



Published in final edited form as:

Cell. 2011 March 18; 144(6): 940–954. doi:10.1016/j.cell.2011.01.032.

Control of Embryonic Stem Cell State

Richard A. Young

Whitehead Institute for Biomedical Research, Cambridge, MA 02142

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142

Abstract

Embryonic stem cells and induced pluripotent stem cells, which can be propagated in culture in an undifferentiated state but induced to differentiate into specialized cell types, hold great promise for regenerative medicine. Moreover, these cells provide a powerful model system for studies of cellular identity and early mammalian development. Insights into the transcriptional control of embryonic stem cell state, including the regulatory circuitry underlying pluripotency and cellular reprogramming, have emerged from the study of these cells. These studies have also revealed fundamental mechanisms that control vertebrate gene expression, connect gene expression to chromosome structure, and contribute to human disease.

Introduction

Embryonic stem cells (ESCs) are pluripotent, self-renewing cells that are derived from the inner cell mass (ICM) of the developing blastocyst. Pluripotency is the capacity of a single cell to generate all cell lineages of the developing and adult organism. Self-renewal is the ability of a cell to proliferate in the same state. The molecular mechanisms that control ESC pluripotency and self-renewal are important to discover because they are key to understanding development. Because defects in development cause many different diseases, improved understanding of control mechanisms in pluripotent cells may lead to new therapies for these diseases.

ESCs have a gene expression program that allows them to self-renew yet remain poised to differentiate into essentially all cell types in response to developmental cues. Recent reviews have discussed developmental potency (Rossant, 2008), the nature of the pluripotent ground state of ESCs (Silva and Smith, 2008), ESC transcriptional regulatory circuitry (Chen et al., 2008a; Jaenisch and Young, 2008; Orkin et al., 2008) and cellular reprogramming into ESC-like states (Yamanaka and Blau, 2010). This review provides a synthesis of key concepts that explain how pluripotency and self-renewal are controlled transcriptionally. These concepts have emerged from genetic, biochemical and molecular studies of the transcription factors, cofactors, chromatin regulators and noncoding RNAs that control the ESC gene expression program.

The regulators of gene expression programs can participate in gene activation, establish a poised state for gene activation in response to developmental cues, or contribute to gene silencing (Figure 1). The molecular mechanisms by which these regulators generally

© 2011 Elsevier Inc. All rights reserved

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

participate in control of gene expression are the subject of other reviews (Bartel, 2009; Fuda et al., 2009; Ho and Crabtree, 2010; Li et al., 2007; Roeder, 2005; Surface et al., 2010; Taatjes, 2010). I describe here the regulators that have been implicated in control of ESC state and discuss how they contribute to pluripotency and self-renewal.

Transcription Factors and Regulatory Circuitry

Transcription factors recognize specific DNA sequences and either activate or prevent transcription. Early studies into the transcriptional control of the *E. coli lac* operon created the framework for understanding gene control (Jacob and Monod, 1961). In the absence of lactose, the lac operon is repressed by the Lac repressor, which binds the *lac* operator and inhibits transcription by RNA polymerase. In the presence of lactose, the Lac repressor is lost and gene expression is activated by a transcription-activating factor that binds a nearby site and recruits RNA polymerase. The fundamental concept that emerged from these studies - that gene control relies on specific repressors and activators and the DNA sequence elements they recognize - continues to provide the foundation for understanding control of gene expression in all organisms.

In mammals, transcription factors make up the largest single class of proteins encoded in the genome, representing approximately 10% of all protein-coding genes (Levine and Tjian, 2003). Transcription factors bind to both promoter-proximal DNA elements and to more distal regions that can be nearby or 100s of kb away. These distal elements that are involved in positive gene regulation are called enhancers, and these elements are generally bound by multiple transcription factors. Transcription factors can activate gene expression by recruiting the transcription apparatus and/or by stimulating release of RNA polymerase II from pause sites (Fuda et al., 2009). They can also recruit various chromatin regulators to promoter regions to modify and mobilize nucleosomes in order to increase access to local DNA sequences (Li et al., 2007).

In ESCs, the pluripotent state is largely governed by the core transcription factors Oct4, Sox2 and Nanog (Table 1) (Chambers and Smith, 2004; Niwa, 2007; Silva and Smith, 2008). Oct4 and Nanog were identified as key regulators based on their relatively unique expression pattern in ESCs and genetic experiments showing that they are essential for establishing or maintaining a robust pluripotent state (Chambers et al., 2003; Chambers and Smith, 2004; Mitsui et al., 2003; Nichols et al., 1998; Niwa et al., 2000). Oct4 functions as a heterodimer with Sox2 in ESCs, thus placing Sox2 among the key regulators (Ambrosetti et al., 2000; Avilion et al., 2003; Masui et al., 2007). Reprogramming of somatic cells into induced pluripotent stem (iPS) cells generally requires forced expression of Oct4 and Sox2, unless endogenous Sox2 is expressed in the somatic cell, consistent with the view that Oct4/Sox2 are key to establishing the ESC state. Although ESCs can be propagated in the absence of Nanog (Chambers et al., 2007), Nanog promotes a stable undifferentiated ESC state (Chambers et al., 2007), is necessary for pluripotency to develop in ICM cells (Silva et al., 2009), and co-occupies most sites with Oct4 and Sox2 throughout the ESC genome (Marson et al., 2008b), so it is included here as a component of the core regulatory circuitry.

Core regulatory circuitry

Two key concepts dominate our understanding of the function of the core transcription factors Oct4, Sox2 and Nanog in control of ESC state (Figure 2): 1) The core transcription factors function together to positively regulate their own promoters, forming an interconnected autoregulatory loop. 2) The core factors co-occupy and activate expression of genes necessary to maintain ESC state, while contributing to repression of genes encoding lineage-specific transcription factors whose absence helps prevent exit from the pluripotent state.

The interconnected autoregulatory loop formed by Oct4, Sox2 and Nanog generates a bi-stable state for ESCs: residence in a positive-feedback-controlled gene expression program when the factors are expressed at appropriate levels, versus entrance into a differentiation program when any one of the master transcription factors is no longer functionally available (Boyer et al., 2005; Loh et al., 2006). This regulatory circuit likely explains the ability to jump-start the ESC gene expression program during reprogramming by forced expression of reprogramming factors (Jaenisch and Young, 2008). Thus, the ectopically expressed factors activate transcription of the endogenous *Pou5f* (*Oct4*), *Sox2* and *Nanog* genes and thereby initiate the positive-feedback loop that sustains ongoing production of these factors from the endogenous genes in the absence of further input from the ectopically expressed factors. Some factors present in reprogramming cocktails, such as c-Myc, appear to facilitate activation of this interconnected autoregulatory circuitry by stimulating gene expression and proliferation more generally (Rahl et al., 2010).

Oct4, Sox2 and Nanog collaborate to activate a substantial fraction of the actively transcribed protein-coding and miRNA genes in ESCs (Chen et al., 2008b; Marson et al., 2008b). Sites co-occupied by the three core regulators generally have enhancer activity, and transcription of genes adjacent to such sites often depends on at least one of the trio (Chen et al., 2008b; Chew et al., 2005; Matoba et al., 2006). Oct4 and Nanog can bind and recruit multiple coactivators, as described below, accounting for their ability to activate genes.

Oct4, Sox2 and Nanog also occupy repressed genes encoding cell lineage-specific regulators, and the repression of these genes is essential for ESCs to maintain a stable pluripotent state and to undergo normal differentiation (Bilodeau et al., 2009; Boyer et al., 2005; Boyer et al., 2006; Lee et al., 2006; Loh et al., 2006; Marson et al., 2008b; Pasini et al., 2008; Pasini et al., 2004). The loss of these core regulators leads to rapid induction of a wide spectrum of genes encoding lineage-specific regulators, indicating that these genes are poised for activation. How might Oct4/Sox2 and Nanog act to repress these genes? The SetDB1 and Polycomb Group (PcG) chromatin regulators have been implicated in repression of these lineage-specific genes. Oct4 can bind sumoylated SetDB1, which catalyzes the repressive histone modification H3K9me3 at these genes (Yeap et al., 2009; Yuan et al., 2009). Polycomb Group (PcG) complexes can associate with nucleosomes with histone H3K9me3 (Margueron et al., 2009) and further contribute to repression through mechanisms described below. It is also possible that Oct4, Sox2 and Nanog activate transcription initiation in the extensive GC-rich promoter regions of these genes, and GC-rich RNA species produced from these regions might then recruit or stabilize Polycomb Group (PcG) complexes (Guenther and Young, 2010).

The ability of Oct4, Sox2 and Nanog to positively regulate genes necessary to maintain ESC state while repressing genes that would enable egress from this state explains, in part, the ability of ESCs to self-renew in an undifferentiated state, yet remain poised to differentiate into all cell types of the body in response to developmental cues. However, the factors and mechanisms that specify positive versus negative regulation by Oct4/Sox2/Nanog are not fully understood. Additional transcription factors are known to collaborate with Oct4, Sox2 and Nanog to regulate the ESC gene expression program, and these make various contributions to selective activation and repression of target genes, as described below.

Control of RNA polymerase II

Transcription factors control at least two major steps in gene expression (Fuda et al., 2009; Peterlin and Price, 2006; Rahl et al., 2010). Some transcription factors recruit RNA polymerase II to promoters, where the enzyme typically transcribes a short distance (approximately 35 bp), and then pauses or terminates. Other transcription factors recruit a cyclin-dependent kinase (Cdk9/cyclinT) called p-TEFb, which phosphorylates the

polymerase and its associated pause control factors, allowing the enzyme to be released from pause sites and fully transcribe the gene. Oct4, Sox2 and Nanog interact with coactivators that bind to RNA polymerase II (Kagey et al., 2010), so the core regulators are involved in RNA polymerase II recruitment. In contrast, c-Myc, which plays important roles in ESC proliferation and self-renewal (Cartwright et al., 2005), does not appear to play an important role in RNA polymerase II recruitment, but rather binds to E-box sequences at core promoter sites and recruits p-TEFb, thus stimulating pause release (Rahl et al., 2010). A large proportion of the actively transcribed genes in ESCs are bound and regulated by both the core transcription factors and c-Myc (Figure 3A). Thus, Oct4/Sox2/Nanog apparently play dominant roles in selecting the set of ESC genes that will be actively transcribed and recruiting RNA polymerase II to these genes, while c-Myc regulates the efficiency with which these selected genes are fully transcribed. This likely explains why forced expression of c-Myc can enhance reprogramming efficiency and why this transcription factor plays such a potent role in proliferation of many cancer cells (Jaenisch and Young, 2008; Rahl et al., 2010).

Multiple enhancers and enhanceosomes

Enhancers are generally bound by multiple transcription factors, forming large nucleoprotein complexes called enhanceosomes, which permit cooperative binding between transcription factors and allow for synergistic and combinatorial effects on gene regulation (Maniatis et al., 1998). The cooperative interactions among transcription factors binding to adjacent DNA sites and to cofactor complexes explains why multiple transcription factors are found together in the genome and why transcription factors bind stably to only a small subset of the millions of DNA sequence motifs present in the vertebrate genome. Many genes have multiple enhancers and thus multiple enhanceosomes (Levine and Tjian, 2003). In *Drosophila*, these multiple, seemingly redundant enhancers have been shown to contribute to phenotypic robustness during embryonic development (Frankel et al., 2010; Hong et al., 2008). That is, normal levels of gene expression are obtained despite environmental and genetic variability so long as genes are equipped with multiple enhancers.

In addition to Oct4, Sox2, Nanog and c-Myc, the transcription factors Tcf3, Smad1, Stat3, Esrrb, Sall4, Tbx3, Zfx, Ronin, Klf2, Klf4, Klf5 and PRDM14 have been shown to play important roles in control of ESC state (Table 1). The ChIP-Seq data that has been obtained for these transcription factors indicate that they can bind to loci occupied by Oct4/Sox2/Nanog as well as other loci (Figure 3B), forming sites that have been called multiple transcription factor binding loci (MTL) (Chen et al., 2008b; Kim et al., 2008). Several lines of evidence indicate that most MTLs are enhancers. Most MTLs are occupied by the p300 cofactor, and the subset of MTLs that are occupied by Oct4/Sox2/Nanog are also occupied by the mediator cofactor (Chen et al., 2008b; Kagey et al., 2010). All Oct4/Sox2/Nanog-containing MTLs tested to date have been shown to exhibit enhancer activity (Chen et al., 2008b). It is therefore likely that functional enhanceosomes are formed at most MTLs.

The evidence obtained thus far suggests that most Oct4/Sox2/Nanog-regulated genes are co-occupied by one or more of the other transcription factors implicated in control of ESCs (Figure 3A–C). Examination of gene tracks for the *Max* gene reveals a typical pattern, where the promoter region contains a site bound by Oct4, Sox2, Nanog, Tcf3 and Esrrb, and various other sites occupied by c-Myc, Zfx, Ronin and Klf4 (Figure 3D). Thus the functions of the core regulators (Oct4, Sox2 and Nanog) are augmented by the functions many of the other transcription factors implicated in control of ESC state at actively transcribed target genes.

Signaling to the core regulatory circuitry

A simple concept has emerged from the study of transcription factors associated with signaling pathways that have roles in control of ESC state: these particular factors tend to co-occupy enhancers bound by Oct4, Sox2 and Nanog, thereby allowing direct control of genes within the core circuitry by these signaling pathways (Figure 4) (Chen et al., 2008b; Cole et al., 2008; Tam et al., 2008). ESCs were initially cultured on a layer of irradiated fibroblasts in order to obtain the necessary factors for self-renewal and pluripotency (Smith, 2001; Smith and Hooper, 1983). LIF, Wnt and ligands of the TGF- β /BMP signaling pathway were among factors supplied by the fibroblasts and found to influence murine ESC state (Okita and Yamanaka, 2006; Sato et al., 2004; Smith et al., 1998; Williams et al., 1988; Ying et al., 2003). Transcription factors associated with LIF, Wnt and BMP4 signaling pathways (Stat3, Tcf3 and Smad1) bind to loci bound by Oct4, Sox2 and Nanog (Chen et al., 2008a; Cole et al., 2008; Tam et al., 2008; Wu et al., 2006; Zhang et al., 2006). Thus, signals mediated by these pathways are delivered directly to the enhancers of genes within the core regulatory circuitry and can thereby have profound effects on pluripotency and self-renewal. This likely explains why manipulation of certain signaling pathways can enhance reprogramming (Lluis et al., 2008; Marson et al., 2008a).

Transcriptional Cofactors

Cofactors include protein complexes that contribute to activation (coactivators) and repression (co-repressors) but do not have DNA-binding properties of their own. Some cofactors mobilize or modify nucleosomes and in these cases they are also considered chromatin regulators. Cofactors are generally expressed in most cell types, but it appears that ESCs are more sensitive than somatic cells to reduced levels of certain cofactors and chromatin regulators (Fazio and Panning, 2010; Kagey et al., 2010).

Transcription factors that occupy active enhancers bind coactivators such as p300 and mediator, which in turn bind and control the activity of the transcription initiation apparatus (Conaway et al., 2005; Malik and Roeder, 2005; Roeder, 1998; Taatjes, 2010). The p300 cofactor occupies most active promoters in ESCs (Chen et al., 2008b). Reduced levels of p300 do not appear to adversely affect ESCs, but rather have a profound effect on ESC differentiation (Chen et al., 2008b; Zhong and Jin, 2009).

Recent studies have shown that mediator physically links Oct4/Sox2/Nanog-bound enhancers to the promoters of active genes in the core regulatory circuitry of ESCs (Figure 5)(Kagey et al., 2010). The form of mediator recruited to the Oct4/Sox2/Nanog-regulated promoters associates with the cohesin loading factor Nipbl, which provides a mechanism for cohesin loading at these sites. The mediator/cohesin complex forms a looped chromosome architecture between enhancers and core promoters that is necessary for normal gene activity. Mediator and cohesin co-occupy different promoters in different cell types, thus generating cell-type-specific DNA loops associated with the gene expression program of each cell.

The CDK8 kinase subunit of the mediator complex can influence the activity of signaling transcription factors (Alarcon et al., 2009; Fryer et al., 2004; Gao et al., 2009; Taatjes, 2010). For example, CDK8 mediated phosphorylation of the linker region within SMAD1/5 or SMAD2/3 complexes can activate these transcription factor complexes, but it also targets them for proteasomal degradation. A dynamic cycle of transcription factor activation and destruction ensures that continuous pathway activation is necessary for continuous gene activation and may facilitate rapid changes in cell state when signaling is altered.

ESCs are also sensitive to changes in the levels of the PAF1 complex, which is associated with RNA polymerase II at active genes (Ding et al., 2009). Based on studies in yeast, the Paf1 complex couples transcription initiation and elongation with histone H3K4 and H3K36 methylation (Krogan et al., 2003). In ESCs, the Paf1 complex may also play this role, as knockdowns lead to reduced levels of histone H3K4me3 at actively transcribed genes (Ding et al., 2009).

Corepressors that have been implicated in control of ES cell state include Dax1, Cnot3 and Trim28 (Hu et al., 2009; Sun et al., 2009). Overexpression of Dax1 causes ESC differentiation, likely due to an inhibitory interaction with Oct4 (Sun et al., 2009). Cnot3 and Trim28 co-occupy many promoters with c-Myc and Zfx, and likely contribute to control of proliferation and self-renewal. They differ somewhat in the additional promoters they occupy, which might explain why loss of Cnot causes ESCs to differentiate into trophectoderm while loss of Trim28 causes cells to differentiate into primitive ectoderm lineage (Hu et al., 2009).

Chromatin Regulators in ESC Gene Activity and Silencing

Eukaryotic genomes are packaged into nucleosomes (Kornberg and Thomas, 1974; Olins and Olins, 1974), which provide a means to compact the genome and to influence gene expression. Early studies showed that nucleosomes can affect transcription in vitro (Knezetic and Luse, 1986; Lorch et al., 1987) and in vivo (Han and Grunstein, 1988; Kayne et al., 1988). Subsequent studies revealed that gene expression can be influenced by proteins that modify histones (Brownell et al., 1996) or mobilize nucleosomes (Cote et al., 1994; Imbalzano et al., 1994; Kwon et al., 1994). Chromatin regulators are generally recruited to genes by DNA binding transcription factors, the transcription apparatus, or specific RNA species (Guenther and Young, 2010; Li et al., 2007; Roeder, 2005; Surface et al. 2010). The chromatin regulators that are known to contribute to the control of ESC state fall into two classes: histone-modifying enzymes and ATP-dependent chromatin remodeling complexes (Table 1).

Histone-Modifying Enzymes

The chromatin regulators known to have the most profound impact on ESC state are histone-modifying enzymes that repress genes encoding lineage-specific developmental regulators. These include the Polycomb group (PcG) protein complexes, SetDB1 and TIP60-p400. Polycomb group (PcG) and Trithorax group (TrxG) genes were discovered in *Drosophila melanogaster* as repressors and activators of Hox genes (Schuettengruber et al., 2007). TrxG proteins catalyze trimethylation of histone H3 lysine 4 (H3K4me3) at the promoters of active genes and facilitate maintenance of active gene states during development, in part by antagonizing the functions of PcG proteins. PcG protein complexes catalyze ubiquitylation of histone H2A lysine 119 (H2AK119u) and trimethylation of histone H3 lysine 27 (H3K27me3) and function in ESCs to help silence genes encoding key regulators of development, yet allow them to remain in a state that is “poised” for activation during differentiation (Azuara et al., 2006; Bernstein et al., 2006; Boyer et al., 2006; Bracken et al., 2006; Endoh et al., 2008; Lee et al., 2006; Li et al., 2010; Pan et al., 2007; Pasini et al., 2010; Peng et al., 2009; Shen et al., 2009; van der Stoep et al., 2008). PcG proteins are thought to inhibit transcription, at least in part, by restraining poised RNA polymerase molecules (Stock et al., 2007; Zhou et al., 2008b). ESCs lacking PcG protein complexes are unstable and tend to differentiate, but fail to execute differentiation programs appropriately (Leeb et al., 2010).

Multiple histone H3 lysine 9 methyltransferases have been implicated in control of ESC state (Bilodeau et al., 2009; Yeap et al., 2009; Yuan et al., 2009). A subset of the silent

genes that encode lineage-specific developmental regulators, including those involved in generating the extraembryonic trophoblast lineage, are occupied and repressed by SetDB1, which catalyzes methylation of histone H3 lysine 9. Thus, multiple repressive mechanisms, involving methylation of H3K27 and H3K9 and ubiquitylation of histone H2A, are used to silence genes encoding lineage-specific developmental regulators.

The Tip60-p400 complex has multiple activities, among which is histone acetylation, and loss of this complex affects ESC morphology and state (Fazzio et al., 2008). It is found associated with active promoters in ESCs and appears to be recruited in two ways, directly by the H3K4me3 mark and indirectly by Nanog. Interestingly, the complex is also associated with nucleosomes with H3K4me3 at PcG-occupied genes encoding lineage-specific regulators, where it apparently facilitates repression of these poised genes. Because Tip60-p400 is generally found associated with active genes, its repressive function may derive from its potential role in facilitating transcription of ncRNAs that recruit or stabilize PcG complexes, as described below.

ATP-dependent nucleosome remodeling

ATP-dependent nucleosome remodeling complexes can be recruited by transcription factors and modified histones to the promoters of genes, where they enhance or reduce the access of transcriptional components to DNA sequences with resulting positive or negative effects on gene activity (Clapier and Cairns, 2009; Ho and Crabtree, 2010). Components of multiple ATP-dependent nucleosome-remodeling complexes have been implicated in control of ESC state (Table 1)(Bilodeau et al., 2009; Gaspar-Maia et al., 2009; Ho and Crabtree, 2010; Klochendler-Yeivin et al., 2000; Schnetz et al.). A complex purified from ESCs called esBAF has been shown to be associated with the promoters of genes under the control of the core regulators Oct4, Sox2 and Nanog, and core subunits of this complex are essential for ESC maintenance (Ho and Crabtree, 2010). Chd1, a member of the chromodomain helicase DNA-binding (CHD) family of ATP-dependent chromatin remodelers, is associated with the promoters of active genes and Chd1-deficient ESCs are incapable of giving rise to primitive endoderm (Gaspar-Maia et al., 2009). Another member of the CHD family, Chd7, is associated with active Oct4/Sox2/Nanog-bound enhancers in ES cells, where it is thought to fine-tune the expression levels of ESC-specific genes (Schnetz et al., 2010). Unlike mutations in esBAF and Chd1, which affect ESC state, the effects of changing Chd7 dosage are subtle and do not appear to affect pluripotency or self-renewal. Thus, multiple ATP-dependent nucleosome remodeling complexes are present at many key ESC genes.

Non-coding RNAs in ESC Regulatory Circuitry

The idea that noncoding RNA (ncRNA) might regulate genes was proposed at the dawn of studies on regulation of gene expression (Britten and Davidson, 1969; Jacob and Monod, 1961). It is now clear that noncoding RNA is involved in regulation of many important biological processes, including X inactivation, dosage compensation, imprinting, polycomb repression and silencing of repeated elements, as described in several recent reviews (Surface et al., 2010; Wilusz et al., 2009; Zariatigui et al., 2007). Indeed, a variety of noncoding RNA species have been implicated in control of ESC state (Table 1). These include miRNAs, which can regulate the stability and translatability of mRNAs and, acting in this fashion, are essential for normal ESC self-renewal and cellular differentiation. They also include longer noncoding RNAs of various types, which have been implicated in recruitment of chromatin regulators such as the PcG complexes (Bracken and Helin, 2009; Guenther and Young, 2010; Surface et al., 2010; Wilusz et al., 2009).

miRNAs and control of ESC identity

Multiple lines of evidence indicate that miRNAs contribute to the control of early development. ESCs deficient in miRNA-processing enzymes show defects in differentiation and proliferation (Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2007). Two key themes have emerged from studying the regulation of miRNA genes in ESCs (Marson et al., 2008b). First, the core regulators Oct4/Sox2/Nanog activate genes for miRNAs that are preferentially expressed in ESCs, and these miRNAs contribute to cell state maintenance and cell state transitions by fine-tuning the expression of key ESC genes and by promoting the rapid clearance of ESC transcripts during differentiation. Second, the core regulators co-occupy repressed lineage-specific miRNA genes with SteDB1 and PcG complexes, thus poising them for expression during differentiation.

The core circuitry controls the expression of miRNAs that fine-tune the expression of key transcripts and promote the rapid clearance of ESC-specific transcripts during differentiation (Figure 6). Several miRNA polycistrons that specify the most abundant miRNAs in ESCs and which are silenced during early differentiation are positively regulated by Oct4/Sox2/Nanog (Marson et al., 2008b). These include the mir-290–295 cluster, and miRNAs with seed sequences in this family have been implicated in cell proliferation (He et al., 2005; O'Donnell et al., 2005; Wang et al., 2008), and have been shown to rescue the proliferation defects observed in miRNA-deficient ES cells (Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2008; Wang et al., 2007). Furthermore, the zebrafish homolog of this miRNA family, miR-430, contributes to the rapid degradation of maternal transcripts in early zygotic development (Giraldez et al., 2006), and this miRNA family also promotes the clearance of transcripts in early mammalian development (Farh et al., 2005).

The core transcription factors and PcG complexes co-occupy genes for miRNAs that are repressed in ESCs, but become selectively expressed in cells of the immune system (mir-155), pancreatic islets (mir-375), neural cells (mir-124 and mir-9), and differentiating ESCs (mir-296) (Figure 6A)(Marson et al., 2008b). This set of miRNA genes is thus poised to contribute to cell-fate decisions during development in the same fashion as genes encoding lineage-specific transcription factors that are co-occupied by the core regulators and PcG complexes. For example, mir-296, which is rapidly induced upon ESC differentiation, targets Nanog mRNA (Tay et al., 2008).

ncRNAs and Polycomb-mediated Silencing

Recent studies indicate that various ncRNA molecules recruit or stabilize PcG complexes at specific sites in the genome. Specific ncRNA molecules have recently been shown to recruit the PcG complexes to the X-inactivation center X(ic), the *kcnq1* domain, the *INK4b/ARF/INK4a* locus, the *HOXD* locus and potentially many other genomic loci in various cell types (Gupta et al., 2010; Pandey et al., 2008; Rinn et al., 2007; Tsai et al., 2010; Yap et al., 2010; Zhao et al., 2008). Short ncRNAs are transcribed bi-directionally by RNA polymerase II from most CpG islands (Core et al., 2008; Guenther et al., 2007; He et al., 2008; Seila et al., 2008), and some of these are able to recruit PcG complexes (Guenther and Young, 2010; Kanhere et al., 2010; Surface et al., 2010). Most PcG-occupied genes in ESCs show some evidence of RNA polymerase II binding or transcripts, but it is challenging to detect low or transient levels of RNA species that serve to initiate Polycomb-mediated silencing at these loci. Thus, much remains to be understood about the role of ncRNAs in PcG complex recruitment.

Transitioning from ES to Specialized States

The ESC regulatory circuitry is reconfigured when cells are stimulated to differentiate (Figure 6). The mechanisms that contribute to the loss of ESC state include the silencing of

the Oct4 gene (Ben-Shushan et al., 1995; Feldman et al., 2006), the proteolytic destruction of Nanog by caspase-3 (Fujita et al., 2008), and miRNA-mediated reduction in Oct4, Nanog and Sox2 mRNA levels (Tay et al., 2008). The mechanisms that influence exit from the ESC state and entrance into new states include the activation of selected cell lineage-specific regulatory genes, downregulation of the miRNA regulator Lin28 with consequent maturation of the Let-7 miRNA (Viswanathan et al., 2008), activation of pioneer transcription factors that have occupied cell type-specific enhancers silently and thus without active promotion of transcription (Zaret et al., 2008), and modifications to the subunit composition of mediator, BAF and TFIID complexes (Deato et al., 2008; Deato and Tjian, 2007; Ho and Crabtree, 2010; Taatjes, 2010).

Insights into Disease Mechanisms

The study of ESC control has provided new insights into mechanisms that are involved in several human diseases. For example, improved understanding of the functions of transcription factors such as c-Myc, cofactor complexes such as mediator and cohesin, and chromatin regulators such as TrxG and PcG has provided new insights into the molecular pathways affected by mutations in these regulators.

Key aspects of the ESC gene expression program are recapitulated in cancer cells (Ben-Porath et al., 2008), and it has been argued that this is largely a consequence of c-Myc (Kim et al., 2010). c-Myc amplification is the most frequent somatic copy-number amplification in tumor cells (Beroukhi et al., 2010). Tumor cells that overexpress c-Myc have enhanced expression of proliferation genes, and this is likely due to the role of c-Myc in recruiting P-TEFb to effect RNA polymerase II pause release at these genes (Rahl et al., 2010). This insight suggests that therapeutic agents that target control of transcription elongation may be valuable for treating tumors that overexpress c-Myc.

Mutations in the genes encoding mediator, cohesin and the cohesin loading factor Nipbl can cause an array of human developmental syndromes and diseases. Mediator mutations have been associated with Opitz-Kaveggia (FG) syndrome, Lujan syndrome, schizophrenia, Transposition of the Great Arteries (TGA) syndrome and colon cancer progression (Ding et al., 2008; Firestein et al., 2008; Muncke et al., 2003; Philibert and Madan, 2007; Risheg et al., 2007; Schwartz et al., 2007). Mutations in Nipbl and cohesin are responsible for most cases of Cornelia de Lange syndrome, which is characterized by developmental defects and mental retardation and appears to be the result of mis-regulation of gene expression rather than chromosome cohesion or mitotic abnormalities (Krantz et al., 2004; Strachan, 2005; Tonkin et al., 2004). Knowledge that mediator, Nipbl and cohesin are linked at active promoters suggests therapies that might compensate for partial loss of transcriptional activity. The CDK8 kinase resides within a subcomplex of mediator that has repressive activities (Knuesel et al., 2009; Taatjes, 2010), so it is conceivable that small molecule antagonists of CDK8 would lead to an increase in transcriptionally active mediator/cohesin assemblies.

Mutations that affect the functions or levels of TrxG and PcG chromatin regulators have been implicated in a variety of cancers (Bracken and Helin, 2009; Krivtsov and Armstrong, 2007). The study of these regulators in ESCs and in cancer cells has revealed how repression of lineage-specific transcription factors and cell cycle regulators may contribute to cancer phenotypes. Chromatin regulators with enzymatic activities are a new class of targets for small molecule drug discovery and we can expect new developments in this field in the near future.

Summary and Outlook

How do regulators of the ESC gene expression program produce a self-renewing cell capable of differentiating into all the cells of the adult? Part of the answer is that the core transcription factors establish autoregulatory loops that help maintain their own expression, activate transcription of a large fraction of the active genes, and contribute to the poised state of lineage-specific genes. The core transcription factors frequently share enhancers with signaling transcription factors, so signal transduction pathways can deliver signals directly to the genes regulated by the core factors. At actively transcribed genes, other transcription factors implicated in proliferation and other aspects of self-renewal bind to sites that can be separate from the core enhancers, and modulate RNA expression levels through mechanisms that include release of paused polymerases. The core factors help create a poised state by recruiting SetDB1 and PcG chromatin regulators to genes encoding lineage-specific factors through mechanisms that are not yet fully understood.

Several of the regulatory features of ESCs probably operate to control cell identity in other cell types. The idea that cells may rely on a small number of master transcription factors for control of cell state is supported by reprogramming and transdifferentiation experiments, and this concept warrants further study in additional cell types (Graf and Enver, 2009; Vierbuchen et al., 2010; Zhou et al., 2008a). Identification of the master transcription factors for all cell types would improve our understanding of cell identity and could facilitate cell-based therapies. Signaling pathways can transmit information to enhancers bound by master regulators in ESCs and it will be useful to know if this extends to other cell types. It will be interesting to determine whether master transcription factors generally recruit mediator and cohesin to cell-type specific enhancers and if they generally function to activate large populations of genes necessary for cell identity. Finally, it will be important to learn if the differentiation potential of cells is simply a function of the set of genes selected for transcriptional activity together with the set of genes encoding lineage-specific regulators that are selected for silencing.

Acknowledgments

Many ideas discussed in this review emerged from conversations with Steve Bilodeau, Laurie Boyer, Megan Cole, Joan and Ron Conaway, Jerry Crabtree, Job Dekker, David Gifford, Amanda Fisher, Garrett Frampton, Matthew Guenther, Kristian Helin, Rudolf Jaenisch, Richard Jenner, Michael Kagey, Tony Lee, Stuart Levine, Charles Lin, John Lis, Alexander Marson, Alan Mullen, Jamie Newman, Huck Ng, Stuart Orkin, David Orlando, Renato Paro, Peter Rahl, Peter Reddien, Robert Roeder, Phillip Sharp, Dylan Taatjes, Ken Zaret, Len Zon, Robert Weinberg and Thomas Zwaka. I am also grateful to David Orlando and Steve Bilodeau for help with data analysis and figures.

References

- Alarcon C, Zaromytidou AI, Xi Q, Gao S, Yu J, Fujisawa S, Barlas A, Miller AN, Manova-Todorova K, Macias MJ, et al. Nuclear CDKs drive Smad transcriptional activation and turnover in BMP and TGF-beta pathways. *Cell*. 2009; 139:757–769. [PubMed: 19914168]
- Ambrosetti DC, Scholer HR, Dailey L, Basilico C. Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. *J Biol Chem*. 2000; 275:23387–23397. [PubMed: 10801796]
- 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]
- Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, et al. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol*. 2006; 8:532–538. [PubMed: 16570078]

- Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009; 136:215–233. [PubMed: 19167326]
- Beattie GM, Lopez AD, Bucay N, Hinton A, Firpo MT, King CC, Hayek A. Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. *Stem Cells*. 2005; 23:489–495. [PubMed: 15790770]
- Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet*. 2008; 40:499–507. [PubMed: 18443585]
- Ben-Shushan E, Sharir H, Pikarsky E, Bergman Y. A dynamic balance between ARP-1/COUP-TFII, EAR-3/COUP-TFI, and retinoic acid receptor:retinoid X receptor heterodimers regulates Oct-3/4 expression in embryonal carcinoma cells. *Mol Cell Biol*. 1995; 15:1034–1048. [PubMed: 7823919]
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*. 2006; 125:315–326. [PubMed: 16630819]
- Beroukhi R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, Barretina J, Boehm JS, Dobson J, Urashima M, et al. The landscape of somatic copy-number alteration across human cancers. *Nature*. 2010; 463:899–905. [PubMed: 20164920]
- Bilodeau S, Kagey MH, Frampton GM, Rahl PB, Young RA. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev*. 2009; 23:2484–2489. [PubMed: 19884255]
- Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*. 2005; 122:947–956. [PubMed: 16153702]
- Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*. 2006; 441:349–353. [PubMed: 16625203]
- Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev*. 2006; 20:1123–1136. [PubMed: 16618801]
- Bracken AP, Helin K. Polycomb group proteins: navigators of lineage pathways led astray in cancer. *Nat Rev Cancer*. 2009; 9:773–784. [PubMed: 19851313]
- Britten RJ, Davidson EH. Gene regulation for higher cells: a theory. *Science*. 1969; 165:349–357. [PubMed: 5789433]
- Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*. 1996; 84:843–851. [PubMed: 8601308]
- Buhler M, Moazed D. Transcription and RNAi in heterochromatic gene silencing. *Nat Struct Mol Biol*. 2007; 14:1041–1048. [PubMed: 17984966]
- Cartwright P, McLean C, Sheppard A, Rivett D, Jones K, Dalton S. LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development*. 2005; 132:885–896. [PubMed: 15673569]
- Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*. 2003; 113:643–655. [PubMed: 12787505]
- Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A. Nanog safeguards pluripotency and mediates germline development. *Nature*. 2007; 450:1230–1234. [PubMed: 18097409]
- Chambers I, Smith A. Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene*. 2004; 23:7150–7160. [PubMed: 15378075]
- Chen X, Vega VB, Ng HH. Transcriptional regulatory networks in embryonic stem cells. *Cold Spring Harb Symp Quant Biol*. 2008a; 73:203–209. [PubMed: 19022762]
- Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*. 2008b; 133:1106–1117. [PubMed: 18555785]

- Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, 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]
- Chia NY, Chan YS, Feng B, Lu X, Orlov YL, Moreau D, Kumar P, Yang L, Jiang J, Lau MS, et al. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature*. 2010; 468:316–320. [PubMed: 20953172]
- Clapier CR, Cairns BR. The biology of chromatin remodeling complexes. *Annu Rev Biochem*. 2009; 78:273–304. [PubMed: 19355820]
- 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]
- Conaway RC, Sato S, Tomomori-Sato C, Yao T, Conaway JW. The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem Sci*. 2005; 30:250–255. [PubMed: 15896743]
- Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*. 2008; 322:1845–1848. [PubMed: 19056941]
- Cote J, Quinn J, Workman JL, Peterson CL. Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science*. 1994; 265:53–60. [PubMed: 8016655]
- Deato MD, Marr MT, Sottero T, Inouye C, Hu P, Tjian R. MyoD targets TAF3/TRF3 to activate myogenin transcription. *Mol Cell*. 2008; 32:96–105. [PubMed: 18851836]
- Deato MD, Tjian R. Switching of the core transcription machinery during myogenesis. *Genes Dev*. 2007; 21:2137–2149. [PubMed: 17704303]
- Dejosez M, Krumenacker JS, Zitur LJ, Passeri M, Chu LF, Songyang Z, Thomson JA, Zwaka TP. Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. *Cell*. 2008; 133:1162–1174. [PubMed: 18585351]
- Dejosez M, Levine SS, Frampton GM, Whyte WA, Stratton SA, Barton MC, Gunaratne PH, Young RA, Zwaka TP. Ronin/Hcf-1 binds to a hyperconserved enhancer element and regulates genes involved in the growth of embryonic stem cells. *Genes Dev*. 2010; 24:1479–1484. [PubMed: 20581084]
- Ding L, Paszkowski-Rogacz M, Nitzsche A, Slabicki MM, Heninger AK, de Vries I, Kittler R, Junqueira M, Shevchenko A, Schulz H, et al. A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. *Cell Stem Cell*. 2009; 4:403–415. [PubMed: 19345177]
- Ding N, Zhou H, Esteve PO, Chin HG, Kim S, Xu X, Joseph SM, Friez MJ, Schwartz CE, Pradhan S, et al. Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. *Mol Cell*. 2008; 31:347–359. [PubMed: 18691967]
- Endoh M, Endo TA, Endoh T, Fujimura Y, Ohara O, Toyoda T, Otte AP, Okano M, Brockdorff N, Vidal M, 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]
- Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, Burge CB, Bartel DP. The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science*. 2005; 310:1817–1821. [PubMed: 16308420]
- Fazio TG, Huff JT, Panning B. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell*. 2008; 134:162–174. [PubMed: 18614019]
- Fazio TG, Panning B. Condensin complexes regulate mitotic progression and interphase chromatin structure in embryonic stem cells. *J Cell Biol*. 2010; 188:491–503. [PubMed: 20176923]
- Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol*. 2006; 8:188–194. [PubMed: 16415856]
- Firestein R, Bass AJ, Kim SY, Dunn IF, Silver SJ, Guney I, Freed E, Ligon AH, Vena N, Ogino S, et al. CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity. *Nature*. 2008; 455:547–551. [PubMed: 18794900]

- Frankel N, Davis GK, Vargas D, Wang S, Payre F, Stern DL. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature*. 2010; 466:490–493. [PubMed: 20512118]
- Fryer CJ, White JB, Jones KA. Mastermind recruits CycC:CDK8 to phosphorylate the Notch ICD and coordinate activation with turnover. *Mol Cell*. 2004; 16:509–520. [PubMed: 15546612]
- Fuda NJ, Ardehali MB, Lis JT. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature*. 2009; 461:186–192. [PubMed: 19741698]
- Fujita J, Crane AM, Souza MK, Dejosez M, Kyba M, Flavell RA, Thomson JA, Zwaka TP. Caspase activity mediates the differentiation of embryonic stem cells. *Cell Stem Cell*. 2008; 2:595–601. [PubMed: 18522852]
- Galan-Cardidad JM, Harel S, Arenzana TL, Hou ZE, Doetsch FK, Mirny LA, Reizis B. Zfx controls the self-renewal of embryonic and hematopoietic stem cells. *Cell*. 2007; 129:345–357. [PubMed: 17448993]
- Gao S, Alarcon C, Sapkota G, Rahman S, Chen PY, Goerner N, Macias MJ, Erdjument-Bromage H, Tempst P, Massague J. Ubiquitin ligase Nedd4L targets activated Smad2/3 to limit TGF-beta signaling. *Mol Cell*. 2009; 36:457–468. [PubMed: 19917253]
- Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ, Heidersbach A, Ramalho-Santos J, McManus MT, Plath K, Meshorer E, et al. Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature*. 2009; 460:863–868. [PubMed: 19587682]
- Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science*. 2006; 312:75–79. [PubMed: 16484454]
- Graf T, Enver T. Forcing cells to change lineages. *Nature*. 2009; 462:587–594. [PubMed: 19956253]
- Grewal SI, Elgin SC. Transcription and RNA interference in the formation of heterochromatin. *Nature*. 2007; 447:399–406. [PubMed: 17522672]
- 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]
- Guenther MG, Young RA. Transcription. Repressive transcription. *Science*. 2010; 329:150–151. [PubMed: 20616255]
- Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*. 2010; 464:1071–1076. [PubMed: 20393566]
- Han J, Yuan P, Yang H, Zhang J, Soh BS, Li P, Lim SL, Cao S, Tay J, Orlov YL, et al. Tbx3 improves the germ-line competency of induced pluripotent stem cells. *Nature*. 2010; 463:1096–1100. [PubMed: 20139965]
- Han M, Grunstein M. Nucleosome loss activates yeast downstream promoters in vivo. *Cell*. 1988; 55:1137–1145. [PubMed: 2849508]
- Hart AH, Hartley L, Ibrahim M, Robb L. Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. *Dev Dyn*. 2004; 230:187–198. [PubMed: 15108323]
- He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, Powers S, Cordon-Cardo C, Lowe SW, Hannon GJ, et al. A microRNA polycistron as a potential human oncogene. *Nature*. 2005; 435:828–833. [PubMed: 15944707]
- He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. The antisense transcriptomes of human cells. *Science*. 2008; 322:1855–1857. [PubMed: 19056939]
- Ho L, Crabtree GR. Chromatin remodelling during development. *Nature*. 2010; 463:474–484. [PubMed: 20110991]
- Hong JW, Hendrix DA, Levine MS. Shadow enhancers as a source of evolutionary novelty. *Science*. 2008; 321:1314. [PubMed: 18772429]
- Hu G, Kim J, Xu Q, Leng Y, Orkin SH, Elledge SJ. A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev*. 2009; 23:837–848. [PubMed: 19339689]
- Imbalzano AN, Kwon H, Green MR, Kingston RE. Facilitated binding of TATA-binding protein to nucleosomal DNA. *Nature*. 1994; 370:481–485. [PubMed: 8047170]

- Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, Schafer X, Lun Y, Lemischka IR. Dissecting self-renewal in stem cells with RNA interference. *Nature*. 2006; 442:533–538. [PubMed: 16767105]
- Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*. 1961; 3:318–356. [PubMed: 13718526]
- Jaenisch R, Young R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell*. 2008; 132:567–582. [PubMed: 18295576]
- James D, Levine AJ, Besser D, Hemmati-Brivanlou A. TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development*. 2005; 132:1273–1282. [PubMed: 15703277]
- Jiang J, Chan YS, Loh YH, Cai J, Tong GQ, Lim CA, Robson P, Zhong S, Ng HH. A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat Cell Biol*. 2008; 10:353–360. [PubMed: 18264089]
- Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature*. 2010
- Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev*. 2005; 19:489–501. [PubMed: 15713842]
- Kanhere A, Viiri K, Araujo CC, Rasaiyaah J, Bouwman RD, Whyte WA, Pereira CF, Brookes E, Walker K, Bell GW, et al. Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Mol Cell*. 2010; 38:675–688. [PubMed: 20542000]
- Kayne PS, Kim UJ, Han M, Mullen JR, Yoshizaki F, Grunstein M. Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. *Cell*. 1988; 55:27–39. [PubMed: 3048701]
- 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]
- Kim J, Woo AJ, Chu J, Snow JW, Fujiwara Y, Kim CG, Cantor AB, Orkin SH. A Myc Network Accounts for Similarities between Embryonic Stem and Cancer Cell Transcription Programs. *Cell*. 2010; 143:313–324. [PubMed: 20946988]
- Klochender Yeivin A, Fiette L, Barra J, Muchardt C, Babinet C, Yaniv M. The murine SNF5/INI1 chromatin remodeling factor is essential for embryonic development and tumor suppression. *EMBO Rep*. 2000; 1:500–506. [PubMed: 11263494]
- Knezetic JA, Luse DS. The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell*. 1986; 45:95–104. [PubMed: 3955658]
- Knuesel MT, Meyer KD, Bernecky C, Taatjes DJ. The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev*. 2009; 23:439–451. [PubMed: 19240132]
- Kornberg RD, Thomas JO. Chromatin structure; oligomers of the histones. *Science*. 1974; 184:865–868. [PubMed: 4825888]
- Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, et al. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of *Drosophila melanogaster* Nipped-B. *Nat Genet*. 2004; 36:631–635. [PubMed: 15146186]
- Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer*. 2007; 7:823–833. [PubMed: 17957188]
- Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, et al. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell*. 2003; 11:721–729. [PubMed: 12667454]
- Kwon H, Imbalzano AN, Khavari PA, Kingston RE, Green MR. Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. *Nature*. 1994; 370:477–481. [PubMed: 8047169]

- Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell*. 2006; 125:301–313. [PubMed: 16630818]
- Leeb M, Pasini D, Novatchkova M, Jaritz M, Helin K, Wutz A. Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev*. 2010; 24:265–276. [PubMed: 20123906]
- Levine M, Tjian R. Transcription regulation and animal diversity. *Nature*. 2003; 424:147–151. [PubMed: 12853946]
- Li B, Carey M, Workman JL. The role of chromatin during transcription. *Cell*. 2007; 128:707–719. [PubMed: 17320508]
- 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]
- Lluis F, Pedone E, Pepe S, Cosma MP. Periodic activation of Wnt/beta-catenin signaling enhances somatic cell reprogramming mediated by cell fusion. *Cell Stem Cell*. 2008; 3:493–507. [PubMed: 18983965]
- Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet*. 2006; 38:431–440. [PubMed: 16518401]
- Lorch Y, LaPointe JW, Kornberg RD. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell*. 1987; 49:203–210. [PubMed: 3568125]
- Malik S, Roeder RG. Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem Sci*. 2005; 30:256–263. [PubMed: 15896744]
- Maniatis T, Falvo JV, Kim TH, Kim TK, Lin CH, Parekh BS, Wathelet MG. Structure and function of the interferon-beta enhanceosome. *Cold Spring Harb Symp Quant Biol*. 1998; 63:609–620. [PubMed: 10384326]
- Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ 3rd, Voigt P, Martin SR, Taylor WR, De Marco V, et al. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature*. 2009; 461:762–767. [PubMed: 19767730]
- Marson A, Foreman R, Chevalier B, Bilodeau S, Kahn M, Young RA, Jaenisch R. Wnt signaling promotes reprogramming of somatic cells to pluripotency. *Cell Stem Cell*. 2008a; 3:132–135. [PubMed: 18682236]
- Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S, Guenther MG, Johnston WK, Wernig M, Newman J, et al. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell*. 2008b; 134:521–533. [PubMed: 18692474]
- Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA, et al. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol*. 2007; 9:625–635. [PubMed: 17515932]
- Matoba R, Niwa H, Masui S, Ohtsuka S, Carter MG, Sharov AA, Ko MS. Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling. *PLoS One*. 2006; 1:e26. [PubMed: 17183653]
- Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*. 2003; 113:631–642. [PubMed: 12787504]
- Muncke N, Jung C, Rudiger H, Ulmer H, Roeth R, Hubert A, Goldmuntz E, Driscoll D, Goodship J, Schon K, et al. Missense mutations and gene interruption in PROSIT240, a novel TRAP240-like gene, in patients with congenital heart defect (transposition of the great arteries). *Circulation*. 2003; 108:2843–2850. [PubMed: 14638541]
- Murchison EP, Partridge JF, Tam OH, Cheloufi S, Hannon GJ. Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A*. 2005; 102:12135–12140. [PubMed: 16099834]
- Niakan KK, Davis EC, Clipsham RC, Jiang M, Dehart DB, Sulik KK, McCabe ER. Novel role for the orphan nuclear receptor Dax1 in embryogenesis, different from steroidogenesis. *Mol Genet Metab*. 2006; 88:261–271. [PubMed: 16466956]

- Nichols J, Zevnik B, Anastasiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Scholer H, Smith A. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*. 1998; 95:379–391. [PubMed: 9814708]
- Niwa H. How is pluripotency determined and maintained? *Development*. 2007; 134:635–646. [PubMed: 17215298]
- Niwa H, Burdon T, Chambers I, Smith A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev*. 1998; 12:2048–2060. [PubMed: 9649508]
- Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet*. 2000; 24:372–376. [PubMed: 10742100]
- Niwa H, Ogawa K, Shimosato D, Adachi K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature*. 2009; 460:118–122. [PubMed: 19571885]
- O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature*. 2005; 435:839–843. [PubMed: 15944709]
- Okita K, Yamanaka S. Intracellular signaling pathways regulating pluripotency of embryonic stem cells. *Curr Stem Cell Res Ther*. 2006; 1:103–111. [PubMed: 18220859]
- Olins AL, Olins DE. Spheroid chromatin units (v bodies). *Science*. 1974; 183:330–332. [PubMed: 4128918]
- Orkin SH, Wang J, Kim J, Chu J, Rao S, Theunissen TW, Shen X, Levasseur DN. The Transcriptional Network Controlling Pluripotency in ES Cells. *Cold Spring Harb Symp Quant Biol*. 2008
- Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, Jonsdottir GA, Stewart R, Thomson JA. 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]
- Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell*. 2008; 32:232–246. [PubMed: 18951091]
- Pasini D, Bracken AP, Agger K, Christensen J, Hansen K, Cloos PA, Helin K. Regulation of stem cell differentiation by histone methyltransferases and demethylases. *Cold Spring Harb Symp Quant Biol*. 2008; 73:253–263. [PubMed: 19022750]
- Pasini D, Bracken AP, Jensen MR, Lazzarini Denchi E, Helin K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J*. 2004; 23:4061–4071. [PubMed: 15385962]
- Pasini D, Cloos PA, Walfridsson J, Olsson L, Bukowski JP, Johansen JV, Bak M, Tommerup N, Rappsilber J, Helin K. JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature*. 2010; 464:306–310. [PubMed: 20075857]
- Peng JC, Valouev A, Swigut T, Zhang J, Zhao Y, Sidow A, Wysocka J. Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell*. 2009; 139:1290–1302. [PubMed: 20064375]
- Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell*. 2006; 23:297–305. [PubMed: 16885020]
- Philibert RA, Madan A. Role of MED12 in transcription and human behavior. *Pharmacogenomics*. 2007; 8:909–916. [PubMed: 17716226]
- Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA. c-Myc regulates transcriptional pause release. *Cell*. 2010; 141:432–445. [PubMed: 20434984]
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*. 2007; 129:1311–1323. [PubMed: 17604720]
- Risheg H, Graham JM Jr, Clark RD, Rogers RC, Opitz JM, Moeschler JB, Peiffer AP, May M, Joseph SM, Jones JR, et al. A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nat Genet*. 2007; 39:451–453. [PubMed: 17334363]
- Roeder RG. Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. *Cold Spring Harb Symp Quant Biol*. 1998; 63:201–218. [PubMed: 10384284]
- Roeder RG. Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett*. 2005; 579:909–915. [PubMed: 15680973]

- Rossant J. Stem cells and early lineage development. *Cell*. 2008; 132:527–531. [PubMed: 18295568]
- Sato N, Meijer L, Skaltsounis L, Greengard P, Brivanlou AH. Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat Med*. 2004; 10:55–63. [PubMed: 14702635]
- Schnetz MP, Handoko L, Akhtar-Zaidi B, Bartels CF, Pereira CF, Fisher AG, Adams DJ, Flicek P, Crawford GE, Laframboise T, et al. CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS Genet*. 2010; 6:e1001023. [PubMed: 20657823]
- Scholer HR, Ruppert S, Suzuki N, Chowdhury K, Gruss P. New type of POU domain in germ line-specific protein Oct-4. *Nature*. 1990; 344:435–439. [PubMed: 1690859]
- Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. *Cell*. 2007; 128:735–745. [PubMed: 17320510]
- Schwartz CE, Tarpey PS, Lubs HA, Verloes A, May MM, Risheg H, Friez MJ, Futreal PA, Edkins S, Teague J, et al. The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. *J Med Genet*. 2007; 44:472–477. [PubMed: 17369503]
- Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, Young RA, Sharp PA. Divergent transcription from active promoters. *Science*. 2008; 322:1849–1851. [PubMed: 19056940]
- Shen X, Kim W, Fujiwara Y, Simon MD, Liu Y, Mysliwiec MR, Yuan GC, Lee Y, Orkin SH. Jumonji modulates polycomb activity and self-renewal versus differentiation of stem cells. *Cell*. 2009; 139:1303–1314. [PubMed: 20064376]
- Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, Wray J, Yamanaka S, Chambers I, Smith A. Nanog is the gateway to the pluripotent ground state. *Cell*. 2009; 138:722–737. [PubMed: 19703398]
- Silva J, Smith A. Capturing pluripotency. *Cell*. 2008; 132:532–536. [PubMed: 18295569]
- Smith AG. Embryo-derived stem cells: of mice and men. *Annu Rev Cell Dev Biol*. 2001; 17:435–462. [PubMed: 11687496]
- Smith SK, Charnock-Jones DS, Sharkey AM. The role of leukemia inhibitory factor and interleukin-6 in human reproduction. *Hum Reprod*. 1998; 13(Suppl 3):237–243. discussion 244–236. [PubMed: 9755426]
- Smith TA, Hooper ML. Medium conditioned by feeder cells inhibits the differentiation of embryonal carcinoma cultures. *Exp Cell Res*. 1983; 145:458–462. [PubMed: 6861901]
- Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koseki H, Brockdorff N, Fisher AG, Pombo A. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol*. 2007; 9:1428–1435. [PubMed: 18037880]
- Strachan T. Cornelia de Lange Syndrome and the link between chromosomal function, DNA repair and developmental gene regulation. *Curr Opin Genet Dev*. 2005; 15:258–264. [PubMed: 15917200]
- Sun C, Nakatake Y, Akagi T, Ura H, Matsuda T, Nishiyama A, Koide H, Ko MS, Niwa H, Yokota T. Dax1 binds to Oct3/4 and inhibits its transcriptional activity in embryonic stem cells. *Mol Cell Biol*. 2009; 29:4574–4583. [PubMed: 19528230]
- Surface LE, Thornton SR, Boyer LA. Polycomb group proteins set the stage for early lineage commitment. *Cell Stem Cell*. 2010; 7:288–298. [PubMed: 20804966]
- Taatjes DJ. The human Mediator complex: a versatile, genome-wide regulator of transcription. *Trends Biochem Sci*. 2010
- Tam WL, Lim CY, Han J, Zhang J, Ang YS, Ng HH, Yang H, Lim B. T-cell factor 3 regulates embryonic stem cell pluripotency and self-renewal by the transcriptional control of multiple lineage pathways. *Stem Cells*. 2008; 26:2019–2031. [PubMed: 18467660]
- 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]
- Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. *Nat Genet*. 2004; 36:636–641. [PubMed: 15146185]

- Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. Long noncoding RNA as modular scaffold of histone modification complexes. *Science*. 2010; 329:689–693. [PubMed: 20616235]
- Vallier L, Alexander M, Pedersen RA. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci*. 2005; 118:4495–4509. [PubMed: 16179608]
- van der Stoop P, Boutsma EA, Hulsman D, Noback S, Heimerikx M, Kerkhoven RM, Voncken JW, Wessels LF, van Lohuizen M. Ubiquitin E3 ligase Ring1b/Rnf2 of polycomb repressive complex 1 contributes to stable maintenance of mouse embryonic stem cells. *PLoS One*. 2008; 3:e2235. [PubMed: 18493325]
- Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Sudhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*. 2010; 463:1035–1041. [PubMed: 20107439]
- Viswanathan SR, Daley GQ, Gregory RI. Selective blockade of microRNA processing by Lin28. *Science*. 2008; 320:97–100. [PubMed: 18292307]
- Wang Y, Baskerville S, Shenoy A, Babiarz JE, Baehner L, Blalock R. Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat Genet*. 2008; 40:1478–1483. [PubMed: 18978791]
- Wang Y, Medvid R, Melton C, Jaenisch R, Blalock R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat Genet*. 2007; 39:380–385. [PubMed: 17259983]
- Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature*. 1988; 336:684–687. [PubMed: 3143916]
- Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev*. 2009; 23:1494–1504. [PubMed: 19571179]
- Wu Q, Chen X, Zhang J, Loh YH, Low TY, Zhang W, Sze SK, Lim B, Ng HH. Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. *J Biol Chem*. 2006; 281:24090–24094. [PubMed: 16840789]
- Yamanaka S, Blau HM. Nuclear reprogramming to a pluripotent state by three approaches. *Nature*. 2010; 465:704–712. [PubMed: 20535199]
- Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell*. 2010; 38:662–674. [PubMed: 20541999]
- Yeap LS, Hayashi K, Surani MA. ERG-associated protein with SET domain (ESET)-Oct4 interaction regulates pluripotency and represses the trophoblast lineage. *Epigenetics Chromatin*. 2009; 2:12. [PubMed: 19811652]
- Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell*. 2003; 115:281–292. [PubMed: 14636556]
- Yuan P, Han J, Guo G, Orlov YL, Huss M, Loh YH, Yaw LP, Robson P, Lim B, Ng HH. Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev*. 2009; 23:2507–2520. [PubMed: 19884257]
- Zaratiegui M, Irvine DV, Martienssen RA. Noncoding RNAs and gene silencing. *Cell*. 2007; 128:763–776. [PubMed: 17320512]
- Zaret KS, Watts J, Xu J, Wandzioch E, Smale ST, Sekiya T. Pioneer factors, genetic competence, and inductive signaling: programming liver and pancreas progenitors from the endoderm. *Cold Spring Harb Symp Quant Biol*. 2008; 73:119–126. [PubMed: 19028990]
- Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY, Soh BS, Lou Y, Yang J, Ma Y, Chai L, et al. Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol*. 2006; 8:1114–1123. [PubMed: 16980957]

- Zhang X, Zhang J, Wang T, Esteban MA, Pei D. Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. *J Biol Chem*. 2008; 283:35825–35833. [PubMed: 18957414]
- Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science*. 2008; 322:750–756. [PubMed: 18974356]
- Zhong X, Jin Y. Critical roles of coactivator p300 in mouse embryonic stem cell differentiation and Nanog expression. *J Biol Chem*. 2009; 284:9168–9175. [PubMed: 19150979]
- Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature*. 2008a; 455:627–632. [PubMed: 18754011]
- Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J, Glass CK, Rosenfeld MG. Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Mol Cell*. 2008b; 29:69–80. [PubMed: 18206970]

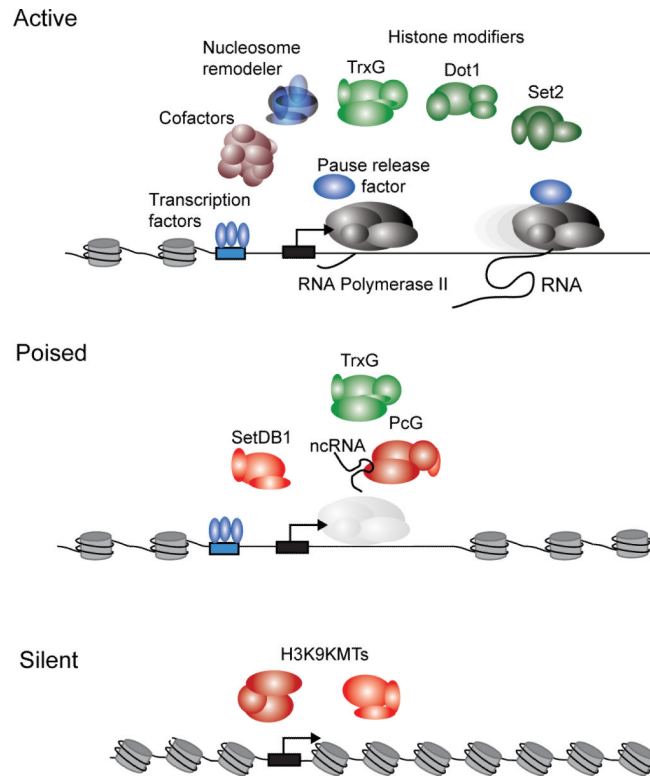


Figure 1. Models for transcriptionally active, poised and silent genes

Transcription factors, cofactors, chromatin regulators and ncRNA regulators can be found at active, poised and silent genes. At active genes, enhancers are typically bound by multiple transcription factors (TFs), which recruit cofactors that can interact with RNA polymerase II at the core promoter. RNA polymerase II generates a short transcript and pauses until pause-release factors and elongation factors allow further transcription. Chromatin regulators, which include nucleosome remodeling complexes such as Swi/Snf complexes and histone modifying enzymes such as TrxG, Dot1 and Set2, are recruited by transcription factors or the transcription apparatus and mobilize or modify local nucleosomes. Poised genes are rapidly activated when ESCs are stimulated to differentiate. At poised genes, transcription initiation and recruitment of TrxG can occur, but pause-release, elongation and recruitment of Dot1 and Set2 do not occur. The PcG and SetDB1 chromatin regulators can contribute to this repression, and these can be recruited by some transcription factors and by ncRNAs. The RNA polymerase II “ghost” in this model of poised genes reflects the low levels of the enzyme that are detected under steady-state conditions. Silent genes show little or no evidence of transcription initiation or elongation and are often occupied by chromatin regulators that methylate histone H3K9 and other residues. Some of these silent genes are probably silenced by mechanisms that depend on transcription of at least a portion of the gene (Buhler and Moazed, 2007; Grewal and Elgin, 2007; Zaratiegui et al., 2007).

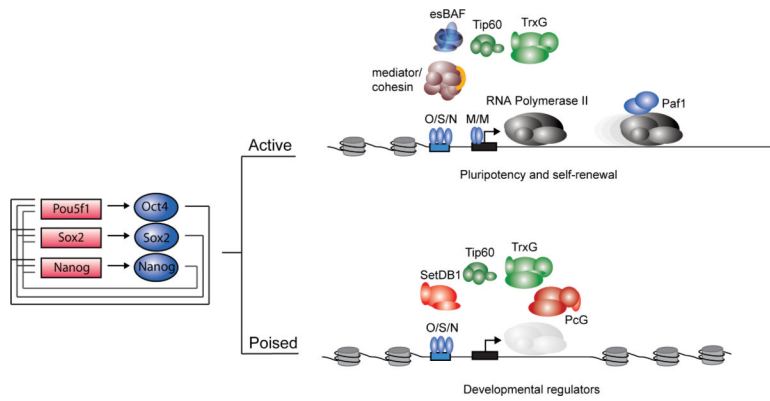


Figure 2. Core regulatory circuitry

Oct4, Sox2 and Nanog collaborate to regulate their own promoters, forming an interconnected autoregulatory loop. The Pou5f (Oct4), Sox2 and Nanog genes are represented as red boxes and proteins as blue balloons. These core transcription factors (O/S/N) function to activate expression of protein-coding and miRNA genes necessary to maintain ESC state, but they also occupy poised genes encoding lineage-specific protein and miRNA regulators whose repression is essential to maintaining that state. Additional transcription factors, such as the c-Myc/Max heterodimer (M/M), cause pause release at actively transcribed genes. A subset of the cofactors and chromatin regulators implicated in control of ES cell state (Table 1) are shown.

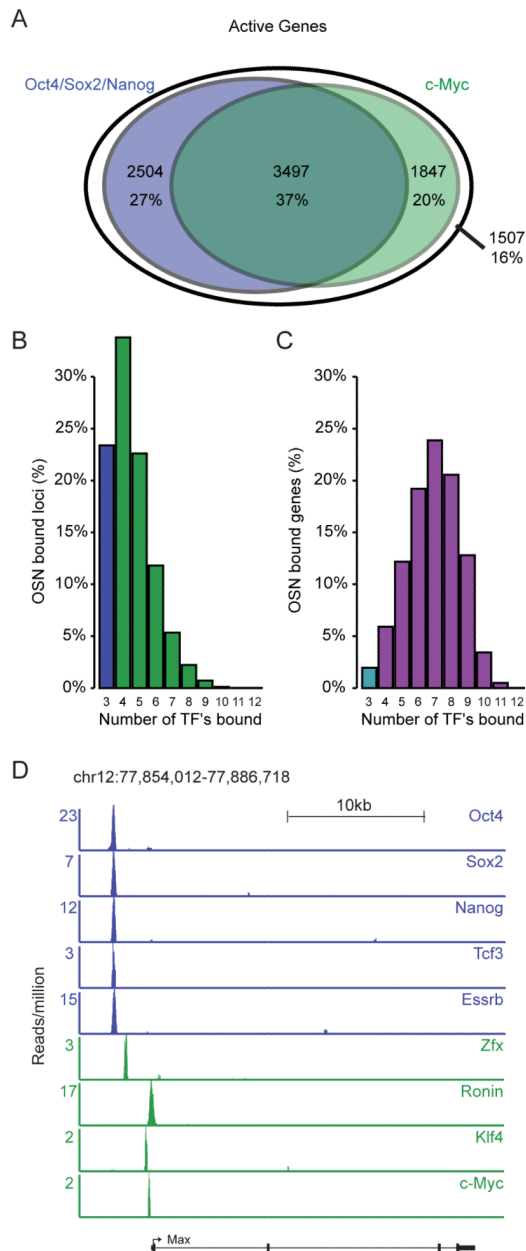


Figure 3. Relationships between core and other transcription factors in regulatory circuitry and gene control

A) Overlap between actively transcribed genes occupied by core transcription factors (union of Oct4, Sox2 and Nanog bound genes) and those of occupied by c-Myc. Active genes were defined as the set of genes occupied by both RNA polymerase II and nucleosomes with histone H3K79me3.

B) Frequency distribution showing how c-Myc, Tcf3, Smad1, STAT3, Esrrb, Tbx3, Zfx, Ronin and Klf4 are associated with Oct4/Sox2/Nanog-occupied loci. Oct4, Sox2 and Nanog are the three transcription factors in the first bin. Binding was called at a high confidence ($p < 10^{-9}$) threshold within a 50bp window, so the actual number of factors bound to Oct4/Sox2/Nanog-occupied loci is somewhat higher.

C) Frequency distribution showing how often c-Myc, Tcf3, Smad1, STAT3, Esrrb, Tbx3, Zfx, Ronin and Klf4 are associated with Oct4/Sox2/Nanog-occupied genes ($p < 10^{-9}$).

D) Gene tracks showing example of an actively transcribed gene (Max) occupied by an Oct4/Sox2/Nanog enhancer and other transcription factors implicated in ESC control. ChIP-Seq data was obtained from GSE11431, GSE11724, GSE12680 and GSE22557.

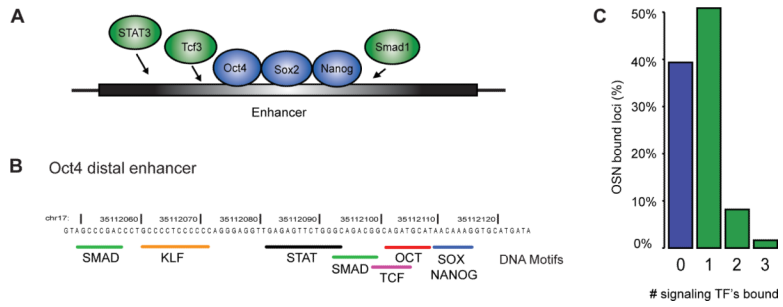


Figure 4. Signaling to core regulatory circuitry

A) Model of an enhancer where transcription factors associated with Wnt, LIF and BMP4 signaling (STAT3, Tcf3 and Smad1) occupy sites near the core regulators.

B) Oct4 distal enhancer provides an example of a DNA element that is bound by the core regulators and signaling transcription factors and contains sequence motifs for each of these factors.

C) Frequency distribution showing how often signaling transcription factors (STAT3, Tcf3 and Smad1) are associated with Oct4/Sox2/Nanog-bound loci throughout genome. Binding was called at a high confidence ($p < 10^{-9}$) threshold within a 50bp window, so the actual number of signaling transcription factors bound to Oct4/Sox2/Nanog-occupied loci is somewhat higher.

ChIP-Seq data was obtained from GSE11431, GSE11724, GSE12680 and GSE22557.

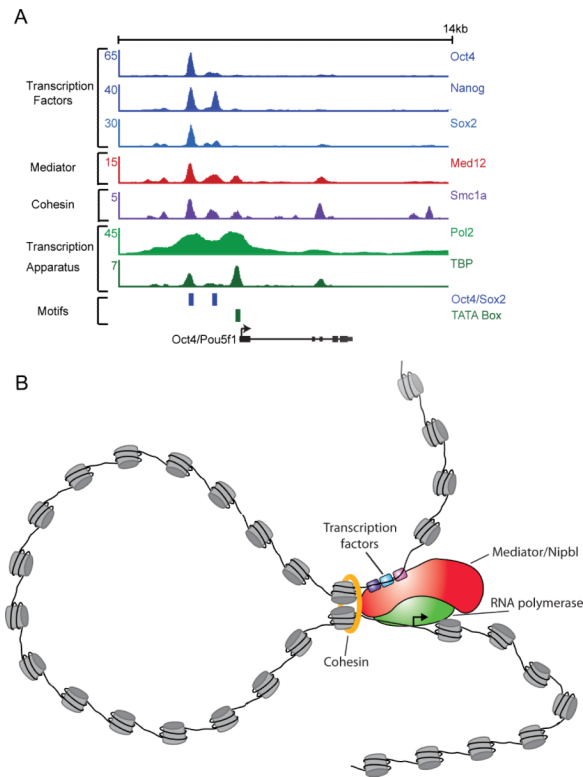


Figure 5. Mediator and cohesin contribute to gene control in core circuitry

A) ChIP-Seq data at the Pou5f gene for transcription factors, mediator and cohesin, and the transcription apparatus (Pol2 and TBP). Note evidence for crosslinking of most components to both enhancer elements and core promoter. The numbers on the Y-axis are reads/million. ChIP-Seq data was obtained from GSE11431, GSE11724, GSE12680 and GSE22557.

B) Model for DNA looping by mediator and cohesin. Oct4, Sox2 and Nanog bind mediator, which binds RNA polymerase II at the core promoter, thus forming a loop between the enhancer and the core promoter. The transcription activator-bound form of mediator binds the cohesin loading factor Nipbl, which provides a means to load cohesin. Both mediator and cohesin are necessary for normal gene activity. This model contains a single DNA loop, but multiple enhancer may be bound simultaneously, generating multiple loops.

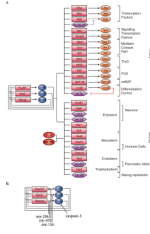


Figure 6. ESC core regulatory circuitry and differentiation

A) This model of core regulatory circuitry incorporates selected protein-coding and miRNA target genes. Oct4, Sox2 and Nanog directly activate transcription of genes whose products include the spectrum of transcription factors, cofactors, chromatin regulators and miRNAs that are known to contribute to ESC state. Oct4, Sox2 and Nanog are also associated with SetDB1 and PcG-repressed protein-coding and miRNA genes that are poised for differentiation.

B) The loss of ESC state during differentiation involves the silencing of the Oct4 gene, the proteolytic destruction of Nanog by caspase-3, and miRNA-mediated reduction in Oct4, Nanog and Sox2 mRNA levels.

Table 1

Transcriptional regulators implicated in control of ESC state*.

Type of Regulator	Function	References
<u>Transcription factors</u>		
Oct4	Core circuitry	1
Sox2	Core circuitry	2
Nanog	Core circuitry	3
Tcf3	Wnt signaling to core circuitry	4
Stat3	Lif signaling to core circuitry	5
Smad1	BMP signaling to core circuitry	6
Smad 2/3	TGF- β /Activin/Nodal signaling	7
c-Myc	Proliferation	8
Esrrb	Steroid hormone receptor	9
Sall4	Embryonic regulator	10
Tbx3	Mediates LIF signaling	11
Zfx	Self-renewal	12
Ronin	Metabolism	13
Klf4	LIF signaling	14
Prdm14	ESC identity	15
<u>Cofactors</u>		
Mediator	Core circuitry	16
Cohesin	Core circuitry	17
Paf1 complex	Couples transcription with histone mod	18
Dax1	Oct4 inhibitor	19
Cnot3	Myc/Zfx cofactor	20
Trim28	Myc/Zfx cofactor	21
<u>Chromatin regulators</u>		
Polycomb Group	Silencing of lineage-specific regulators	22
SetDB1 (ESET)	Silencing of lineage-specific regulators	23
esBAF	Nucleosome mobilization	24
Chd1	Nucleosome mobilization	25
Chd7	Nucleosome mobilization	26
Tip60-p400	Histone acetylation	27
<u>ncRNA regulators</u>		
miRNAs	Fine tuning of pluripotency transcripts	28
GC-rich ncRNAs	PcG complex recruitment	29

LIF-Stat3 signaling is important for maintenance of murine ESCs and Activin-Smad2/3 signaling has been demonstrated to be important for human ESCs.

References:

* The vast majority of these regulators were identified in murine ES cells, but most appear to play similar roles in human ES cells.

¹(Chambers et al., 2003; Chambers and Smith, 2004; Hart et al., 2004; Mitsui et al., 2003; Nichols et al., 1998; Niwa et al., 2000; Scholer et al., 1990);

- ²(Chambers and Smith, 2004; Masui et al., 2007);
- ³(Chambers et al., 2003; Hart et al., 2004; Mitsui et al., 2003);
- ⁴(Cole et al., 2008; Marson et al., 2008b);
- ⁵(Niwa et al., 1998);
- ⁶(Ying et al., 2003);
- ⁷(Beattie et al., 2005; James et al., 2005; Vallier et al., 2005),
- ⁸(Cartwright et al., 2005);
- ⁹(Ivanova et al., 2006; Zhang et al., 2008);
- ¹⁰(Wu et al., 2006; Zhang et al., 2006);
- ¹¹(Han et al.; Ivanova et al., 2006; Niwa et al., 2009),
- ¹²(Galan-Caridad et al., 2007);
- ¹³(Dejosez et al., 2008; Dejosez et al., 2010);
- ¹⁴(Jiang et al., 2008; Niwa et al., 2009);
- ¹⁵(Chia et al.,2010);
- ¹⁶(Hu et al., 2009; Kagey et al., 2010);
- ¹⁷(Fazzio et al.,2008; Hu et al., 2009; Kagey et al., 2010);
- ¹⁸(Ding et al., 2009);
- ¹⁹(Kim et al., 2008; Niakan et al., 2006; Sun et al., 2009);
- ²⁰(Hu et al., 2009);
- ²¹(Fazzio et al., 2008; Hu et al., 2009);
- ²²(Azura et al., 2006; Bernstein et al., 2006; Boyer et al.,2006; Bracken et al., 2006; Lee et al., 2006; Leeb et al., 2010; Li et al.,2010; Pasini et al.; Peng et al., 2009; Shen et al., 2009; Stock et al., 2007; van der Stoop et al., 2008);
- ²³(Bilodeau et al., 2009; Yeap et al., 2009; Yuan et al., 2009);
- ²⁴(Ho and Crabtree, 2010; Schnetz et al., 2010);
- ²⁵(Gaspar-Maia et al., 2009);
- ²⁶(Schnetz et al., 2010);
- ²⁷(Fazzio et al., 2008);
- ²⁸(Marson et al., 2008b);
- ²⁹(Guenther and Young, 2010; Surface et al.,2010).