Published in final edited form as:
Annu Rev Genomics Hum Genet. 2011 September 22; 12: 165–185. doi:10.1146/annurev-genom-082410-101506.

Genomic Approaches to Deconstruct Pluripotency

Yuin-Han Loh[1,2,*], **Lin Yang**[1,2,*], **Jimmy Chen Yang**[1,2,**], **Hu Li**[3,4,**], **James J. Collins**[3,4,5], and **George Q. Daley**[1,2,5,6,7]

George Q. Daley: george.daley@childrens.harvard.edu

[1]Stem Cell Transplantation Program, Division of Pediatric Hematology/Oncology, Children's Hospital Boston; Dana-Farber Cancer Institute; and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

[2]Harvard Stem Cell Institute, Cambridge, Massachusetts 02115

[3]Department of Biomedical Engineering and Center for BioDynamics, Boston University, Boston, Massachusetts 02215

[4]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts 02115

[5]Howard Hughes Medical Institute, Boston, Massachusetts 02115

[6]Division of Hematology, Brigham and Women's Hospital, Boston, Massachusetts 02115

[7]Manton Center for Orphan Disease Research, Boston, Massachusetts 02115

Abstract

Embryonic stem cells (ESCs) first derived from the inner cell mass of blastocyst-stage embryos have the unique capacity of indefinite self-renewal and potential to differentiate into all somatic cell types. Similar developmental potency can be achieved by reprogramming differentiated somatic cells into induced pluripotent stem cells (iPSCs). Both types of pluripotent stem cells provide great potential for fundamental studies of tissue differentiation, and hold promise for disease modeling, drug development, and regenerative medicine. Although much has been learned about the molecular mechanisms that underlie pluripotency in such cells, our understanding remains incomplete. A comprehensive understanding of ESCs and iPSCs requires the deconstruction of complex transcription regulatory networks, epigenetic mechanisms, and biochemical interactions critical for the maintenance of self-renewal and pluripotency. In this review, we will discuss recent advances gleaned from application of global “omics” techniques to dissect the molecular mechanisms that define the pluripotent state.

INTRODUCTION

Embryonic stem cells (ESCs) utilize highly complex genomic networks to maintain pluripotency. Recent work has begun to unravel these intricacies, which involve interactions between transcriptional regulatory networks and epigenetic factors. Much of these “omics” data (e.g., transcriptomics, epigenomics, and genome-wide DNA-protein interactions) have been generated through a combination of experimental strategies that utilize both

*These authors contributed equally to this work.
**These authors contributed equally to this work.

Disclosure statement

G.Q.D. is a member of the scientific advisory boards of and holds equity in the following companies: Epizyme, iPierian, Solasia KK, Verastem, and MPM Capital LLP.

experimental studies that examine functional interactions and theoretical computational approaches that infer interactions (69). These data are typically illustrated as a complex interaction network and have led to the development of numerous models, such as the attractor landscape, that describe how cells maintain or change their state (40, 114).

The ESC transcriptional network is strongly influenced by the transcription factors Oct4, Sox2, and Nanog (9, 68). Multiple experimental studies have shown that these factors regulate a plethora of target genes and participate in autoregulatory, feed-forward, and feedback interactions. By examining downstream targets, an extensive genome-wide map has been established in ESCs that outlines the interactions that preserve ESC pluripotency, centered on these factors (52). Additional elements such as downstream effectors of extrinsic cytokines, noncoding RNAs (ncRNAs), and transposable elements add complexity to this transcriptional circuitry (18, 21, 57, 70).

Epigenetic modification plays an important role in ESCs, generating a unique genomic landscape that influences regulatory networks (82). DNA methylations and histone modifications have been intensively explored, and studies have implicated these epigenetic processes in ESC fate and pluripotency. For example, epigenetic control of self-renewal has been revealed in studies that link DNA methylation at promoters of pluripotency genes to levels of transcription. Work in reprogramming has also expanded our knowledge of how transcriptional networks and epigenetic modifications affect cell fate, such as the extent to which pluripotent cells may tolerate epigenetic modifications characteristic of differentiated cells (54, 94).

In this review, we focus on genomic approaches to understanding ESC pluripotency, emphasizing recent work that has refined our understanding of established mechanisms. We assume throughout that ESCs and induced pluripotent stem cells (iPSCs) represent comparable states of pluripotency, but we will review where evidence suggests epigenetic differences between these two cell types, which likely reflects the technical limitations inherent in the reprogramming process. First, we discuss the dominant transcriptional regulatory networks that play important roles in ESC and iPSC identity and highlight the relationship of core transcriptional factors with other regulatory processes. Second, we describe the epigenetic landscape of ESCs and iPSCs, stressing how epigenetic modifications complement the pluripotent regulatory network.

TRANSCRIPTION REGULATORY NETWORKS IN EMBRYONIC STEM CELLS

Master Regulators of Pluripotency

Several transcription factors are known to be preferentially expressed and to play essential roles in both early embryonic development and maintenance of ESCs. Oct4 is a POU homeodomain transcription factor encoded by the *Pou5f1* gene that plays an essential role in the establishment and maintenance of pluripotency; *Oct4*-null mouse embryos fail to develop an inner cell mass (ICM) and contain only trophectoderm cells (81). Suppression of *Oct4* expression in ESCs likewise leads to differentiation along the extraembryonic trophoblast lineage (20, 46). Mechanistically, Oct4 is the master regulator of the pluripotency network, and acts to control embryonic cell fates by regulating a broad range of downstream target genes.

Sox2 was originally implicated in pluripotency by its capacity to heterodimerize with Oct4 to regulate the pluripotency-related gene *Fgf4* (139). *Sox2* has an expression pattern similar to *Oct4*, and *Sox2* knock-out results in defective epiblasts and differentiation into trophectoderm lineages (1). Subsequent studies revealed the presence of an octamer–sox

motif in the regulatory elements of downstream target genes to which the OCT4/SOX2 complex recognizes and cooperatively binds (8, 20).

Nanog was first identified through a functional screen for novel pluripotency regulators that can maintain mouse ESCs in the absence of LIF and a feeder layer (14, 79). Overexpression of *Nanog* bypasses the need for FGF and TGFβ signaling in human ESCs (137). Mouse embryos with a *Nanog* deletion fail to develop an epiblast, but Nanog is dispensable in the maintenance of pluripotency in cultured ESCs (15, 79). Thus, Nanog may be crucial as the gateway for the acquisition and establishment of pluripotency but not for its maintenance (107).

Core Transcriptional Circuitry in Embryonic Stem Cells

To interrogate the interactions among the Oct4, Sox2, and Nanog transcription factors and to catalog their downstream genomic targets, two groups initially undertook genome-wide mapping studies of these master regulators in human ESCs, using chromatin immunoprecipitation (ChIP) coupled with a promoter array (ChIP-on-chip; 9) and, in mouse ESCs, using ChIP with paired end tag sequencing (ChIP-PET; 68). An important principle to emerge from these studies is the autoregulatory binding of these transcription factors to their own promoters, and reciprocal regulation with other core members (9, 68). Both studies also reported the synergistic co-occupancy by the core factors of large ensembles of downstream target genes—in particular, Nanog bound to the regulatory sequences of more than 90% of the genes bound by the Oct4-Sox2 heterodimer. The core transcription regulatory network is highly enriched with tight feed-forward and autoregulatory loops, which are believed to confer stability to the system while allowing for rapid response to developmental switching depending on environmental stimuli (Figure 1*a*). Comparison with expression data of Oct4- and Nanog-depleted ESCs revealed that the core factors regulate both active and inactive genes (Figure 1*a*; 68). These observations support a model that implicates these core factors in the maintenance of the pluripotent state by promoting the expression of downstream self-renewal genes while simultaneously repressing the activity of differentiation-promoting genes (Figure 1*a*; 9, 68).

To generate a broader perspective on the transcriptional network governing pluripotency beyond the core set of transcription factors, Kim et al. (52) used in vivo biotinylation-mediated ChIP (bioChIP) coupled to a promoter array to investigate the genomic targets of Oct4, Sox2, Klf4, c-Myc, Nanog, Dax1, Rex1, Nac1, and Zfp281, all of which are associated with either pluripotency or somatic cell reprogramming (52). Another independent study used ChIP coupled to massively parallel DNA sequencing technology (ChIP-seq) to determine the target binding profiles of 13 transcription factors (Nanog, Oct4, STAT3, Smad1, Sox2, Zfx, c-Myc, n-Myc, Klf4, Esrrb, Tcfcp2l1, E2f1, and CTCF) and two coregulators (p300 and Suz12) in ESCs (Figure 2; 18). Both studies discovered combinatorial binding by multiple transcription factors on common hot-spot loci, an indication of extensive coregulatory mechanisms in the maintenance of ESCs. Kim et al. found colocalization of at least four transcription factors at 800 gene promoters, and found that 50% of the 6,632 target genes were bound by more than one transcription factor. Chen et al. (18) uncovered a total of 3,583 multiple transcription-factor-binding loci bound by four or more transcription factors across the entire genome. Moreover, the number of bound transcription factors was shown to correlate with the level of gene expression. Genes whose promoters were bound by multiple factors were preferentially expressed in ESCs, while singly bound promoters were often inactive in ESCs but activated upon differentiation (18, 52).

An important insight from these studies was the segmentation of ESC regulatory networks into Oct4-centric and Myc-centric modules. In Chen et al. (18), the Oct4-centric clusters

encompassed Sox2, Nanog, STAT3, Esrrb, Klf4, Tcfcp2l1, and Smad1, while the Myc-centric clusters consisted of c-Myc, n-Myc, E2f1, Zfx, and CTCF. Similarly, Kim et al. (52), uncovered high degrees of overlap among Oct4, Sox2, Nanog, Klf4, Dax1, Nac1, and Zfp281, while c-Myc and Rex1 binding sites formed a separate cluster. Notably, several of the genomic loci from the Oct4-centric cluster showed characteristic ESC-specific enhancer activity.

Enhanceosomes are nucleoprotein complexes of multiple transcription factors binding to enhancer DNA elements (122). ChIP-seq analysis of histone acetyl transferase p300, a transcription coactivator commonly found at enhancer regions, showed preferential localization to the Oct4-centric clusters. Importantly, the recruitment of p300 to the genomic sites is dependent on Oct4, Sox2, and Nanog, as their depletion led to reduction in p300 binding. It is likely that the cobinding of these transcription factors creates an interface for the recruitment of p300. Furthermore, several genomic fragments of loci from the Oct4-centric clusters exhibited ESC-specific enhancer activity when tested with the luciferase reporter assay (18). Target genes of the c-Myc cluster, on the other hand, were implicated more frequently in protein metabolism than in developmental processes. It was hypothesized that c-Myc and Rex1 maintain the expression of housekeeping genes that play key roles in sustaining the high proliferative capacities of ESCs. Evidence from a recent study identifies Myc target gene modules as the explanation for the common gene expression signatures of ESCs and proliferative cancer cells (53).

Given the large number of transcription-factor-bound sites identified from numerous studies using various platforms, it is imperative to cross-validate the list of binding sites to confidently identify the bona fide target genes. Additional confirmation of functional significances inferred from the interactions can also be achieved by integrating the binding data with loss-of-function studies. This will help to further refine the transcriptional regulatory network.

In summary, the core transcriptional network is characterized by extensive coregulation, as well as specialized and segregated modules in which regulatory factors have different responsibilities in maintenance of the pluripotent state. These features confer robustness to the transcriptional regulation of pluripotency and differentiation genes and facilitate a stable system that can rapidly respond to extrinsic differentiation cues.

Connecting Noncoding RNAs to the Transcriptional Circuitry

MicroRNAs (miRNAs) are a class of small ncRNAs that repress gene expression at the posttranscriptional level by base pairing to complementary sites of target messenger RNAs (mRNAs). Experiments that genetically ablate *Dicer* or *DGCR8*, key components in miRNA biogenesis, indicate that miRNAs are crucial for early development, proliferation, and differentiation of ESCs (6, 51, 131). In a study that implicated the ESC core network in regulation of ncRNAs, Marson et al. (70) demonstrated that ESC-specific transcription factors bind to promoters of ESC-specific miRNAs and induce their expression. The key factors also occupy a set of miRNA genes that are transcriptionally silenced in ESCs. The promoter regions of these differentiation-associated miRNA genes are occupied by polycomb group (PcG) proteins, which are postulated to poise the miRNA genes for expression during lineage specification (70). Interestingly, feedback loops in which miRNAs are responsible for controlling the expression of pluripotency factors were established. The differentiation-associated miRNA let-7 is known to target c-*Myc, Sall4*, and *Lin28* (76). Other differentiation-related miRNAs—miR-134, miR-296, and miR-470—were recently found to mediate regulation of *Oct4, Sox2*, and *Nanog* by targeting their coding regions (120). In addition, miR-200c, miR-203, and miR-183 cooperate to repress the pluripotency

factors *Sox2* and *Klf4* in mouse ESCs (133). Likewise, miR-145 represses the pluripotency machinery of human ESCs (136).

In addition to miRNAs, the mammalian genome also encodes large intergenic noncoding RNAs (lincRNAs) with lengths greater than 200 nucleotides. The evolutionarily conserved lincRNAs are associated with diverse biological processes. Work by Guttman et al. (38) identified 1,000 lincRNAs in murine ESCs, out of which 118 lincRNAs are bound and transcriptionally regulated by the core transcription factors Oct4 and Nanog. Interestingly, one of the ESC-specific lincRNAs identified by Guttman et al. was shown to be necessary for maintaining ESCs by RNA-interference knockdown experiment (181 kb from Enc1; 46), thus providing functional confirmation that lincRNAs play a direct role in ESC maintenance. Another study that combined the mouse ESC transcriptome with genomic location mapping identified four ESC-specified lincRNAs that are bound by Oct4 and Nanog (105). Importantly, knockdown and overexpression of these transcripts led to changes in the pluripotency of mouse ESCs and expression level of *Oct4, Nanog*, and lineage-specific genes. The authors further characterized one of the lincRNAs (AK028326) as a coactivator of Oct4 in a regulatory feedback loop. Recent work in iPSCs has revealed that OCT4, SOX2, and NANOG colocalize at the promoters of three lincRNAs whose expressions are highly enriched in reprogrammed cells (Figure 1*b*; 66). These lincRNAs were downregulated upon *OCT4* depletion as well as during differentiation of iPSCs and ESCs, indicating coregulation of specific lincRNAs by key pluripotency factors. The authors further demonstrated that lincRNA–regulator of reprogramming (lincRNA-RoR) regulates defined factor reprogramming, as knockdown of lincRNA-RoR significantly reduced the efficiency of iPSC formation. These results suggest that key pluripotency factors induce expression of lincRNA-RoR, which in turn modulates the establishment and maintenance of pluripotency (Figure 1*b*). Collectively, these discoveries underscore the intricate and complex network of regulatory loops through which ncRNAs (miRNAs and lincRNAs) are part of the integrated transcriptional circuitry that regulates the expression of genes that define ESC fate and behavior.

Rewiring the Core Transcriptional Network by Transposable Elements

Curiously, comparison of the OCT4- and NANOG-bound sites between the mouse and human genomes has revealed only modest similarity (9, 68). There are merely 32 common genes that are cobound by the two transcription factors in both mouse and human ESCs (68). Part of the dissimilarities could be attributed to differences in scope and breadth of the mapping platforms, such as the wider genome coverage offered by ChIP-PET technology (68). On the other hand, numerous studies have identified differences between human and mouse ESCs in terms of their morphologies, growth rates, marker expression, and growth factor requirements. For example, human ESCs depend on bFGF for self-renewal, whereas their mouse counterparts depend on the Lif/Stat3 pathway (97, 123). Recent studies have shown greater resemblance between human ESCs with mouse epiblast stem cells (11, 121). It is conceivable that the observed differences between mouse and human ESC networks could be due to the fact that they are pluripotent cells derived from different developmental stages.

A new study has provided key insights into the species divergences of the core circuitry. Kunarso et al. (57) performed ChIP-seq profiling of OCT4, NANOG, and CTCF binding in human ESCs and made direct comparison with existing ChIP-seq data sets previously generated for mouse ESCs (18, 57). Consistent with previous studies, limited conservation between mouse and human was found in OCT4 and NANOG binding (57). Of the most enriched 10% OCT4 and NANOG binding sites in human ESCs, barely 5% have homologous occupancy in mouse ESCs. This is in great contrast to the 50% conservation for CTCF binding between the two species (57). Intriguingly, part of the differences between

the binding sites in both species is due to the insertion of transposable elements, which comprise up to 25% of the OCT4- and NANOG-binding sites in both humans and mice. In humans, 20% of OCT4 and 15% of NANOG binding sites are associated with transposon elements, whereas in mice, it accounted for 7% of the OCT4 sites and 17% of NANOG sites. In another study, Xie and coworkers (135) used gene expression profiling of preimplantation embryos from three mammalian species (human, mouse, and bovine), integrating with comparative genomic data to predict the gene regulatory network that controls preimplantation embryonic development. Notably, the authors found widespread rewiring of the regulatory network between the species. Most of these observations can be attributed to single-nucleotide mutations and species-specific transposon insertion in the *cis*-regulatory modules of transcription factor binding sites.

The fact that species-specific transposable elements can rewire the binding landscape of pivotal factors in mammalian ESCs suggests a striking plasticity in the core transcriptional network. It also reveals transposable elements as a novel class of regulatory elements in the transcriptional circuitry governing ESC pluripotency and early embryonic development.

Integrating the Embryonic Stem Cell Protein Interaction Network with the Transcriptional Circuitry

Several groups have interrogated the protein interaction network of key pluripotency factors using affinity purification coupled to mass spectrometry. Using in vivo biotinylated Nanog protein as bait, Wang et al. (130) identified associated protein complexes and further applied the same tagging strategy to several Nanog partners to elucidate the protein interactome of ESC pluripotency. More recently, two groups using improved affinity FLAG or FTAP tagging methodologies reported a more extensive interactome of transcription factor Oct4 (89, 125). Several interesting features of how pluripotency is regulated have emerged from these protein-interaction network studies.

First, Nanog and Oct4 are connected to other critical pluripotency factors through protein interactions; these networks are enriched for transcription factors or proteins that are critical for ESC pluripotency or early mouse development (52). Second, chromatin modifiers such as histone deacetylase NuRD (MBD3, HDAC1/2, and CHD4), PcG proteins (YY1, RNF2, and RYBP), SWI/SNF chromatin remodeling complexes (BAF155 and BRG1) (reviewed in 42), DNA methyltransferases (DNMT3A/DNMTI), histone demethylase (AOF2), and the co-repressor KAP1 (TIF1) have been found to interact either directly or indirectly with Oct4 and Nanog (89, 125, 130). This suggests that Oct4 and Nanog may regulate transcription of downstream targets through the fine-tuning of chromatin states. Notably, the interaction of Oct4 with PcG proteins could be implicated in transcriptional repression of differentiation-promoting genes in the pluripotent state (27). Third, the expression level of the majority of interacting proteins is controlled by Nanog, Oct4, or other pluripotency transcription factors. For example, 56% of the genes in the Nanog interactome (130) and 51% of genes encoding Oct4 partners (89) are targets of at least one key ESC transcription factor (9, 18, 52, 68). Previous genome-wide location studies reported extensive colocalization of pluripotency factors and formation of enhanceosome sites (18, 52). Indeed, Dax1, Tcfcp2l1, and Esrrb are targeted in an Oct4-dependent manner to several of their shared binding sites with Oct4 (125). This observation suggests that protein-protein interactions may initiate the formation and stabilization of enhanceosomes at the multiple-factors-bound sites within the ESC transcriptional circuitry (Figure 1*c*; 125). This has provided mechanistic insights into previous findings where the consensus binding motifs of Dax1, Tcfcp2l1, Esrrb, and Nanog were found to be almost identical to the Oct4/Sox2 binding sites (18, 52). The high degree of interconnectivity between transcription and protein interaction networks appears to maintain ESCs in an internally stable and self-sustaining pluripotent state.

EPIGENETIC REGULATION OF THE PLURIPOTENT STATE

In recent years, it has become increasingly clear that pluripotency in ESCs is not only governed by extensive transcriptional regulatory networks, but also determined by the complement of histone modifications, DNA methylations, and chromatin remodelers that are responsible for establishing a unique ESC chromatin signature. In the second part of this review, we summarize the known features of the epigenetic landscape in ESCs and iPSCs, and how these marks add to our appreciation of the complexity of the pluripotency regulatory circuitry.

DNA Methylation in Embryonic Stem Cells

In mammalian systems, the addition of methyl groups to the 5′ position of cytosines at selected locations in the genome, usually where cytosine is immediately adjacent to guanine (known as a CpG site), is an integral part of early development. Indeed, the creation of the totipotent zygote from the fusion of terminally differentiated gametes is intimately interconnected and dependent upon dynamic changes in the methylome of the parental pronuclei. Analysis of DNA methylation profiles in gametes has demonstrated that both nuclei undergo extensive loss of DNA methylation marks—the male pronucleus is subjected to active demethylation prior to nuclear fusion (50, 73, 86, 98), while the female counterpart loses DNA methylation passively over the first few cell cleavages owing to a lack of maintenance methyltransferases (98). The hypomethylated genomic state persists until implantation, following which a wave of de novo methylation deposits epigenetic marks that are believed to correlate with increasing cell lineage specification and decreasing developmental potential (7, 50, 98). This epigenetic transition is conserved in cattle but occurs at an earlier time point—bovine embryos undergo demethylation around embryonic day 8 and de novo methylation between days 8 and 16. Analyses of zygotic methylomes of rats and pigs have also provided evidence for demethylation of parental genomes in these mammals (23).

The role of DNA methylation is of obvious importance in mammalian development: knockout studies of key DNA methyltransferases—namely, the maintenance methyltransferase *Dnmt1* and the de novo methyltransferases *Dnmt3a* and *Dnmt3b* (17, 83, 84)—have demonstrated that any misregulation of DNA methylation results in early embryonic lethality (60, 84). Given the nature of development as essentially the spatial and temporal interplay of pluripotency and differentiation cell programs, the severe in vivo repercussions of erroneous DNA methylation programs points strongly toward a role for DNA methylation in regulating the progression from ground-state pluripotency toward more lineage-specified cell states. This idea has been pursued through surveys and comparisons of 5′-methylcytosine distribution in the genomes of both pluripotent and differentiated cells.

Through such studies, some basic insights into the nature of methylation patterns have emerged. First, cytosine methylations demonstrate bimodality in which regions that are densely packed with CpG dinucleotides (also known as CpG islands) tend to be protected from methylation, whereas regions with low-density CpG dinucleotides are not (50). This is by no means absolute, because different cell types are known to demonstrate differential methylation at certain CpG islands. A closer look at CpG sites in the human genome has also revealed that many CpG sites fall within putative gene promoter regions (33). These promoters have been divided into those that are CpG rich (high CpG promoters) or CpG poor (low CpG promoters) (26, 100, 132). Interestingly, the integration of DNA methylation maps with gene expression profiles and RNA polymerase II binding sites suggested a negative correlation between methylation marks and transcription. High CpG promoters often marked active genes and were mostly associated with ubiquitously expressed housekeeping genes, whereas low CpG promoters usually corresponded to tissue-specific

genes, of which only a small subset is expressed in any given cell type so as to maintain cell identity (132).

In the context of ESCs, the connection made between DNA methylation and transcriptional regulation has garnered great interest, primarily because it hinted at a new layer of transcriptional control for the pluripotent cell state (12). Numerous studies utilizing either methylcytosine affinity enrichment (28, 31) or bisulfite treatment (3, 24, 74, 75) followed by a variety of targeted approaches (e.g., user-defined arrays, library generation using custom-designed padlock probes or reduced representation sequencing) have been performed to define an ESC-specific DNA methylation profile that provides insight into the epigenetic control of pluripotency.

Using methylated DNA immunoprecipitation coupled to microarray analysis, two groups have independently mapped the methylome of mouse ESCs (Figure 2; 28, 31). Gene ontology analysis of genes with methylated promoters (which were often transcriptionally silent) or unmethylated promoters (mostly transcriptionally active) demonstrated that although methylated genes generally corresponded to those implicated in differentiation, over half of unmethylated genes were regulators of transcription. In addition, the promoter regions of important pluripotency genes expressed in ESCs—such as *Oct4, Nanog, Sox2, Stat3, Tdgf1, Lefty1*, and *Rex1*—are unmethylated in ESCs but methylated in somatic cells. Interestingly, a comparison of DNA methylation profiles with Oct4 or Nanog binding sites in the mouse genome revealed that transcription levels of genes activated by Oct4 or Nanog binding were also weakly affected by the presence of methylation marks at promoter regions. Thus, although the link between DNA methylation and the regulation of pluripotency is still not fully understood, the above observations present evidence for a model in which the CpG methylation machinery, through highly selective deposition of DNA methylations, functions both as an upstream regulator of key pluripotency genes and as a mediator for fine-tuning the regulatory effects mediated by transcription factors.

One issue with the use of targeted approaches for DNA methylation analysis is the inherent bias introduced by the need to select regions of the genome for analysis. A study by Lister et al. (63) overcame this limitation with the Methyl-seq approach, in which deep sequencing was applied to bisulfite-treated genomic DNA from human ESCs and fibroblasts. The team was able to map, on an unprecedented scale, the human methylome at single-base resolution. Their data revealed that apart from CpG methylations, ESCs were also exclusively methylated at non-CpG sites where methyl groups were appended to cytosines at CHG or CHH trinucleotides (where H = A, C, or T). This is in agreement with earlier reports in which non-CpG marks were found in mouse ESCs (95). Further characterization of CHG and CHH methylations demonstrated that these marks are enriched in exons of highly expressed genes such as *OCT4* (63). Although little is currently known about the mechanisms of depositing non-CpG methylations and their exact role in ESCs, the fact that they are observed in iPSCs but rarely in fibroblasts argues for their importance in the establishment of the pluripotent cell state.

Methyl-seq has also been used to compare methylomes across iPSC lines generated with different reprogramming methods and human ESC lines (64). The study revealed that differentially methylated regions (DMRs) iniPSCs that are shared by the different iPSC lines (e.g., close to centromeres and telomeres) could define loci that are less amenable to changes in the methylation (64). The molecular basis for these hot spots of differential methylation is intriguing and provides new insights into the manner in which methyl groups are deposited or lost from cytosine residues.

The traditional emphasis on studying DNA methylation at CpG islands has also been challenged in another study where DMRs between iPSCs and fibroblasts were compiled (25). Contrary to what was expected, the majority of DMRs found from the comparison of the two cell types were located at CpG island "shores"—regions located in the vicinity of CpG islands that have lower CpG density, rather than within actual CpG islands—and overlapped with DMRs that distinguished different tissue types (25). Tissue-specific DMRs are critical for cell fate specification; the fact that resetting the terminally differentiated epigenome back to the pluripotent state involves a very similar set of DMR loci is indicative of the DNA methylation-mediated pathway in which developmental gene expression is modulated for the recapitulation of pluripotency.

Besides the modification of cytosines to methylcytosines, recent studies have also found that methylated cytosines can be further modified to 5′-hydroxymethylcytosines (5hmC) almost exclusively by the TET (ten-eleven translocation) family of proteins (45, 55, 117). Of the three TET enzymes, *Tet1* and *Tet2* are expressed in mouse ESCs, and emerging evidence has implicated them in both ESC maintenance and differentiation. Work by Ito et al. (45) has suggested that *Tet1* knockdown affects *Nanog* expression, dampens ESC renewal rates, and drives the transcription of trophectoderm-specific markers such as *Cdx2* and *Hand1. TET1* depletion also correlated with an increase in methylation at *Nanog* promoters. Data from Koh et al. (55) demonstrated a role for Tet1 and Tet2 in ESC differentiation. In particular, the two enzymes were both downregulated upon differentiation but upregulated over the course of iPSC genesis from fibroblasts. Koh et al. proposed that correlations between 5hmC epigenetic marks and the pluripotent cell state could arise from the regulation of the TET proteins by Oct4 and Sox2 given the presence of binding sites for these two master regulators at *Tet1* and *Tet2*, and their downregulation under Oct4/Sox2 knockdown conditions. Although the role of the 5hmC epigenetic mark in stabilizing ESCs in the pluripotent state is still under debate and awaits more conclusive experimental testing, the fresh insights provided by these studies establish an additional link between the epigenetic state of chromatin and the core transcriptional regulatory network controlling pluripotency. In addition, the extension of the findings from mouse to human ESCs is an exciting possibility that remains to be studied.

Although DNA methylation has been observed to interact with and modulate the master regulators of the ESC fate, whether DNA methylation is absolutely necessary to maintain ESC in stable equilibrium remains controversial; ESCs in which *Dnmt1, Dnmt3a*, or *Dnmt3b* are knocked out can be maintained in a self-perpetuating state without differentiating, and express the characteristic suite of pluripotency markers (83, 123). However, these cells also tend to be trapped in their pluripotent states and undergo rapid apoptosis when induced to differentiate (88). On the other hand, DNA methyltransferase inhibitors such as 5-azacytidine (106) and RG108 (77) have been known to enhance direct reprogramming of mouse fibroblasts by as much as 10–30-fold, arguing for a strong influence of DNA methylation status on the establishment of the pluripotent cell state. Given that both processes in which the methylome landscape is critical for pluripotency involve cell state transitions, one could speculate that the accurate establishment of DNA methylation across the genome is the gateway to a fully functional developmental program. As such, although it is true that DNA methylation is not essential for maintaining self-renewal in ESCs, it is crucial for fulfilling the other defining characteristic of ESC—the ability to give rise to all tissues in the adult organism.

Histone Modifications in Embryonic Stem Cells

Besides DNA methylations, the vertebrate genome is also populated by a variety of histone modifications (methylations, acetylations, phosphorylations, etc.). These epigenetic marks tend to occur at conserved amino acid residues of histone tail domains and have the ability

to fine-tune gene expression (115). Histone modifications for which correlations between the modifications and gene activity are well established have been mapped in ESCs (4, 42). Although the information generated from these studies has been helpful in sketching out key features of the ESC-specific epigenome, a true appreciation of the mechanisms by which histone modifications influence pluripotent cell identity is currently lacking. However, such studies have been instrumental in driving home the idea of the combinatorial nature of histone modifications. The proposal of a histone code (48, 113), in which histone modifications gather in different permutations and combinations to regulate gene activity, allows for a more sensitive and versatile epigenetic regulatory mechanism.

H3K4 and H3K27 methylation. The mapping of histone 3 lysine-4 and lysine-27 trimethylation (H3K4me3 and H3K27me3, respectively) at highly conserved noncoding elements and later on a genome-wide scale in mouse and human ESCs has revealed that the majority of protein-coding genes surveyed are associated with H3K4me3 (2, 4, 5, 35, 36, 42, 78, 87, 140). This observation has been confirmed in vivo in early embryos of zebra fish (126). H3K4 modifications usually concentrate at transcription start sites, and, in agreement with previously established notions of H3K4 being an activating mark (99, 101), are observed in association with key pluripotency factors such as *Nanog, Oct4*, and *Sox2*. Interestingly, H3K4me3 often falls within larger swaths of genome populated by repressive H3K27me3 marks, resulting in the formation of bivalent domains (5), which are found with high incidence at transcription start sites of transcription factors crucial for cell fate specification, such as the Sox, Hox, and Pax gene families. Interestingly, the comparison of histone modification domains in ESCs and somatic cells revealed that bivalent domains are rare in differentiated cells. Indeed, bivalent loci present in ESCs usually resolved into patches of H3K4 and H3K27 methylation that exist independently as monovalent domains. Notably, methylation of H3K27 is epistatic to that of H3K4, such that genes marked with bivalent promoters tend to be minimally expressed. As such, bivalency in histone modifications has been regarded as a reflection of an ESC-specific chromatin state in which factors critical to differentiation and cell fate specification are repressed, yet kept in a dynamic state poised for rapid activation upon induction of differentiation.

Bivalent domains are critical for development—knocking out genes encoding the trithorax group proteins [which deposit H3K4 methylation marks (10)] and the PcG proteins (which deposit H3K27 methylation marks) is lethal for gastrulation and embryonic survival (90, 102, 128). Given their importance, genome-wide location analysis has been undertaken to characterize bivalent domains in finer detail. Binding analysis for the two polycomb repressive complexes (PRC1 and PRC2) has revealed that although they colocalize to a large extent with each other and with H3K27 methylation marks, there exists a significant population of bivalent domains in ESCs that associate with PRC2 alone (56). Unlike bivalent domains occupied by both complexes, the correspondence between this new class of bivalent domains and transcription factor promoters was weak and tended to lose H3K27 methylations upon differentiation. Further analysis of the activity of PRC1 and PRC2 is required before we can better appreciate the implications of having two varying classes of bivalent domains in ESCs.

Given the role of H3K27me3 as a repressor of cell lineage specification, there has been great interest in understanding the molecular mechanisms regulating the recruitment of PcG proteins to gene promoters. PcG localization in *Drosophila* occurs via the recognition of specific DNA sequences in the genome called polycomb response elements (16, 108, 116), but no comparable recognition systems have been found in mammals. Although PcG localization sites corresponded to genomic regions of high CG content (56), this property alone was not sufficiently stringent for the prediction of PcG distribution. A comparison of OCT4, SOX2, and NANOG binding sites with PcG target genes in human ESCs has also

revealed a significant colocalization of PcG proteins with the core transcriptional regulatory network. Although direct interactions between the two groups of proteins have not been found, it is possible to conceive a PcG recruitment mechanism in which binding of the transcription factors provides some sort of signal that flags genes for PcG targeting.

Recently, Jarid2—a member of the Jumonji family of lysine demethylases—has also been recognized as a modulator of PRC2 activity in mouse ESCs (61, 91, 92). This model is attractive because JARID2 is a DNA binding protein (61) known to colocalize to genomic loci occupied by PRC2. The promoter of *Jarid2* is occupied by OCT4, SOX2, and NANOG, suggesting that Jarid2 could be the mediator through which PcG proteins are wired into an extended transcriptional network in ESCs. Apart from JARID2, the polycomb-like 2 (PCL2) protein was also found to associate with PRC2 in mouse ESCs (129) and could play a role in PcG recruitment and stable propagation of the ESC state. The various threads of evidence presented above suggest the existence of several mechanisms through which bivalent domains are established in ESCs.

H3K9 methylation. In addition to the distinctive H3K4me3 and H3K27me3 co-occupancy domains, ESCs also contain the repressive H3K9 di- or trimethylation histone modifications. Both types of H3K9 methylations have been mapped in ESCs using either ChIP-on-chip (134) or single-molecule sequencing (77). Notably, the dynamics of H3K9me2 and H3K9me3 deposition are distinctly different. H3K9me2-populated genomic domains seem to undergo spreading as ESCs differentiate, such that they eventually occupy large blocks of chromatin, but this phenomenon was not observed for H3K9me3. Beyond the idea that H3K9me2 mediates gene silencing, it is not clear how methylation spreading is carried out and how it is kept in check in undifferentiated cells. The two lysine methylations also colonize different regions of the genome, suggesting nonredundancy in the epigenetic effects mediated by each mark. Interestingly, H3K9me3 is mostly detected in partnership with H4K20me3, providing further evidence in support of the histone code hypothesis (77).

Crosstalk Between Histone Modification and DNA Methylation

DNA methylation and histone modification have long been known to engage in extensive crosstalk (Figure 3). Compelling evidence for their intimate relationship has arisen from studies in *Neurospora* and *Arabidopsis* (47, 119). Similar observations have been made in mammalian systems; large-scale mapping of DNA methylation and H3K4me2 showed a strong negative correlation between these two marks in both ESCs and more differentiated cells (75), an observation that is in agreement with earlier studies (41, 58). Notably, the exchange of information between epigenetic regulators is bidirectional: deposition of histone modifications may be both a result of DNA methylation and an inducer of the epigenetic mark, and the same applies for DNA methylation. Indeed, cases in which CpG methylations appear to induce histone deacetylation or H3K9 methylations have been reported.

The presence of DNA methylation on genomic segments is detected by methylcytosine binding proteins—in particular, MECP2 and MBD2. MECP2 has been found to associate with histone deacetylases (49, 80) and histone H3 methyltransferases (32). Likewise, mediators of H3K9 and H3K27 methylation also demonstrate a propensity for interaction with DNA methyltransferases (DNMTs) (Figure 3*a*). G9A and ESET, both methyltransferases for H3K9, are able to interact with DNMT3A/3B to direct the deposition of DNA methylations in their vicinity (Figure 3*b*; 29, 62). G9A along with SUV39H1/2 also associate with the heterochromatin protein HP1 to bring about the recruitment of DNMT1 and DNMT3B to pericentric repeat regions of the genome (Figure 3*b*; 59, 110). There is also evidence that the PRC2 component EZH2 engages in direct interaction with DNMTs to establish stable repression of target genes (Figure 3*c*; 127).

The idea of a bidirectional pipeline through which epigenetic information can be propagated might explain how the wave of de novo methylation that occurs early in mammalian development is precisely imposed on specific genomic loci. It has been proposed that the presence of H3K4 methylations, which is deposited prior to de novo methylation during maternal-to-zygotic transition, might occlude sites of histone modification and thus protect specific regions of the genome from acquiring methylation (13). Because methylations at H3K4 are activating and normally found at CpG islands in ESCs, the model also explains how most CpG islands are accorded an unmethylated status.

DEFINING STEM CELL PLURIPOTENCY: INSIGHTS FROM REPROGRAMMING

In mammalian systems, the pluripotent cell state exists transiently as the embryo proceeds through the phases of development to give rise to an organism with a diverse array of specialized cell types. This progressive and inexorable loss of developmental plasticity that occurs with embryonic maturation is notoriously difficult to reverse. Indeed, the few methods available for artificial reprogramming—namely, somatic cell nuclear transfer (22, 37, 44), cell fusion (111), and defined cell culture conditions (72, 103)—require very specific cell types and specialized manipulation conditions, often relying heavily on a complement of unknown factors present in oocytes or pluripotent cells before the rare reprogrammed cell can be derived. Groundbreaking work by Takahashi & Yamanaka (118) in 2006 revolutionized the field with the demonstration that the ectopic introduction of merely four transcription factors known to be implicated in the regulation of pluripotency could result in a dramatic switching of cell potency and revert differentiated cells back to an ESC-like state. The implications this has for regenerative medicine, disease modeling, and drug discovery are immense, and the development of such a controlled in vitro system—one in which key molecular events along the path toward the pluripotency can be dissected—has expanded our understanding of ESC biology.

Immediate insights into pluripotency arose from efforts to refine and enhance the efficiency with which iPSCs could be generated. The discovery that small chemicals that demethylate DNA or alter histone acetylation can enhance iPSC derivation indicates that pluripotent cell states can be attained only with an open and permissive genomic architecture (30). This notion is reinforced by evidence that introducing factors engaged in remodeling chromatin complexes, such as BAF (109) and CHD1 (34, 93), can help differentiated cells scale the epigenetic barrier to attain an ESC-like state.

More importantly, analysis of “omics” data generated from cells at different stages along the path toward pluripotency has identified critical prerequisites for the establishment of pluripotency—features that might otherwise have been lost in the sea of information generated from genome-wide surveys of ESC transcriptomes and epigenomes. For example, comparisons of global gene and small RNA expression profiles between fully and partially reprogrammed iPSCs have implicated expression of protein-coding genes and miRNAs at the *Dlk1-Dio3* locus as essential to the most faithful acquisition of pluripotency (65, 112).

Recent genome-scale studies of distinct pluripotent stem cell types indicate enough heterogeneity to suggest that we need to refine our definition of pluripotency. Comparisons of the DNA methylation profiles of pluripotent murine cells derived by somatic cell nuclear transfer or factor-based reprogramming from different somatic cell types indicate that reprogramming fails to fully erase past epigenetic marks, even in iPSCs that pass high stringency tests for pluripotency, such as the ability for germline transmission (54). Whole-genome bisulfite sequence analysis of human ESCs and iPSCs has revealed that specific regions of the genome, particularly peri-centromeric and telomeric regions, can retain

aberrant methylation (64). The concept of persistent epigenetic memory in iPSCs is intriguing, because it suggests that the pluripotent state of the genome is governed by a restricted network of loci and remains tolerant of a wide array of epigenetic marks outside that network. These observations also beg the question of how stringent the genomic definition for pluripotency needs to be; the answer is essential to defining how we use genomic data as a benchmark to assess cellular pluripotency, either for basic research purposes or as a quality control measure in the selection of clinical-grade human iPSCs (for which rigorous functional tests are limiting). Further complicating the notion of pluripotency is the existence of metastable, interchangeable states that exist in different pluripotent stem cell isolates (19). For example, it is well known that important distinctions in cell morphology, signaling pathways, and X inactivation status exist between ESCs derived from mouse and human blastocysts (71, 123; reviewed in 96). Notably, human ESCs have been observed to bear significant biological resemblance to mouse epiblast stem cells (121) and were thus thought to be representative of a “primed” pluripotent state, in contrast to “naive” mouse ESCs, which reside in a more primitive ground state of pluripotency and show greater propensities for chimera tissue contribution (138). The notion of alternative and interchangeable states of pluripotency has been verified recently by Hanna et al. (39), who demonstrated that human ESCs could undergo transitions to a more murine ESC-like state. It is possible that subtle differences in the transcriptome and epigenome exist even within the naive or primed states, given that several degrees of variability in the differentiation potential of human ESCs have been observed in a manner independent of passage number (85).

From these studies, it appears that the definition of the genome-wide transcriptional and epigenetic profiles that define the pluripotency of ESCs are complex and subject to revision and refinement. Although ESCs possess characteristic transcriptional network, DNA methylation, and histone modification signatures, it would be wise to remain mindful that the presence of metastable pluripotency states challenges the notion of a singular, uniform genomic state of pluripotency.

CONCLUDING REMARKS

Using various “omics” studies, stem cell scientists have generated vast quantities of data that have begun to illustrate the epigenetic landscape of pluripotent stem cells and the transitions that occur during commitment to differentiation. This information has undoubtedly been crucial for describing genomic networks, but our understanding of the functional role and significance of particular epigenomic differences remains limited, and more mechanistic studies, together with ongoing efforts to map genomic features of ESCs, iPSCs, and the growing number of alternative cell states, are needed.

As discussed by Loh & Lim (67), lack of standardization in microarray data analysis may lead to different conclusions in comparing pluripotent states of iPSCs and native ESCs. Ultimately, as “omics” technologies become increasingly accessible, the stem cell community must agree on and adopt common statistical standards of data analysis, as well as precise molecular definitions of pluripotency in all its varieties.

Acknowledgments

G.Q.D. is supported by grants from the NIH (RO1-DK70055, RO1-DK59279, UO1-HL100001, and special funds from the ARRA stimulus package—RC2-HL102815), the Roche Foundation for Anemia Research, Alex's Lemonade Stand, and the Harvard Stem Cell Institute, and is an affiliate member of the Broad Institute, a recipient of Clinical Scientist Awards in Translational Research from the Burroughs Wellcome Fund and the Leukemia and Lymphoma Society, and an investigator of the Howard Hughes Medical Institute and the Manton Center for Orphan Disease Research. J.J.C. is supported by the Systems-Based Consortium for Organ Design and Engineering

(SysCODE), NIH grant RL1DE019021, the Ellison Medical Foundation, and the Howard Hughes Medical Institute. Y.-H.L. and L.Y. are supported by the Agency of Science, Technology, and Research (A*STAR), Singapore.

REFERENCES

1. Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev. 2003; 17:126–140. [PubMed: 12514105]
2. Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, et al. Chromatin signatures of pluripotent cell lines. Nat. Cell Biol. 2006; 8:532–538. [PubMed: 16570078]
3. Ball MP, Li JB, Gao Y, Lee J-H, LeProust EM, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat. Biotechnol. 2009; 27:361–368. [PubMed: 19329998]
4. Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, et al. High-resolution profiling of histone methylations in the human genome. Cell. 2007; 129:823–837. [PubMed: 17512414]
5. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006; 125:315–326. [PubMed: 16630819]
6. Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, et al. Dicer is essential for mouse development. Nat. Genet. 2003; 35:215–217. [PubMed: 14528307]
7. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002; 16:6–21. [PubMed: 11782440]
8. Botquin V, Hess H, Fuhrmann G, Anastassiadis C, Gross MK, et al. New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. Genes Dev. 1998; 12:2073–2090. [PubMed: 9649510]
9. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell. 2005; 122:947–956. [PubMed: 16153702]
10. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature. 2006; 441:349–353. [PubMed: 16625203]
11. Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, et al. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature. 2007; 448:191–195. [PubMed: 17597762]
12. Busslinger M, Hurst J, Flavell RA. DNA methylation and the regulation of globin gene expression. Cell. 1983; 34:197–206. [PubMed: 6883509]
13. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nat. Rev. Genet. 2009; 10:295–304. [PubMed: 19308066]
14. Chambers I, Colby D, Robertson M, Nichols J, Lee S, et al. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell. 2003; 113:643–655. [PubMed: 12787505]
15. Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, et al. Nanog safeguards pluripotency and mediates germline development. Nature. 2007; 450:1230–1234. [PubMed: 18097409]
16. Chan CS, Rastelli L, Pirrotia V. A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. EMBO J. 1994; 13:2553–2564. [PubMed: 7912192]
17. Chen T, Ueda Y, Dodge JE, Wang Z, Li E. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol. Cell. Biol. 2003; 23:5594–5605. [PubMed: 12897133]
18. Chen X, Xu H, Yuan P, Fang F, Huss M, et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell. 2008; 133:1106–1117. [PubMed: 18555785]
19. Cherry A, Daley GQ. Another horse in the meta-stable state of pluripotency. Cell Stem Cell. 2010; 7:641–642. [PubMed: 21112555]
20. Chew JL, Loh YH, Zhang W, Chen X, Tam WL, et al. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. Mol. Cell. Biol. 2005; 25:6031–6046. [PubMed: 15988017]

21. Cole MF, Johnstone SE, Newman JJ, Kagey MH, Young RA. Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev. 2008; 22:746–755. [PubMed: 18347094]
22. Cowan CA, Atienza J, Melton DA, Eggan K. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science. 2005; 309:1369–1373. [PubMed: 16123299]
23. Dean W, Santos F, Stojkovic M, Zakhartchenko V, Walter J, et al. Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc. Natl. Acad. Sci. USA. 2001; 98:13734–13738. [PubMed: 11717434]
24. Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, et al. Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. Nat. Biotechnol. 2009; 27:353–360. [PubMed: 19330000]
25. Doi A, Park I-H, Wen B, Murakami P, Aryee MJ, et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat. Genet. 2009; 41:1350–1353. [PubMed: 19881528]
26. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. Nat. Genet. 2006; 38:1378–1385. [PubMed: 17072317]
27. Endoh M, Endo TA, Endoh T, Fujimura Y, Ohara O, et al. Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. Development. 2008; 135:1513–1524. [PubMed: 18339675]
28. Farthing CR, Ficz G, Ng RK, Chan C-F, Andrews S, et al. Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. PLoS Genet. 2008; 4:e1000116. [PubMed: 18584034]
29. Feldman N, Gerson A, Fang J, Li E, Zhang Y, et al. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat. Cell Biol. 2006; 8:188–194. [PubMed: 16415856]
30. Feng B, Ng J-H, Heng J-CD, Ng H-H. Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. Cell Stem Cell. 2009; 4:301–312. [PubMed: 19341620]
31. Fouse SD, Shen Y, Pellegrini M, Cole S, Meissner A, et al. Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell Stem Cell. 2008; 2:160–169. [PubMed: 18371437]
32. Fuks, Fo; Hurd, PJ.; Wolf, D.; Nan, X.; Bird, AP.; Kouzarides, T. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. J. Biol. Chem. 2003; 278:4035–4040. [PubMed: 12427740]
33. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J. Mol. Biol. 1987; 196:261–282. [PubMed: 3656447]
34. Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ, et al. Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature. 2009; 460:863–868. [PubMed: 19587682]
35. Guenther MG, Frampton GM, Soldner F, Hockemeyer D, Mitalipova M, et al. Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. Cell Stem Cell. 2010; 7:249–257. [PubMed: 20682450]
36. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. Cell. 2007; 130:77–88. [PubMed: 17632057]
37. Gurdon JB, Elsdale TR, Fischberg M. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. Nature. 1958; 182:64–65. [PubMed: 13566187]
38. Guttman M, Amit I, Garber M, French C, Lin MF, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature. 2009; 458:223–227. [PubMed: 19182780]
39. Hanna J, Cheng AW, Saha K, Kim J, Lengner CJ, et al. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc. Natl. Acad. Sci. USA. 2010; 107:9222–9227. [PubMed: 20442331]

40. Hanna JH, Saha K, Jaenisch R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell. 2010; 143:508–525. [PubMed: 21074044]
41. Hashimshony T, Zhang J, Keshet I, Bustin M, Cedar H. The role of DNA methylation in setting up chromatin structure during development. Nat. Genet. 2003; 34:187–192. [PubMed: 12740577]
42. Hawkins RD, Hon GC, Lee LK, Ngo Q, Lister R, et al. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell Stem Cell. 2010; 6:479–491. [PubMed: 20452322]
43. Ho L, Crabtree GR. Chromatin remodelling during development. Nature. 2010; 463:474–484. [PubMed: 20110991]
44. Hochedlinger K, Jaenisch R. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. Nature. 2002; 415:1035–1038. [PubMed: 11875572]
45. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature. 2010; 466:1129–1133. [PubMed: 20639862]
46. Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, et al. Dissecting self-renewal in stem cells with RNA interference. Nature. 2006; 442:533–538. [PubMed: 16767105]
47. Jackson JP, Johnson L, Jasencakova Z, Zhang X, PerezBurgos L, et al. Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. Chromosoma. 2004; 112:308–315. [PubMed: 15014946]
48. Jenuwein T, Allis CD. Translating the histone code. Science. 2001; 293:1074–1080. [PubMed: 11498575]
49. Jones PL, Jan Veenstra GC, Wade PA, Vermaak D, Kass SU, et al. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat. Genet. 1998; 19:187–191. [PubMed: 9620779]
50. Kafri T, Ariel M, Brandeis M, Shemer R, Urven L, et al. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. Genes Dev. 1992; 6:705–714. [PubMed: 1577268]
51. Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev. 2005; 19:489–501. [PubMed: 15713842]
52. Kim J, Chu J, Shen X, Wang J, Orkin SH. An extended transcriptional network for pluripotency of embryonic stem cells. Cell. 2008; 132:1049–1061. [PubMed: 18358816]
53. Kim J, Woo AJ, Chu J, Snow JW, Fujiwara Y, et al. A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. Cell. 2010; 143:313–324. [PubMed: 20946988]
54. Kim K, Doi A, Wen B, Ng K, Zhao R, et al. Epigenetic memory in induced pluripotent stem cells. Nature. 2010; 467:285–290. [PubMed: 20644535]
55. Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, et al. Tet1 and Tet2 Regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell. 2011; 8:200–213. [PubMed: 21295276]
56. Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet. 2008; 4:e1000242. [PubMed: 18974828]
57. Kunarso G, Chia NY, Jeyakani J, Hwang C, Lu X, et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat. Genet. 2010; 42:631–634. [PubMed: 20526341]
58. Lande-Diner L, Zhang J, Ben-Porath I, Amariglio N, Keshet I, et al. Role of DNA methylation in stable gene repression. J. Biol. Chem. 2007; 282:12194–12200. [PubMed: 17311920]
59. Lehnertz B, Ueda Y, Derijck AAHA, Braunschweig U, Perez-Burgos L, et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr. Biol. 2003; 13:1192–1200. [PubMed: 12867029]
60. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 1992; 69:915–926. [PubMed: 1606615]

61. Li G, Margueron R, Ku M, Chambon P, Bernstein BE, Reinberg D. Jarid2 and PRC2, partners in regulating gene expression. Genes Dev. 2010; 24:368–380. [PubMed: 20123894]
62. Li H, Rauch T, Chen ZX, Szabo PE, Riggs AD, Pfeifer GP. The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. J. Biol. Chem. 2006; 281:19489–19500. [PubMed: 16682412]
63. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009; 462:315–322. [PubMed: 19829295]
64. Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature. 2011; 471:68–73. [PubMed: 21289626]
65. Liu L, Luo G-Z, Yang W, Zhao X, Zheng Q, et al. Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. J. Biol. Chem. 2010; 285:19483–19490. [PubMed: 20382743]
66. Loewer S, Cabili MN, Guttman M, Loh Y-H, Thomas K, et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nat. Genet. 2010; 42:1113–1117. [PubMed: 21057500]
67. Loh KM, Lim B. Recreating pluripotency? Cell Stem Cell. 2010; 7:137–139. [PubMed: 20682438]
68. Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. 2006; 38:431–440. [PubMed: 16518401]
69. Macarthur BD, Ma'ayan A, Lemischka IR. Systems biology of stem cell fate and cellular reprogramming. Nat. Rev. Mol. Cell Biol. 2009; 10:672–681. [PubMed: 19738627]
70. Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, et al. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell. 2008; 134:521–533. [PubMed: 18692474]
71. Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. USA. 1981; 78:7634–7638. [PubMed: 6950406]
72. Matsui Y, Zsebo K, Hogan BL. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell. 1992; 70:841–847. [PubMed: 1381289]
73. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Embryogenesis: demethylation of the zygotic paternal genome. Nature. 2000; 403:501–502. [PubMed: 10676950]
74. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res. 2005; 33:5868–5877. [PubMed: 16224102]
75. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008; 454:766–770. [PubMed: 18600261]
76. Melton C, Judson RL, Blelloch R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature. 2010; 463:621–626. [PubMed: 20054295]
77. Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, et al. Dissecting direct reprogramming through integrative genomic analysis. Nature. 2008; 454:49–55. [PubMed: 18509334]
78. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature. 2007; 448:553–560. [PubMed: 17603471]
79. Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell. 2003; 113:631–642. [PubMed: 12787504]
80. Nan X, Ng H-H, Johnson CA, Laherty CD, Turner BM, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature. 1998; 393:386–389. [PubMed: 9620804]
81. Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, et al. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell. 1998; 95:379–391. [PubMed: 9814708]

82. Niwa H. Open conformation chromatin and pluripotency. Genes Dev. 2007; 21:2671–2676. [PubMed: 17974911]
83. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999; 99:247–257. [PubMed: 10555141]
84. Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat. Genet. 1998; 19:219–220. [PubMed: 9662389]
85. Osafune K, Caron L, Borowiak M, Martinez RJ, Fitz-Gerald CS, et al. Marked differences in differentiation propensity among human embryonic stem cell lines. Nat. Biotechnol. 2008; 26:313–315. [PubMed: 18278034]
86. Oswald J, Engemann S, Lane N, Mayer W, Olek A, et al. Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. 2000; 10:475–478. [PubMed: 10801417]
87. Pan G, Tian S, Nie J, Yang C, Ruotti V, et al. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell. 2007; 1:299–312. [PubMed: 18371364]
88. Panning B, Jaenisch R. DNA hypomethylation can activate Xist expression and silence X-linked genes. Genes Dev. 1996; 10:1991–2002. [PubMed: 8769643]
89. Pardo M, Lang B, Yu L, Prosser H, Bradley A, et al. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell. 2010; 6:382–395. [PubMed: 20362542]
90. Pasini D, Bracken AP, Jensen MR, Denchi EL, Helin K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. 2004; 23:4061–4071. [PubMed: 15385962]
91. Pasini D, Cloos PAC, Walfridsson J, Olsson L, Bukowski J-P, et al. JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature. 2010; 464:306–310. [PubMed: 20075857]
92. Peng JC, Valouev A, Swigut T, Zhang J, Zhao Y, et al. Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. Cell. 2009; 139:1290–1302. [PubMed: 20064375]
93. Persson J, Ekwall K. Chd1 remodelers maintain open chromatin and regulate the epigenetics of differentiation. Exp. Cell Res. 2010; 316:1316–1323. [PubMed: 20211173]
94. Polo JM, Liu S, Figueroa ME, Kulalert W, Eminli S, et al. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat. Biotechnol. 2010; 28:848–855. [PubMed: 20644536]
95. Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc. Natl. Acad. Sci. USA. 2000; 97:5237–5242. [PubMed: 10805783]
96. Rao M. Conserved and divergent paths that regulate self-renewal in mouse and human embryonic stem cells. Dev. Biol. 2004; 275:269–286. [PubMed: 15501218]
97. Reubinoff BE, Pera MF, Fong CY, Trounson A, Bongso A. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat. Biotechnol. 2000; 18:399–404. [PubMed: 10748519]
98. Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev. Biol. 2002; 241:172–182. [PubMed: 11784103]
99. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, et al. Active genes are trimethylated at K4 of histone H3. Nature. 2002; 419:407–411. [PubMed: 12353038]
100. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl. Acad. Sci. USA. 2006; 103:1412–1417. [PubMed: 16432200]
101. Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat. Cell Biol. 2004; 6:73–77. [PubMed: 14661024]
102. Schumacher A, Faust C, Magnuson T. Positional cloning of a global regulator of anterior-posterior patterning in mice. Nature. 1996; 383:250–253. [PubMed: 8805699]

103. Shamblott MJ, Axelman J, Wang S, Bugg EM, Littlefield JW, et al. Derivation of pluripotent stem cells from cultured human primordial germ cells. Proc. Natl. Acad. Sci. USA. 1998; 95:13726–13731. [PubMed: 9811868]
104. Sharov AA, Masui S, Sharova LV, Piao Y, Aiba K, et al. Identification of *Pou5f1, Sox2*, and *Nanog* downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data. BMC Genomics. 2008; 9:269. [PubMed: 18522731]
105. Sheik Mohamed J, Gaughwin PM, Lim B, Robson P, Lipovich L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. RNA. 2010; 16:324–337. [PubMed: 20026622]
106. Shi Y, Desponts C, Do JT, Hahm HS, Schöler HR, Ding S. Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell. 2008; 3:568–574. [PubMed: 18983970]
107. Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, et al. Nanog is the gateway to the pluripotent ground state. Cell. 2009; 138:722–737. [PubMed: 19703398]
108. Simon J, Chiang A, Bender W, Shimell MJ, O'Connor M. Elements of the *Drosophila* bithorax complex that mediate repression by Polycomb group products. Dev. Biol. 1993; 158:131–144. [PubMed: 8101171]
109. Singhal N, Graumann J, Wu G, Araúzo-Bravo MJ, Han DW, et al. Chromatin-remodeling components of the BAF complex facilitate reprogramming. Cell. 2010; 141:943–955. [PubMed: 20550931]
110. Smallwood A, Esteve PO, Pradhan S, Carey M. Functional cooperation between HP1 and DNMT1 mediates gene silencing. Genes Dev. 2007; 21:1169–1178. [PubMed: 17470536]
111. Solter D. From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research. Nat. Rev. Genet. 2006; 7:319–327. [PubMed: 16534514]
112. Stadtfeld M, Apostolou E, Akutsu H, Fukuda A, Follett P, et al. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. Nature. 2010; 465:175–181. [PubMed: 20418860]
113. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000; 403:41–45. [PubMed: 10638745]
114. Strogatz, SH. Nonlinear Dynamics and Chaos: With Applications to Physics, Chemistry, Biology and Engineering. Cambridge, MA: Westview; 2000.
115. Struhl K. Histone acetylation and transcriptional regulatory mechanisms. Genes Dev. 1998; 12:599–606. [PubMed: 9499396]
116. Strutt H, Cavalli G, Paro R. Co-localization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression. EMBO J. 1997; 16:3621–3632. [PubMed: 9218803]
117. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009; 324:930–935. [PubMed: 19372391]
118. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006; 126:663–676. [PubMed: 16904174]
119. Tamaru H, Selker EU. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. Nature. 2001; 414:277–283. [PubMed: 11713521]
120. Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature. 2008; 455:1124–1128. [PubMed: 18806776]
121. Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, et al. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature. 2007; 448:196–199. [PubMed: 17597760]
122. Thanos D, Maniatis T. Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. Cell. 1995; 83:1091–1100. [PubMed: 8548797]
123. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, et al. Embryonic stem cell lines derived from human blastocysts. Science. 1998; 282:1145–1147. [PubMed: 9804556]

124. Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M, et al. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells. 2006; 11:805–814. [PubMed: 16824199]
125. Van Den Berg DLC, Snoek T, Mullin NP, Yates A, Bezstarosti K, et al. An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell. 2010; 6:369–381. [PubMed: 20362541]
126. Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, et al. Chromatin signature of embryonic pluripotency is established during genome activation. Nature. 2010; 464:922–926. [PubMed: 20336069]
127. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, et al. The Polycomb group protein EZH2 directly controls DNA methylation. Nature. 2006; 439:871–874. [PubMed: 16357870]
128. Voncken JW, Roelen BA, Roefs M, de Vries S, Verhoeven E, et al. Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proc. Natl. Acad. Sci. USA. 2003; 100:2468–2473. [PubMed: 12589020]
129. Walker E, Chang WY, Hunkapiller J, Cagney G, Garcha K, et al. Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation. Cell Stem Cell. 2010; 6:153–166. [PubMed: 20144788]
130. Wang J, Rao S, Chu J, Shen X, Levasseur DN, et al. A protein interaction network for pluripotency of embryonic stem cells. Nature. 2006; 444:364–368. [PubMed: 17093407]
131. Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat. Genet. 2007; 39:380–385. [PubMed: 17259983]
132. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. 2007; 39:457–466. [PubMed: 17334365]
133. Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, et al. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. Nat. Cell Biol. 2009; 11:1487–1495. [PubMed: 19935649]
134. Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat. Genet. 2009; 41:246–250. [PubMed: 19151716]
135. Xie D, Chen CC, Ptaszek LM, Xiao S, Cao X, et al. Rewirable gene regulatory networks in the preimplantation embryonic development of three mammalian species. Genome Res. 2010; 20:804–815. [PubMed: 20219939]
136. Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell. 2009; 137:647–658. [PubMed: 19409607]
137. Xu RH, Sampsell-Barron TL, Gu F, Root S, Peck RM, et al. NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. Cell Stem Cell. 2008; 3:196–206. [PubMed: 18682241]
138. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, et al. The ground state of embryonic stem cell self-renewal. Nature. 2008; 453:519–523. [PubMed: 18497825]
139. Yuan H, Corbi N, Basilico C, Dailey L. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes Dev. 1995; 9:2635–2645. [PubMed: 7590241]
140. Zhao XD, Han X, Chew JL, Liu J, Chiu KP, et al. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell. 2007; 1:286–298. [PubMed: 18371363]

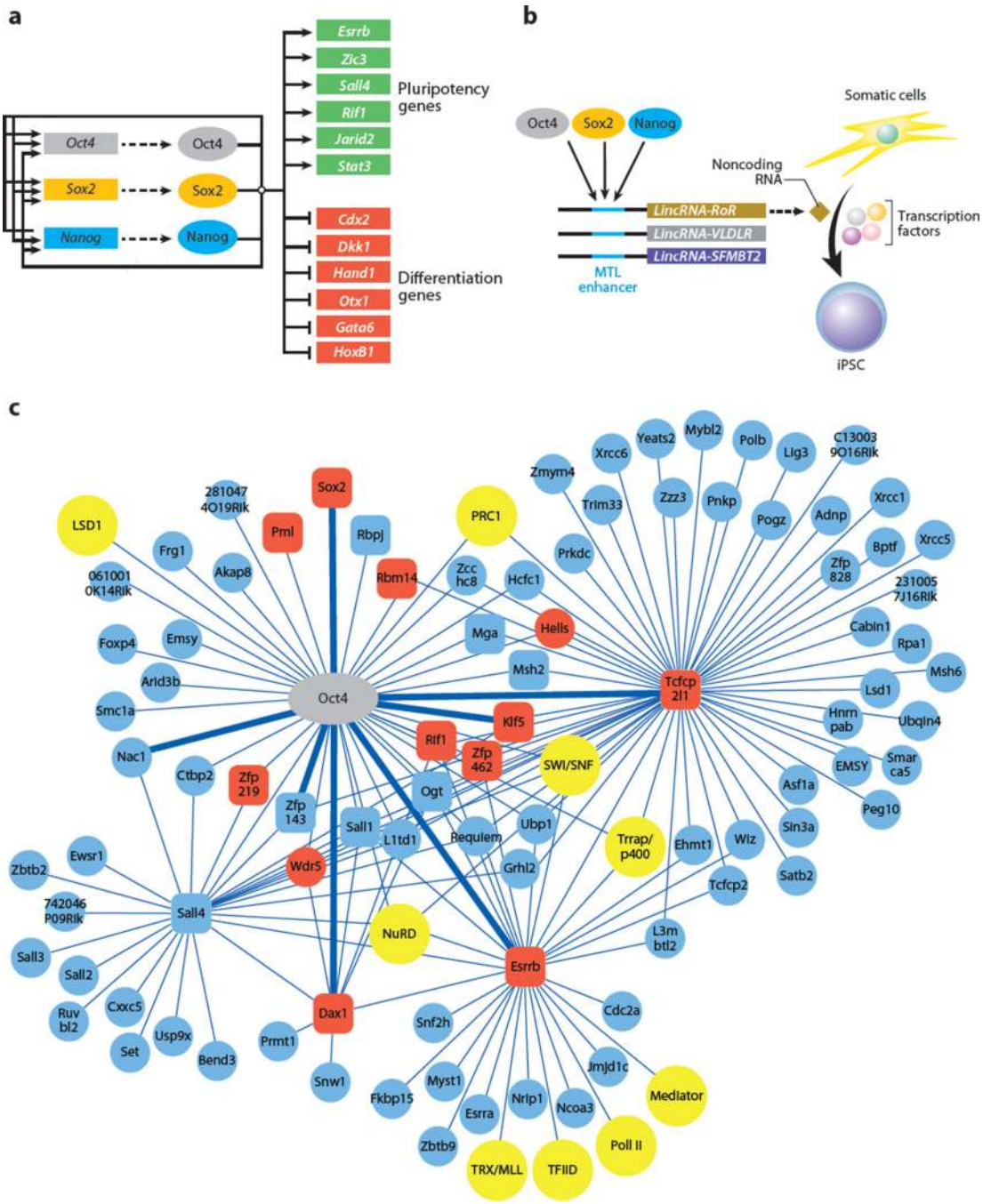

Figure 1.

Transcriptional circuitry that maintains pluripotency. (*a*) The key transcription factors of pluripotency form positive reciprocal and autoregulatory loops that maintain the expression of *Oct4, Sox2*, and *Nanog*. The key factors also synergistically co-occupy numerous downstream target genes that promote self-renewal and maintain pluripotency, while repressing developmentally regulated genes that drive differentiation. Transcription factors are represented by ovals, and the genes are represented by rectangles. (*b*) OCT4, SOX2, and NANOG co-occupy the multiple transcription-factor-binding loci (MTL) enhancer and positively regulate numerous noncoding RNAs in human induced pluripotent stem cells (iPSCs). Large intergenic noncoding RNA–regulator of reprogramming (LincRNA-RoR), a

downstream effector of the core network, is important for establishing pluripotency during iPSC reprogramming. (*c*) An Oct4-centered network in ESCs. A schematic network, constructed based on Reference 125, consists of Oct4-interacting proteins and interacting partners of Oct4-associated proteins. Complexes consisting of several protein subunits are indicated by large yellow circles. The Oct4 interactome was further wired to the transcription regulatory network through integration of data sets from microarray profiling and transcription factor binding. Rectangular nodes represent genes that are bound by Oct4 as reported by previous ChIP-on-chip or ChIP-seq studies (18, 52). Red indicates functional regulation, as the expressions of respective genes were repressed with reduced levels of Oct4 in ZHBTc4 ESCs (104). Thick blue lines connect Oct4 with transcription factors that synergistically co-occupy downstream target genes with Oct4.

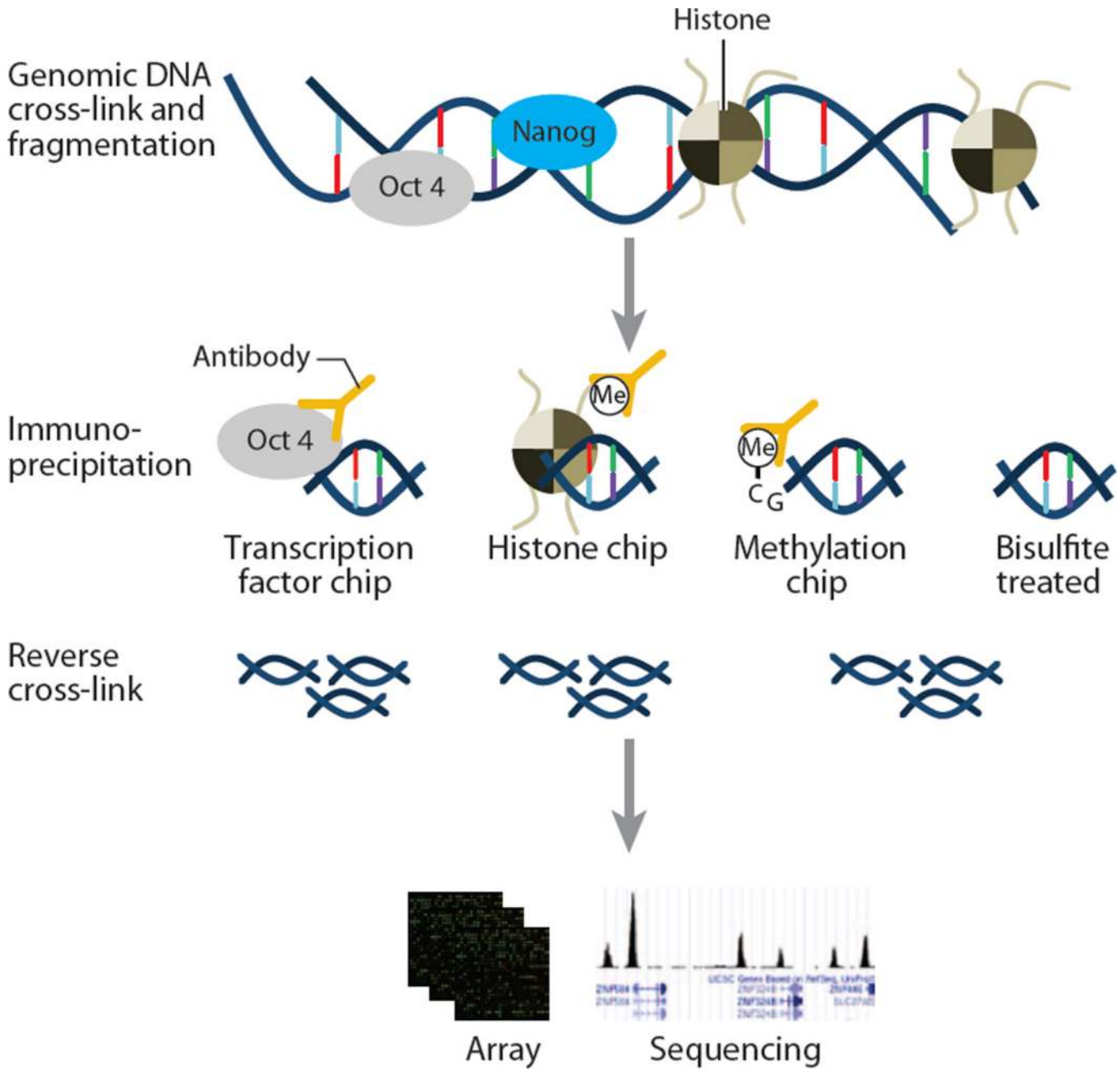

Figure 2.

Genomic-wide mapping of protein-DNA interactions, histone modifications, and DNA methylation. Transcriptional networks and characteristics of the epigenome such as histone modifications and DNA methylation can be uncovered by highly condensed microarray chips or by next-generation sequencing technologies, respectively.

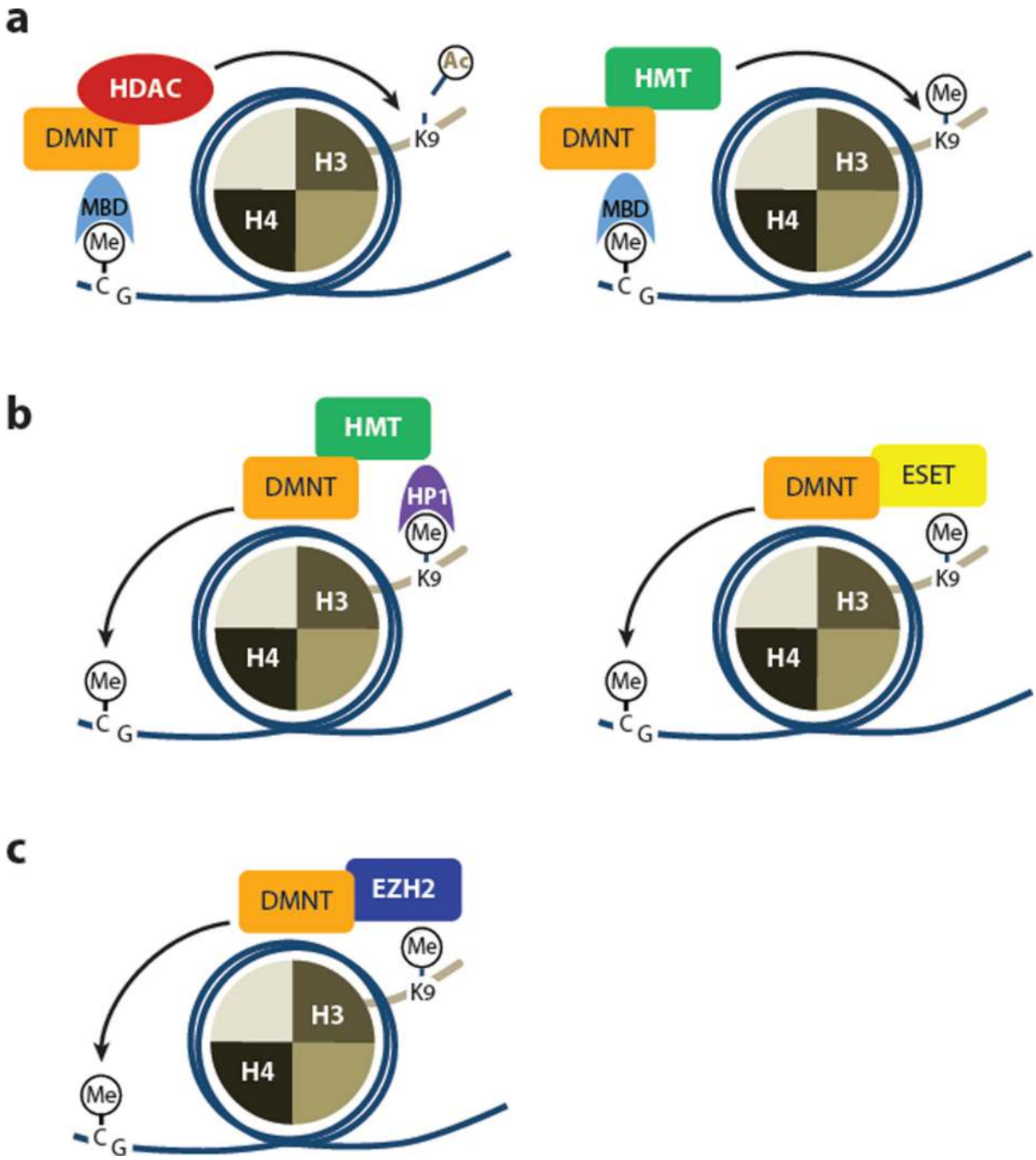

Figure 3.

Crosstalk between histone and DNA modifications. (*a*) DNA methylation can direct either acetylation or H3K9 methylation. DNA methyltransferases (DNMTs) have been found to associate with both histone deacetylases (HDACs, *left*) as well as H3K9 methyltransferases (*right*—e.g., G9a). DNA methyl binding proteins (e.g., MECP2) associate with both HDACs and H3K9 methyltransferases (e.g., ESET). (*b*) H3K9 methyltransferases can direct DNA methylation. The SUV39H1/2 and G9a histone methyltransferases (HMTs), when complexed with HP1 (adapter), can recruit DNMTs (*left*). ESET can complex with DNMTs directly (*right*). (*c*) H3K27 methyltransferase EZH2 can direct DNA methylation. DNMTs have been found to bind directly to EZH2.