCs were washed twice with PBS and released from the dish using trypsin. ESCs were fixed and permeabilized on ice for 20 minutes using kit reagents, treated with DNase to expose incorporated BrdU, and then stained with anti-BrdU and fluorochrome-conjugated Ki-67 (BD Biosciences #558616 1:20) for 20 minutes at room temperature. Stained cells were resuspended in 20 uL of the included

7-AAD solution and the volume brought up to 1 mL with kit staining buffer. Data was acquired on a LSR II Flow Cytometer (BD Biosciences). Proliferation was determined by quantifying the % of Ki67 positive cells. For cell cycle analysis, doublet discrimination was performed, and the correlation of 7AAD staining to BrdU incorporation was used to identify cell cycle positions.

High-throughput Sequencing and Data Analysis

Sequencing and Mapping of GRO-seq Samples

Libraries were paired-end sequenced on an Illumina MiSeq. GRO-seq reads were trimmed for quality and adapter sequence using Cutadapt 1.2.1 , discarding pairs where either mate was trimmed shorter than 15 nt (-m 15 -q 10 --match-read-wildcards). Read pairs originating from rRNA were filtered by aligning against an index consisting of the 45S and 5S transcripts using Bowtie 0.12.8, allowing two mismatches (-m1 -v2 -X1000 --un --max). Unmapped pairs were subsequently aligned to the mm9 index, allowing two mismatches and retaining only uniquely aligned pairs (-m1 -v2 -X1000). BedGraph files were generated from the combined alignments of all technical and biological replicates, reduced to their strand-specific 5'-end mapping locations (total uniquely mappable reads=59,099,712: individual samples had mappable read depths of 33,123,684 for Library#1, and 25,976,028 for Library #2. Agreement between replicates was strong (R^2=0.96, Figure S1A), thus the data was combined for further analysis.

RefSeq annotations were downloaded from the UCSC table browser (February 2012) and filtered to minimize overlapping search spaces. Annotations from chrY, chrM, and random chromosomes were also excluded. Given the difficulty of determining levels of Pol II near promoters vs. within genes for very short genes (e.g. those with gene body windows <750 nt in length), these genes were removed from most analyses, resulting in 22,080 unique RefSeq transcription start sites for which we could characterize Pol II distribution.

Promoter and gene body windows were defined respectively as 'promoter=TSS±150 nt' and 'gene body= +250 to 2250' with respect to the TSS (or gene end when less than 2250 nt long). For each of these windows, the count of intersecting 5'-ends of GRO-seq read pairs was determined. Pausing indices were calculated as the ratio of promoter window density (read pairs/kb) to gene body density (also calculated as read pairs/kb). For display of the data as heatmaps, and creation of metagene analyses, counts of same-strand read pairs intersecting eighty 50 nt bins (from -1500 to +2500 nt relative to each TSS), were determined using an internally developed tool. Data was plotted as heatmaps using Partek Genomics Suite version 6.5.

Statistically significant promoter window count (i.e. “Pol II-occupied”) was defined as ≥13 read pairs in the promoter region (FDR < 0.001), using the method described in (Jothi et al., 2008). Thus, of the 22,080 transcription start sites examined, we defined 11,592 “Pol II- occupied” TSSs.

Analysis of previously published GRO-seq and ChIP-seq data

Published ESC GRO-seq (Jonkers et al., 2014) and Pol II ChIP-seq data (Marks et al., 2012; Rahl et al., 2010; Shen et al., 2012; Tee et al., 2014) were downloaded from the following NCBI GEO and SRA accessions: GSM758173, GSM758174, GSM1186440, GSM1186441, GSM515667, GSM515670, GSM918749, and SRP028688. Files were converted from SRA to .FASTQ format using the SRA toolkit utility fastq-dump. GRO-seq data analysis was performed as described above. To allow direct comparisons between all data sets, ChIP-seq reads were all mapped identically, using Bowtie version 0.12.8 to the mouse mm9 index maintaining only unique hits with a maximum mismatch of 2 (-m1 -v2). Replicates were combined, and hit locations were shifted to the estimated fragment center (as determined using SISSRS (Jothi et al., 2008), and counts per strand merged. TSS-centric binned read counts were determined as for GRO-seq data above, and heatmaps generated as described.

RNA-seq data analysis

Read pairs were filtered, requiring a mean quality score greater than or equal to 20, for both mates, then aligned to the mm9 reference genome using TopHat 2.0.4, reporting up to 10 alignments per read pair (-g 10). Mean fragment sizes and standard deviations were determined using Picard Tools 1.86 CollectInsertSizeMetrics, based on a bowtie 0.12.8 alignment (-m1 -v2 -X10000) of a subset of one million reads to an index of RefSeq transcripts, and passed to TopHat using the --mate-inner-dist and --mate-std-dev parameters. A total of 113,817,860 and 108,809,117 read pairs were successfully aligned for the Control and NELF-B KO samples, respectively. Read counts were calculated per Entrez gene ID using htseq-count 0.6.0, and differentially expressed genes identified using DESeq 1.12.1 under R 3.0.1, with default parameters, for RefSeq annotations included in the GRO-seq analysis described above. Histone genes were removed to mitigate challenges associated with multi-mapping and artifactual cross-cluster splicing. At an adjusted p-value threshold of <0.01, 771 genes (Entrez IDs) were identified as differentially expressed. UCSC Browser tracks displaying read coverage normalized per million mappable fragments were generated from the combined replicates per condition.

Start-RNA data analysis

Read pairs were first trimmed for adapter sequence using Cutadapt 1.2.1, discarding those containing reads of trimmed length shorter than 20 nt (-m 20 -q 10). Pairs were aligned to an index composed of the 15 synthetic spike-in RNAs, and then to the mm9 reference genome using bowtie 0.12.8 (-m1 -v2 -X1000 --best). The total number of mappable reads for control ESCs was 82,685,356 (Clone 1, 27,484,164; Clone 2, 32,563,316; Clone 3, 22,637,876) and for NELF-B KO ESCs was 77,635,525 (Clone 1, 25,031,185; Clone 2, 18,888,508; Clone 3, 33,715,832).

For each of the three conditional NELF-B ESC clones used in this experiment (+/- 4OHT), reads were normalized based on sequencing depth (i.e. per million read pairs uniquely aligned to the mouse genome), scaling the sample with the higher uniquely mappable read pair count to match the lower. Read pairs aligning to each spike-in were determined for each ESC clone (+/- 4OHT), based on the depth normalized read counts. Ratios of these values were calculated for each NELF-B ESC clone (+4OHT/-4OHT). All the ratios were very close to 1 (Clone 1: 1.01, Clone 2: 1.09, Clone 3: 1.01), thus, no further normalization was performed based on the spike-in RNAs. Combined bedGraph files were generated for the control and NELF-B KO samples containing the count of Start-RNA 5-prime ends, at single nucleotide resolution.

Distributions of Pol II elongation complex positions relative to TSSs were determined per-gene, based on the length of Start-RNAs from 5-prime end (designating the TSS) to the 3-prime end (designating the position of Pol II at the time the RNA was isolated). Read pairs included in this analysis were those with 5-prime ends mapping near TSSs (within a ±100 nt window about each TSS), from the combined N=3 data sets for each condition. Mean counts were plotted across a range of distances, applying a 10nt smoothing window, for genes up-regulated, down-regulated, and unchanged upon NELF-B KO in RNA-seq experiments.

Fold changes of Start-RNA reads associated with genes affected by RNA-seq were determined based on the count of normalized 5-prime ends falling within 100 nt of a TSS, summed per-condition. Significance of these changes was assessed using the Mann-Whitney U-test. TSSs showing fewer Start-RNAs (down-regulated TSSs), increased Start-RNAs (up-regulated TSSs) or unchanged Start-RNA levels (unchanged) were identified using a two-tailed t-test, assuming unequal variance and a p-value threshold of 0.1. Significance of the overlap between up-regulated and down-regulated genes identified via total RNA-seq and Start- RNA-seq were determined using a hypergeometric test, implemented in R.

Gene lists, Gene Ontology and Pathway analysis

Promoters of RefSeq transcripts were identified as bivalent based on published data from the Bernstein lab (Ku et al., 2008). Inclusion in GO categories of Cell Cycle (GO:0007049) or Development (GO:0045165, GO:0048864, GO:0007498), was determined using MGI accession as an intermediary, and matching to lists downloaded manually from geneontology.org. MGI accession numbers were attributed to each RefSeq transcript via the MGI associations to sequence table (ftp://ftp.informatics.jax.org/pub/reports/MRK_Sequence.rpt), pairing by gene symbol in cases where no RefSeq ID match was available.

Gene Ontology analysis was performed using an online tool and standard parameters (Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (http:david.abcc.ncifcrf.gov/home.jsp). Functional analyses were generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA®, Version: 18488943, QIAGEN Redwood City, www.qiagen.com/ingenuity), last updated March 2014. The full list of down- and up-regulated genes (Table S1) was uploaded into IPA to identify the top Molecular and Cellular Functions categories perturbed in NELF KO ESCs.

Supplemental References

Gertsenstein, M., Nutter, L.M., Reid, T., Pereira, M., Stanford, W.L., Rossant, J., and Nagy, A. (2010). Efficient generation of germ line transmitting chimeras from C57BL/6N ES cells by aggregation with outbred host embryos. PloS one *5*, e11260.

Jothi, R., Cuddapah, S., Barski, A., Cui, K., and Zhao, K. (2008). Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. Nucleic acids research *36*, 5221-5231.