CellPress

Molecular Cell
Article

Histone-Fold Domain Protein NF-Y Promotes Chromatin Accessibility for Cell Type-Specific Master Transcription Factors

Andrew J. Oldfield,[1,4] Pengyi Yang,[1,2,4] Amanda E. Conway,[1] Senthilkumar Cinghu,[1] Johannes M. Freudenberg,[1] Sailu Yellaboina,[3] and Raja Jothi[1,2,*]
[1]Systems Biology Section, Laboratory of Molecular Carcinogenesis
[2]Biostatistics Branch
National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, USA
[3]CR Rao Advanced Institute of Mathematics, Statistics & Computer Science, Hyderabad, Andhra Pradesh 500 046, India
[4]Co-first authors
*Correspondence: jothi@mail.nih.gov
http://dx.doi.org/10.1016/j.molcel.2014.07.005

SUMMARY

Cell type-specific master transcription factors (TFs) play vital roles in defining cell identity and function. However, the roles ubiquitous factors play in the specification of cell identity remain underappreciated. Here we show that the ubiquitous CCAAT-binding NF-Y complex is required for the maintenance of embryonic stem cell (ESC) identity and is an essential component of the core pluripotency network. Genome-wide studies in ESCs and neurons reveal that NF-Y regulates not only genes with housekeeping functions through cell type-invariant promoter-proximal binding, but also genes required for cell identity by binding to cell type-specific enhancers with master TFs. Mechanistically, NF-Y's distinct DNA-binding mode promotes master/pioneer TF binding at enhancers by facilitating a permissive chromatin conformation. Our studies unearth a conceptually unique function for histone-fold domain (HFD) protein NF-Y in promoting chromatin accessibility and suggest that other HFD proteins with analogous structural and DNA-binding properties may function in similar ways.

INTRODUCTION

Master transcription factors (TFs) establish and/or maintain cellular identity by orchestrating distinct profiles of gene expression that are faithfully transmitted through cell division. The simplified paradigm for master TFs relies on the premise that their expression is spatiotemporally restricted to one or few cell types or lineages that depend on their activity (Oestreich and Weinmann, 2012). This assumption may have unintentionally led to the underappreciation of the roles "housekeeping" or other essential TFs, expressed ubiquitously or in a multitude of cell types, might play in cell specification. Recent studies, however, have shown that many ubiquitous factors, previously thought to have an exclusive housekeeping function, have additional cell type-specific roles (Chia et al., 2010; Cinghu et al., 2014; Golan-Mashiach et al., 2012; Kagey et al., 2010; Kim et al., 2010a; Pijnappel et al., 2013). This emerging body of evidence indicates that we are yet to fully appreciate the significance of ubiquitously expressed proteins in many settings, and argues for a comprehensive reassessment of the roles other housekeeping proteins may play in cell specification.

NF-Y, also known as the CCAAT-binding factor CBF, is a ubiquitously expressed heterotrimeric TF composed of NF-YA, NF-YB, and NF-YC subunits, all of which are conserved from yeast to human (Maity and de Crombrugghe, 1998). NF-Y binds to the CCAAT box, which occurs at ~30% of all eukaryotic promoters (Dolfini et al., 2012a). NF-YB and NF-YC dimerize via their histone-fold domains (HFDs) before associating with NF-YA (Romier et al., 2003), which harbors both DNA-binding and transactivation domains. The crystal structure of NF-Y bound to DNA shows that while NF-YA makes sequence-specific DNA contacts, NF-YB/NF-YC interacts with DNA via nonspecific HFD-DNA contacts (Nardini et al., 2013). The key structural feature of the NF-Y/DNA complex is the minor-groove interaction of NF-YA, which induces an ~80° bend in the DNA. The structure and DNA-binding mode of NF-YB/NF-YC HFDs are similar to those of the core histones H2A/H2B, TATA-binding protein (TBP)-associated factors (TAFs), the TBP/TATA-binding negative cofactor 2 (NC2α/β), and the CHRAC15/CHRAC17 subunits of the nucleosome remodeling complex CHRAC (Nardini et al., 2013). Yet, unlike H2A/H2B, which lack sequence specificity, NF-YB/NF-YC interaction with NF-YA provides the NF-Y complex with sequence-specific targeting capability as well as nucleosome-like properties of nonspecific DNA binding, a combination that allows for stable DNA binding.

NF-Y, largely described as a transcription activator via its promoter-proximal binding, is a key regulator of cell-cycle progression in proliferating cells (Benatti et al., 2011; Bungartz et al., 2012; Hu and Maity, 2000), with its activity often downregulated during cellular differentiation and senescence (Bungartz et al., 2012; Farina et al., 1999). In addition to binding core promoters, NF-Y has also been shown to bind enhancer elements away from transcription start sites (TSSs) (Fleming et al., 2013; Testa et al.,

2005), but its function and mechanism of action at these distal regulatory elements remain to be elucidated.

Consistent with its role in cell-cycle regulation, NF-Y is required for ESC and hematopoietic stem cell proliferation (Bungartz et al., 2012; Dolfini et al., 2012b; Grskovic et al., 2007). While *NF-YA* heterozygous mice are normal and fertile, *NF-YA* null mice die prior to 8.5 dpc (Bhattacharya et al., 2003), suggesting an essential role for NF-Y in early mouse embryonic development. Interestingly, conditional deletion of NF-YA in postmitotic mouse neurons induces progressive neurodegeneration (Yamanaka et al., 2014), which suggests a role for NF-Y that is independent of its role in cell-cycle regulation, as has also been shown in hepatocytes (Luo et al., 2011).

Given the relatively high expression of one or more NF-Y subunits in mouse oocytes (Su et al., 2004) and the inner cell mass (ICM) of the mouse blastocyst (Yoshikawa et al., 2006), we set out to determine NF-Y's role in ESCs, derivatives of the ICM. Although NF-Y has been predominantly studied in the context of its promoter-proximal binding at key cell-cycle genes, given the prevalence of NF-Y targeting sites at greater distances from TSSs (Fleming et al., 2013), we investigated NF-Y's function and mechanism of action at distal regulatory elements. We demonstrate a requirement for all three NF-Y subunits in the expression of core ESC self-renewal and pluripotency genes, and in the maintenance of ESC identity. Through genome-wide occupancy and transcriptomic analyses in ESCs and neurons, we show that not only does NF-Y regulate genes with housekeeping functions through cell type-invariant promoter-proximal binding, but also genes required for cell identity by binding to cell type-specific enhancers with master TFs. We present evidence that NF-Y's distinct DNA-binding mode facilitates a permissive chromatin conformation and promotes enhanced binding of master ESC TFs at enhancers. Our studies unearth a function for NF-Y in promoting chromatin accessibility and specification of cell identity.

RESULTS

Genomic Profiling of NF-YA, NF-YB, and NF-YC Binding Sites in Mouse ESCs

To gain insight into NF-Y-mediated transcriptional regulation in early embryogenesis, we investigated the genome-wide occupancy of all three subunits of the NF-Y complex in mouse ESCs using chromatin immunoprecipitation (ChIP) followed by sequencing (ChIP-Seq). ChIP-Seq analyses revealed enrichment for NF-Y occupancy near TSSs of annotated genes (Figure 1A), consistent with NF-Y's preference for binding and recruiting RNA polymerase II and general TFs to various CCAAT motif-containing promoters (Kabe et al., 2005). Binding of all three subunits to the promoters of known NF-Y targets including *Cdc25c*, *Rnf5*, and *Zic3* (Grskovic et al., 2007; Yamanaka et al., 2014) attested to the sensitivity of the ChIP-Seq data (Figure 1B). ChIP followed by quantitative polymerase chain reaction (qPCR) analysis of certain sites, either previously demonstrated as NF-Y-bound regions (*Cdc25c*, *Rnf5*, *Zic3*) or highly enriched for NF-Y binding (*Khsrp*), further validated the quality and the reproducibility of the ChIP-Seq data (Figure 1C). Analysis of ChIP-Seq data sets using the SISSRs peak-calling algorithm (Jothi et al., 2008; Narlikar and Jothi, 2012) identified a total of 4,642, 2,774, and 2,675 binding sites (peaks) for NF-YA, NF-YB, and NF-YC, respectively, with at least 7-fold or more ChIP over input enrichment ($p < 10^{-3}$). As expected, de novo sequence motif analysis identified the known RRCCAATVR consensus motif within the binding sites for all three subunits (Figure 1D).

NF-Y Binding Requires All Three Subunits and Correlates with Gene Expression

Although a vast majority of the binding sites for NF-YA, NF-YB, or NF-YC overlap with those for one or both of the other two subunits (Figure 1E), about one-third of the NF-YA binding sites did not show statistically significant evidence for the cobinding of the other two subunits. To determine whether NF-YA alone binds certain genomic loci, we knocked down *NF-YA* or *NF-YC* using siRNAs for 48 hr (see Figures S1A and S1B available online) and analyzed their binding patterns at sites bound by all three NF-Y subunits (Figure 1B) as well as those bound only by NF-YA (Figure S1C). ChIP-qPCR analyses show that upon NF-YC knockdown (KD), NF-YA occupancy on DNA is compromised even at sites that seem to bind only NF-YA (Figure S1D), indicating that NF-YC is essential for NF-YA binding. Low NF-YB and NF-YC occupancy at sites defined as bound only by NF-YA compared to those bound by all three subunits (Figure S1E) suggest that the failure to call peaks for NF-YB/NF-YC at these sites is likely due to weaker ChIP-Seq signals that did not pass the statistical threshold for peak calling, and/or less immunoefficient antibodies for NF-YB/NF-YC compared to that of NF-YA. Together with the fact that nearly all of the NF-YB-only and NF-YC-only sites contain the CCAAT motif, we conclude that NF-Y binding requires all three subunits. Although we cannot completely rule out that a very low and/or transient occupancy of NF-YA alone or NF-YB/NF-YC heterodimer may exist, our data show that such events are highly unlikely if not sparse, consistent with in vitro biochemical studies showing NF-Y binding requiring all three subunits (Sinha et al., 1995). Therefore, hereafter, we will refer to the union of all NF-YA, NF-YB, and NF-YC binding sites (5,359 in total) as NF-Y binding sites.

Relative to annotated RefSeq genes, a vast majority of all NF-Y binding sites are located near genic regions (Figure 1F), suggesting a role for NF-Y in regulation of gene expression. Indeed, the levels of NF-Y binding at gene promoters correlate positively with gene expression (Figure 1G), which is consistent with NF-Y's established role as a transcription activator (Dolfini et al., 2012a).

NF-Y Co-occupies Enhancers with ESC-Specific Master TFs

NF-Y has predominantly been reported to bind in close proximity to TSSs. While we confirm a strong enrichment for NF-Y binding sites around TSSs (within 500 bp of TSSs), with a clear preference for binding to the region immediately upstream of the TSS (Figure 2A), nearly half of all NF-Y binding sites are located at distances greater than 500 bp from the TSSs (Figure 2B). These data suggest a role for NF-Y in ESC transcription regulation by binding to sites distal to TSSs. To investigate the role distal sites may play in the regulation of gene expression, and

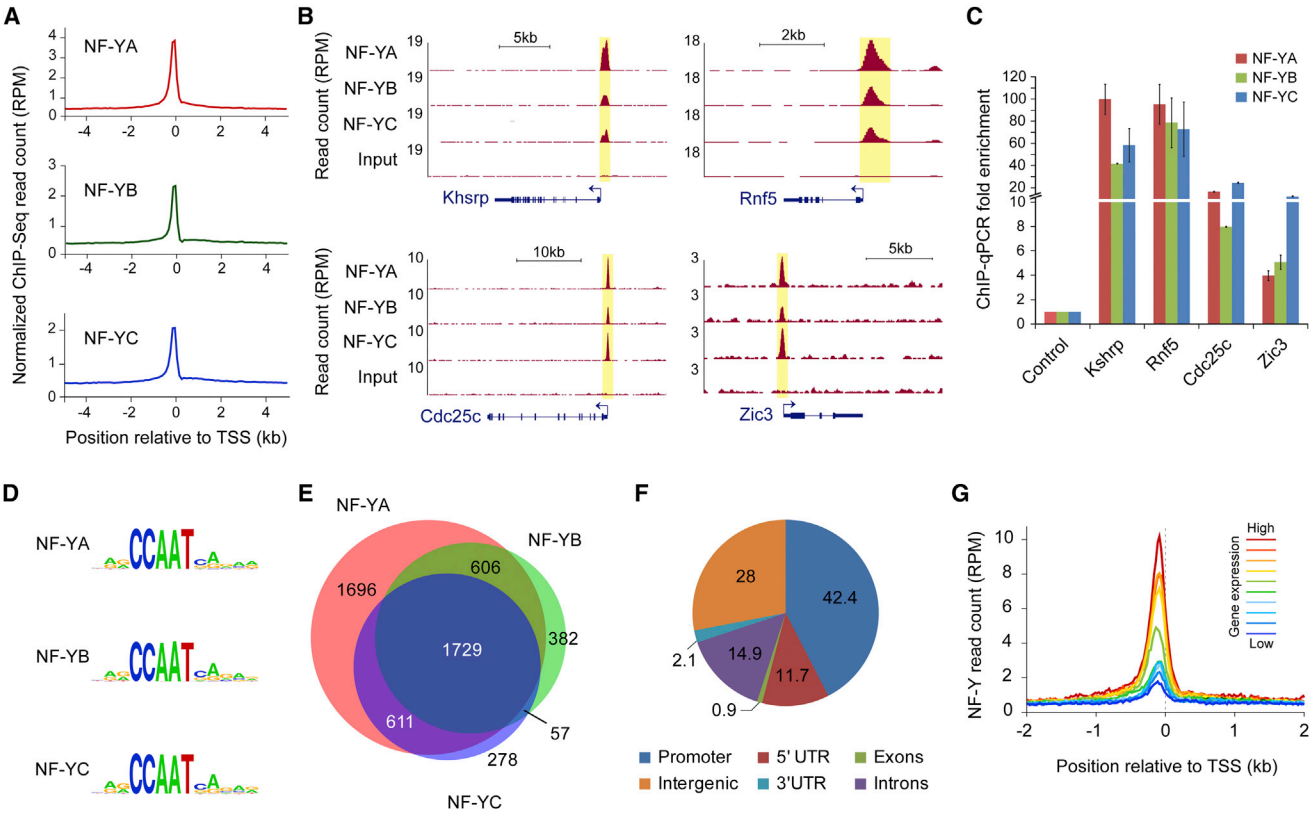


Figure 1. Genomic Profiling of NF-YA, NF-YB, and NF-YC Binding Sites in Mouse ESCs

(A) ChIP-Seq read density plot showing NF-YA, NF-YB, and NF-YC occupancy near TSSs of all mouse RefSeq genes.

(B) Genome browser shots showing NF-YA, NF-YB, and NF-YC occupancy at gene promoters.

(C) ChIP-qPCR validation of the NF-Y occupancy at sites highlighted in (B). Error bars represent SEM of three experiments.

(D) Consensus sequence motifs enriched within NF-YA, NF-YB, and NF-YC sites using de novo motif analysis.

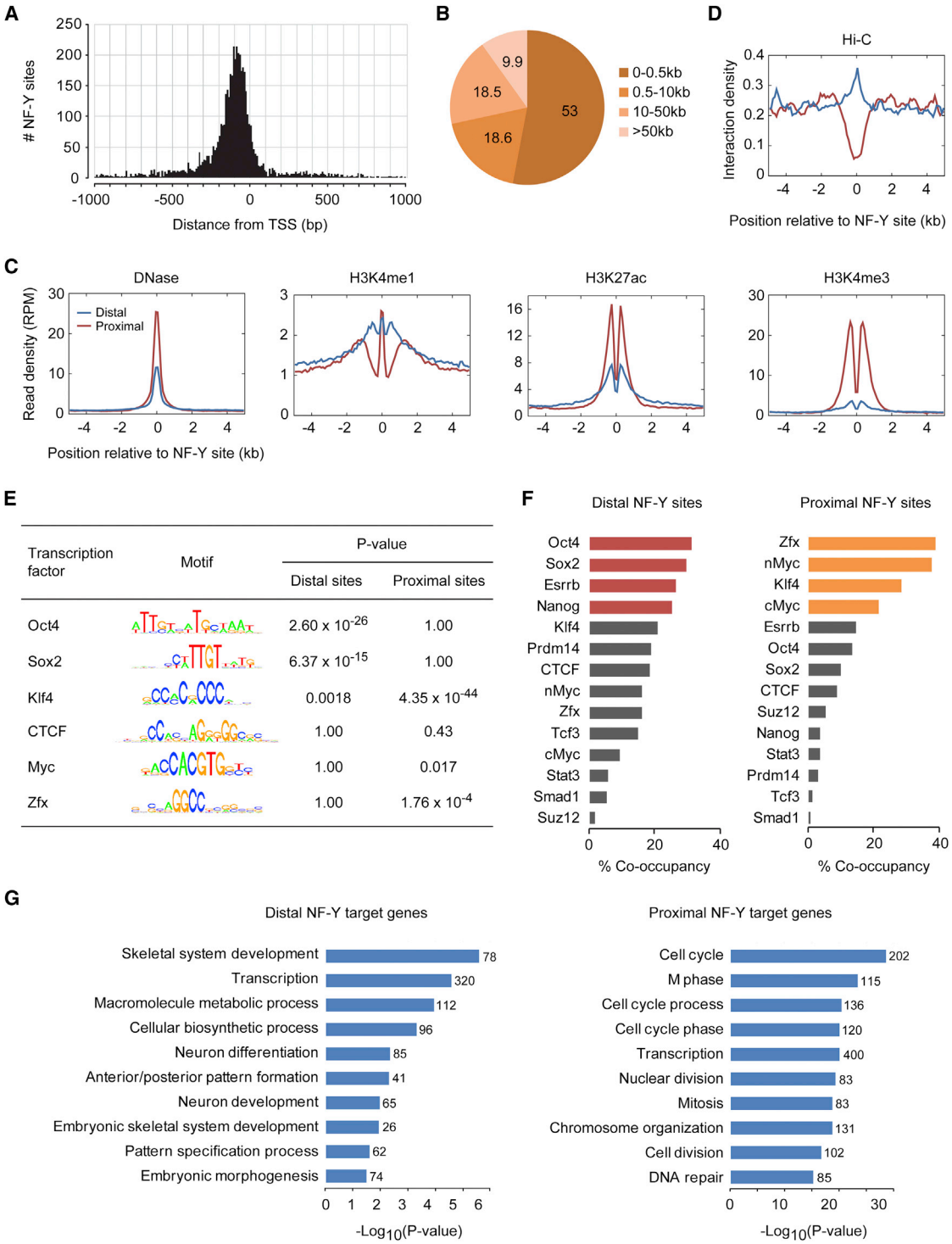
(E) Overlap among NF-YA, NF-YB, and NF-YC sites (peaks).

(F) Genome-wide distribution of the NF-Y sites (union of the NF-YA, NF-YB, and NF-YC sites). Promoter, 500 bp upstream of TSS.

(G) Correlation between gene expression levels and NF-Y occupancy near TSSs. See also Figure S1.

how they might differ in function compared to that of promoter-proximal NF-Y sites, we examined published ChIP-Seq data sets for histone modifications, DNase I hypersensitivity, Hi-C interaction profiles, 14 different TFs, and other genomic features. As one would expect from their proximity to TSSs, nearly 80% of all promoter-proximal NF-Y sites overlap with CpG islands (Figure S1F), consistent with ~85% of all promoters overlapping CpG islands. In contrast, only about one-tenth of distal NF-Y sites overlap with CpG islands (Figure S1F), a fraction similar to that observed for enhancers. Furthermore, ChromHMM annotation of chromatin states (Ernst and Kellis, 2012) shows a majority of distal NF-Y sites with chromatin features reminiscent of enhancers (Figure S1G). Chromatin marks at distal NF-Y sites revealed open chromatin architecture marked by high DNase I activity, accompanied by enhancer marks H3K4me1 and H3K27ac but not the promoter mark H3K4me3 (Figure 2C). Additionally, enrichment of Hi-C interaction density at distal but not proximal NF-Y strongly supports the notion that distal NF-Y sites are within enhancers (Figure 2D).

To gain insight into the mechanism underlying NF-Y's function at enhancers, we studied the frequency of known TF binding motifs within distal and proximal NF-Y sites. We found that distal NF-Y sites are significantly enriched for sequence motifs bound by master ESC TFs known to bind distal enhancers, including Oct4 and Sox2, but not for motifs bound by ubiquitous TFs, such as CTCF (Figure 2E). In contrast, promoter-proximal NF-Y sites are significantly enriched for sequence motifs bound by Klf4, cMyc, and Zfx, all known to preferentially bind to promoters (Figure 2E) (Chen et al., 2008). Furthermore, analysis of NF-Y binding in conjunction with other TF binding data revealed an extensive colocalization of NF-Y with ESC-specific master TFs Oct4, Sox2, Nanog, and Esrrb at distal NF-Y sites, and with TFs Zfx, cMyc, and Klf4 at promoter-proximal NF-Y sites (Figure 2F). Consistent with NF-Y's co-occupancy with master ESC TFs at distal NF-Y sites, gene ontology analysis revealed enrichment for genes with roles in early embryonic development among those targeted by distal NF-Y sites (Figure 2G). Interestingly, genes with NF-Y sites at proximal promoters are enriched for housekeeping functions including cell cycle and proliferation. Taken together, these data suggest a potential role for NF-Y in the regulation of the core pluripotency transcription network through distal enhancer binding.



Transcription factor	Motif	P-value	
		Distal sites	Proximal sites
Oct4		2.60×10^{-26}	1.00
Sox2		6.37×10^{-15}	1.00
Klf4		0.0018	4.35×10^{-44}
CTCF		1.00	0.43
Myc		1.00	0.017
Zfx		1.00	1.76×10^{-4}

Figure 2. NF-Y Co-occupies Enhancers with ESC-Specific Master TFs

(A) Frequency distribution of NF-Y sites within 1 kb of TSSs.

(B) Proximity of NF-Y sites to TSSs.

(C) Levels of DNase I hypersensitivity, H3K4me1, H3K27ac, and H3K4me3 near proximal and distal NF-Y sites.

(D) Hi-C interaction density near proximal and distal NF-Y sites.

(E) TF binding motifs enriched within distal versus proximal NF-Y sites. CTCF is not enriched at either set of NF-Y sites.

(F) Co-occupancy between ESC TFs (Chen et al., 2008; Ho et al., 2011; Ma et al., 2011; Marson et al., 2008) and NF-Y at distal and proximal sites.

(G) Gene ontology categories enriched among distal and proximal NF-Y target genes.

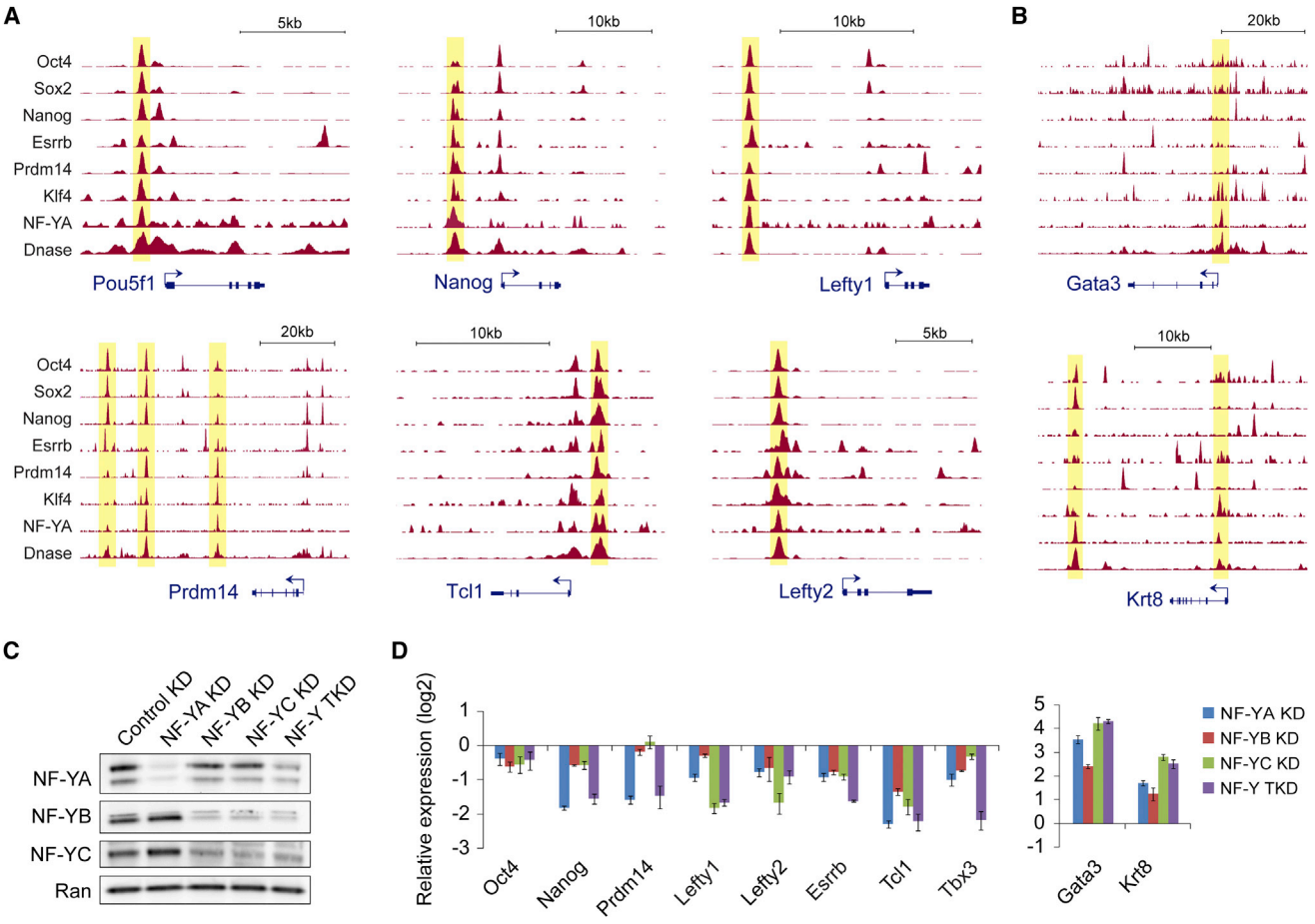


Figure 3. NF-Y Binds to and Regulates Core ESC Self-Renewal and Pluripotency Genes

(A and B) Genome browser shots showing NF-Y co-occupancy with master ESC TFs (Chen et al., 2008; Ma et al., 2011; Marson et al., 2008) at the enhancers of ESC identity genes (A), and at the promoters/enhancers of differentiation/developmental genes (B).

(C) Western blot analysis of NF-YA, NF-YB, and NF-YC in *NF-YA* knockdown (KD); *NF-YB* KD, *NF-YC* KD, or NF-Y triple KD (TKD; KD of all three NF-Y subunits) ESCs 96 hr after siRNA transfection. Ran is used as a loading control.

(D) RT-qPCR analysis of relative mRNA levels of ESC identity genes (left), and differentiation markers (right) in *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and NF-Y TKD ESCs compared to control KD ESCs 96 hr after siRNA transfection. Data normalized to *Actin*, *HAZ*, and *TBP*. Error bars, SEM. See also Figure S2.

NF-Y Is Required for the Expression of Core ESC Self-Renewal and Pluripotency Genes

Genes targeted by NF-Y include many key pluripotency-associated genes that have been implicated in the control of ESC identity (Figure 3A). These include genes encoding the master ESC TFs Oct4, Nanog, Prdm14, Esrrb, Tcl1, Tbx3, Rex1, Lefty1, and Lefty2. More importantly, in a majority of cases, NF-Y colocalizes with master ESC TFs at distal enhancers previously demonstrated (Kagey et al., 2010; Zhang et al., 2013) to interact with promoters of these ESC identity genes (Figures 3A and S2A). This suggests that NF-Y might be regulating these genes via enhancer-promoter interactions. In addition to its binding at the enhancers of self-renewal and pluripotency genes, NF-Y binds the promoters of genes with known roles in cell cycle (Figure S2B) as well as development (including *Gata3*, *Krt18*, *Krt8*, *Srf*, *Jun*, *Prom1*, and *Lef1*) (Figures 3B and S2C).

To establish a functional role for NF-Y in the regulation of these genes, we used RNAi to knock down subunits of the NF-Y complex individually or in combination (Figures 3C and S2D) and profiled gene expression changes using qPCR following reverse transcription (qRT-PCR). As a positive control, we observed significant changes in the expression of key cell-cycle genes in cells depleted of individual NF-Y subunits (Figure S2E) accompanied by severe proliferation defects and growth arrest (Figures S2F and S2G), corroborating NF-Y's established role in proliferation (Dolfini et al., 2012a). Depletion of NF-Y subunits decreased the expression of pluripotency-associated factors by 2- to 5-fold, and increased the expression of differentiation genes by up to an order of magnitude (Figure 3D). Together, these findings show that the NF-Y complex is required for the expression of key ESC self-renewal and pluripotency genes, and support the conclusion that NF-Y regulates not only genes with housekeeping functions through its canonical promoter-proximal binding but also genes required for ESC identity by cobinding distal enhancers with cell type-specific master regulators.

All Three NF-Y Subunits Are Required for ESC Self-Renewal and Pluripotency

To investigate whether NF-Y is essential for the maintenance of ESCs in an undifferentiated state, we examined colony morphology and alkaline phosphatase (AP) staining 96 hr after RNAi-mediated silencing of individual or all subunits of the NF-Y complex (Figure S2D). We observed notable self-renewal defects accompanied by loss of characteristic ESC colony morphology and AP staining (ESC marker), all consistent with ESC differentiation (Figure 4A). *NF-YA* KD ESCs showed profound proliferation defects compared to *NF-YB*/*NF-YC* KD ESCs, perhaps due to NF-YA being the limiting subunit of the NF-Y complex with relatively low expression compared to the other two subunits (Figure S2H). Nevertheless, these cellular changes are consistent with NF-Y's control over both housekeeping and ESC-specific functions through distinct modes of regulation (Figures 3 and S2). At the molecular level, immunostaining experiments after *NF-YA* silencing revealed a drastic loss of Nanog protein levels, while NF-YC protein levels remain unchanged (Figure 4B). Furthermore, RT-qPCR analysis of ESCs depleted of NF-Y subunits revealed a significant increase in the expression of differentiation and lineage markers (Figure 4C) in addition to a significant decrease in the expression of pluripotency genes including *Oct4*, *Nanog*, *Prdm14*, *Esrrb*, *Tcl1*, *Tbx3*, and nodal antagonists *Lefty1* and *Lefty2* (Figure 3D), which are among the earliest to be downregulated during ESC differentiation. Based on these data, we conclude that the NF-Y complex is essential to maintain ESC identity, and that the depletion of even one of the three NF-Y subunits leads to the loss of the pluripotent state.

NF-Y Is an Essential Component of the Core Pluripotency Network

To determine the extent to which NF-Y regulates gene expression programs influencing ESC self-renewal and pluripotency, we used microarrays to profile global gene expression in ESCs transfected with a control siRNA or siRNA(s) targeting individual or all of the NF-Y subunits (Figures 3C and S3A). Consistent with gene expression changes observed in the RT-qPCR analysis (Figures 3D and 4C), global gene expression changes upon silencing of all three NF-Y subunits correlated highly with those after individual knockdowns (Figure S3B). Notably, many of the NF-Y targets lie at a "crossroads" of cell-fate determination during embryonic development (Figure S4). Using a stringent criteria (FDR < 0.05; fold change ≥ 2.0), we identified 847 genes that were differentially expressed in cells depleted of all three NF-Y subunits (NF-Y TKD) (Figure 4D), a vast majority of which were also differentially expressed in individual knockdowns (Figure 4E). Notably, while a majority of the downregulated genes harbor NF-Y binding sites either within 5 kb upstream of the TSS or the gene body, only ~18% of the upregulated genes have NF-Y binding sites (Figure 4D). With much of the upregulation attributable to indirect effects, this is a clear indication that NF-Y plays a predominantly activating role in ESCs.

We observed a strong positive correlation between gene expression changes upon NF-Y TKD ESCs and those during the normal course of embryoid body differentiation (Figure 4F). Consistent with this observation, principle component analysis of gene expression profiles during differentiation of ESCs into three different lineages showed ESCs depleted of NF-Y cluster away from undifferentiated ESCs (Figure 4G). Analysis of published gene expression microarray data from ESCs depleted of key pluripotency-associated factors revealed a strong correlation between the global gene expression changes upon loss of individual or all NF-Y subunits and those observed after the loss of *Oct4*, *Sox2*, *Nanog*, *Ncl*, *Tet1*, or *Klf*s (Figure 4H). These data, together with NF-Y's co-occupancy with and regulation of master ESC TFs (Figures 2F, 3A, and 3D), place NF-Y as a member of the core pluripotency network.

NF-Y Binding at Enhancers Is Cell Type Specific, whereas NF-Y Binding at Promoters Is Cell Type Invariant

To investigate whether NF-Y's function at enhancers is broadly conserved in other cell types, we compared NF-Y binding sites in ESCs with previously published NF-Y ChIP-Seq data from ESC-derived neuronal progenitors (NPCs) and terminal neurons (Tiwari et al., 2012). Remarkably, while promoter-proximal NF-Y binding is largely conserved across the three cell types, distal NF-Y binding is specific to cell type (Figure 5A). Panther pathway analysis of genes associated with either proximal or distal NF-Y binding sites in ESCs or neurons revealed that while genes associated with proximal NF-Y binding are enriched for housekeeping functions in both cell types, genes associated with distal NF-Y binding in ESCs or neurons are highly enriched for pathways known to have critical roles in ESC maintenance or neuronal function, respectively (Figure 5B). This suggests that while promoter-proximal NF-Y binding is largely cell type invariant and regulates several housekeeping functions, distal NF-Y binding is largely cell type specific and controls pathways characteristic of cell identity. Hence, we reasoned that we should observe loss of ESC-specific NF-Y binding at distal sites immediately after ESCs exit from the pluripotent state. As predicted, NF-YA ChIP followed by qPCR in differentiating ESCs revealed a significant loss of NF-Y binding at distal but not cell type-invariant proximal sites (Figure 5C).

NF-Y Co-occupies Active Enhancers with Cell Type-Specific Master TFs

To gain additional insight into mechanisms underlying NF-Y's cell type-specific binding, we performed de novo motif analysis of 200 bp regions centered on distal NF-Y binding sites in ESCs and neurons, in search of sequence motifs for potential cell type-specific cofactors. We found that distal NF-Y binding sites in ESCs are significantly enriched for sequence motifs known to bind Oct4, Sox2, and Nanog (Figure 5D), consistent with their colocalization at distal NF-Y sites (Figure 2F). In contrast, distal NF-Y binding sites in neurons are enriched for a sequence motif known to bind CTCF, an essential master regulator of neural development (Golan-Mashiach et al., 2012). In good agreement, Oct4, Sox2, Nanog, and Prdm14 colocalize specifically at distal but not proximal NF-Y binding sites in ESCs (Figures 5E, S5A, and S5B), whereas these factors are not enriched at distal or proximal sites bound by NF-Y in neurons. In contrast, CTCF co-localizes specifically at distal NF-Y binding sites in neurons, but not in ESCs (Figures 5E and S5B), and CTCF is not enriched at

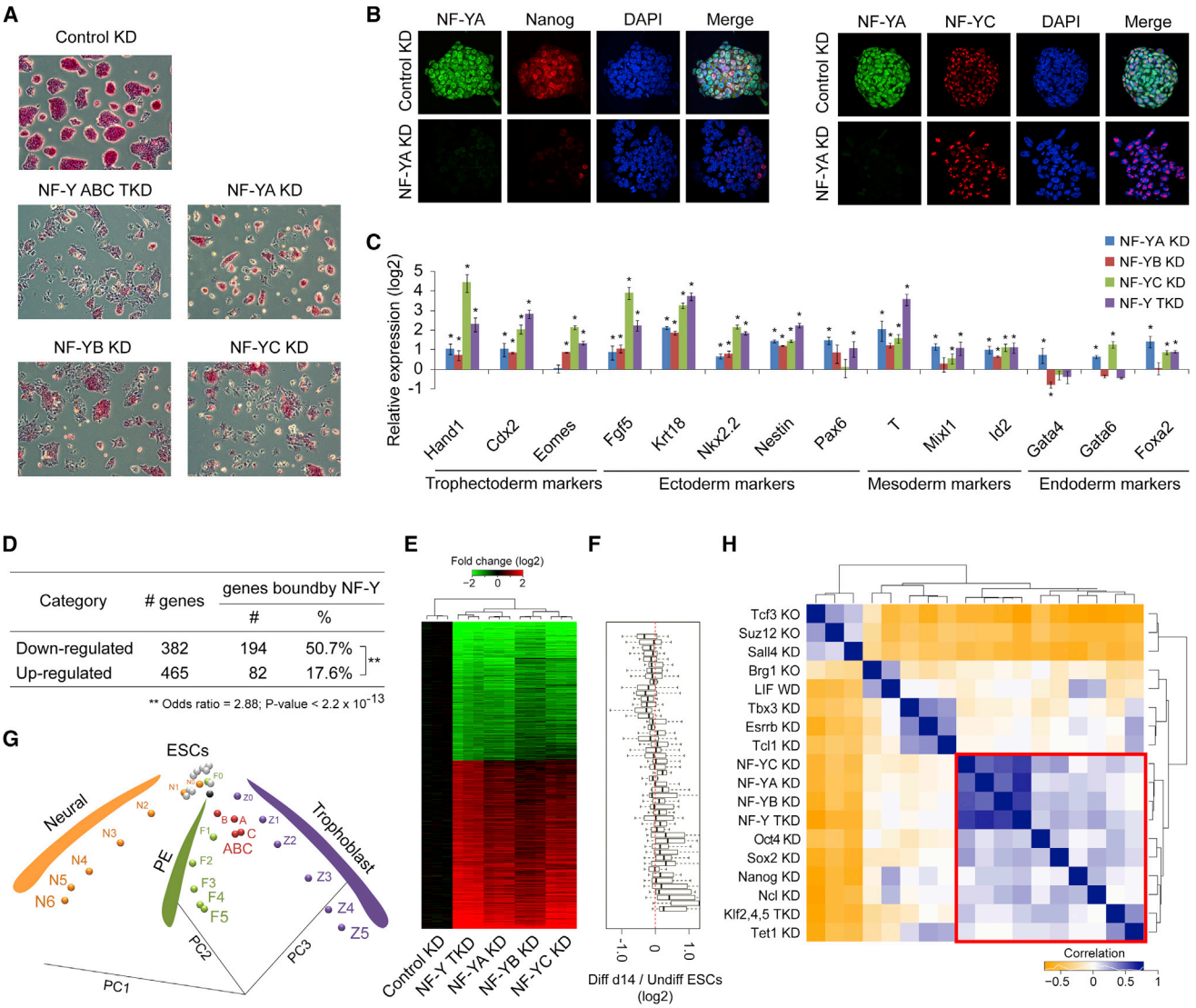


Figure 4. NF-Y Is Required for ESC Self-Renewal and Is an Essential Component of the Core Pluripotency Network

(A) Morphology and alkaline phosphatase staining of *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and *NF-Y* TKD ESCs 96 hr after siRNA transfection.

(B) Coimmunostaining of NF-YA and Nanog (left)/NF-YC (right) in control and *NF-YA* KD ESCs 96 hr after siRNA transfection. Nuclei counterstained by DAPI.

(C) RT-qPCR analysis of relative mRNA levels of differentiation markers in *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and NF-Y TKD ESCs compared to control KD ESCs 96 hr after siRNA transfection. Data normalized to *Actin*, *HAZ*, and *TBP*. Error bars, SEM.

(D) Number of genes up- and downregulated in NF-Y TKD ESCs 96 hr after siRNA transfection, and percent of those that are also bound by NF-Y.

(E) Gene expression fold changes upon *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, or NF-Y TKD, measured 96 hr after siRNA transfection. Experiments performed in triplicates. Only genes that were differentially expressed (FDR $\leq$ 0.05 and fold change $\geq$ 2) in NF-Y TKD are shown.

(F) Relative gene expression changes of NF-Y-regulated genes during the normal course of embryoid body formation (day 14) compared to undifferentiated ESCs. NF-Y-regulated genes, ordered as in Figure 4E, were grouped into 50 bins.

(G) Principal component analysis of gene expression profiles showing *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and NF-Y TKD ESCs (all in red) alongside differentiation of ESCs into three different lineages (Nishiyama et al., 2009). Z0–Z5, trophoblast lineage (ESCs differentiating into trophoblast cells from day 0 to day 5; purple); N0–N6, neural lineage (ESCs differentiating into neural lineage from day 0 to day 6; orange); F0–F5, primitive endoderm (PE; embryonal carcinoma cells differentiating into PE from day 0 to day 5; green); ESCs, white; control KD ESCs, black; A, *NF-YA* KD; B, *NF-YB* KD; C, *NF-YC* KD; ABC, NF-Y TKD.

(H) Correlation between global gene expression changes upon *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and NF-Y TKD and those observed after KD or knockout (KO) of other ESC-associated factors, as previously reported. Rows/columns ordered based on unsupervised hierarchical clustering. TKD, Triple KD; WD, Withdrawal. See also Figures S3 and S4.

proximal NF-Y sites in neurons or ESCs (Figure 5E). Examination of published ChIP-Seq data for various TFs in ESCs and neurons further confirmed NF-Y binding to distal enhancer sites with cell type-specific master TFs: NF-Y colocalizes with master ESC TFs Oct4, Sox2, Nanog, and Prdm14 in ESCs, and with master neuronal regulators CTCF, NPAS4, CBP, and CREB in neurons

(Kim et al., 2010b) (Figure 5F). Notably, in neurons, NF-Y cobinds with CTCF, NPAS4, CBP, and CREB at distant enhancer elements controlling stochastic promoter choice and alternative isoform expression of clustered protocadherin genes (*Pcdhα*, *Pcdhβ*, and *Pcdhγ*) (Figures S6A and S6B), which are essential for neuronal diversity (Hirayama et al., 2012).

Examination of chromatin features at ESC-specific distal NF-Y binding sites revealed an open chromatin architecture marked by high levels of H3K4me1 and H3K27ac but lacking H3K27me3 in ESCs, all hallmarks of active enhancers, as opposed to high levels of H3K27me3 and no H3K4me1 and H3K27ac in neurons (Figure 5G, left). Interestingly, however, neuron-specific distal NF-Y binding sites, despite not bound by NF-Y in ESCs, had surprisingly high levels of H3K4me1 and H3K27ac but no H3K27me3 in ESCs, comparable to those observed in neurons, indicating that neuron-specific NF-Y binding sites are "primed" in ESCs (Figure 5G, right). This, in combination with low to moderate levels of Oct4/Sox2 occupancy (Figures 5F and S5A), suggests a potential role for Oct4/Sox2 in priming lineage-specific enhancers in addition to regulating ESC-specific gene expression programs. Together, these data support a model whereby NF-Y uses distinct mechanisms to regulate housekeeping and cell identity genes: NF-Y regulates housekeeping genes by binding to their proximal promoters, and regulates genes required for cell identity by binding to active enhancers along with cell type-specific master TFs (Figure 5H).

NF-Y Enhances Oct4, Sox2, Nanog, and Prdm14 Binding on DNA, but Not Vice Versa

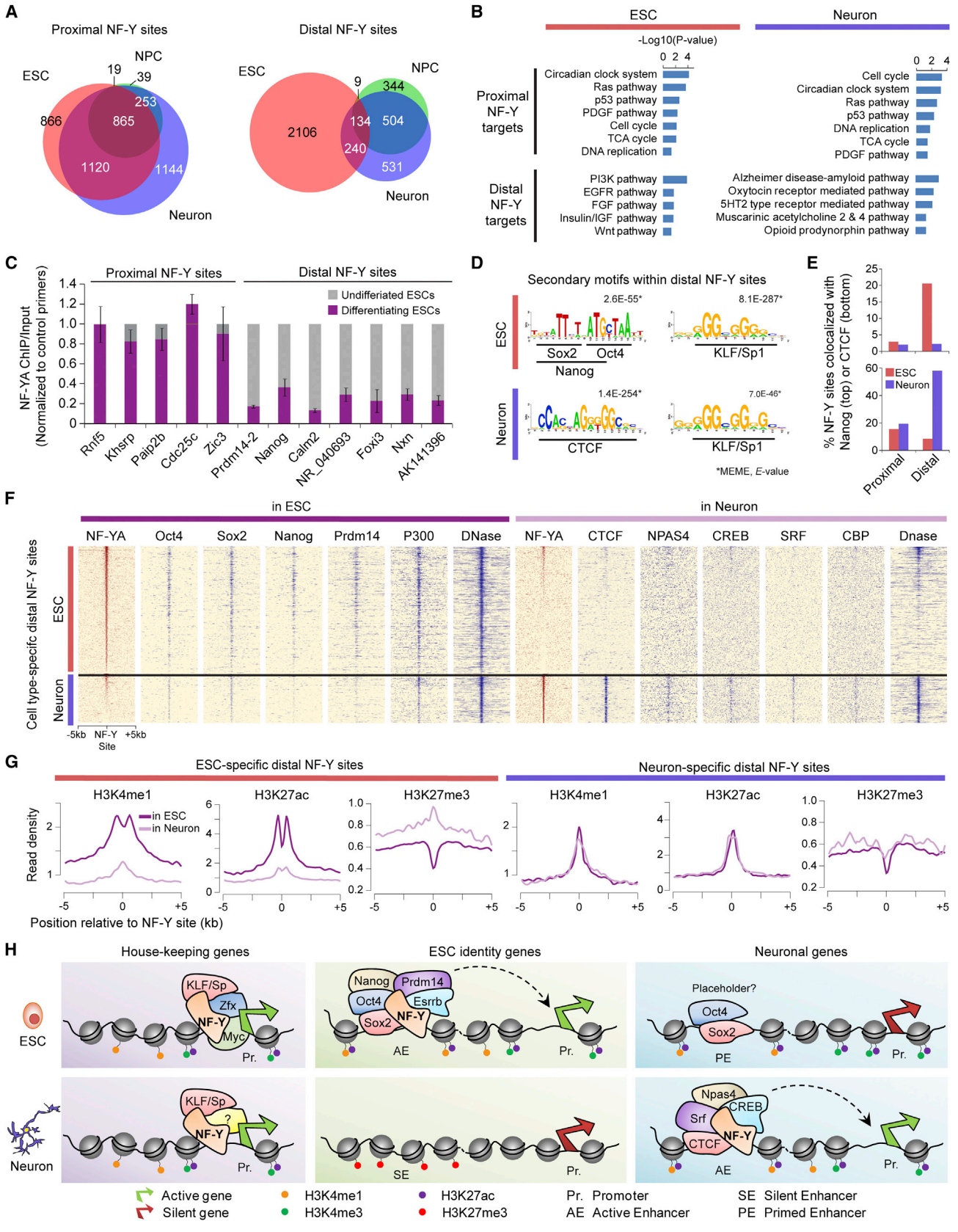
To gain insight into the mechanisms underlying NF-Y's function at active enhancers, we examined DNase I hypersensitivity at sites bound by NF-Y and various other TFs in ESCs. Intriguingly, NF-Y binding sites are the most hypersensitive to DNase I digestion compared to other TF sites analyzed, and are about 2-fold more hypersensitive than those bound by master ESC TFs Oct4, Sox2, Nanog, and Prdm14 (OSNP) (Figure 6A). In agreement with this observation, NF-Y binds to sites that are relatively nucleosome depleted (Figure 6B). In contrast, OSNP bind to sites within/near nucleosomal DNA, which is consistent with previous reports showing their penchant for binding "closed chromatin" (Soufi et al., 2012; Teif et al., 2012).

To investigate whether the increased accessibility at NF-Y binding sites is due to NF-Y potentially recruiting a chromatin remodeler to facilitate permissive chromatin, we analyzed our previously published ChIP-Seq data for Brg1 and Ino80 (Ho et al., 2011; Wang et al., 2014), the ATPase subunits of the chromatin remodeling complexes esBAF and INO80, respectively. We found that the occupancy levels of chromatin remodeling factors at the NF-Y binding sites are relatively low compared to those observed at many other TF binding sites (Figures S7A and S7B), suggesting that increased hypersensitivity at NF-Y sites cannot be simply explained by higher levels of chromatin remodeling activity. Examination of NF-Y and Brg1 (esBAF) occupancy at their sites of colocalization and at sites where they bind without the other reveals that the levels of esBAF occupancy at NF-Y/esBAF colocalized sites are no higher than those at esBAF-only sites (Figure 6C), suggesting that it is less likely that NF-Y recruits esBAF. Conversely, the levels of NF-Y occupancy are the same whether or not NF-Y colocalizes with esBAF (Figure S7C), suggesting that NF-Y occupancy is independent of esBAF. Taken together, these data argue against interdependency between NF-Y and esBAF and for a direct role for NF-Y in promoting enhanced accessibility.

NF-Y interacts with DNA through insertion of the NF-YA A2 helix into the DNA minor groove, which induces a significant 78°–80° bend in the DNA (Nardini et al., 2013). Because NF-YA utilizes such a unique DNA-binding mode, and because the resulting bend in DNA may promote binding of other TFs in the adjacent major grooves, we hypothesized that NF-Y might foster binding of master ESC TFs at their sites of colocalization. Toward testing this hypothesis, we examined NF-Y and OSNP occupancy at their sites of colocalization and at sites where they bind without the other. Strikingly, the levels of OSNP occupancy are significantly higher at sites where they colocalize with NF-Y compared to those at sites where they do not (Figure 6D). In contrast, levels of NF-Y occupancy are the same whether or not NF-Y colocalizes with OSNP (Figure 6E), indicating that NF-Y promotes enhanced binding of OSNP at their sites of colocalization, but not vice versa. A similar phenomenon was also observed in neurons at sites co-localized by NF-Y and CTCF (Figure S7D). Consistent with this observation, target genes of Oct4/Sox2 with NF-Y co-occupancy have significantly higher expression compared to those of Oct4/Sox2 without NF-Y co-occupancy, and are significantly downregulated in NF-Y TKD cells (Figure S7E). Interestingly, de novo motif analysis of Oct4/Sox2 sites with or without NF-Y co-occupancy revealed identical Oct4/Sox2 motifs (Figure S7F), suggesting that enhanced Oct4/Sox2 binding at sites of NF-Y co-occupancy is likely due to NF-Y and not the underlying sequence motifs.

NF-Y Interacts with Oct4, and Oct4/Sox2 Binding Is Dependent on NF-Y

To establish the mode of functional interaction between NF-Y and TFs whose binding it might promote, we performed coimmunoprecipitation experiments using protein extracts from ESCs and antibodies against NF-YA or NF-YC subunits of the NF-Y complex. As expected, NF-YA coimmunoprecipitated with NF-YC and vice versa (Figure 6F). More importantly, we found that Oct4 coimmunoprecipitated with NF-YC, but not NF-YA (Figure 6F). Conversely, NF-YC, but not NF-YA, coimmunoprecipitated with Oct4. Given that Oct4 interacts with the NF-YC subunit of the NF-Y complex and that Oct4 binding is enhanced at sites of NF-Y cobinding (Figure 6D), we next asked whether Oct4 binding is dependent on NF-Y. To explore this connection, we studied NF-YA KD ESCs 48 hr after siRNA transfection, when NF-YA is depleted but the cells appear normal, with no obvious reduction in Oct4 or Sox2 levels compared to control KD ESCs (Figures 6G, S7G, and S7H). ChIP using antibodies against Oct4 or Sox2 followed by qPCR revealed a significant loss of Oct4/Sox2 binding specifically at Oct4/Sox2 sites cobound by NF-Y, but not at sites where Oct4/Sox2 bind without NF-Y (Figure 6H). These data indicate that Oct4/Sox2 binding is NF-Y dependent at sites where Oct4/Sox2 colocalize with NF-Y.



(legend on next page)

NF-Y Promotes Oct4/Sox2 Binding by Facilitating a Favorable Chromatin Conformation

To probe the link between NF-Y deficiency and diminished Oct4/Sox2 binding at their sites of colocalization, we next examined the DNase I hypersensitivity at sites bound by OSNP either with or without NF-Y. We found that those sites cobound by NF-Y and OSNP are 2- to 3-fold more hypersensitive to DNase I digestion than those that are bound only by OSNP (Figure 7A). DNase I hypersensitivity assay followed by qPCR amplification confirmed that NF-Y-independent Oct4/Sox2 sites in ESCs are less accessible compared with NF-Y-dependent Oct4/Sox2 sites, as shown by their greater resistance to digestion by DNase I (Figure 7B). Consistent with the higher DNase I hypersensitivity at NF-Y-dependent OSNP sites (Figures 7A and 7B), we find lower nucleosome occupancy at sites where OSNP colocalize with NF-Y (Figure 7C), which is especially pronounced on the side of NF-Y occupancy. Hence, we reasoned that diminished Oct4/Sox2 binding upon NF-Y deficiency might be due to loss of intrinsic chromatin accessibility.

To investigate this possibility, we depleted NF-Y and examined the accessibility of NF-Y-dependent versus NF-Y-independent Oct4/Sox2 sites using a DNase I hypersensitivity assay followed by qPCR amplification of the target region to quantitatively assess the relative "openness" of the region. Depletion of NF-Y results in a significant reduction in the accessibility at NF-Y-dependent but not NF-Y-independent Oct4/Sox2 sites (Figures 7D, 7E, S7I, and S7J). Because sites that only bind NF-Y (and not Oct4/Sox2) also lose accessibility (Figure S7K) but not control sites devoid of both NF-Y and Oct4/Sox2 binding (Figure S7L), we conclude that the observed reduction in accessibility is a direct consequence of NF-Y deficiency.

The observation that NF-Y is required to induce enhanced chromatin accessibility for Oct4/Sox2 binding prompted us to examine whether increased nucleosome occupancy accompanies reduced chromatin accessibility upon NF-Y depletion. Examination of nucleosome occupancy, as assessed using histone H3 ChIP followed by qPCR, reveals elevated levels of nucleosome presence at NF-Y-dependent but not NF-Y-independent Oct4/Sox2 sites upon NF-Y KD (Figure 7F). Together, these data support the conclusion that NF-Y promotes Oct4/Sox2 binding by facilitating enhanced chromatin accessibility.

DISCUSSION

NF-Y's function has almost exclusively been studied in relation to its promoter-proximal binding (Dolfini et al., 2012a), likely driven by the prevalence of CCAAT motif within core promoters and promoter-focused arrays. In this study, by focusing on the surprisingly high proportion of NF-Y sites at distal regulatory elements, we report a cell type-specific function for NF-Y, whereby NF-Y promotes chromatin accessibility for cell type-specific master TFs at active enhancers.

NF-Y's Distinct Modes of Binding

Through comprehensive genome-wide occupancy analysis of NF-YA, NF-YB, and NF-YC subunits of the NF-Y complex in mouse ESCs, we find that nearly half of all NF-Y binding sites are within enhancers located at least 500 bp from TSSs. At these distal sites, NF-Y preferentially colocalizes with master ESC TFs known to bind active enhancers. In addition to confirming its regulation of key cell-cycle genes via promoter-proximal binding, we show that NF-Y regulates many key pluripotency-associated genes including *Oct4*, *Nanog*, and *Prdm14* by binding to their enhancers, where it colocalizes with master ESC TFs. Overall, genes targeted by distal NF-Y binding are enriched for processes related to early embryonic development, whereas those bound by NF-Y at the proximal promoter are enriched for housekeeping functions, including cell cycle.

NF-Y Is Required for ESC Identity

In line with NF-Y's regulation of master ESC TFs, global gene expression changes upon NF-Y KD are similar to those observed after *Oct4*, *Sox2*, *Nanog*, *Ncl*, or *Tet1* KD. ESCs lacking one or more of the NF-Y subunits exhibit severe self-renewal defects, undergo differentiation, and have a global expression profile that is comparable to that of cells differentiating toward the trophoblast lineage, all consistent with the propensity of ESCs lacking many of these master ESC factors to differentiate toward trophoectoderm (Cinghu et al., 2014; Freudenberg et al., 2012; Li et al., 2007; Nichols et al., 1998). In light of overexpression of NF-YA counteracting ESC differentiation induced by Leukemia inhibitory factor (LIF) withdrawal (Dolfini et al., 2012b), NF-Y's broad control of key ESC identity genes including master ESC TFs places NF-Y as part of the core pluripotency network. While we have not yet established whether the proclivity of ESCs lacking NF-Y to favor differentiation toward trophectoderm (TE) is a direct effect of NF-Y deficiency, we postulate that this outcome is perhaps indirect due to NF-Y's positive regulation of genes including *Oct4* and *Sox2*, which are known to inhibit master TE regulator(s). An understanding of NF-Y's role in pluripotency, lineage specification/commitment, and developmental potential will require analysis of NF-Y-deficient mice during early embryonic development.

Figure 5. NF-Y Binding at Enhancers Is Cell Type Specific, whereas NF-Y Binding at Promoters Is Cell Type Invariant

(A) Overlap among NF-Y sites in ESCs (red), neural progenitor cells (NPCs; green), and neurons (blue).

(B) Panther pathway enrichment analysis of proximal and distal NF-Y target genes in ESCs and neurons.

(C) ChIP-qPCR analysis of NF-YA sites in undifferentiated ESCs and differentiating ESCs (induced by retinoic acid, 96 hr). Error bars, SEM.

(D) Sequence motifs, other than that of NF-Y, enriched within distal NF-Y sites in ESCs or neurons, identified using de novo motif analysis.

(E) Nanog and CTCF (ENCODE, GSE49847) co-occupancy at promoter-proximal and ESC-specific (red) and neuron-specific (purple) distal NF-Y binding sites.

(F) TF occupancy (ChIP-Seq read density) near ESC-specific and neuron-specific distal NF-Y sites in ESCs and neurons.

(G) Levels of active enhancer marks H3K4me1 and H3K27ac, and repressive mark H3K27me3 near ESC-specific and neuron-specific distal NF-Y binding sites in ESCs (Creyghton et al., 2010; Ho et al., 2011) and neurons (ENCODE, GSE49847).

(H) Model for NF-Y-mediated transcriptional regulation of housekeeping and cell identity genes. See also Figures S5 and S6.

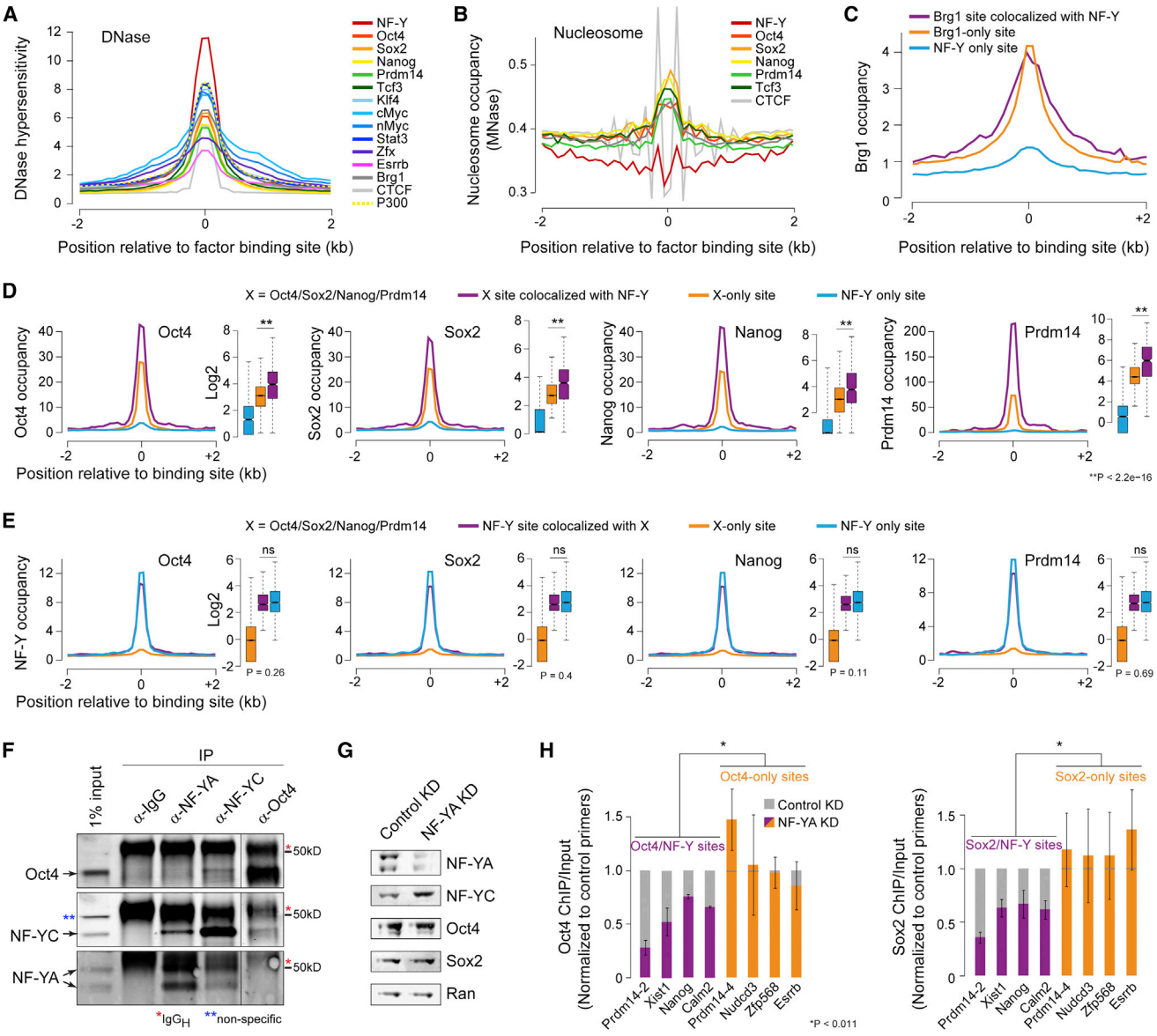


Figure 6. Oct4/Sox2 Binding Is Dependent on NF-Y

(A) DNase I hypersensitivity, as measured by DNase-Seq (ENCODE, GSE37074), at distal binding sites for various TFs in ESCs.

(B) Nucleosome occupancy, as measured by MNase-Seq (Teif et al., 2012), at distal binding sites for various TFs in ESCs.

(C) Brg1 occupancy at distal Brg1 sites colocalized with NF-Y (purple), distal Brg1-only sites (orange), distal NF-Y sites colocalized with Brg1 (cyan).

(D) Oct4/Sox2/Nanog/Prdm14 (X) occupancy (Ma et al., 2011; Marson et al., 2008) at distal X sites colocalized with NF-Y (purple), distal X-only sites (orange), and distal NF-Y only sites (blue).

(E) NF-Y occupancy at distal NF-Y sites colocalized with X (purple), distal NF-Y only sites (blue), and distal X-only sites (orange).

(F) Coimmunoprecipitation and western blot analysis showing NF-YC and Oct4 coimmunoprecipitating with Oct4 and NF-YC, respectively. As a positive control, NF-YC and NF-YA, two subunits of NF-Y complex, coimmunoprecipitate with NF-YA and NF-YC, respectively.

(G) Western blot analysis of NF-YA, NF-YC, Oct4, and Sox2 in *NF-YA* KD ESCs 48 hr after siRNA transfection. Ran used as a loading control.

(H) ChIP-qPCR analysis of sites cobound by Oct4/Sox2 and NF-Y (purple) or bound by Oct4/Sox2 but not NF-Y (orange) in control or NF-YA KD ESCs. Error bars, SEM. See also Figure S7.

Cell Type-Specific Function for NF-Y

A comparison of NF-Y occupancy in ESCs, and ESC-derived NPCs and terminal neurons revealed that while promoter-proximal NF-Y binding is largely conserved across cell types, distal NF-Y binding is cell type specific. More importantly, as in ESCs where NF-Y colocalizes with master ESC TFs, NF-Y preferentially colocalizes with master neuronal regulators at active enhancers of neuronal genes. These findings, along with those linking NF-Y to neuronal identity (Tiwari et al., 2012), point to a critical role for NF-Y in the specification of neuronal cells that is

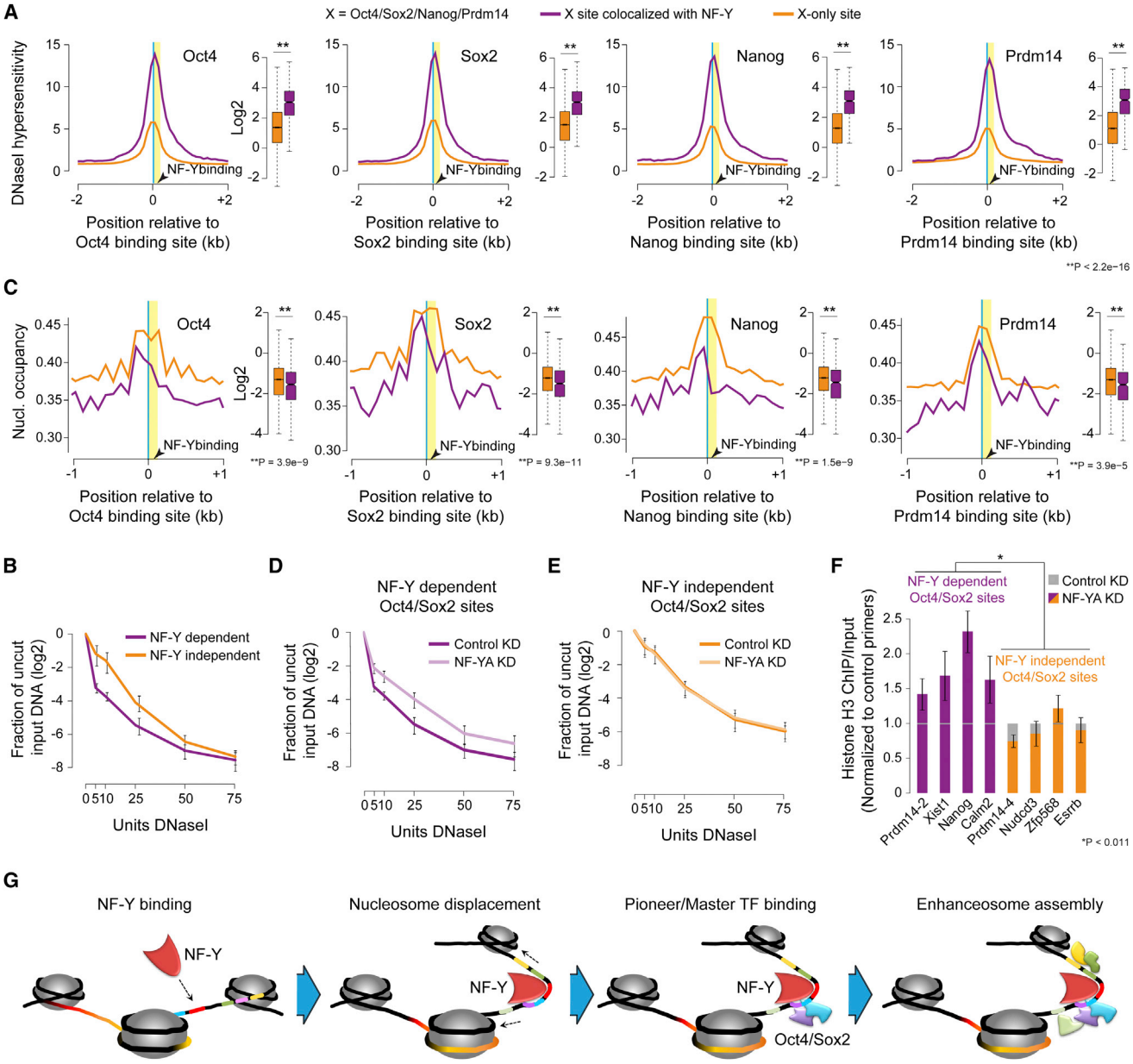


Figure 7. NF-Y Facilitates Oct4/Sox2 Binding by Promoting Chromatin Accessibility

(A) DNase I hypersensitivity (ENCODE, GSE37074) at NF-Y-dependent (purple) and independent (orange) Oct4/Sox2/Nanog/Prdm14 sites in ESCs.

(B) DNase I and qPCR analysis of NF-Y-dependent and NF-Y-independent Oct4/Sox2 sites in ESCs. Error bars, SEM.

(C) Nucleosome occupancy (Teif et al., 2012) at NF-Y-dependent (purple) and independent (orange) Oct4/Sox2/Nanog/Prdm14 sites in ESCs.

(D and E) DNase I and qPCR analysis of NF-Y-dependent (D) and NF-Y-independent (E) Oct4/Sox2 sites (n = 3) in control and NF-YA KD ESCs. Error bars, SEM.

(F) ChIP-qPCR analysis of NF-Y-dependent (purple) and NF-Y-independent (orange) Oct4/Sox2 sites in Control and NF-YA KD ESCs. Error bars, SEM.

(G) Proposed model for NF-Y-mediated chromatin accessibility. See also Figure S7.

independent of its role in cell-cycle progression. This supposition is in accordance with a recent study showing that conditional deletion of *NF-YA* in postmitotic mouse neurons induces progressive neurodegeneration (Yamanaka et al., 2014).

Consistent with NF-Y's role in ESC specification, ESCs lose NF-Y binding at distal but not proximal sites immediately after they exit the pluripotent state, with ESC-specific distal NF-Y sites epigenetically silenced in neurons. Interestingly, however, neuron-specific NF-Y binding sites, despite their not binding NF-Y in ESCs, are within "open" chromatin regions marked by DNase I hypersensitivity, histone marks characteristic of active enhancers, and low to moderate levels of Oct4/Sox2 occupancy. Although further studies are required to ascertain the potential role for Oct4/Sox2 at these sites, we postulate that Oct4/Sox2

binding at lineage-specific enhancers might serve as placeholders for lineage-specific TFs, "priming" these sites for rapid initiation of lineage-specific gene expression programs during differentiation. Such a role for Oct4/Sox2 in priming the chromatin would be different from, but consistent with, their pioneer activity during induced pluripotent stem cell (iPSC) reprogramming (Soufi et al., 2012). Further investigation is needed to precisely characterize the functional significance of these "primed enhancers."

NF-Y Promotes Chromatin Accessibility for Master TFs

The levels of Oct4, Sox2, Nanog, and Prdm14 occupancy in ESCs, and CTCF occupancy in neurons are significantly higher at sites where they colocalize with NF-Y compared to those sites where they bind without NF-Y. This enhancement cannot be attributed exclusively to cooperative binding of TFs, because NF-Y occupancy levels are invariant whether or not it colocalizes with other factors. This inequitable augmentation suggests that NF-Y promotes binding of master TFs at their sites of colocalization. Supporting this line of argument, NF-Y depletion in ESCs results in reduced Oct4/Sox2 occupancy at NF-Y-dependent but not NF-Y-independent Oct4/Sox2 sites, accompanied by reduced chromatin accessibility.

NF-Y's Role in Nucleosome Displacement, Pioneer Factor Binding, and Enhanceosome Assembly

As pioneer factors, Oct4 and Sox2 bind nucleosomal DNA in ESCs (Teif et al., 2012), as also observed here and during reprogramming (Soufi et al., 2012). Given the enhanced binding of master TFs at sites of NF-Y colocalization, our observations raise a tantalizing possibility that Oct4/Sox2 and perhaps other pioneer factors with a preference for binding nucleosomal DNA view the NF-Y/DNA complex, with its nucleosome-like properties, as an attractive target. This supposition may not be farfetched given that the structure and DNA-binding mode of the NF-YB/NF-YC dimer are nearly identical to that of histones H2A/H2B. Additionally, Oct4's interaction with NF-Y, as shown here, would add to a stronger and a more stable Oct4/Sox2 binding, enabling Oct4/Sox2 to recruit other TFs for an ordered enhanceosome assembly (Chen et al., 2014), consistent with the proposed role for pioneer factors (Zaret and Carroll, 2011). Given that in vitro studies have previously demonstrated NF-Y to displace nucleosomes (Coustry et al., 2001), it will be interesting to explore whether and how NF-Y mediates pioneer activity or even how NF-Y itself could play the role of a pioneer factor.

Based on our findings, we propose a general model wherein NF-Y facilitates enhanced chromatin accessibility via two modes. First, NF-Y's nucleosome-like properties provide for a stable and a stronger binding of pioneer factors Oct4/Sox2. Second, NF-Y's unique DNA-binding mode, which induces an ~80° bend in the DNA, imposes sufficient spatial constraints to induce local nucleosome repositioning (such as flanking nucleosomes sliding out), thus allowing and/or promoting binding of other TFs, whose recognition sequences become more accessible (Figure 7F). While we could not find evidence supporting increased chromatin remodeling activity at NF-Y binding sites by known chromatin remodelers, we cannot rule out the possibility that the increased accessibility at NF-Y binding sites could be due to NF-Y potentially recruiting a bona fide chromatin remodeler to facilitate permissive chromatin. Given that CHRAC15 and CHRAC17, two small subunits of the nucleosome remodeling complex CHRAC with HFDs structurally similar to that of NF-YB/NF-YC and histones H2A/H2B, have previously been shown to facilitate nucleosome sliding (Kukimoto et al., 2004), it is tempting to speculate that NF-Y, with its sequence-specific binding ability, may be a "pseudo chromatin remodeler" mimicking the role of a bona fide ATP-dependent chromatin remodeler, albeit at a very small and a limited scale. As in the case of chromatin-remodeling enzymes, a formal demonstration of NF-Y's role in chromatin remodeling, direct or indirect, will require in vitro biochemical experiments involving reconstituted nucleosomes.

Although NF-Y's role in transcription is widely accepted to be through transactivation, mechanistic details remain undetermined. Based on our findings, it is conceivable that NF-Y's primary mode of action, both at enhancers and promoters, is to promote binding of other TFs including basal transcription machinery by facilitating an accessible chromatin. This supposition is consistent with a strong positive correlation between NF-Y binding and gene expression.

Our findings establish NF-Y as an essential regulator of the pluripotent state in ESCs, and suggest a role for NF-Y in cell specification. It remains to be seen how NF-Y selectively targets cell type-specific enhancers. Perhaps this is accomplished through direct recruitment by a cell type-specific TF, alternative isoform usage, and/or posttranslational modifications. In summary, our studies unearth a cell type-specific function for NF-Y in promoting chromatin accessibility. Our data suggest that other HFD proteins, with structural and DNA-binding properties analogous to NF-Y, may function in similar ways.

EXPERIMENTAL PROCEDURES

See Supplemental Experimental Procedures.

Mouse ESC Culture, RNAi, and AP Staining

Mouse ESC culture, siRNA transfection, and AP staining were performed as previously described (Freudenberg et al., 2012) and as detailed in the Supplemental Experimental Procedures.

Chromatin Immunoprecipitation

Mouse ESCs (1×10^7) were crosslinked with 1% formaldehyde, and the reaction was quenched with glycine for 5 min. Cells were lysed and then sonicated for 16 cycles (30 s on, 50 s off) to obtain ~200–500 bp fragments. ChIP was peformed as described in Supplemental Experimental Procedures. See Table S1 for a list of ChIP primers used.

DNase I Hypersensitivity

Mouse ESCs treated with nontargeting control siRNA or *NF-YA* siRNA were collected 48 hr posttransfection. Nuclei isolation and DNase I digestion were performed as previously described (Burch and Weintraub, 1983), with minor modifications as detailed in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

ChIP-Seq and microarray data generated for this study have been deposited in the GEO repository under the accession number GSE56840.

SUPPLEMENTAL INFORMATION

Supplemental information includes seven figures, two tables, and Supplemental Experimental Procedures and can be found with this article at http://dx.doi.org/10.1016/j.molcel.2014.07.005.

AUTHOR CONTRIBUTIONS

A.J.O. and R.J. conceived the study. A.J.O., A.E.C., and S.C. performed the experiments. P.Y., J.E.M., S.Y., and R.J. analyzed the data. A.J.O., P.Y., and R.J. wrote the manuscript.

ACKNOWLEDGMENTS

We thank K. Adelman, G. Hu, and P.A. Wade for critical comments on the manuscript, and R.S. Williams and T.K. Archer for useful discussion. We thank G. Hu and X. Zheng for guidance on RNAi experiments, S. Peddada for guidance on statistical tests, and J. Hoffman for guidance on DNase experiments. We thank NIEHS Microarray, Next-Generation Sequencing, Confocal Microscopy, and Flow Cytometry cores for their support. This work was supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (R.J., 1ZIAES102625).

Received: May 7, 2014
Revised: June 13, 2014
Accepted: July 9, 2014
Published: August 14, 2014

REFERENCES

Benatti, P., Dolfini, D., Viganò, A., Ravo, M., Weisz, A., and Imbriano, C. (2011). Specific inhibition of NF-Y subunits triggers different cell proliferation defects. Nucleic Acids Res. *39*, 5356–5368.

Bhattacharya, A., Deng, J.M., Zhang, Z., Behringer, R., de Crombrugghe, B., and Maity, S.N. (2003). The B subunit of the CCAAT box binding transcription factor complex (CBF/NF-Y) is essential for early mouse development and cell proliferation. Cancer Res. *63*, 8167–8172.

Bungartz, G., Land, H., Scadden, D.T., and Emerson, S.G. (2012). NF-Y is necessary for hematopoietic stem cell proliferation and survival. Blood *119*, 1380–1389.

Burch, J.B., and Weintraub, H. (1983). Temporal order of chromatin structural changes associated with activation of the major chicken vitellogenin gene. Cell *33*, 65–76.

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.

Chen, J., Zhang, Z., Li, L., Chen, B.C., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E., et al. (2014). Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. Cell *156*, 1274–1285.

Chia, N.Y., Chan, Y.S., Feng, B., Lu, X., Orlov, Y.L., Moreau, D., Kumar, P., Yang, L., Jiang, J., Lau, M.S., et al. (2010). A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature *468*, 316–320.

Cinghu, S., Yellaboina, S., Freudenberg, J.M., Ghosh, S., Zheng, X., Oldfield, A.J., Lackford, B.L., Zaykin, D.V., Hu, G., and Jothi, R. (2014). Integrative framework for identification of key cell identity genes uncovers determinants of ES cell identity and homeostasis. Proc. Natl. Acad. Sci. USA *111*, E1581–E1590.

Coustry, F., Hu, Q., de Crombrugghe, B., and Maity, S.N. (2001). CBF/NF-Y functions both in nucleosomal disruption and transcription activation of the chromatin-assembled topoisomerase IIalpha promoter. Transcription activation by CBF/NF-Y in chromatin is dependent on the promoter structure. J. Biol. Chem. *276*, 40621–40630.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. USA *107*, 21931–21936.

Dolfini, D., Gatta, R., and Mantovani, R. (2012a). NF-Y and the transcriptional activation of CCAAT promoters. Crit. Rev. Biochem. Mol. Biol. *47*, 29–49.

Dolfini, D., Minuzzo, M., Pavesi, G., and Mantovani, R. (2012b). The short isoform of NF-YA belongs to the embryonic stem cell transcription factor circuitry. Stem Cells *30*, 2450–2459.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat. Methods *9*, 215–216.

Farina, A., Manni, I., Fontemaggi, G., Tiainen, M., Cenciarelli, C., Bellorini, M., Mantovani, R., Sacchi, A., and Piaggio, G. (1999). Down-regulation of cyclin B1 gene transcription in terminally differentiated skeletal muscle cells is associated with loss of functional CCAAT-binding NF-Y complex. Oncogene *18*, 2818–2827.

Fleming, J.D., Pavesi, G., Benatti, P., Imbriano, C., Mantovani, R., and Struhl, K. (2013). NF-Y coassociates with FOS at promoters, enhancers, repetitive elements, and inactive chromatin regions, and is stereo-positioned with growth-controlling transcription factors. Genome Res. *23*, 1195–1209.

Freudenberg, J.M., Ghosh, S., Lackford, B.L., Yellaboina, S., Zheng, X., Li, R., Cuddapah, S., Wade, P.A., Hu, G., and Jothi, R. (2012). Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. Nucleic Acids Res. *40*, 3364–3377.

Golan-Mashiach, M., Grunspan, M., Emmanuel, R., Gibbs-Bar, L., Dikstein, R., and Shapiro, E. (2012). Identification of CTCF as a master regulator of the clustered protocadherin genes. Nucleic Acids Res. *40*, 3378–3391.

Grskovic, M., Chaivorapol, C., Gaspar-Maia, A., Li, H., and Ramalho-Santos, M. (2007). Systematic identification of cis-regulatory sequences active in mouse and human embryonic stem cells. PLoS Genet. *3*, e145.

Hirayama, T., Tarusawa, E., Yoshimura, Y., Galjart, N., and Yagi, T. (2012). CTCF is required for neural development and stochastic expression of clustered Pcdh genes in neurons. Cell Rep. *2*, 345–357.

Ho, L., Miller, E.L., Ronan, J.L., Ho, W.Q., Jothi, R., and Crabtree, G.R. (2011). esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. Nat. Cell Biol. *13*, 903–913.

Hu, Q., and Maity, S.N. (2000). Stable expression of a dominant negative mutant of CCAAT binding factor/NF-Y in mouse fibroblast cells resulting in retardation of cell growth and inhibition of transcription of various cellular genes. J. Biol. Chem. *275*, 4435–4444.

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 Res. *36*, 5221–5231.

Kabe, Y., Yamada, J., Uga, H., Yamaguchi, Y., Wada, T., and Handa, H. (2005). NF-Y is essential for the recruitment of RNA polymerase II and inducible transcription of several CCAAT box-containing genes. Mol. Cell. Biol. *25*, 512–522.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

Kim, J., Woo, A.J., Chu, J., Snow, J.W., Fujiwara, Y., Kim, C.G., Cantor, A.B., and Orkin, S.H. (2010a). A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. Cell *143*, 313–324.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010b). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182–187.

Kukimoto, I., Elderkin, S., Grimaldi, M., Oelgeschläger, T., and Varga-Weisz, P.D. (2004). The histone-fold protein complex CHRAC-15/17 enhances nucleosome sliding and assembly mediated by ACF. Mol. Cell *13*, 265–277.

Li, J., Pan, G., Cui, K., Liu, Y., Xu, S., and Pei, D. (2007). A dominant-negative form of mouse SOX2 induces trophectoderm differentiation and progressive polyploidy in mouse embryonic stem cells. J. Biol. Chem. *282*, 19481–19492.

Luo, R., Klumpp, S.A., Finegold, M.J., and Maity, S.N. (2011). Inactivation of CBF/NF-Y in postnatal liver causes hepatocellular degeneration, lipid deposition, and endoplasmic reticulum stress. Sci. Rep. *1*, 136.

Ma, Z., Swigut, T., Valouev, A., Rada-Iglesias, A., and Wysocka, J. (2011). Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. Nat. Struct. Mol. Biol. *18*, 120–127.

Maity, S.N., and de Crombrugghe, B. (1998). Role of the CCAAT-binding protein CBF/NF-Y in transcription. Trends Biochem. Sci. *23*, 174–178.

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.

Nardini, M., Gnesutta, N., Donati, G., Gatta, R., Forni, C., Fossati, A., Vonrhein, C., Moras, D., Romier, C., Bolognesi, M., and Mantovani, R. (2013). Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. Cell *152*, 132–143.

Narlikar, L., and Jothi, R. (2012). ChIP-Seq data analysis: identification of protein-DNA binding sites with SISSRs peak-finder. Methods Mol. Biol. *802*, 305–322.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379–391.

Nishiyama, A., Xin, L., Sharov, A.A., Thomas, M., Mowrer, G., Meyers, E., Piao, Y., Mehta, S., Yee, S., Nakatake, Y., et al. (2009). Uncovering early response of gene regulatory networks in ESCs by systematic induction of transcription factors. Cell Stem Cell *5*, 420–433.

Oestreich, K.J., and Weinmann, A.S. (2012). Master regulators or lineage-specifying? Changing views on CD4+ T cell transcription factors. Nat. Rev. Immunol. *12*, 799–804.

Pijnappel, W.W., Esch, D., Baltissen, M.P., Wu, G., Mischerikow, N., Bergsma, A.J., van der Wal, E., Han, D.W., Bruch, Hv., Moritz, S., et al. (2013). A central role for TFIID in the pluripotent transcription circuitry. Nature *495*, 516–519.

Romier, C., Cocchiarella, F., Mantovani, R., and Moras, D. (2003). The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. J. Biol. Chem. *278*, 1336–1345.

Sinha, S., Maity, S.N., Lu, J., and de Crombrugghe, B. (1995). Recombinant rat CBF-C, the third subunit of CBF/NFY, allows formation of a protein-DNA complex with CBF-A and CBF-B and with yeast HAP2 and HAP3. Proc. Natl. Acad. Sci. USA *92*, 1624–1628.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994–1004.

Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., et al. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. Proc. Natl. Acad. Sci. USA *101*, 6062–6067.

Teif, V.B., Vainshtein, Y., Caudron-Herger, M., Mallm, J.P., Marth, C., Höfer, T., and Rippe, K. (2012). Genome-wide nucleosome positioning during embryonic stem cell development. Nat. Struct. Mol. Biol. *19*, 1185–1192.

Testa, A., Donati, G., Yan, P., Romani, F., Huang, T.H., Viganò, M.A., and Mantovani, R. (2005). Chromatin immunoprecipitation (ChIP) on chip experiments uncover a widespread distribution of NF-Y binding CCAAT sites outside of core promoters. J. Biol. Chem. *280*, 13606–13615.

Tiwari, V.K., Stadler, M.B., Wirbelauer, C., Paro, R., Schübeler, D., and Beisel, C. (2012). A chromatin-modifying function of JNK during stem cell differentiation. Nat. Genet. *44*, 94–100.

Wang, L., Du, Y., Ward, J.M., Shimbo, T., Lackford, B., Zheng, X., Miao, Y.L., Zhou, B., Han, L., Fargo, D.C., et al. (2014). INO80 facilitates pluripotency gene activation in embryonic stem cell self-renewal, reprogramming, and blastocyst development. Cell Stem Cell *14*, 575–591.

Yamanaka, T., Tosaki, A., Kurosawa, M., Matsumoto, G., Koike, M., Uchiyama, Y., Maity, S.N., Shimogori, T., Hattori, N., and Nukina, N. (2014). NF-Y inactivation causes atypical neurodegeneration characterized by ubiquitin and p62 accumulation and endoplasmic reticulum disorganization. Nat. Commun. *5*, 3354.

Yoshikawa, T., Piao, Y., Zhong, J., Matoba, R., Carter, M.G., Wang, Y., Goldberg, I., and Ko, M.S. (2006). High-throughput screen for genes predominantly expressed in the ICM of mouse blastocysts by whole mount in situ hybridization. Gene Expr. Patterns *6*, 213–224.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes Dev. *25*, 2227–2241.

Zhang, Y., Wong, C.H., Birnbaum, R.Y., Li, G., Favaro, R., Ngan, C.Y., Lim, J., Tai, E., Poh, H.M., Wong, E., et al. (2013). Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. Nature *504*, 306–310.

Molecular Cell, Volume *55*

Supplemental Information

Histone-Fold Domain Protein NF-Y Promotes Chromatin Accessibility for Cell Type-Specific Master Transcription Factors

Andrew J. Oldfield, Pengyi Yang, Amanda E. Conway, Senthilkumar Cinghu, Johannes M. Freudenberg, Sailu Yellaboina, and Raja Jothi

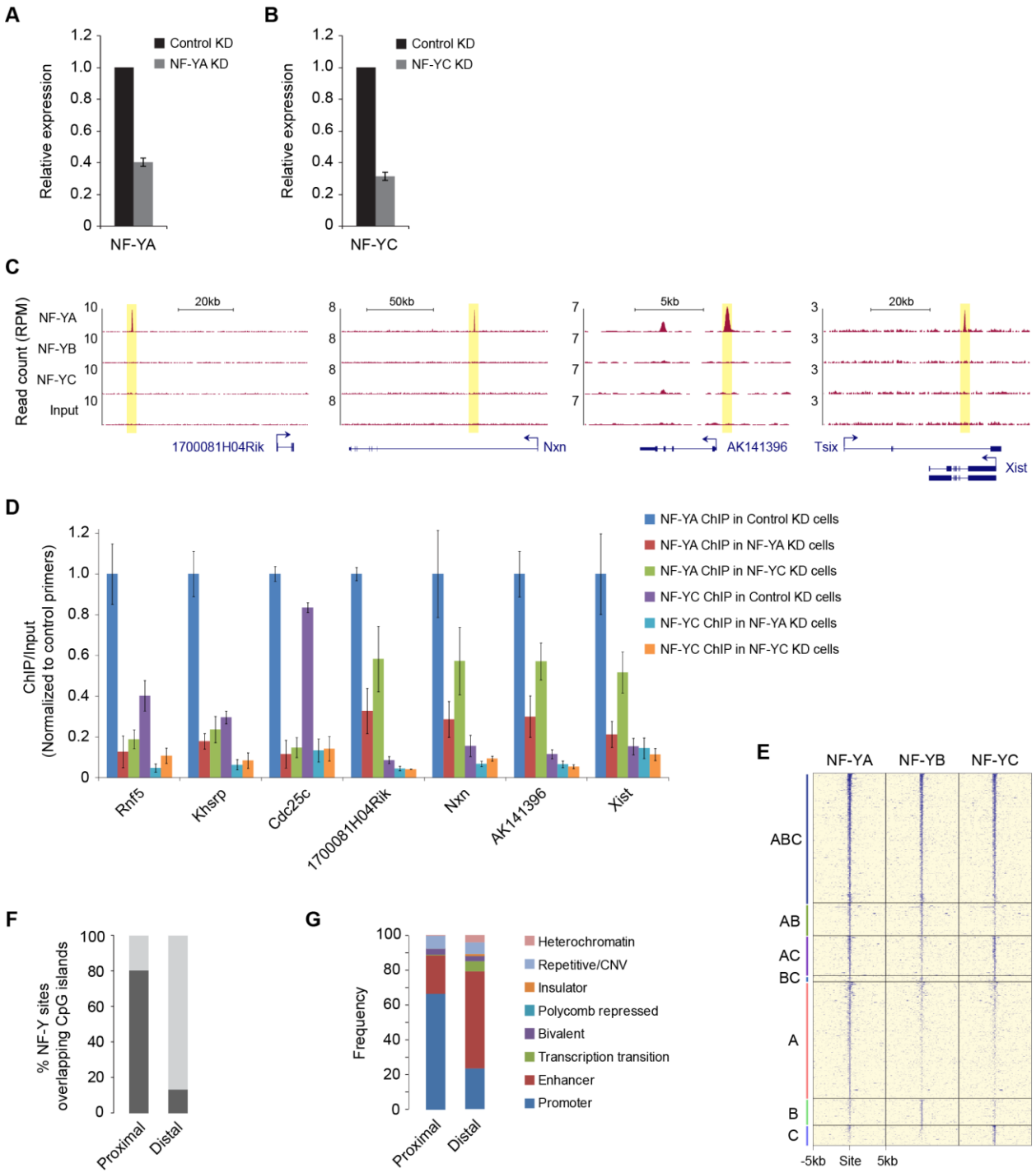


Figure S1 (related to Figure 1). NF-Y binding requires all three subunits

(A-B) RT-qPCR analysis showing relative mRNA levels of *NF-YA* and *NF-YC* in *NF-YA* KD (A) and *NF-YC* KD (B) ESCs (48h), respectively. Data are normalized to *Actin*, *HAZ*, and *TBP*. Error bars represent S.E.M. of three experiments.

(C) Genome browser shots showing NF-YA occupancy at the promoters of *1700081H04Rik*, *Nxn*, *AK141396*, and *Xist* genes, but not NF-YB and NF-YC, the other two subunits of the NF-Y complex.

(D) ChIP-qPCR analysis of the NF-YA and NF-YC occupancy in Control KD, *NF-YA* KD, and *NF-YC* KD ESCs 48h after siRNA transfection. Error bars represent S.E.M. of three experiments.

(E) Heatmap representation of NF-YA, NF-YB, and NF-YC occupancy (ChIP-Seq read density) near called sites bound by NF-YA, NF-YB, NF-YC, or a combination of the three.

(F) Percent of NF-Y sites overlapping CpG islands. Proximal, NF-Y sites within 500 bp of TSSs; Distal, NF-Y sites that are >500 bp away from TSSs.

(G) Annotation of proximal and distal NF-Y sites using chromatin state maps (Ernst and Kellis, 2012).

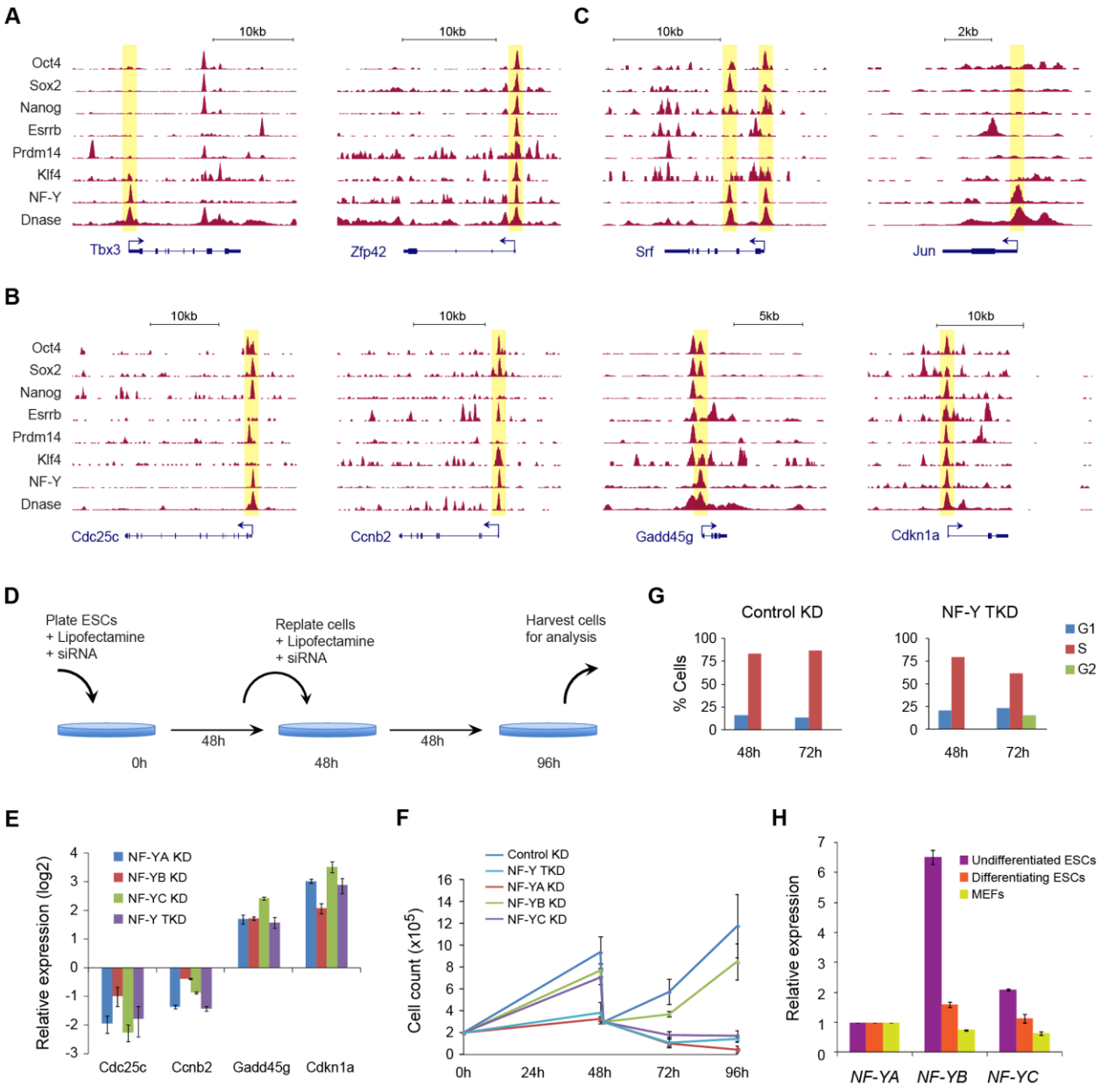


Figure S2 (related to Figure 3). NF-Y controls ESC proliferation by regulating the expression of key cell cycle genes

(A) Genome browser shots showing TF occupancy at pluripotency-associated genes *Tbx3* and *Zfp42*.

(B) Genome browser shots showing TF occupancy at cell cycle genes *Cdc25c*, *Ccnb2*, *Gadd45g*, and *Cdkn1a* (*p21*).

(C) Genome browser shots showing TF occupancy at differentiation genes *Srf* and *Jun*.

(D) Schematic showing siRNA transfection strategy.

(E) RT-qPCR analysis of relative mRNA levels of cell cycle genes shown in (B) in *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and NF-Y TKD ESCs compared to control KD ESCs 96h after siRNA transfection. Data are normalized to *Actin*, *HAZ*, and *TBP*. Error bars represent S.E.M. of three experiments.

(F) Cell proliferation growth curves were determined by counting the cell numbers in control KD, *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and NF-Y TKD ESCs at 48h, 72h, and 96h after siRNA transfection. Cells were replated and retransfected at 48h after the initial transfection. The average cell numbers from three experiments are shown. Error bars represent with S.E.M. of three experiments.

(G) Cell cycle distribution measured by flow cytometry analysis in control KD and NF-Y TKD ESCs 48h and 72h after siRNA transfection.

(H) RT-qPCR analysis of relative mRNA levels of NF-Y subunits *NF-YA*, *NF-YB*, and *NF-YC* in wild-type mouse ESCs. Error bars represent S.E.M. of three experiments.

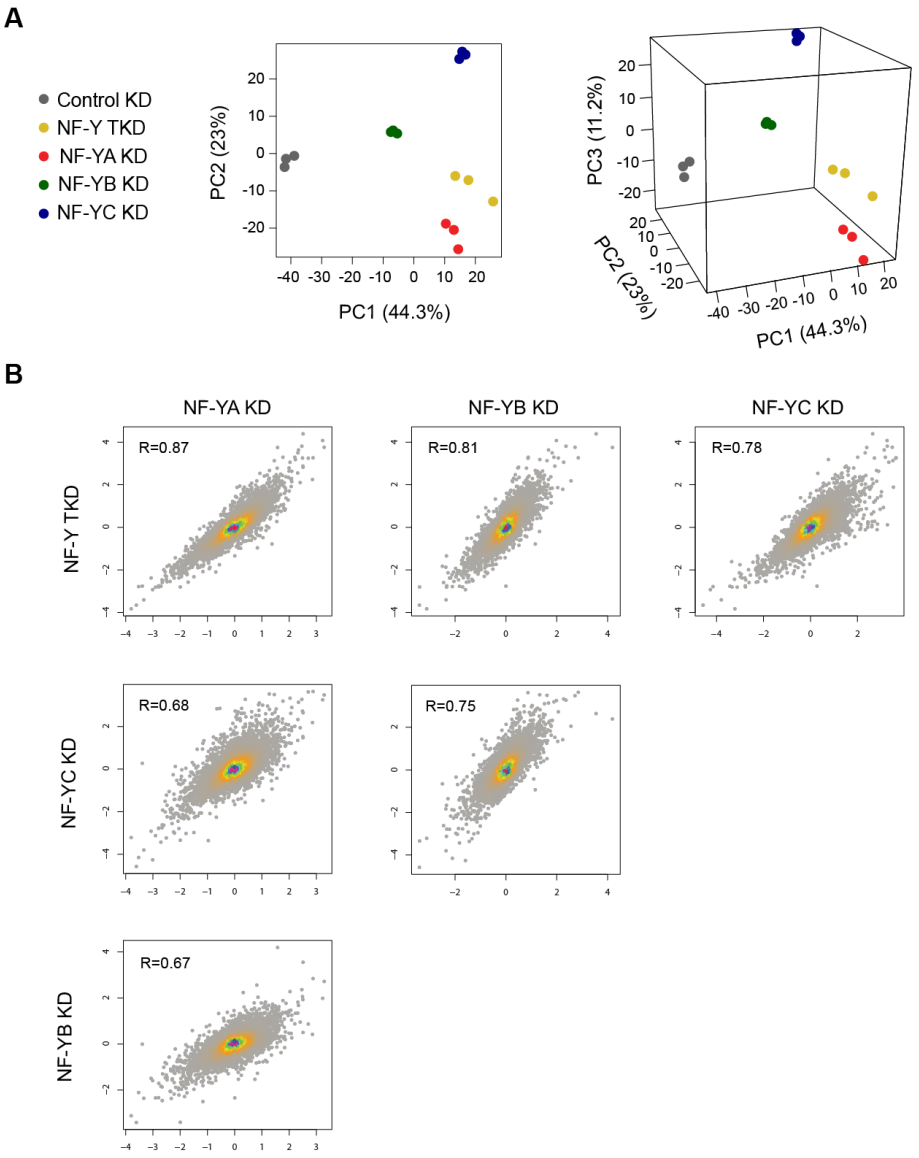


Figure S3 (related to Figure 4). Global gene expression changes upon depletion of individual or all NF-Y subunits are correlated.

(A) Principal component analysis (PCA) of gene expression profiles showing Control KD, *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and NF-Y TKD ESCs. Each dot represents a biological replicate.

(B) Scatter plots showing pair-wise correlation of global gene expression changes upon *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and NF-Y TKD ESCs compared to control KD ESCs.

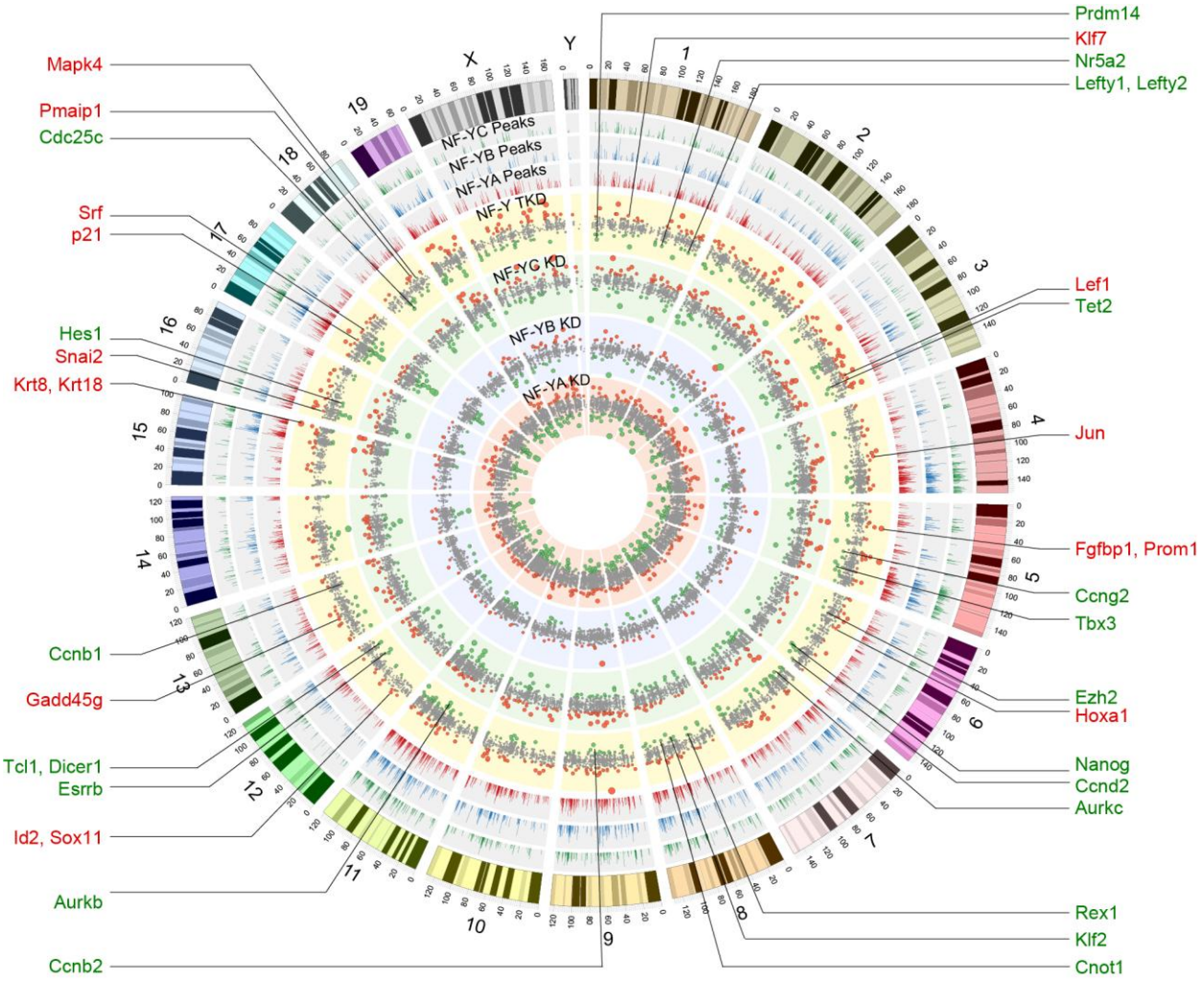


Figure S4 (related to Figure 4). NF-Y regulates key ESC identity and cell cycle genes. Circos plot depicting NF-YA, NF-YB, and NF-YC binding sites across the genome (outer-most tracks). Y-axis represents the peak height for binding sites. Also shown are gene expression fold changes upon *NF-YA* KD, *NF-YB* KD, *NF-YC* KD, and NF-Y TKD (inner-most tracks). Genes whose expression increases (or decreases) by two or more fold are denoted by red (green, respectively) circles, with the size of the circle denoting the magnitude of the change. Selected genes that bind NF-Y and whose expression ≥2-fold are highlighted.

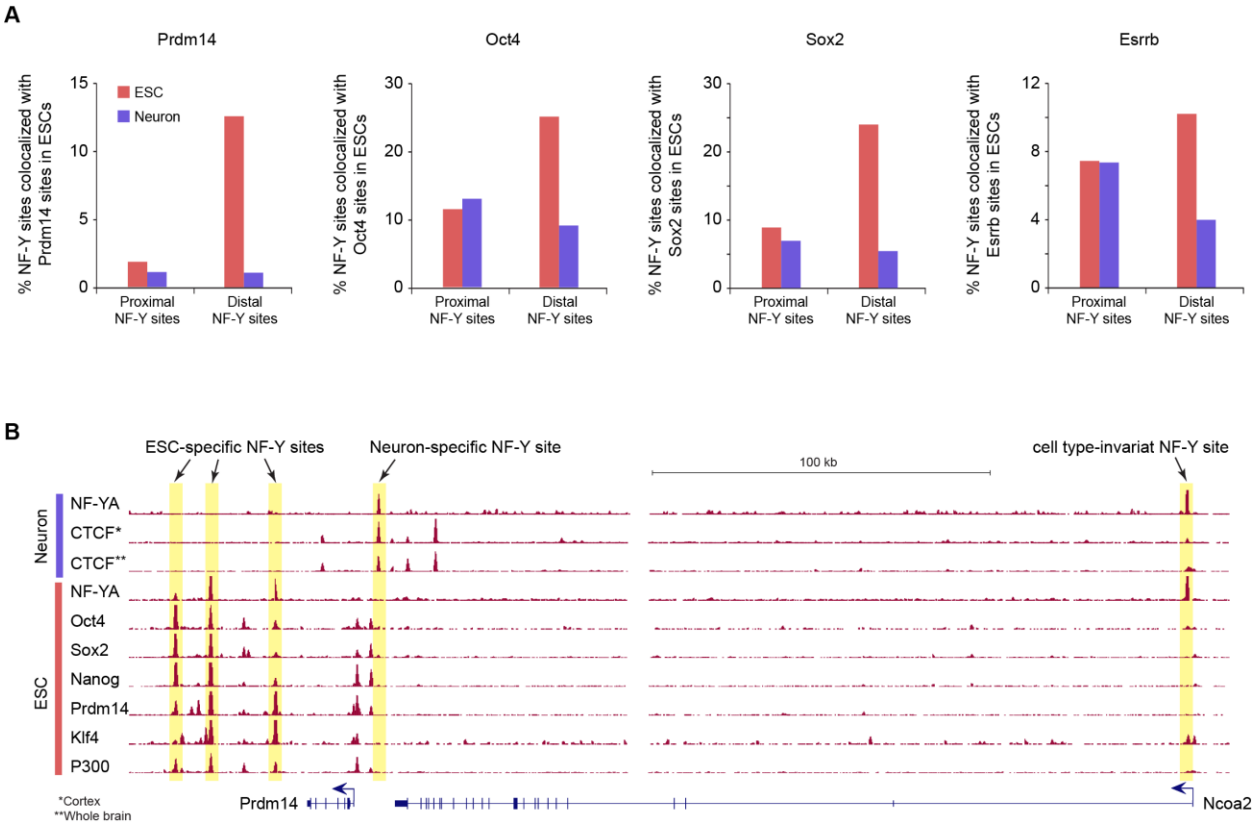


Figure S5 (related to Figure 5). TF co-occupancy at proximal and ESC-specific and neuron-specific distal NF-Y sites.

(A) Master ESC TF co-occupancy at proximal and ESC-specific (red) and neuron-specific (purple) distal NF-Y sites in ESCs

(B) Genome browser shot showing TF occupancy in ESCs and neurons at *Prdm14-Ncoa2* locus. ESC-specific, neuron-specific, and cell type-invariant NF-Y sites are highlighted.

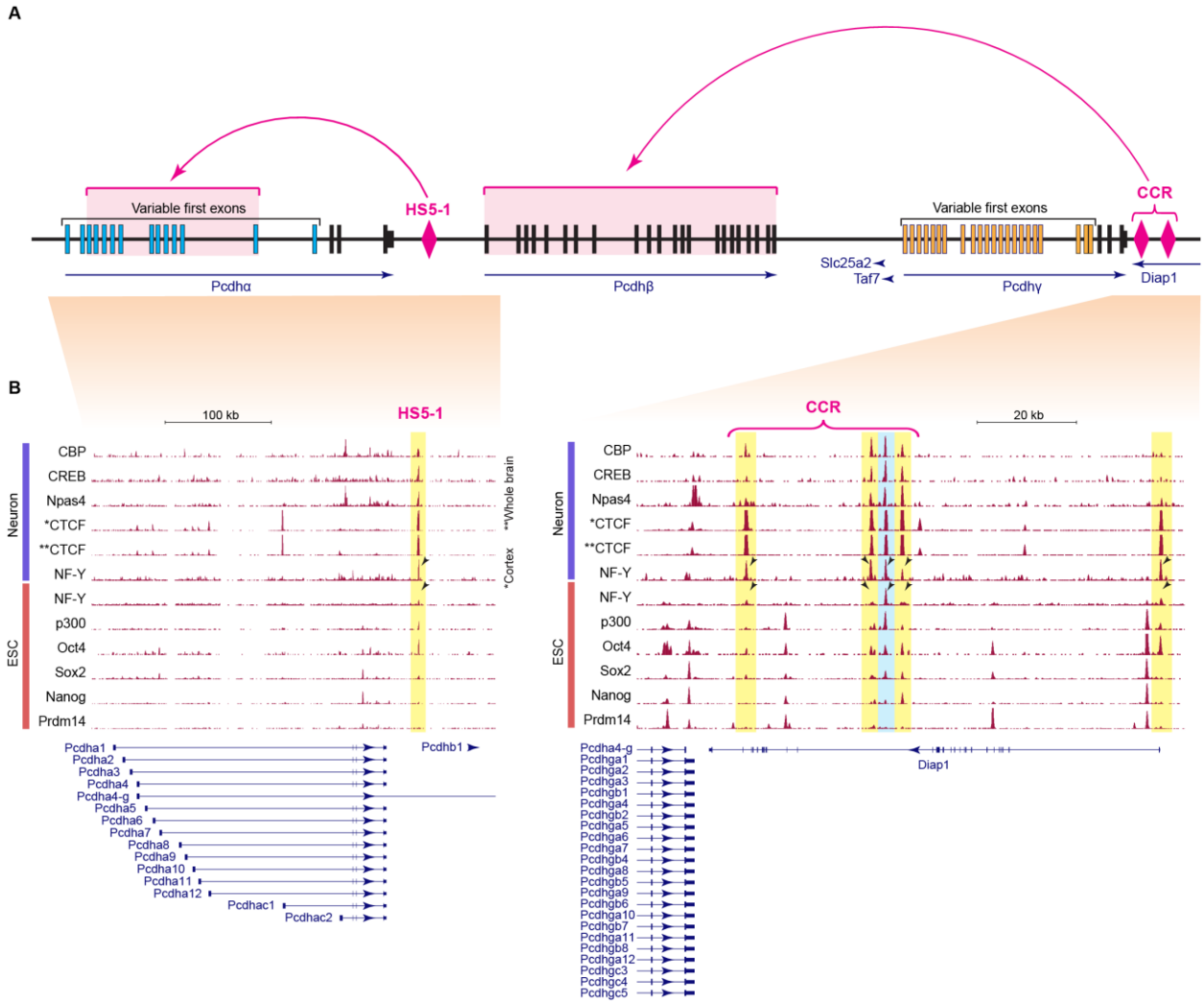


Figure S6 (related to Figure 5). NF-Y colocalizes, in a neuronal specific manner, with CTCF at sites controlling clustered Protocadherin expression.

(A) Genomic structure of the clustered Pcdh gene locus, adapted from Hirayama et al (Hirayama et al., 2012). Rectangles represent exons. Pink diamonds represent enhancer regions (HS5-1 enhances Pcdhα3–α12 and Pcdhαc1; CCR enhances Pcdhaβ1–β22).

(B) Genome browser shot showing TF occupancy in ESCs and neurons at the *Pcdh* gene locus. Neuron-specific and cell type-invariant NF-Y sites are highlighted in yellow and blue, respectively.

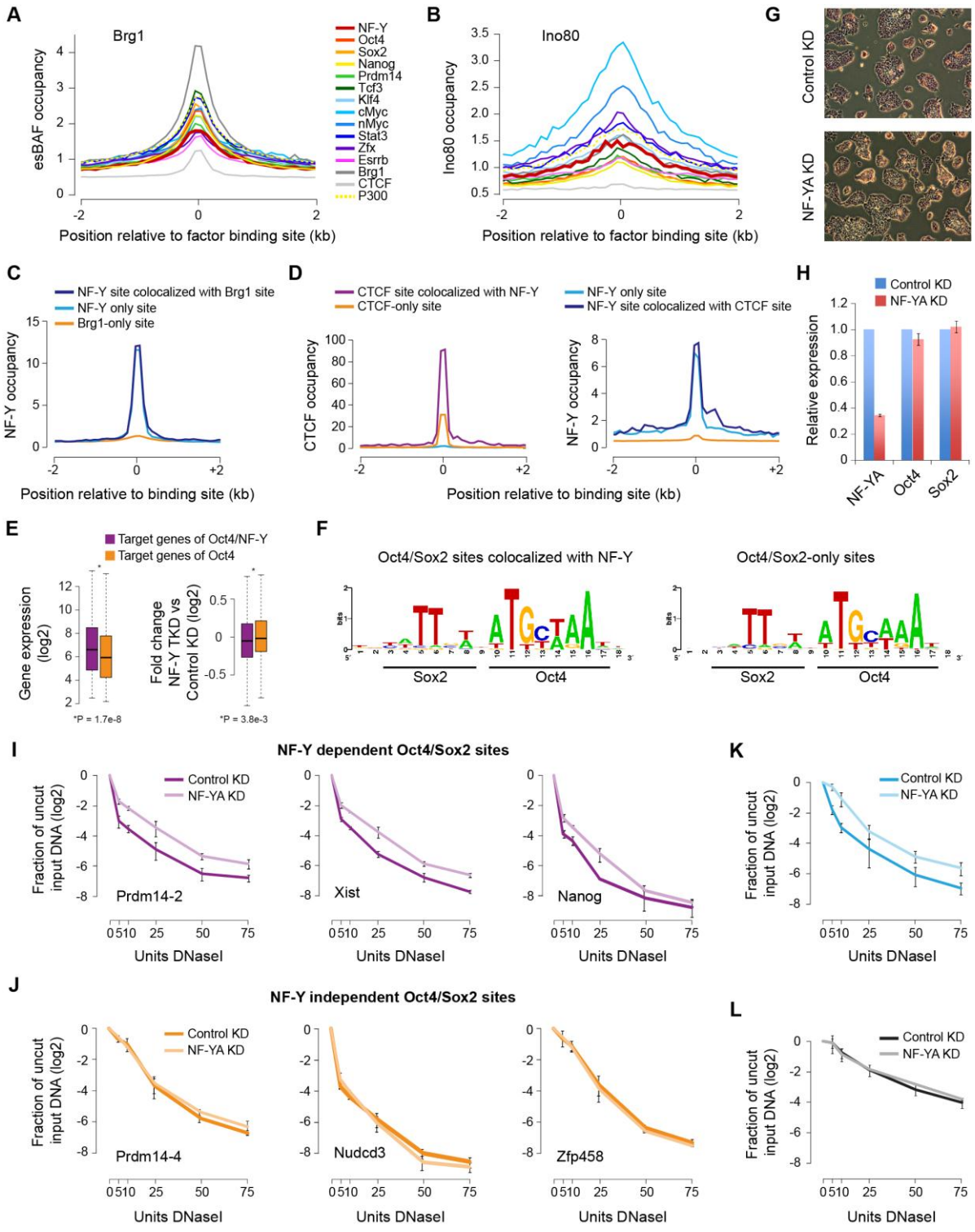


Figure S7 (related to Figures 6 and 7). Depletion of NF-Y diminishes chromatin accessibility at sites co-bound by NF-Y and Oct4/Sox2 but not at sites bound only by Oct4/Sox2.

(A) Chromatin remodeling complex esBAF occupancy, as measured by Brg1 ChIP-Seq (Ho et al., 2009) at various TF (distal) binding sites in ESCs.

(B) Chromatin remodeling complex Ino80 occupancy, as measured by Ino80 ChIP-Seq (Wang et al., 2014), at various TF (distal) binding sites in ESCs. Colors and labels same as in Figure S7A.
(C) NF-Y occupancy at distal NF-Y sites colocalized with Brg1 (blue), distal NF-Y-only sites (cyan), and distal Brg1 only sites (orange).
(D) CTCF (left) and NF-Y (right) occupancy at distal CTCF sites colocalized with NF-Y (purple), distal CTCF-only sites (orange), distal NF-Y sites colocalized with CTCF (cyan), and distal NF-Y only sites (blue).
(E) *Left*: Expression levels of target genes of Oct4 with and without NF-Y co-occupancy in control ESCs. *Right*: Expression fold change of target genes of Oct4 with and without NF-Y co-occupancy in NF-Y TKD vs Control KD ESCs.
(F) Consensus sequence motif enriched within Oct4/Sox2 binding sites with and without NF-Y co-occupancy using *de novo* motif analysis.
(G) Morphology and alkaline phosphatase staining of control KD and *NF-YA* KD ESCs 48h after siRNA transfection.
(H) RT-qPCR analysis of relative mRNA levels of NF-Y subunits *NF-YA*, *Oct4*, and *Sox2* in Control and NF-YA KD ESCs. Data are normalized to *Actin*, *HAZ*, and *TBP*. Error bars represent S.E.M. of three experiments.
(I) DNase I and qPCR analysis of NF-Y dependent Oct4/Sox2 sites (n = 3 each) in Control and NF-YA KD ESCs. Error bars represent S.E.M. of four experiments.
(J) DNase I and qPCR analysis of NF-Y independent Oct4/Sox2 sites (n = 3 each) in Control and NF-YA KD ESCs. Error bars represent S.E.M. of four experiments.
(K) DNase I and qPCR analysis of a distal NF-Y site within the Srf gene, bound by NF-Y (but not Oct4, Sox2, Nanog, Prdm14, Esrrb or Klf4), in Control and NF-YA KD ESCs. Error bars represent S.E.M. of four experiments.
(L) DNase I and qPCR analysis of a control site, that does not bind NF-Y, in Control and NF-YA KD ESCs. Error bars represent S.E.M. of four experiments.

Table S1 (related to Figures 1, 5, 6, and 7). Primers used for ChIP-qPCR and DNase I hypersensitivity analysis

Primer name	Forward primer	Reverse primer
1700081H04Rik	AGGGTCCGGTAACTCCTCTC	GGCTCATACCTGTCGTTCCA
AK141396	CAGCACACAGGAGGAGGACT	CGGCTGGTCTCTAACTGCAT
Calm2	AGGGCCATTCACAACAAAAG	TGTTAGGATGCGTGCTTGTC
Cdc25c	GGAACAGAAGCGATTCTTCG	TCATGTTCTGGTGCGATGAT
Esrrb	TGGGAAGTGTTGCTATTCCA	TCAGAGCTCCAGATCCCCTA
Foxi3	GGCCTCGAGGTTAGAGTCCT	AGGGAACTAGCGTGTTCAGC
Khsrp	CGGCAAGATAGTCGTCAACA	AGTCACGTCCAAATCCAAGC
Nanog	CCGCTCCTTTTCAGCACTAA	CAGGGAAGCGGTTTGAATAG
NR_040693	AGGGTCCGGTAACTCCTCTC	GGCTCATACCTGTCGTTCCA
Nudcd3	GGATGTGTGCAATGTTGAGG	CTCACTTCCGCTCCTTCTGT
Nxn	CAACTGTCGACTGTGCGTTT	TAGCTTTGACTGCCCTGACC
Paip2b	GAGAATGCCGCTCACTTACC	TCAGAGCTGGGAGTCTCCAT
Prdm14-2	AAGCAGCAGGGTGGAGATAA	CAAACGGATTGGAGGTTGAT
Prdm14-4	GCAAATTAGGCTCATTCGTG	AATCCTCGGGAGTTCTGGTT
Rnf5	ACCCTTGCACGATGATGATT	CGAGTTCTGTAGGCCTGAGC
Srf	AGGCCTTGAGAACCAAGCTA	AACGCCTTTACCAATCAACG
Xist1	AACCCTTTTAAGTCCACTGTAAATTCC	TAGAGAGCCAGACAATGCTAAGCC
Zfp568	TCCGCCCCTATATCTTGTTG	CTGGGTCTCTGGACTTCAGG
Zic3	CCGCAGCTACCCAATCAG	AATCACTCACTCCTCGCACA
Ager (control #1)	ACCCCACTCAGACATGAACC	TGGCAATTCCCCTCAGTTAG
chr8 (control #2)	AAGGGGCCTCTGCTTAAAAA	AGAGCTCCATGGCAGGTAGA
Pdp2 (control #3)	CTGGACCTCTCTGGTTCTGG	TTGGCCTCTTAGCGACAAGT
Tmem179 (control #4)	TTCCGTGTCCCCAGAATAAG	TTAAGCCATCCACTCCCTTG
Tpg (control #5)	TGGAGCTCTTTCATGTTCTTCCTT	ATGAATGGGCTTCTTGAATTTCTACT

Table S2 (related to Figures 3 and 4). Gene specific primers used for RT-qPCR analysis

Primer name	**Forward primer**	**Reverse primer**
Brachyury	ACCCAGACTCGCCCAATTT	CACGATGTGAATCCGAGGTT
Ccnb2	ACTGGCTGGTCCAAGTCCAT	ATATTTGGAAGCCAGGAGCA
Cdc25c	TTCAGAAGACCCAATGGAGTG	GAATGGCGTTCATGTCACAG
Cdkn1a	TTGCCAGCAGAATAAAAGGTG	TTTGCTCCTGTGCGGAAC
Cdx2	AGGCTGAGCCATGAGGAGTA	CGAGGTCCATAATTCCACTCA
Eomes	ACCGGCACCAAACTGAGA	AAGCTCAAGAAAGGAAACATGC
Esrrb	CAGGCAAGGATGACAGACG	GAGACAGCACGAAGGACTGC
Fgf5	AAAACCTGGTGCACCCTAGA	CATCACATTCCCGAATTAAGC
Foxa2	CGAGCTAAAGGGAGCACCT	TAATGGTGCTCGGGCTTC
Gadd45g	GGATAACTTGCTGTTCGTGGA	AAGTTCGTGCAGTGCTTTCC
Gata3	TTATCAAGCCCAAGCGAAG	TGGTGGTGGTCTGACAGTTC
Gata4	GGAAGACACCCCAATCTCG	CATGGCCCCACAATTGAC
Gata6	GGTCTCTACAGCAAGATGAATGG	TGGCACAGGACAGTCCAAG
Hand1	CAAGCGGAAAAGGGAGTTG	GTGCGCCCTTTAATCCTCTT
Id2	GACAGAACCAGGCGTCCA	AGCTCAGAAGGGAATTCAGATG
Krt18	AGATGACACCAACATCACAAGG	TCCAGACCTTGGACTTCCTC
Krt8	AGTTCGCCTCCTTCATTGAC	GCTGCAACAGGCTCCACT
Lefty1	ACTCAGTATGTGGCCCTGCTA	AACCTGCCTGCCACCTCT
Lefty2	CACAAGTTGGTCCGTTTCG	GGTACCTCGGGGTCACAAT
Mixl1	CATGTACCCAGACATCCACTTG	ACTCTGGCGCCTGGACTT
Nanog	AAGCAGAAGATGCGGACTGT	ATCTGCTGGAGGCTGAGGTA
Nestin	TCCCTTAGTCTGGAAGTGGCTA	GGTGTCTGCAAGCGAGAGTT
NF-YA	GGCACAATTCTCCAGCAAG	GGCTCCTGTCTGAACGATCT
NF-YB	GCTTCATCACGTCGGAAGCAAGCG	GTCCATCTGTGGCGGAGACTGC
NF-YC	CCCACTGGCTCGTATTAAGAA	GGCTCGAAGAGTCAGCTCAG
Nkx2.2	GCAGCGACAACCCCTACA	ATTTGGAGCTCGAGTCTTGG
Oct4	CCAATCAGCTTGGGCTAGAG	CCTGGGAAAGGTGTCCTGTA
Pax6	GTTCCCTGTCCTGTGGACTC	ACCGCCCTTGGTTAAAGTCT
Prdm14	GGCCATACCAGTGCGTGTA	TGCTGTCTGATGTGTGTTCG
Tbx3	AGATCCGGTTATCCCTGGGAC	CAGCAGCCCCCACTAACTG
Tcl1	ACCTTGGGGGAAGCTATGTC	CTTGGAGCCCAGTGTAGAGG
Actin (control #1)	AAGGCCAACCGTGAAAAGAT	GTGGTACGACCAGAGGCATAC
HAZ (control #2)	CGTTGTAGGAGCCCGTAGGTCAT	TCTGGTTGCGGAAGCATTGGG
TBP (control #3)	CTGAAGAAAGGGAGAATCATGG	TGTCTTTGTTGCTCTTCCAAAA

Supplemental Experimental Procedures

Mouse ES Cell Culture, RNAi, and AP staining

Mouse ESC culture, siRNA transfection, and alkaline phosphatase (AP) staining were performed as previously described (Freudenberg et al., 2012). Briefly, E14Tg2a ESCs were maintained on gelatin-coated plates in the ESGRO complete plus clonal grade medium (Millipore). For siRNA transfections, ESCs were cultured in M15 medium: DMEM (Invitrogen) supplemented with 15%FBS, 10µM 2-mercaptoethanol, 0.1mM nonessential amino acids (Invitrogen), 1x EmbryoMax nucleosides (Millipore), 1U/ml of ESGRO mLIF (Millipore). ESCs (~$25x10^3$) were transfected with siRNAs at 50nM using lipofectamine 2000 (Invitrogen) at day 0, re-plated and re-transfected at 48h, and collected after 96h (see Figure S2A). AP staining was performed using Alkaline Phosphatase Detection Kits from Stemgent (00-0055) according to the manufacturer's instructions. Gene specific siRNAs used: NF-YA (Invitrogen, MSS247473), NF-YB (Qiagen, SI01327193), NF-YC (Qiagen, SI05348217), non-targeting control (Dharmacon, D-001810-02-50).

Chromatin Immunoprecipitation (ChIP)

Mouse ESCs ($1x10^7$) were cross-linked with 1% formaldehyde in DMEM for 10 min, and the reaction was quenched by the addition of glycine at a final concentration of 125 mM for 5 min. Cells were scraped, pelleted, washed twice with PBS, and resuspended in 1ml of lysis buffer A [50mM Hepes pH7.5; 140mM NaCl; 1mM EDTA; 10% Glycerol; 0.5% IGEPAL CA-630; 0.25% Triton X-100; 1x Complete protease inhibitor mixture (Roche), 200nM PMSF]. After 10 min on ice, the cells were pelleted and resuspended in 1ml of lysis buffer B [10mM Tris-HCl pH8.0; 200mM NaCl; 1mM EDTA; 0.5mM EGTA; 1x protease inhibitors, 200nM PMSF]. After 10min at room temperature, cells were sonicated in lysis buffer C [10mM Tris-HCl pH8.0; 100mM NaCl; 1mM EDTA; 0.5mM EGTA; 0.1% sodium deoxycholate; 0.5% N-lauroylsarcosine; 1x protease inhibitors, 200nM PMSF] using Diagenode Bioruptor for 16 cycles (30sec ON; 50sec OFF) to obtain ~200–500 bp fragments. Cell debris were pre-cleared by centrifugation at 14,000 rpm for 20 min, and 25µg of chromatin was incubated with either NF-YA (Santa Cruz, G-2, sc-17753X), NF-YB (Santa Cruz, FL-207X, sc-13045X), NF-YC (Santa Cruz, N-19X, sc-7715X), Oct4 (Santa Cruz, N-19, sc-8628X), Sox2 (Santa Cruz, Y-17, sc-17320X), or histone H3 (Abcam, ab1791) antibodies overnight at 4˚C. Protein A/G-conjugated magnetic beads (Pierce Biotech) were added the next day for 2 hours. Subsequent washing and reverse cross-linking were performed as previously described (Heard et al., 2001). ChIP enrichment for a primer-set was evaluated using quantitative PCR as percentage of input, and normalized to a negative primer-set. See **Table S1** for a list of ChIP primers used.

Quantitative RT-PCR

Total RNAs were prepared from cells using Trizol (Invitrogen), and cDNAs were generated using the iScript kit (Bio-Rad) according to the manufacturer's instructions. Quantitative PCRs were performed on the Bio-rad CFX-96 or CFX-384 Real-Time PCR System using the Bio-rad SsoFast EvaGreen supermix. Three or more biological replicates were performed for each experiment. Data are normalized to *Actin*, *Haz* and *TBP* expression, and plotted as mean +/- S.E.M. See **Table S2** for a list of gene specific primers used.

Western Blot

Cell pellets, lysed in RIPA buffer (25mM Tris-HCl, pH 7.4, 150mM NaCl, 1% NP-40, 1% Sodium deoxycholate) with protease inhibitors, were sonicated using Bioruptor (Diagenode) for three cycles (30 seconds ON; 50 seconds OFF). The lysate was boiled with SDS-PAGE sample buffer, loaded onto NuPAGE gel, and transferred to 0.22µM PVDF membranes. Each membrane was treated with appropriate primary and secondary antibodies. Antibodies used: NF-YA (Santa Cruz, H-209, sc-10779X), NF-YB (Santa Cruz, FL-207X, sc-13045X), NF-YC (Santa Cruz, N-19X, sc-7715X), and Ran (BD Bioscience, 610341). The membrane was then incubated with a horseradish-peroxidase-conjugated secondary antibody and developed with enhanced chemi-luminescence PLUS reagent (Amersham). Loading was normalized based on Ran.

Immunoflourescence

Mouse ESCs transfected with *NF-YA* or control siRNA were grown on gelatin-coated glass cover-slips, and stained for appropriate antibodies for 1hr at 37°C as previously described (Cinghu et al., 2012)(Cinghu et al., 2012), followed by washing in phosphate buffered saline (PBS). Antibodies used: NF-YA (Santa Cruz, G-2X, sc-13057X), NF-YC (Santa Cruz, N-19X, sc-7715X), Nanog (BD Bioscience, 560259). The cover-slips were then treated with appropriate secondary antibody (Alexa 488 or Alexa 594; Invitrogen) for another 1hr at 37°C. Staining of nuclei was accomplished by incubation with DAPI (5mg/ml) for 10min. The slides were then washed extensively in PBS and mounted using Prolong gold anti-fade (Invitrogen). Specimens were viewed using a Zeiss N710 confocal microscope.

Cell Cycle Analysis

Cell cycle analysis was performed using flow cytometry as previously described (Cinghu et al., 2014). Briefly, 48 or 72 hours after siRNA transfection, cells were fixed with 70% ethanol overnight. The next day, cells were washed with PBS, treated with 10µg/ml RNase A (Roche, 10109169001) for 1 hour at room temperature and then stained with 50µg/ml propidium iodide (Sigma, P4170) for at least 10 minutes before loading on a FACSCalibur flow cytometer

(BD Biosciences). Raw data was analyzed using the CellQuest Pro software, and the percentage of each cell cycle phase was calculated with the Flow Jo software.

Co-Immunoprecipitation

ESCs ($1.2x10^7$) were lysed with 1 mL lysis buffer [20 mM Tris-HCl, 2 mM EDTA, 137 mM NaCl, 10% glycerol, 1% Triton-X] for 30 min at 4^oC. Lysates were incubated with antibodies against NF-YA (Santa Cruz, H-209, sc-10779X), NF-YC (Santa Cruz, N-19X, sc-7715X), Oct4 (Santa Cruz, N-19, sc-8628X), or IgG (Santa Cruz, sc-2027) overnight at 4^oC. Magnetic Dynabeads (Life Technologies) were added for 3h at 4^oC. Beads were washed three times with lysis buffer, and proteins were eluted with Laemmli sample buffer at 100^oC for 15 min. Western blots were performed according to standard protocols and were developed using the Odyssey Infrared Imaging System (Li-Cor Biosciences).

DNase I Hypersensitivity

Mouse ESCs treated with non-targeting control siRNA or *NF-YA* siRNA were collected 48 hours post-transfection in cold PBS. Nuclei isolation and DNase I digestion were performed as previously described (Burch and Weintraub, 1983), with minor modifications. Nuclei were isolated by incubation of 10^7 cells for 10 min on ice with 5 ml RSB buffer [10 mM Tris–HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.15 mM spermine and 0.5 mM spermidine, 1mM PMSF, 0.5% IGEPAL], and pelleted by centrifugation at 300*g* and 4˚C for 10 min. Nuclei were then resuspended in 1 ml DNase reaction buffer [40mM Tris-HCl pH7.4, 10mM NaCl, 6mM $MgCl_2$, 1mM $CaCl_2$, 0.15mM Spermine, 0.5mM Spermidine] and counted. Additional resuspension buffer was used to generate equal concentrations of nuclei between samples. Nuclei from $5x10^5$ cells were aliquoted into microcentrifuge tubes and incubated at 37°C for 5 min with varying amounts of DNase (0U to 75U, Worthington). Digestion was stopped by addition of an equal volume of termination buffer [10 mM Tris pH 7.4, 50 mM NaCl, 100 mM EDTA, 2% SDS, 10 µg/ml RNAse cocktail]. The nuclei were then incubated at 55˚C for 15 min, followed by addition of 2 µl of 20 mg/ml Proteinase K. Reaction mixtures were incubated overnight at 55°C, followed by a phenol–chloroform extraction followed by chloroform extraction of the DNA. The DNA was then precipitated and resuspended in 100 µl H2O.

Microarray Analysis

Raw microarray data files were processed and normalized using RMA and Affymetrix Mouse Genome 430 2.0 Array annotation using applicable R/Bioconductor packages to generate single expression intensity measure per gene per sample. All subsequent analyses were carried out on

the $\log_2$ scale. Differential expression analysis (NF-Y TKD vs. control) was performed using limma R package (Smyth, 2005), followed by Benjamin-Hochberg multiple testing correction to control the false discovery rate (FDR). Genes were considered differentially expressed if they had an FDR≤0.05 and were at least 2.0 fold up- or down-regulated. To compare gene expression changes upon *NF-Y* depletion to those observed after KD or knockout (KO) of other pluripotency-associated factors in previously published studies, respective reference datasets were downloaded from GEO and processed as previously described (Freudenberg et al., 2012). Correlation between global gene expression changes upon KD/KO of various factors were computed and visualized as a heatmap. Published mircoarray datasets used for comparative analysis: Oct4 KD and Nanog KD (Loh et al., 2006); Sox2 KD, Tbx3 KD, Esrrb KD, and Tcl1 KD (Ivanova et al., 2006); Ncl KD (Cinghu et al., 2014); Brg1 KO and LIF withdrawal (Ho et al., 2011); Klf2/4/5 TKD (Jiang et al., 2008); Tet1 KD (Freudenberg et al., 2012); Tcf3 KO (Yi et al., 2011); Suz12 KO (Pasini et al., 2007); Sall4 KD (Lim et al., 2008); embryoid body differentiation (Hailesellasse Sene et al., 2007). Principal component analysis (PCA) of expression profiles from control and NF-Y depleted ESCs and previously published data from wild-type and differentiating ESCs (Aiba et al., 2009; Nishiyama et al., 2009) was performed using R and visualized using R package "rgl."

ChIP-Seq Data Analysis

Single-end 36 bp reads generated from NF-YA, NF-YB, and NF-YC ChIP-Seq were aligned to the mouse reference genome (mm9 assembly) using Bowtie version 0.12.8 (Langmead et al., 2009), and only those reads that mapped to unique genomic locations with at most two mismatches were retained for further analysis. For visualization on the UCSC Genome Browser, and generation of screenshots and read density plots, the data was normalized to reads per million (RPM) and plotted as histograms. For binding site definition (peak calls), aligned reads were processed using SISSRs (Jothi et al., 2008; Narlikar and Jothi, 2012) using default settings. Genome-wide distribution of binding sites was determined with reference to RefSeq gene annotations. Binding sites located withing 500 bp of transcription start sites (TSSs) were defined as promoter-proximal or proximal sites, with the rest defined as distal sites. Genes binding NF-Y within 500 bp of their TSSs were defined as proximal NF-Y target genes, and those that bind NF-Y within 50 Kb but not within 500 bp of their TSSs were defined as distal NF-Y target gene. Binding sites for two transcription factors were defined to be colocalized if the centers of the corresponding peaks are within 500 bp of each other. CpG island annotations were downloaded from UCSC genome (Karolchik et al., 2014). Published ESC ChIP-Seq datasets used for comparative analysis: Oct4, Sox2, Nanog, Tcf3, and Suz12 (Marson et al., 2008); Prdm14 (Ma et al., 2011); Esrrb, Klf4, CTCF, cMyc, nMyc, Zfx, Smad1 (Chen et al., 2008); Stat3 and H3K27me3 (Ho et al., 2011); H3K4me1 and P300 (Creyghton et al., 2010) , H3K4me3 (Agarwal and Jothi,

2012); H3K27ac (ENCODE, GSE31039); DNase (ENCODE, GSE37074); Hi-C (Dixon et al., 2012); Brg1 (Ho et al., 2009); and Ino80 (Wang et al., 2014). Published neuron ChIP-Seq datasets used for comparative analysis: NF-YA and H3K27me3 (Tiwari et al., 2012); CTCF (ENCODE, GSE49847); NPAS4, CREB, SRF, CBP, H3K4me1 and H3K27ac (Kim et al., 2010), DNase (ENCODE, GSE37074).

Functional and pathway enrichment analysis

Gene Ontology (GO) functional enrichment analysis was performed using DAVID (Huang da et al., 2009), and pathway enrichment analysis was performed using PANTHER classification system (Mi et al., 2013).

Motif Analysis

Relevant sequences spanning 200 nucleotides around the centers of NF-Y binding sites were retrieved from the reference genome, and *De-novo* motif search was performed using MEME (Bailey and Elkan, 1994) using zoops (zero or one motif occurrence per sequence) option. Search for known TF motifs, obtained from TRANSFAC (Matys et al., 2006) and JASPAR (Mathelier et al., 2014), was performed using the MAST (Bailey and Gribskov, 1998).

Supplemental References

Agarwal, S.K., and Jothi, R. (2012). Genome-wide characterization of menin-dependent H3K4me3 reveals a specific role for menin in the regulation of genes implicated in MEN1-like tumors. PLoS One *7*, e37952.

Aiba, K., Nedorezov, T., Piao, Y., Nishiyama, A., Matoba, R., Sharova, L.V., Sharov, A.A., Yamanaka, S., Niwa, H., and Ko, M.S. (2009). Defining developmental potency and cell lineage trajectories by expression profiling of differentiating mouse embryonic stem cells. DNA Res *16*, 73-80.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol *2*, 28-36.

Bailey, T.L., and Gribskov, M. (1998). Combining evidence using p-values: application to sequence homology searches. Bioinformatics *14*, 48-54.

Burch, J.B., and Weintraub, H. (1983). Temporal order of chromatin structural changes associated with activation of the major chicken vitellogenin gene. Cell *33*, 65-76.

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.

Cinghu, S., Goh, Y.M., Oh, B.C., Lee, Y.S., Lee, O.J., Devaraj, H., and Bae, S.C. (2012). Phosphorylation of the gastric tumor suppressor RUNX3 following H. pylori infection results in its localization to the cytoplasm. J Cell Physiol *227*, 1071-1080.

Cinghu, S., Yellaboina, S., Freudenberg, J.M., Ghosh, S., Zheng, X., Oldfield, A.J., Lackford, B.L., Zaykin, D.V., Hu, G., and Jothi, R. (2014). Integrative framework for identification of key cell identity genes uncovers determinants of ES cell identity and homeostasis. Proc Natl Acad Sci U S A.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A.*, et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A *107*, 21931-21936.

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

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat Methods *9*, 215-216.

Freudenberg, J.M., Ghosh, S., Lackford, B.L., Yellaboina, S., Zheng, X., Li, R., Cuddapah, S., Wade, P.A., Hu, G., and Jothi, R. (2012). Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. Nucleic Acids Res *40*, 3364-3377.

Hailesellasse Sene, K., Porter, C.J., Palidwor, G., Perez-Iratxeta, C., Muro, E.M., Campbell, P.A., Rudnicki, M.A., and Andrade-Navarro, M.A. (2007). Gene function in early mouse embryonic stem cell differentiation. BMC Genomics *8*, 85.

Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C.D., and Spector, D.L. (2001). Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. Cell *107*, 727-738.

Hirayama, T., Tarusawa, E., Yoshimura, Y., Galjart, N., and Yagi, T. (2012). CTCF is required for neural development and stochastic expression of clustered Pcdh genes in neurons. Cell Rep *2*, 345-357.

Ho, L., Jothi, R., Ronan, J.L., Cui, K., Zhao, K., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. Proc Natl Acad Sci U S A *106*, 5187-5191.

Ho, L., Miller, E.L., Ronan, J.L., Ho, W.Q., Jothi, R., and Crabtree, G.R. (2011). esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. Nat Cell Biol *13*, 903-913.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533-538.

Jiang, J., Chan, Y.S., Loh, Y.H., Cai, J., Tong, G.Q., Lim, C.A., Robson, P., Zhong, S., and Ng, H.H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol *10*, 353-360.

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 Res *36*, 5221-5231.

Karolchik, D., Barber, G.P., Casper, J., Clawson, H., Cline, M.S., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M.*, et al.* (2014). The UCSC Genome Browser database: 2014 update. Nucleic Acids Res *42*, D764-770.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S.*, et al.* (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182-187.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lim, C.Y., Tam, W.L., Zhang, J., Ang, H.S., Jia, H., Lipovich, L., Ng, H.H., Wei, C.L., Sung, W.K., Robson, P.*, et al.* (2008). Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. Cell Stem Cell *3*, 543-554.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J.*, et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet *38*, 431-440.

Ma, Z., Swigut, T., Valouev, A., Rada-Iglesias, A., and Wysocka, J. (2011). Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. Nat Struct Mol Biol *18*, 120-127.

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.

Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C.Y., Chou, A., Ienasescu, H.*, et al.* (2014). JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res *42*, D142-147.

Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K.*, et al.* (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res *34*, D108-110.

Mi, H., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2013). Large-scale gene function analysis with the PANTHER classification system. Nat Protoc *8*, 1551-1566.

Narlikar, L., and Jothi, R. (2012). ChIP-Seq data analysis: identification of protein-DNA binding sites with SISSRs peak-finder. Methods Mol Biol *802*, 305-322.

Nishiyama, A., Xin, L., Sharov, A.A., Thomas, M., Mowrer, G., Meyers, E., Piao, Y., Mehta, S., Yee, S., Nakatake, Y.*, et al.* (2009). Uncovering early response of gene regulatory networks in ESCs by systematic induction of transcription factors. Cell Stem Cell *5*, 420-433.

Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol Cell Biol *27*, 3769-3779.

Smyth, G.K. (2005). Limma: linear models for microarray data. Bioinformatics and Computational Biology Solutions using R and Bioconductor, 397-420.

Tiwari, V.K., Stadler, M.B., Wirbelauer, C., Paro, R., Schubeler, D., and Beisel, C. (2012). A chromatin-modifying function of JNK during stem cell differentiation. Nat Genet *44*, 94-100.

Wang, L., Du, Y., Ward, J.M., Shimbo, T., Lackford, B., Zheng, X., Miao, Y.-l., Zhou, B., Han, L., Fargo, D.C.*, et al.* (2014). INO80 Facilitates Pluripotency Gene Activation in Embryonic Stem Cell Self-Renewal, Reprogramming, and Blastocyst Development. Cell Stem Cell *14*, 575-591.

Yi, F., Pereira, L., Hoffman, J.A., Shy, B.R., Yuen, C.M., Liu, D.R., and Merrill, B.J. (2011). Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. Nat Cell Biol *13*, 762-770.