Cell Stem Cell
Resource

Distinct Epigenomic Landscapes of Pluripotent and Lineage-Committed Human Cells

R. David Hawkins,[1,9] Gary C. Hon,[1,9] Leonard K. Lee,[1] QueMinh Ngo,[1] Ryan Lister,[2] Mattia Pelizzola,[2] Lee E. Edsall,[1] Samantha Kuan,[1] Ying Luu,[1] Sarit Klugman,[1] Jessica Antosiewicz-Bourget,[3] Zhen Ye,[1] Celso Espinoza,[1] Saurabh Agarwahl,[1] Li Shen,[4] Victor Ruotti,[3,6] Wei Wang,[4] Ron Stewart,[3,6] James A. Thomson,[3,6,7,8] Joseph R. Ecker,[2] and Bing Ren[1,5,*]

[1]Ludwig Institute for Cancer Research, La Jolla, CA 92093, USA
[2]Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
[3]Morgridge Institute for Research, Madison, WI 53707, USA
[4]Department of Chemistry and Biochemistry
[5]Department of Cellular and Molecular Medicine, Institute of Genomic Medicine, and UCSD Moores Cancer Center
University of California San Diego, La Jolla, CA 92093, USA
[6]Genome Center of Wisconsin, Madison, WI 53706, USA
[7]Wisconsin National Primate Research Center, University of Wisconsin-Madison, Madison, WI 53706, USA
[8]Department of Anatomy, University of Wisconsin-Madison, Madison, WI 53715, USA
[9]These authors contributed equally to this work
*Correspondence: biren@ucsd.edu
DOI 10.1016/j.stem.2010.03.018

SUMMARY

Human embryonic stem cells (hESCs) share an identical genome with lineage-committed cells, yet possess the remarkable properties of self-renewal and pluripotency. The diverse cellular properties in different cells have been attributed to their distinct epigenomes, but how much epigenomes differ remains unclear. Here, we report that epigenomic landscapes in hESCs and lineage-committed cells are drastically different. By comparing the chromatin-modification profiles and DNA methylomes in hESCs and primary fibroblasts, we find that nearly one-third of the genome differs in chromatin structure. Most changes arise from dramatic redistributions of repressive H3K9me3 and H3K27me3 marks, which form blocks that significantly expand in fibroblasts. A large number of potential regulatory sequences also exhibit a high degree of dynamics in chromatin modifications and DNA methylation. Additionally, we observe novel, context-dependent relationships between DNA methylation and chromatin modifications. Our results provide new insights into epigenetic mechanisms underlying properties of pluripotency and cell fate commitment.

INTRODUCTION

Pluripotent embryonic stem cells (ESCs) possess the ability to differentiate into multiple cell lineages in the body (Thomson et al., 1998). The underlying molecular mechanisms of pluripotency and cell fate commitment have not been completely understood. At least two models account for differences between pluripotent stem cells and lineage-committed cells. In the first, "master switches" activate distinct networks of transcription factors that govern the transcriptional program of each cell type and dictate cellular properties (Marson et al., 2008). For example, the OCT4/SOX2/NANOG network enables self-renewal properties of ESCs, and ectopic expression of these factors together with additional transcription factors reprogram somatic cells into pluripotent cells (iPSCs) (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007). But the efficiency of reprogramming is low, suggesting the involvement of additional factors or mechanisms. Additionally, lineage-committed cells are stable over many cell divisions, after the initial master switch transcription factors are no longer expressed. Such a phenomenon, referred to as "cellular memory," has been difficult to explain by the transcription factor network model (Ringrose and Paro, 2004).

Another model for cellular memory involves the cell's epigenomic landscape consisting of covalent modifications to DNA or histones (Ringrose and Paro, 2004), which either enable or prevent parts of the genome to be active in different cell types. In stem cells, the epigenome is highly malleable and responsive (Meshorer and Misteli, 2006), unlike that of somatic cells. Multiple lines of evidence support this model. First, ESCs have a higher number of "bivalent" or "poised" promoters marking important developmental regulators compared to differentiated cells, as indicated by the repressive mark histone H3 lysine 27 trimethylation (H3K27me3) and the active chromatin modification H3K4me3 (Azuara et al., 2006; Bernstein et al., 2006; Pan et al., 2007). Second, immunofluorescent imaging showed that after differentiation, mouse ESCs display increased heterochromatin in the nucleus (Meshorer and Misteli, 2006). Additionally, depletion of the Jmjd1a and Jmjd2c demethylases for the heterochromatin modification H3K9me3 results in stem cell differentiation (Loh et al., 2007). Third, DNA methylation is found at the promoters of critical pluripotency genes such as *Oct4* during differentiation and is responsible to keep such genes silent in differentiated cells (Ben-Shushan et al., 1993; Deb-Rinker et al., 2005).

Recent large-scale analyses of DNA methylation and histone modifications revealed dynamic chromatin states and DNA methylation status at promoters and most CpG islands (Brunner et al., 2009; Meissner et al., 2008), showing that the methylation state of H3K4 is a good indicator of promoter DNA methylation levels in mammalian cells. This is consistent with prior studies indicating that H3K4 methylation disrupts DNA methylation by inhibiting contact of DNMTs with histones, whereas promoters marked only with H3K27me3 in mouse ESCs are more likely to exhibit de novo DNA methylation after differentiation (Meissner et al., 2008; Mohn et al., 2008). Although these insights suggest potential mechanisms of epigenetic regulation at promoters, the epigenetic regulatory mechanisms outside of promoters remains largely unclear.

To obtain a better understanding of the epigenomic landscapes in the pluripotent and differentiated cell states and explore the links between histone modifications and DNA methylation throughout the genome, we conducted a comprehensive analysis of 11 histone modifications and a recently acquired genome-wide nucleotide resolution map of DNA methylation in H1 human embryonic stem cells (hESCs) and fetal lung fibroblasts (IMR90) (Lister et al., 2009). We show a large-scale expansion of H3K9me3 and H3K27me3 domains in differentiated cells relative to hESCs, which selectively affects genes related to pluripotency, development, and lineage-specific functions. We also find multiple epigenetic mechanisms by which key pluripotent transcription factors are silenced in somatic cells. Finally, our analysis of both cell types reveals context-dependent, complex relationships between chromatin modifications and DNA methylation.

RESULTS AND DISCUSSION

Chromatin Modification Landscapes

We performed ChIP-Seq (Johnson et al., 2007) experiments to identify global histone modification patterns for H2BK5ac, H3K4me1, H3K4me2, H3K4me3, H3K9ac, H3K9me3, H3K18ac, H3K27ac, H3K27me3, H3K36me3, and H4K5ac. Antibodies were validated for specificity with peptide dot blot assays and western blotting (Figure S1 available online). ChIP-Seq experiments produced 5 to 40 million monoclonal, uniquely mapped tags per chromatin modification and input DNA for well-correlated biological replicates in H1 hESCs and IMR90 fibroblasts (Figure 1; Table S1). We chose acetylation marks, which are indicative of open chromatin, as well as H3K4 methylation marks featured at active enhancers and promoters (Heintzman et al., 2007, 2009). We contrasted this architecture with the repressive structures of H3K27me3 and H3K9me3 and the genic localization of H3K36me3, a transcription elongation-coupled chromatin mark (Krogan et al., 2003; Li et al., 2002).

Visualization of these maps indicates that histone modifications show two distinct types of spatial distributions: small, punctuated peaks and large, spreading domains (Figure 1). Modifications such as H3K4 methylation and acetylation are typically found in peak-like structures (Figures 1A and 1B), whereas H3K36me3, H3K27me3, and H3K9me3 are usually seen in broad domains of varying widths (Figures 1A, 1C, and 1D). It is possible that these two types of chromatin-modification patterns correspond to distinct physical structures of chromatin fibers in vivo (Figure 1E).

Most peak-finding programs rely on the assumption of stranded distributions of tags at peaks, and as such cannot detect the broad domains typical of histone modifications (Kharchenko et al., 2008; Zhang et al., 2008). To systematically enumerate genomic regions enriched for chromatin modifications, we devised a computational method called ChromaBlocks, which is tailored to identify both peak-like structures and large chromatin domains from ChIP-Seq data (see Experimental Procedures). This analysis yielded between 9,414 and 43,536 blocks for each chromatin modification in hESC or IMR90, occupying between 68 million and 510 million base pairs (Table S2). Analysis with permuted data indicates that this method is highly specific (median FDR = 0.97%, Table S2) and approaches saturation around 10 million unique monoclonal reads (Figures S2D–S2G). With this tool in place, we are able to assess varying chromatin structures and compare how the chromatin landscape of hESCs differs from that of differentiated cells.

Redistribution of Large Domains of Chromatin Modifications

Comparing the domains in hESC and IMR90, we find that the chromatin architecture at a surprisingly large fraction of the genome is different (Figures 2A and 2B). Many of the H3K9me3 domains in hESCs appear small and interspersed (median length = 6.9 kb) but are expanded significantly in IMR90 cells (median length = 11.4 kb, $p < 1E\text{-}15$) (Figures 2A, 2C, and 2D; Tables S3 and S4). Correspondingly, the number of base pairs spanned by H3K9me3 in IMR90 (510 Mb) is 3.4 times larger than in hESCs (148 Mb) (Figure 2C). Similarly, the number of base pairs spanned by H3K27me3 domains in IMR90 (394 Mb) is 3.3 times more than in hESCs (119 Mb) (Tables S5 and S6). The median size of an H3K27me3 domain in IMR90 (16,400 bp) expands to almost twice that of hESCs (8600 bp, $p < 1E\text{-}308$, Wilcoxon rank sum test; Figure 2C), whereas coverage of H3K36me3 domains remains consistent (hESC median = 15,700 bp; IMR90 median size = 16,900 bp, $p = 4.6E\text{-}7$) (Figures 1A, 2C, and 2D). In total, 50% of the IMR90 genome is spanned by chromatin domains from the 11 marks profiled, whereas only 32% of the hESC genome is in chromatin domains (Figure 2E). The difference is largely accounted for by H3K27me3 and H3K9me3 dispersal (Figures 2D, 2F, and 2G). Each of these marks covers 4% of the hESC genome, but expands to 12% (H3K27me3) and 16% (H3K9me3) in IMR90 cells (Figures 2F and 2G).

To confirm whether expansion of repressive domains is a key characteristic of differentiated cell fate, we tested whether expanded domains exist in other cells. We analyzed genome-wide profiles of H3K27me3 in $CD4^+$ T cells (Barski et al., 2007; Cuddapah et al., 2009). Indeed, much of the $CD4^+$ T cell genome (324 Mb) is spanned by H3K27me3 domains, nearly three times as large as in hESCs. Because of insufficient read depth for H3K9me3 in this cell type, we did not include H3K9me3 blocks in this analysis. The expansion of H3K27me3 holds true in various differentiated cell types. ChromaBlocks analysis of H3K27me3 shows expansion in nontransformed HUVEC (347 Mb) and NHEK (312 Mb) cells, as well as transformed GM12878 lymphoblasts (286 Mb) and K562 leukemia cells (353 Mb) (Figure S3A; data produced and released from the

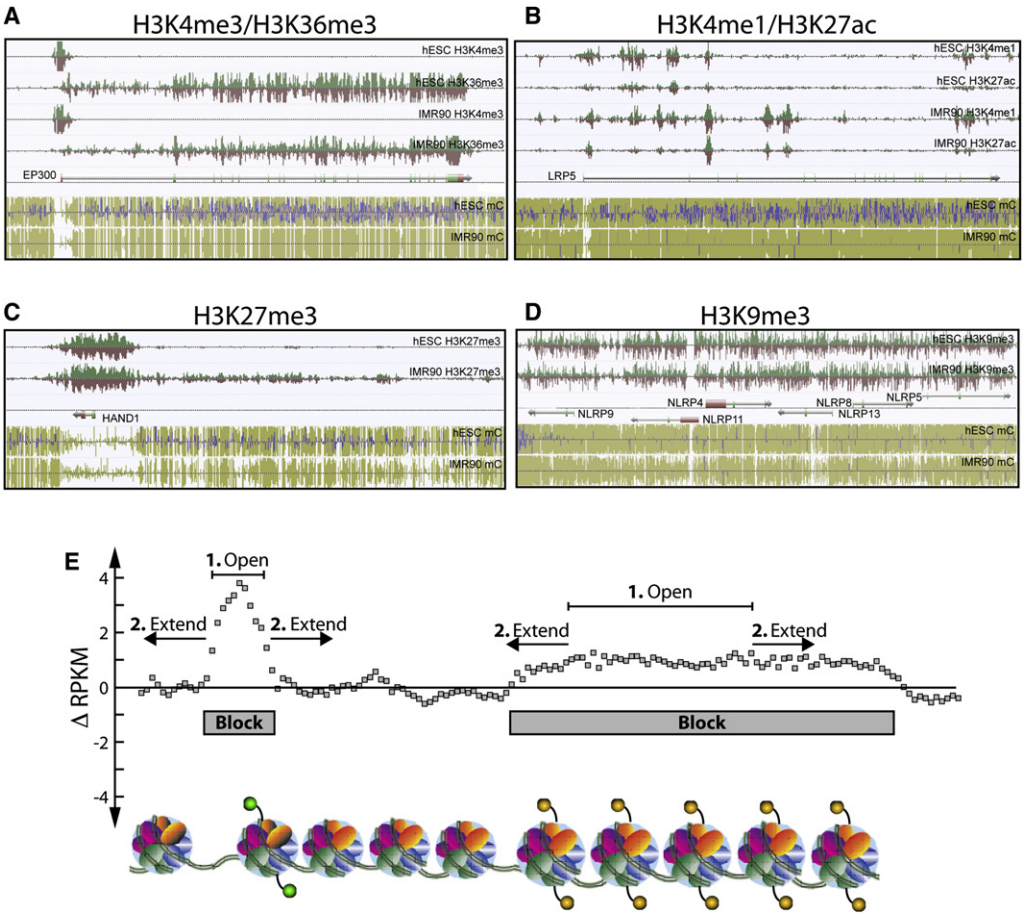

Figure 1. Histone Modifications Form Block Structures

(A–D) AnnoJ Browser snapshots of representative examples of the relationship between histone modifications and DNA methylation. Strand-specific mapped tags are displayed above (green) and below (red) the line. Strand-specific DNA methylation (mC): mCG (yellow-green bars), mCHG (blue bars), mCHH (pink bars).

(A) The *p300* (*EP300*) locus illustrating depleted mCG in the promoter containing H3K4me3, and H3K36me3 in the gene body of both cell types.

(B) H3K4me1 and H3K27ac occur at both the promoter of *LRP5* and distal sites that show differences in mCG.

(C) H3K27me3 is present at the mCG-depleted promoter of the developmental transcription factor *HAND1*. Spreading of H3K27me3 in IMR90 cells accounts for correlated differences with mCG.

(D) The NLRP gene cluster is encompassed by H3K9me3 in both cells types, a reduction in mCG is evident in IMR90 cells.

(E) A schematic of the ChromaBlocks strategy for identifying chromatin domains from input-normalized ChIP-Seq data.

See also Figure S1, Table S1.

ENCODE Project [Birney et al., 2007]). All lineage-committed cell types examined, normal and cancer cell types, have a clear increase in repressive chromatin structure compared to hESCs. In contrast, modifications associated with gene activity including H3K4me1, H3K4me2, H3K4me3, H3K27ac, and H3K36me3 are not expanded in these differentiated cell lines compared to hESCs (data not shown).

Expansion of H3K27me3 and H3K9me3 Selectively Affects Developmental Genes

We investigated the extent to which genes are silenced in fibroblasts compared to hESCs by broad domains. Despite expansion of H3K27me3 domains, we observed similar numbers of promoters covered by H3K27me3 in hESCs (4736) and IMR90 cells (4279) (Figures S3B and S3C). However, 83% of H3K27me3-marked promoters are H3K4me3/H3K27me3 bivalent promoters in hESCs, compared to 50% in IMR90 cells (Figure 3A). Of the 2485 promoters commonly marked by H3K27me3 in hESCs and IMR90 cells, 1512 show greater than 50% expansion in IMR90 cells. The median H3K27me3 domain spanning a promoter in hESCs is 10 kb, compared to 28 kb in IMR90 and 22 kb in $CD4^+$ T cells, suggesting that expansion of H3K27me3 domains is a common feature of differentiated cells (Figure S3).

Gene ontology (GO) analysis via DAVID (Dennis et al., 2003) of genes marked by H3K27me3 expansion in IMR90 cells reveals significant enrichment in developmental processes (Figure S3D), including genes from the BMP, FGF, FOX, SOX, and GATA families (Table S7). Additionally, GO analysis on the 314 promoters marked by H3K27me3 expansion in hESCs compared to IMR90 cells revealed significant enrichment of a small class of developmental proteins including brain-specific

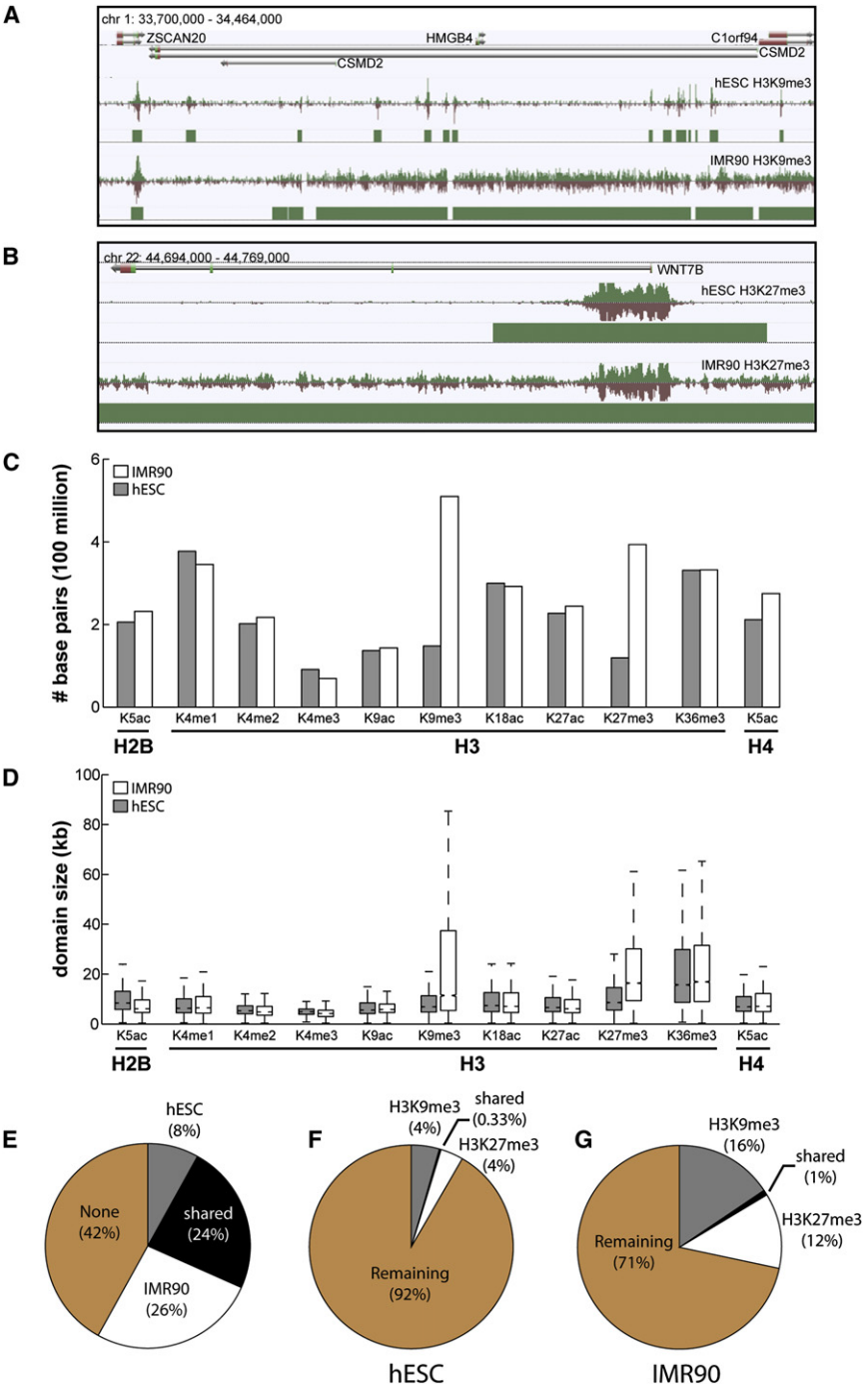

Figure 2. Features of Chromatin Domains

(A and B) Examples of (A) H3K9me3 and (B) H3K27me3 domains expanded in IMR90 cells relative to hESCs.

(C and D) The total number of base pairs spanned by domains (C) and the distribution of domain sizes for all mapped chromatin modifications (D). IMR90, white; hESC, gray.

(E) The fraction of the human genome spanned by the 11 chromatin modifications profiled in both cell types.

(F and G) The fraction of the (F) hESC and (G) IMR90 genomes spanned by H3K9me3 and H3K27me3 domains.

See also Figures S2 and S3, Tables S2–S6.

genes such as EMX2 and BAI1, as well as several HOX genes. The bivalent ESC-chromatin state often transitions to H3K27me3 and expands to cover the entire gene locus and frequently neighboring gene loci. An additional 1742 promoters are marked by H3K27me3 only in IMR90 cells but not in hESCs and are enriched for nondevelopmentally related biological processes including immune and defense response (Figure S3D). Interestingly, distinct genes are marked by expanded H3K27me3 domains in IMR90 and $CD4^+$ T cells. There are 2138 promoters marked by H3K27me3 in hESCs that expand in IMR90 or $CD4^+$ cells. Of these, 929 (44%) are IMR90 specific and 626 (29%) exhibit $CD4^+$-specific expansion, whereas 583 (27%) expand in both cells. IMR90-specific expanded domains are enriched for T cell, B cell, and lymphocyte-related genes, whereas $CD4^+$-specific expanded domains are enriched for extracellular matrix proteins, such as collagens. These results suggest that cell fate commitment is also a process of epigenetic gene repression that is unique to different lineages.

Similar to H3K27me3, there are 45% more H3K9me3-only marked domains in IMR90 cells (659) than in hESCs (456). However, a smaller fraction of gene promoters are spanned by H3K9me3 domains in hESCs (932, 5%) and IMR90 cells (1448, 8%) (overlap 456 promoters) (Figures 2C, 2F, 2G and 3A). In several instances, gene family clusters are marked by H3K9me3 domains, including olfactory receptors (OR) and the NLRP (Figure 1D), KRTAP, LCE, and zinc finger (ZNF) families. Notably, H3K9me3 spreading covers the gene encoding the early developmental factor *DPPA3* in IMR90 cells, and this spreading ends immediately before *NANOG*. Interestingly, we observe five times as many promoters simultaneously marked by H3K9me3 and H3K27me3 in IMR90 cells (359) as in hESCs (71) (Figure 3A). For example, the Simpson-Golabi-Behmel overgrowth syndrome gene, *GPC3* (Pilia et al., 1996), is bivalent (H3K4me3/H3K27me3) and expressed in hESCs, yet is repressed by H3K9me3 and H3K27me3 in IMR90 cells (Figure 3B). Additionally, we observe key developmental transcription factors associated with both H3K9me3 and H3K27me3 in IMR90 cells. As examples, *POU3F4* and *PAX3* are marked by H3K27me3 in hESCs and gain H3K9me3 in IMR90 cells to become dually marked for repression (Figures 3C and 3D). POU class transcription factors establish cell type-specific gene expression and cell

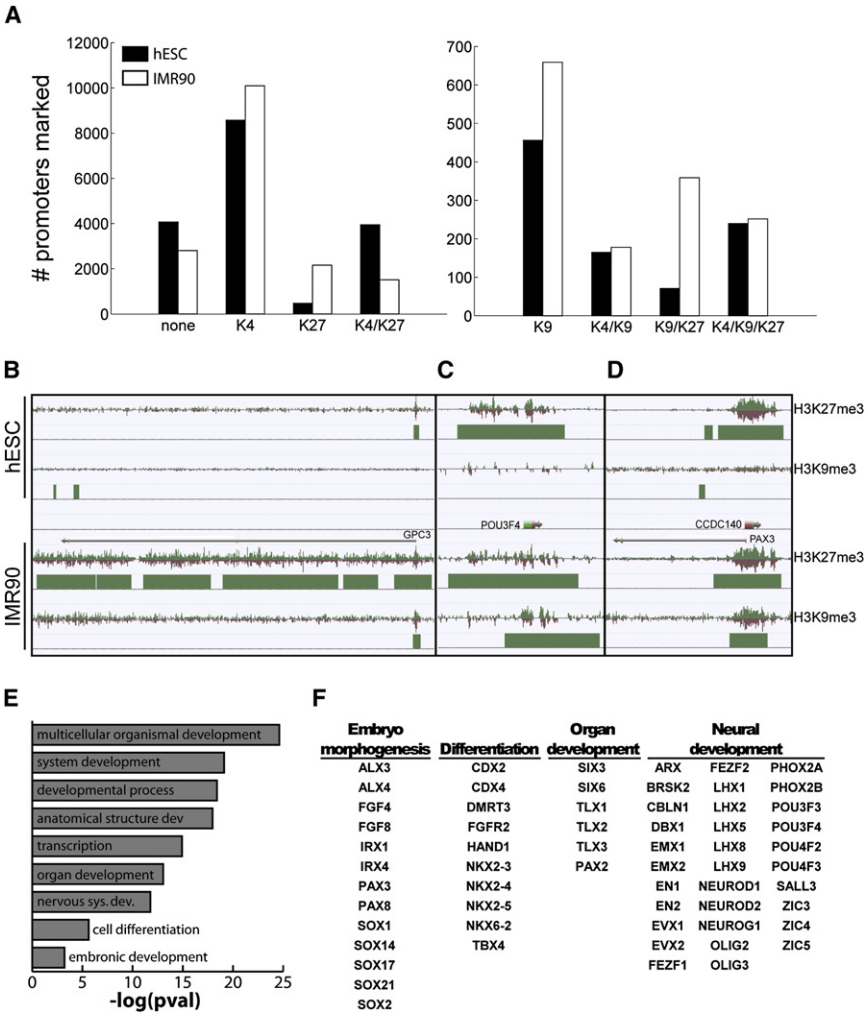

Figure 3. Expansion of Distinct Chromatin Blocks Marks Developmental Genes

(A) The number of promoters marked by various combinations of chromatin modifications in hESCs (black) and IMR90 cells (white). Exact combinations, including H3K4me3, can be found for RefSeq genes in Table S7. Abbreviations: K4, H3K4me3; K9, H3K9me3; K27, H3K27me3; none, lacking K4/K9/K27.

(B–D) Snapshots of domains and chromatin structure around promoters that gain H3K9me3 and H3K27me3 in IMR90 cells, for (B) *GPC3*, (C) *POU3F4*, and (D) *PAX3*.

(E and F) Gene ontology - biological processes enriched in promoters that are marked by both H3K9me3 and H3K27me3 in IMR90 cells (E), along with (F) examples of developmentally relevant genes having this chromatin structure.

See also Figure S3, Table S7.

fate decisions (Ryan and Rosenfeld, 1997), and PAX genes are critical for cell fate, organogenesis, and proliferation, leading to an important role in cancer (Loh et al., 2006). In light of this, we asked whether H3K9me3/H3K27me3-marked promoters in IMR90 cells share similar functions. GO analysis revealed significant enrichment for developmental processes, including multicellular organismal development, system development, and anatomical structure development (Figure 3E), noting that many dual-marked promoters include key developmental regulators (Figure 3F), such as HOX genes. The colocalization of H3K27me3 and H3K9me3 at genes was recently noted by others (Bilodeau et al., 2009). Not all H3K27me3-marked sites are associated with H3K9me3. This dual repression at specialized regulators may point to the importance of maintaining their silencing in differentiated cells. These genes serve as likely candidates for cell fate decisions.

With distinct spatial modes of H3K9me3- and H3K27me3-marking genes, we asked how these modes affect gene repression. We considered four cases: (1) promoters not marked in hESCs but marked in IMR90 cells (appear), (2) promoters covered by blocks in hESCs but expanded by at least 50% in IMR90 cells (expand), (3) promoters unmarked in both cells (unmarked), and (4) promoters marked by similarly sized blocks in both cells (marked). Table S7 shows the H3K4-, K9-, and K27-me3 states for all genes and expression changes. The acquisition of H3K9me3 at promoters showed the most significant decrease in gene expression ($p = 0.011$) compared to the other H3K9me3 scenarios (Figures 4A, 4C, and 4D). In contrast, the greatest repression for H3K27me3 cases occurred upon domain expansion (Figures 4B, 4E, and 4F). This may imply that the presence of H3K9me3 alone is enough to suppress gene expression and expansion of domains has a smaller effect on gene repression. By contrast, the presence of narrowly marked H3K27me3 promoters is frequently insufficient to silence a gene (Figure 4F), as they are often in a bivalent state, but the expansion seems to maintain stable silencing of gene expression in the differentiated cells ($p = 1.3E-19$).

Global Analysis of Associations between Chromatin Modifications and DNA Methylation

DNA methylation is a critical component of the epigenome that represses gene expression through promoter CG methylation, in addition to being localized at heterochromatin and repetitive elements in the genome. This mark is thought to maintain long-term repression and to be less dynamic than histone modifications. There is also evidence that some crosstalk between these epigenetic mechanisms exists, suggesting a direct link between H3K9 methyltransferases (HKMT) and DNA methyltransferases (DNMTs) (Fuks et al., 2003; Jackson et al., 2002; Lehnertz et al., 2003; Li et al., 2006; Tamaru and Selker, 2001). However, mutations in the SET domain of HKMT G9a do not alleviate DNA methylation (Dong et al., 2008; Tachibana et al., 2008). To date, it remains to be determined whether other chromatin structures or regulatory elements have an association with DNA methylation.

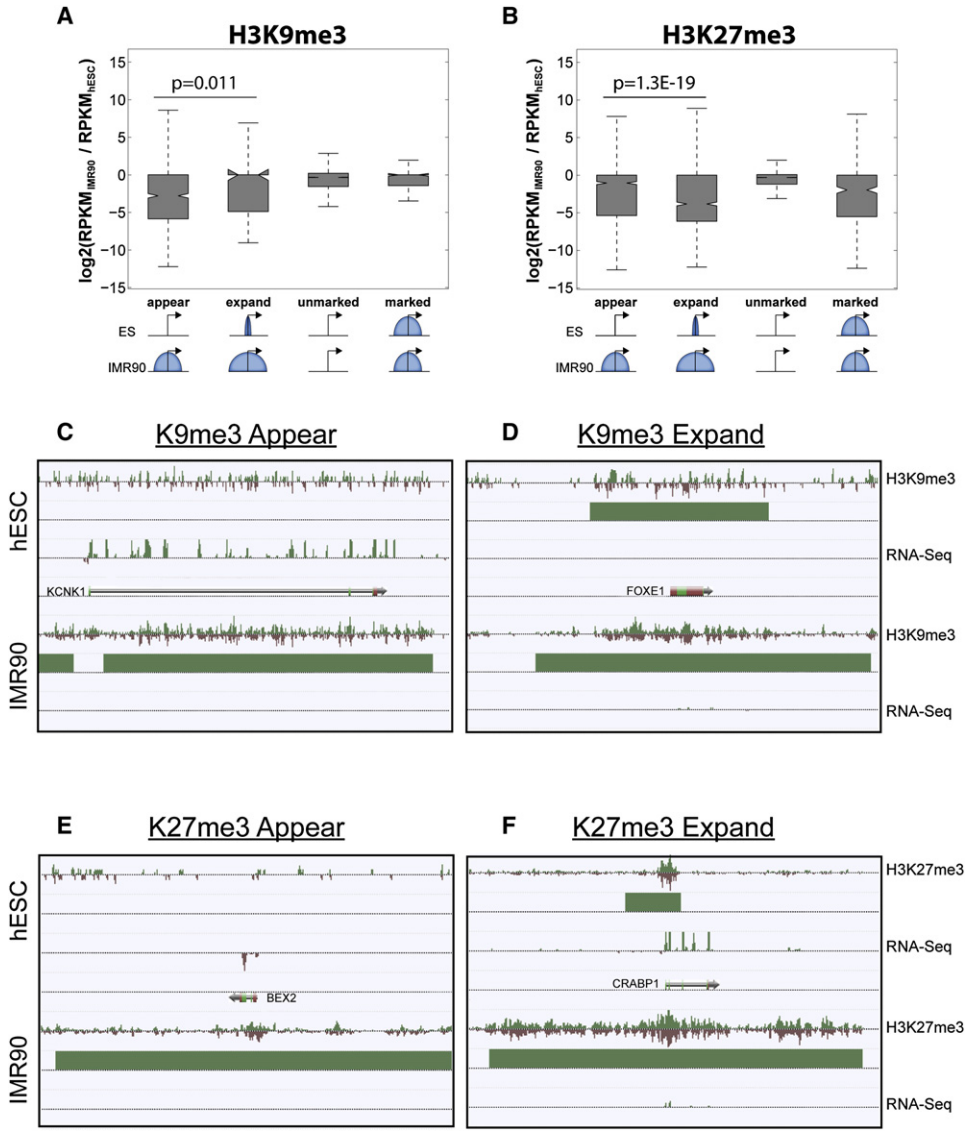

Figure 4. Expression of Genes in Different Types of Chromatin Domains

(A) Change in gene expression for promoters where H3K9me3 appears, expands, remains unmarked, or remains similarly marked in IMR90 cells compared to hESCs.

(B) As in (A), but for H3K27me3.

(C–F) Snapshots of genes exhibiting (C) H3K9me3 appearance, (D) H3K9me3 expansion, (E) H3K27me3 appearance, and (F) H3K27me3 expansion, as observed in IMR90 cells relative to hESCs.

See also Table S7.

By using genome-wide, nucleotide-resolution maps of DNA methylation (mC) generated in H1 hESCs and IMR90 cells (Lister et al., 2009), we systematically investigated the relationships between the 11 chromatin modifications described above and DNA methylation. By using a sliding window approach, we plotted the presence of ChIP-Seq tags for each histone modification normalized to input, against the average percentage of CG methylation (mCG) in the same window, to determine the association of the two epigenetic components in hESCs and IMR90 (Figures 5A–5D; Figure S4). Consistent with previous findings, we observe that H3K4me3 is inversely correlated with mCG (Figures 1A and 5A). Similarly, H3K27me3 is associated with hypomethylation of DNA at promoters, consistent with previous observations that bivalent promoters are typically hypomethylated (Brunner et al., 2009; Meissner et al., 2008). Outside of promoters, H3K27me3 is also positively associated with DNA methylation across a broad range of mCG. Interestingly, the degree of association between H3K9me3 and DNA methylation is less prevalent than that between mCG and H3K36me3, which is always associated with fully methylated DNA in both hESCs and IMR90 cells (Figures 1A, 5B, and 5D). These results suggest cell type-specific relationships between DNA methylation and

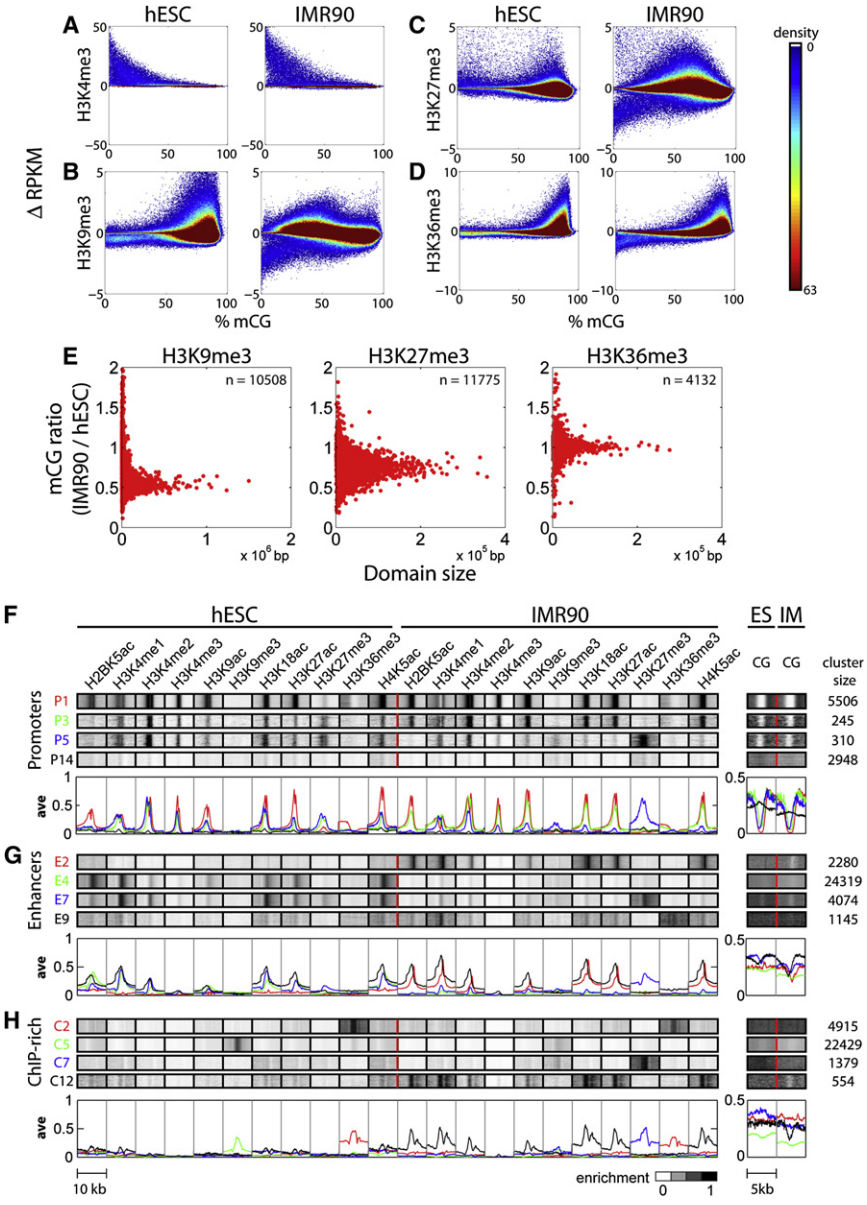

Figure 5. Genome-wide Correlation Analysis between Histone Modifications and Methylated CG

(A–D) Change in RPKM for several histone modifications relative to input control were calculated within 2.5 kb window for (A) H3K4me3 and 5 kb for (B) H3K9me3, (C) H3K27me3, and (D) H3K36me3, and plotted against %mCG in the same window throughout the genome. Density of spots in the plot is displayed by the heatmap (right).

(E) For H3K9me3 (left), H3K27me3 (middle), and H3K36me3 (right), spreading chromatin domains were identified as IMR90 domains having at most 20% overlap with hESC domains. Shown is the IMR90 to hESC ratio of %mCG for each of these domains, as a function of domain size.

(F–H) Representative punctuated chromatin signatures. ChromaSig was used to simultaneously cluster and align all 11 chromatin modifications mapped from both cell types at (F) promoters, (G) predicted enhancers, and (H) regions of ChIP-Seq enriched sequences. On the left are the chromatin signatures recovered for each cluster, and on the right are %mCG tracks appended after clustering. Enrichment of histone modifications and %mCG are indicated by the heatmap (bottom). The number of genomic loci in each cluster is indicated on the far right. Average profiles for each modification are plotted beneath the representative clustergrams. Profiles are color-coordinated with the cluster ID, e.g., P1 (red), P3 (green), P5 (blue), P14 (black).

See also Figures S4–S7, Tables S8–S12.

histone modifications (Figure 1C; also see Figure 6 and Figure S4).

H3K36me3 and DNA methylation have been described in exons of transcribed gene bodies (Ball et al., 2009; Hellman and Chess, 2007; Kolasinska-Zwierz et al., 2009). Examining the enrichment of H3K36me3 and mCG across 231,984 human exons, we observe that there is concordance between H3K36me3 and mCG (Figures S5A and S5B). However, H3K36me3 is more positively correlated with gene expression (Figures S5C and S5D), as indicated by the fact that most exons are marked with DNA methylation (Figures S5A and S5B). A possible mechanism may involve H3K36me3 in the recruitment of DNA methyltransferases to maintain DNA methylation at transcribed gene bodies. This would fit the "methylation paradox," noting that although promoter CpG islands remain unmethylated, transcribed CpG islands show an increase in DNA methylation (Jones, 1999). However, DNA methylation is also found outside transcribed regions. Therefore, whereas H3K36me3 predicts DNA methylation, the converse is not true.

To assess how changes in chromatin blocks correlate with changes in DNA methylation, we analyzed the genomic regions that were differentially marked in IMR90. Regions of the genome uniquely marked by H3K9me3 in IMR90 cells contain roughly half as much mCG as in hESCs (Figure 5E; Figures S5E–S5G). This depletion is most evident for larger domains, whereas small domains have a wide variance of mCG. Similarly, H3K27me3 domains unique to IMR90 cells are also depleted of mCG relative to hESCs, though at levels intermediate to H3K9me3 and H3K36me3 blocks, which exhibit equivalent levels of DNA methylation in each cell type (Figures 5E and 5F; Figures S5E–S5G). This intermediate level of methylation is not due to the contribution of promoters (Figure S5H). These relationships hold when comparing domains marked in both hESCs and IMR90 cells. Together, these results suggest that, on a global scale, gain of the repressive modifications H3K27me3 or H3K9me3 is associated with a corresponding decrease in DNA methylation level in the differentiated IMR90 cells. This observation suggests a much more complex relationship between

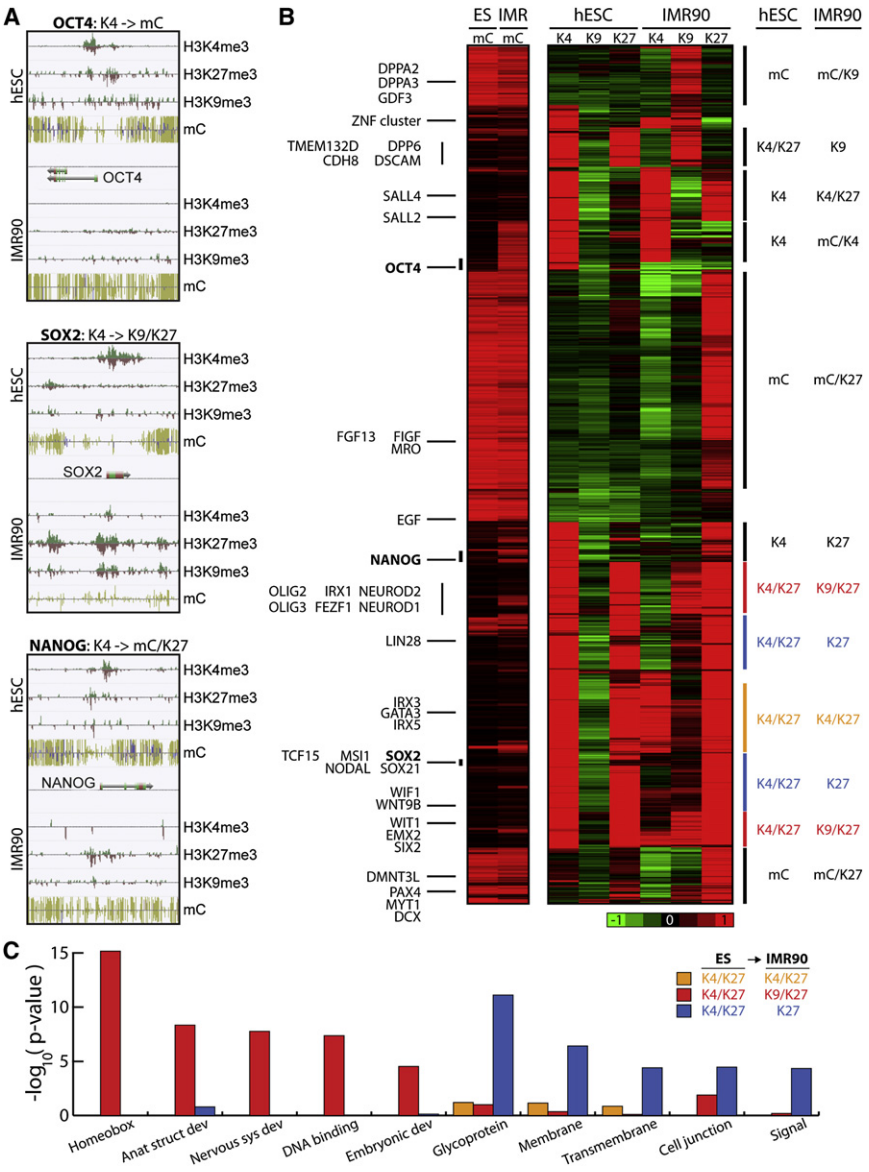

Figure 6. Distinct Modes of Epigenetic Repression at Developmental Genes

(A) Snapshots of chromatin modifications and DNA methylation around *OCT4*, *SOX2*, and *NANOG* in hESCs and IMR90 cells, illustrating distinct transitions to repressive epigenetic states. Abbreviations: K4, H3K4me3; K9, H3K9me3; K27, H3K27me3; mC, DNA methylation.

(B) Unsupervised clustering of H3K4me3, H3K9me3, H3K27me3, and mCG at promoters of genes downregulated by 2-fold in IMR90 cells relative hESCs. Relevant genes are indicated to the left. Distinct clusters are noted on the right. The enrichment scale is shown below the heatmap.

(C) GO analysis of genes whose promoters remain bivalent (H3K4/K27me3), acquire H3K9me3 (H3K4/K27me3 to H3K9/K27me3), or lose H3K4me3 (H3K4/K27me3 to H3K4me3) in IMR90 cells.

See also Table S7.

H3K9me3 or H3K27me3 and DNA methylation than previously proposed (Cedar and Bergman, 2009; Suzuki and Bird, 2008).

DNA Methylation at Potential Regulatory Sequences

A complex relationship also exists between mCG and H3K4me1, H3K4me2, and histone acetylation, which mark active regulatory sequences such as promoters and enhancers (Figures 5A–5D; Figure S7). To relate complex patterns of histone modifications with mCG at functional elements, we utilized an unbiased clustering algorithm, ChromaSig (Hon et al., 2008). By focusing on modifications with a punctuated footprint (Figures 5F–5H; Figure S6, all clusters), our analysis revealed 42 frequently occurring patterns, grouped into three categories corresponding to known promoters (Figure 5F; Figure S6A, Table S8), H3K4me1-predicted enhancers (Figure 5G; Figure S6B, Table S9), and other regions with enriched chromatin modification signals, ChIP-rich regions (Figure 5H; Figure S6C, Table S10).

Although most promoters have a strong inverse correlation between H3K4me3 and mCG, a subset of gene promoters lacking chromatin modifications are hypermethylated and typically silenced (Figure 5F; Figures S6A, S7A, and S7D). At enhancers (Tables S11 and S12), cell-specific relationships exist. Acetylated, IMR90-specific enhancers (Figure 5G, clusters E1-E3, E10-E11) show depletion of mCG in IMR90 cells, but not in hESCs (cluster E2). Nonacetylated IMR90-specific enhancers have minimal mCG depletion (cluster E9). At shared enhancers (E5, E7, E8), both cell types have mCG depletion (Figure 5G, E7) whereas hESC-specific enhancers (E4, E6, E12) show no obvious depletion of mCG (E4 in Figure 5G; Figures S6B, S7B, and S7E). Instead, hESC enhancers and transcription factor binding sites for OCT4, SOX2, NANOG, and others are depleted of non-CG methylation (Lister et al., 2009). Therefore, the chromatin modifications at these elements are largely cell specific and are inversely related to DNA methylation in a cell-specific manner. However, the mechanism for the role of CG versus nonCG methylation is unclear.

ChromaSig analysis of ChIP-rich regions outside of known promoters and predicted enhancers confirmed our block analysis, illustrating H3K36me3 regions enriched for mCG (Figure 5H, cluster C2; Figure S6C, cluster C1–C3) and the presence of peak-like H3K9me3 sites in hESCs (Figure 5H, cluster C5). ChromaSig also detects patterns missed by peak-finding or prediction algorithms. Cluster C12 illustrates likely IMR90 cell-specific enhancers previously missed (Figure 5H), and clusters C6-C7 with H3K27me3 in hESCs are unannotated promoters enriched for DNA methylation, similar to promoter clusters P16-P17 (Figure 5H, cluster C7). These ChIP-rich regions

outside of known annotations provide interesting avenues for future study. Much like promoters, ChIP-rich regions show little variation in DNA methylation across cell types (Figure S6C). Therefore, our analysis reveals that H3K4me1-predicted enhancer regions, which we previously showed to be cell type specific (Heintzman et al., 2009), exhibit the most dynamic DNA methylation changes.

Distinct Mechanisms of Epigenetic Repression of Pluripotency Genes in Fibroblasts

The core pluripotent transcription factors *OCT4*, *SOX2*, and *NANOG* are each marked in IMR90 cells by distinct combinations of repressive modifications (DNA methylation, H3K9me3, and H3K27me3). The *OCT4* promoter is repressed through DNA methylation, as previously demonstrated (Barrand and Collas, 2010; Ben-Shushan et al., 1993; Deb-Rinker et al., 2005). The *SOX2* promoter harbors both H3K27me3 and H3K9me3, and *NANOG* shows an increase in promoter mCG and acquires H3K27me3 (Figure 6A). To determine whether other genes are regulated by these epigenetic modes of repression, we enumerated nearly 1400 genes with >2-fold reduction in expression and that acquired at least one repressive modification in IMR90 cells relative to hESCs. From an unbiased hierarchical clustering of these genes, we observe that only small groups of genes follow the OCT4-, SOX2-, and NANOG-like patterns (Figure 6B). Because of their limited expression, it is not surprising that few genes are repressed in the same manner. However, these small, uniquely marked sets of genes may provide insight to stem cell biology or diseases such as cancer. For example, the NANOG-like class contains *KLF8*, a known regulator of the reprogramming factor KLF4. The OCT4-like class contains the cell cycle regulator *RAB25*, which has a role in cell cycle regulation and tumor invasion (Caswell et al., 2007). Because a common mechanism of epigenetic repression exists, these genes may also be positively coregulated in a tissue-specific manner. This is most evident with SOX2, a marker of neural stem cells. Many of the SOX2-like class of genes have known roles in neurobiology, such as *CADPS, RIC3, ZIC5*, and *KIF1A*.

Other combinations are more common. A large number of DNA methylated promoters gain H3K27me3 (mC to mC/H3K27me3), which may reflect a functional link between these marks for a subset of genes. Previous studies have suggested that H3K27me3-marked promoters are linked to DNA hypermethylation in cancer cells (Ohm et al., 2007; Schlesinger et al., 2007; Viré et al., 2006; Widschwendter et al., 2007). Here, we note the reciprocal relationship also occurs. Not all statically marked mC genes acquire H3K27me3, a smaller class gain H3K9me3 (mC to mC/H3K9me3) (Figure 6B), suggesting different mechanisms. hESC bivalent promoters, which are enriched for developmental regulators, are equally distributed among three states in IMR90 cells: remaining H3K4/27me3 or becoming H3K9/K27me3 or mC/H3K27me3 marked. GO analysis showed that the developmental regulators tend to be marked by H3K9/27me3 in IMR90 cells (Figure 6C). Moreover, genes dually marked for repression (mC/K27me3 and K27/K9me3) constitute the majority of the 1400 genes examined. Presumably, these added levels of epigenetic repression decrease the likelihood of escaping gene repression and altering cell fate. The multiple modes of repression suggest that several mechanisms are in place to restrict dedifferentiation in lineage-committed cells.

Repressive Chromatin Domains in Induced Pluripotent Cells

iPSCs and hESCs are both pluripotent and share the ability to self-renew. Furthermore, the global gene expression profiles of IMR90-derived iPSCs are more similar to hESCs than the original IMR90 cells (Yu et al., 2007). However, it remains to be determined whether this is true of repressive chromatin structure. By using iPSCs reprogrammed from IMR90 cells by Yu et al., we asked whether repressive domains are a key feature of cell fate, and therefore are broad domains remodeled during reprogramming, especially outside a few key gene promoter regions. We observed that reprogramming results in a noticeably reduced distribution of repressive domains (Figure S3), with H3K9me3 reduced to 275 MB and H3K27me3 reduced to 195 MB, between the values observed in IMR90 cells (510 MB, 394 MB) and hESCs (148 MB, 119 MB). Despite the wider coverage of H3K9me3 in IMR90 cells, there is more overlap between iPSCs and hESCs (H3K9me3 = 105 MB, H3K27me3 = 102 MB) than between IMR90 cells and hESCs (H3K9me3 = 76 MB, H3K27me3 = 68 MB), indicating that reprogramming has resulted in an epigenome that is more similar to the hESC epigenome (Figure 7A; Figure S3), consistent with previous observations (Maherali et al., 2007; Mikkelsen et al., 2008; Sridharan et al., 2009).

On a global scale, the profiles of H3K9me3 and H3K27me3 in iPSCs are similar to those in hESCs (Figure 7C). However, regions of discordance exist, at times spanning genic regions (Figure 7C). To dissect these differences, we defined three interesting chromatin structures related to reprogramming. The first, denoted iPS reprogrammed, are domains shared between iPSCs and hESCs. The second, denoted iPS unchanged, are marked in iPSCs and IMR90 cells, therefore not in a hESC-like state. The third, denoted iPS unique, are domains not marked in either hESCs or IMR90 cells. The iPS unchanged and iPS unique groups have comparable genomic distributions, generally spanning intergenic regions of the genome, with only 12% of H3K9me3 and 21% of H3K27me3 iPS domains intragenic on average (Figure 7B). Thus, although there are subtle differences in repressive chromatin structure between iPSCs and hESCs, most of these differences are confined to regions outside of genes.

iPS-reprogrammed H3K27me3 domains cover nearly five times as many gene promoters as the other groups (iPS reprogrammed, 1736; iPS unchanged, 325; iPS unique, 362) (Figure 7D). This indicates that reprogramming has generally reestablished the same H3K27me3 structure as hESCs at promoters, and only a small fraction are incorrectly reprogrammed. We do not observe the same phenomenon for H3K9me3, with more genes in iPS unique (417) than iPS reprogrammed (201) or iPS unchanged (173). These iPS/hESC differences in H3K9me3 structure at gene promoters is evident at several important epigenetic and developmental genes. For example, the H3K9me3 demethylase JMJD1A is marked by H3K9me3 in both hESCs and IMR90 cells but not iPSCs (Figure 7E). Also interesting, the WNT receptor FZD10 retains the incorrect repressive mark: H3K9/K27me3 in IMR90 cells, H3K27me3 in hESCs, yet

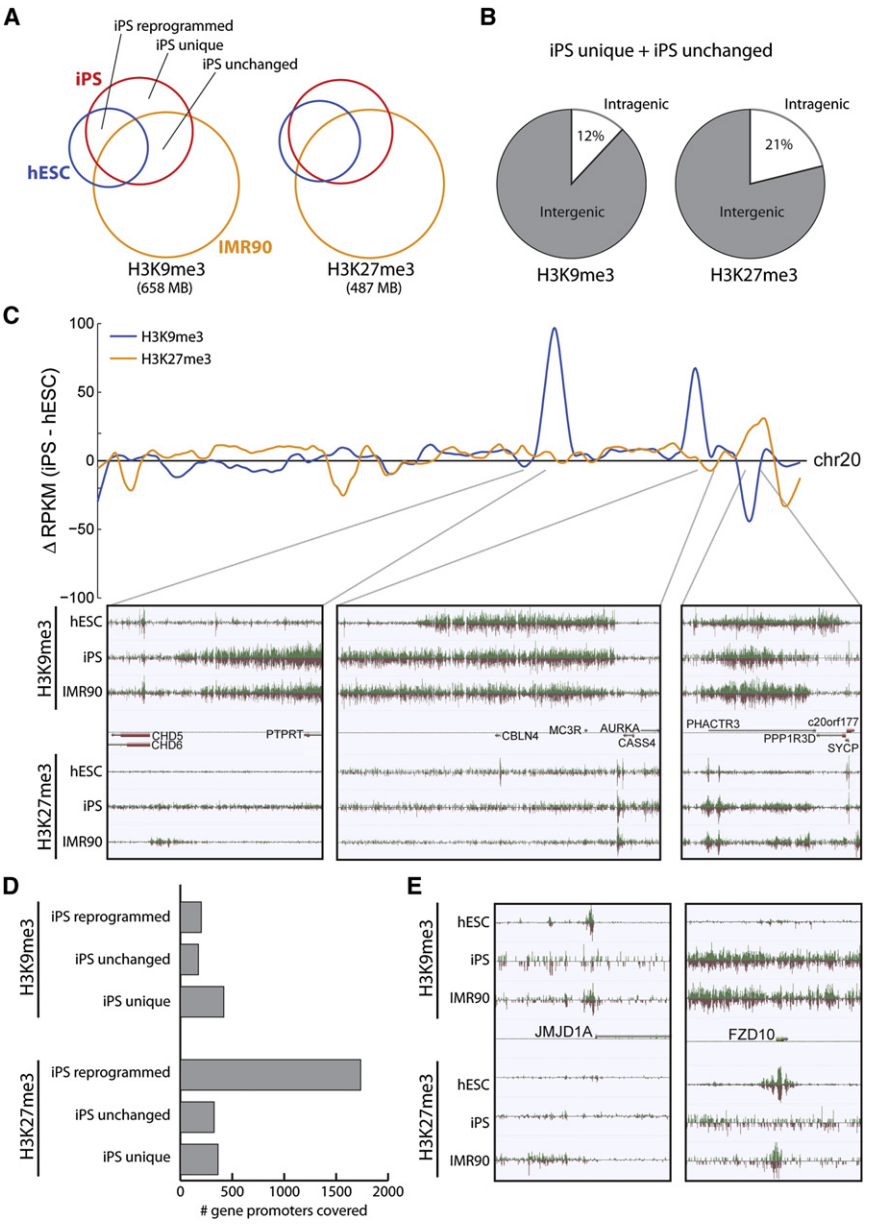

Figure 7. Comparison of Repressive Chromatin Structure in hESCs and iPSCs

(A) Overlap of repressive chromatin domains in hESCs (blue), iPSCs (red), and IMR90 cells (yellow) for H3K9me3 (left) and H3K27me3 (right). The total coverage of these modifications is indicated below. Regions related to reprogramming are labeled: iPS reprogrammed regions, defined as overlap of iPSCs and hESC but not IMR90 cells; iPS unchanged regions, defined as overlap of iPSCs and IMR90 cells but not hESCs; and iPS unique regions.

(B) The average genic distribution of iPS unchanged and iPS unique regions.

(C) Top: Difference in RPKM of iPS profiles compared to hESC profiles for H3K9me3 (blue) and H3K27me3 (yellow) throughout the length of chr20. Bottom: Zoomed in views of genomic regions having differences between iPSCs and hESCs.

(D) For the three types of iPS regions, the number of gene promoters covered by H3K9me3 (top) or H3K27me3 domains (bottom).

(E) Snapshots of repressive chromatin structure at JMJD1A (left) and FZD10 (right) promoters, illustrating differences in iPSCs and hESCs.

H3K9me3 in iPSCs (Figure 7E). Recently, Chin et al. found 3947 genes differentially expressed between several iPSC and hESC lines (Chin et al., 2009). We tested whether H3K9me3 or H3K27me3 was differentially enriched at these gene promoters between IMR90-derived iPSCs and H1 hESCs. From the 3947 genes, 770 promoters are marked by H3K27me3 in iPSCs, of which only 14.5% are not marked in hESCs ($p = 1$, Binomial, expectation is 15.2%). In contrast, 160 promoters are marked by H3K9me3 in iPSCs, 68% of which are not marked in hESC ($p = 0.0023$, Binomial, expectation is 56.8%). Thus, it appears that H3K9me3 contributes more than H3K27me3 to the differences in gene expression between iPSCs and hESCs. Perhaps these observations indicate flexibility in gene repression during reprogramming.

Our findings suggest that expanded repressive domains are a key aspect of differentiated cell fates, perhaps reducing the plasticity seen in stem cells. Reprogramming of somatic cells involves remodeling of these domains to reflect the hESC epigenome. Therefore, it is likely that partially reprogrammed iPSCs (piPSCs) still harbor expanded chromatin domains that prevent them from achieving pluripotency. Our results call for future studies of these chromatin marks in additional hESCs, iPSCs, and piPSCs. Moreover, expediting chromatin remodeling would probably improve reprogramming efficiency.

Conclusion

A fundamental question is how the identical genome sequence gives rise to a diversity of cell types with different gene expression profiles and cellular functions. By comparing the epigenome of pluripotent stem cells to that of a differentiated cell, we provide evidence that lineage-committed cells are characterized by significantly expanded repressive chromatin domains that selectively affect genes involved in pluripotency and development, suggesting that epigenetic mechanisms play a critical role in cellular differentiation and maintenance of differentiated cellular state. Recently, a similar conclusion was reached by examining H3K9me2 in mouse ESCs and differentiated cells (Wen et al., 2009). However, an independent analysis of the data via a different block finding method has reached at a different

conclusion, leading to a debate on the genomic distribution of H3K9me2 domain in mammalian cells (Filion and van Steensel, 2010; Wen et al., 2010). In the present work, the expansion of H3K27me3 or H3K9me3-marked repressive domains was confirmed by several measures. First, the expansion of H3K27me3 was illustrated in independently derived data from multiple cell types. Second, expanded H3K27me3 and H3K9me3 show reduction in IMR90-derived iPSCs. Finally, we applied an independent block-finding algorithm, TileHMM (Humburg et al., 2008), used to reanalyze the Wen et al. H3K9me2 data (Filion and van Steensel, 2010), and confirm the expansion of H3K27me3 and H3K9me3 in IMR90-relative hESCs (Figure S3E). Our result provides additional support to the general model proposed by Wen et al. (2009).

Knockout studies of H3K9 methyltransferases and H3K27 methyltransferases lead to differentiation or developmental defects (Dodge et al., 2004; Faust et al., 1998; O'Carroll et al., 2001; Pasini et al., 2007; Peters et al., 2001; Tachibana et al., 2002), suggesting that epigenetic mechanisms play a critical role in cell fate determination. The expansion of H3K9me3 and H3K27me3 domains in differentiated cells relative to hESCs support this model. Reprogrammed cells reorganize their chromatin architecture to reflect the less repressive state of stem cells. This suggests that inhibiting repressive chromatin structure would facilitate reprogramming either through inhibition of HKMTs or overexpression of demethylases. Early evidence supports this hypothesis (Shi et al., 2008a, 2008b; Wendt et al., 2008). Collectively, this suggests that one aspect of epigenetic cell fate lies in the structure of repressive or compacted chromatin.

EXPERIMENTAL PROCEDURES

ChIP-Seq

ChIP was carried out as previously described with 500 μg chromatin and 5 μg antibody (antibodies are listed in Supplemental Information) (Heintzman et al., 2007; Kim et al., 2007). ChIP libraries for sequencing were prepared according to Illumina protocols with minor modifications (see Supplemental Information) (Illumina, San Diego, CA). Libraries were sequenced with the Illumina GAII machine as per manufacturer's protocols. After sequencing, cluster imaging, base calling, and mapping were conducted with the Illumina pipeline (Illumina).

Cell Growth

IMR90 cells were grown as previously described (Kim et al., 2007). H1 embryonic stem cells were grown as previously described with Matrigel (BD Biosciences) and mTeSR (Ludwig et al., 2006a, 2006b).

Data Normalization

For each chromatin mark and the input control, we divided the genome into 100 bp bins, counted the number of reads falling into each bin i, and computed the number of reads per kilobase of bin per million reads sequenced (denoted $RPKM_{mark,i}$), with the exception of the input where RPKM is computed over the five consecutive bins centered at i. Because input represents the entirety of the genome and is not sequenced to saturation like ChIP data, a 500 bp bin is used to reduce noise that occurs in a 100 bp window. Finally, normalized ChIP enrichment is computed as $\Delta RPKM = RPKM_{mark,i} - RPKM_{input,i}$.

ChromaBlocks

To identify both peak-like and domain-like enrichment, a block is started for (1) regions of 10 consecutive bins having average ΔRPKM $\geq$ 1.5 with at least 5 bins above this threshold or (2) regions of 50 consecutive bins having average ΔRPKM $\geq$ 0.5 with at least 25 above this threshold. Once a block is started, its boundaries are extended in both directions until reaching a region of 50 bins having average ΔRPKM < 0.1 (Figure 1E). A full description is provided in the Supplemental Experimental Procedures.

Quantifying DNA Methylation

For a given cytosine residue with CG context at position i in the genome, its measure of mCG M_i is defined as the fraction of methylated cytosines called out of all methylC-Seq reads spanning i. Given a region, we then use %mCG as a measure of DNA methylation enrichment, defined as the sum of the M_i for all CG dinucleotides in the region, divided by the total number of CG dinucleotides in the region. The only exception is when calling promoters enriched with DNA methylation for a given cell. Here, we use an absolute measure of mCG with $M_i = 1$, and we define enriched promoters as those with %mCG $\geq$ 50% for a 1 kb window centered at the TSS.

Identifying Punctuated Chromatin Signatures

RefSeq promoters are clustered by chromatin signature via ChromaSig as described (Hon et al., 2008) (motif width w = 4000 bp, wandering distance d = 1000 bp, $\sigma_{another}$ = 2.5, and p_a = 0.01). Putative enhancers in hESCs and IMR90 cells were predicted on the basis of chromatin signatures as described (Heintzman et al., 2007) and clustered by ChromaSig (w = 4000 bp, d = 2000 bp, $\sigma_{another}$ = 2.5, and p_a = 0.001). Regions of significant ChIP enrichment of histone modifications were identified as described (ChromaSig) (p value cutoff 1E-3), filtered to remove promoters, predicted enhancers, gene 3′ ends, and CTCF binding sites, and finally clustered by ChromaSig (w = 4000 bp, d = 1000 bp, $\sigma_{another}$ = 2.5, and p_a = 0.001). In all instances, ChromaSig performed the clustering on all 11 chromatin modification maps from both cell types simultaneously.

ACCESSION NUMBERS

All data have been deposited to the Sequence Read Archive (SRA), accession SRP000941.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, 7 figures, and 12 tables and can be found with this article online at doi:10.1016/j.stem.2010.03.018.

ACKNOWLEDGMENTS

We acknowledge Bradley Bernstein and the ENCODE Chromatin Group at Broad Institute and MGH for generating the H3K27me3 data sets for HUVEC, NHEK, GM, and K562 cell types. These data were used for validation during the revision and were under embargo until October 7, 2009. R.D.H. is supported by an American Cancer Society Postdoctoral Fellowship. R.L. is supported by a Human Frontier Science Program Long-term Fellowship. This work was supported by the following: NIH Epigenomics Roadmap Project (to B.R., J.A.T., J.R.E., and W.W.), the California Institute for Regenerative Medicine (B.R.), the Ludwig Institute for Cancer Research (B.R.), the Mary K. Chapman Foundation (J.R.E.), and the J.T.F. Morgridge Institute for Research (J.A.T.). This study was carried out as part of the NIH Roadmap Epigenomics Program (http://nihroadmap.nih.gov/epigenomics/referenceepigenomeconsortium.asp).

Received: September 22, 2009
Revised: January 6, 2010
Accepted: March 12, 2010
Published: May 6, 2010

REFERENCES

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., and Fisher, A.G. (2006). Chromatin signatures of pluripotent cell lines. Nat. Cell Biol. *8*, 532–538.

Ball, M.P., Li, J.B., Gao, Y., Lee, J.H., LeProust, E.M., Park, I.H., Xie, B., Daley, G.Q., and Church, G.M. (2009). Targeted and genome-scale strategies reveal

gene-body methylation signatures in human cells. Nat. Biotechnol. *27*, 361–368.

Barrand, S., and Collas, P. (2010). Chromatin states of core pluripotency-associated genes in pluripotent, multipotent and differentiated cells. Biochem. Biophys. Res. Commun. *391*, 762–767.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823–837.

Ben-Shushan, E., Pikarsky, E., Klar, A., and Bergman, Y. (1993). Extinction of Oct-3/4 gene expression in embryonal carcinoma x fibroblast somatic cell hybrids is accompanied by changes in the methylation status, chromatin structure, and transcriptional activity of the Oct-3/4 upstream region. Mol. Cell. Biol. *13*, 891–901.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes Dev. *23*, 2484–2489.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., et al. ENCODE Project Consortium; NISC Comparative Sequencing Program Baylor College of Medicine Human Genome Sequencing Center; Washington University Genome Sequencing Center; Broad Institute; Children's Hospital Oakland Research Institute. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

Brunner, A.L., Johnson, D.S., Kim, S.W., Valouev, A., Reddy, T.E., Neff, N.F., Anton, E., Medina, C., Nguyen, L., Chiao, E., et al. (2009). Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. Genome Res. *19*, 1044–1056.

Caswell, P.T., Spence, H.J., Parsons, M., White, D.P., Clark, K., Cheng, K.W., Mills, G.B., Humphries, M.J., Messent, A.J., Anderson, K.I., et al. (2007). Rab25 associates with alpha5beta1 integrin to promote invasive migration in 3D microenvironments. Dev. Cell *13*, 496–510.

Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: Patterns and paradigms. Nat. Rev. Genet. *10*, 295–304.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J., et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell *5*, 111–123.

Cuddapah, S., Jothi, R., Schones, D.E., Roh, T.Y., Cui, K., and Zhao, K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res. *19*, 24–32.

Deb-Rinker, P., Ly, D., Jezierski, A., Sikorska, M., and Walker, P.R. (2005). Sequential DNA methylation of the Nanog and Oct-4 upstream regions in human NT2 cells during neuronal differentiation. J. Biol. Chem. *280*, 6257–6260.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. *4*, 3.

Dodge, J.E., Kang, Y.K., Beppu, H., Lei, H., and Li, E. (2004). Histone H3-K9 methyltransferase ESET is essential for early development. Mol. Cell. Biol. *24*, 2478–2486.

Dong, K.B., Maksakova, I.A., Mohn, F., Leung, D., Appanah, R., Lee, S., Yang, H.W., Lam, L.L., Mager, D.L., Schübeler, D., et al. (2008). DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. EMBO J. *27*, 2691–2701.

Faust, C., Lawson, K.A., Schork, N.J., Thiel, B., and Magnuson, T. (1998). The Polycomb-group gene eed is required for normal morphogenetic movements during gastrulation in the mouse embryo. Development *125*, 4495–4506.

Filion, G.J., and van Steensel, B. (2010). Reassessing the abundance of H3K9me2 chromatin domains in embryonic stem cells. Nat. Genet. *42*, 4, author reply 5–6.

Fuks, F., Hurd, P.J., Deplus, R., and Kouzarides, T. (2003). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res. *31*, 2305–2312.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. *39*, 311–318.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108–112.

Hellman, A., and Chess, A. (2007). Gene body-specific methylation on the active X chromosome. Science *315*, 1141–1143.

Hon, G., Ren, B., and Wang, W. (2008). ChromaSig: A probabilistic approach to finding common chromatin signatures in the human genome. PLoS Comput. Biol. *4*, e1000201.

Humburg, P., Bulger, D., and Stone, G. (2008). Parameter estimation for robust HMM analysis of ChIP-chip data. BMC Bioinformatics *9*, 343.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature *416*, 556–560.

Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science *316*, 1497–1502.

Jones, P.A. (1999). The DNA methylation paradox. Trends Genet. *15*, 34–37.

Kharchenko, P.V., Tolstorukov, M.Y., and Park, P.J. (2008). Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nat. Biotechnol. *26*, 1351–1359.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell *128*, 1231–1245.

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. Nat. Genet. *41*, 376–381.

Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C., et al. (2003). Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. Mol. Cell. Biol. *23*, 4207–4218.

Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr. Biol. *13*, 1192–1200.

Li, J., Moazed, D., and Gygi, S.P. (2002). Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. J. Biol. Chem. *277*, 49383–49388.

Li, H., Rauch, T., Chen, Z.X., Szabó, P.E., Riggs, A.D., and Pfeifer, G.P. (2006). The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. J. Biol. Chem. *281*, 19489–19500.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G.C., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

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.

Loh, Y.H., Zhang, W., Chen, X., George, J., and Ng, H.H. (2007). Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. Genes Dev. *21*, 2545–2557.

Ludwig, T.E., Bergendahl, V., Levenstein, M.E., Yu, J., Probasco, M.D., and Thomson, J.A. (2006a). Feeder-independent culture of human embryonic stem cells. Nat. Methods *3*, 637–646.

Ludwig, T.E., Levenstein, M.E., Jones, J.M., Berggren, W.T., Mitchen, E.R., Frane, J.L., Crandall, L.J., Daigh, C.A., Conard, K.R., Piekarczyk, M.S., et al. (2006b). Derivation of human embryonic stem cells in defined conditions. Nat. Biotechnol. *24*, 185–187.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., et al. (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell *1*, 55–70.

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.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766–770.

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

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49–55.

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and Schübeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol. Cell *30*, 755–766.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol. Cell. Biol. *21*, 4330–4336.

Ohm, J.E., McGarvey, K.M., Yu, X., Cheng, L., Schuebel, K.E., Cope, L., Mohammad, H.P., Chen, W., Daniel, V.C., Yu, W., et al. (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat. Genet. *39*, 237–242.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and Thomson, J.A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell *1*, 299–312.

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.

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schöfer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., et al. (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323–337.

Pilia, G., Hughes-Benzie, R.M., MacKenzie, A., Baybayan, P., Chen, E.Y., Huber, R., Neri, G., Cao, A., Forabosco, A., and Schlessinger, D. (1996). Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. Nat. Genet. *12*, 241–247.

Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu. Rev. Genet. *38*, 413–443.

Ryan, A.K., and Rosenfeld, M.G. (1997). POU domain family values: Flexibility, partnerships, and developmental codes. Genes Dev. *11*, 1207–1225.

Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B.E., et al. (2007). Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat. Genet. *39*, 232–236.

Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Schöler, H.R., and Ding, S. (2008a). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell *3*, 568–574.

Shi, Y., Do, J.T., Desponts, C., Hahm, H.S., Schöler, H.R., and Ding, S. (2008b). A combined chemical and genetic approach for the generation of induced pluripotent stem cells. Cell Stem Cell *2*, 525–528.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell *136*, 364–377.

Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: Provocative insights from epigenomics. Nat. Rev. Genet. *9*, 465–476.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., and Shinkai, Y. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev. *16*, 1779–1791.

Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H., and Shinkai, Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. EMBO J. *27*, 2681–2690.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Tamaru, H., and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. Nature *414*, 277–283.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871–874.

Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat. Genet. *41*, 246–250.

Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2010). Reply to [ldquo]Reassessing the abundance of H3K9me2 chromatin domains in embryonic stem cells. Nat. Genet. *42*, 5–6.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796–801.

Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D.J., Campan, M., Young, J., Jacobs, I., and Laird, P.W. (2007). Epigenetic stem cell signature in cancer. Nat. Genet. *39*, 157–158.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917–1920.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.