The chromatin signature of pluripotent cells*

Ky Sha¹ and Laurie A. Boyer¹,§, ¹Massachusetts Institute of Technology, Department of Biology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Table of Contents

1. Introduction .............................................................................2
2. The chromatin landscape of the pluripotent genome .................................................................3
2.1. Histone modification patterns .................................................................3
2.2. Bivalent chromatin domains .......................................................................4
2.3. Polycomb and Trithorax group proteins .........................................................5
2.4. Histone variants ......................................................................................6
3. Histone modifiers and chromatin remodelers shape the stem cell genome ................................7
3.1. Histone lysine methyltransferases and demethylases ..............................................7
3.2. Histone acetyltransferases (HATs) and deacetylases (HDACs) ..................................9
3.3. ATP-dependent chromatin remodeling .....................................................................10
4. Chromatin and reprogramming cell fate .............................................................................11
5. Perspectives and Future Directions ...............................................................................11
6. Acknowledgements ..............................................................................12
7. References ......................................................................................12

Abstract

Pluripotency, a property ascribed to the cells that constitute the early embryo as well as to embryonic stem (ES) cells, is progressively lost during differentiation as the activation of lineage-specific programs steer the cell towards a particular fate. The regulation of chromatin structure has emerged as a key mechanism to modulate developmental gene expression patterns by contributing to the activation or silencing of subsets of genes and through maintenance of expression states during subsequent cell divisions. Recent global analyses have revealed key differences in the chromatin landscape in pluripotent embryonic stem (ES) cells as compared to lineage-committed cells, suggesting that chromatin states may be linked to cell fate. Moreover, the molecular and biochemical characterization of a large group of enzymes that regulate chromatin structure and organization has revealed roles for these factors in early development and stem cell function. Such studies have begun to provide critical insights into how the ES cell genome may remain uncommitted, yet poised for differentiation. These findings have broad implications for understanding development as well as the process of re-programming the somatic genome into a pluripotent-like state and of the progression from normal to disease states.


Copyright: © 2009 Ky Sha and Laurie Boyer. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

§To whom correspondence should be addressed. E-mail: lboyer@mit.edu
1. Introduction

In metazoans, the zygotic genome represents the founding blueprint for all cell types in the developing organism. Despite an identical genotype, each cell type displays a distinct gene expression profile that codes for its unique functional characteristics. Embryonic stem (ES) cells, which are derived from the inner cell mass of the developing blastocyst, possess the dual capacities of self-renewal and pluripotency (Chen and Daley, 2008; Jaenisch and Young, 2008). The first property allows stem cells to generate an unlimited number of equivalent descendants while pluripotency is the capacity to differentiate into any cell type in the organism. These properties underscore the potential for the use of ES cells in therapeutic applications (Murry and Keller, 2008). The transition from an uncommitted to a differentiated state requires the rapid and global execution of specific gene programs, whereby pluripotency factors are silenced while lineage-specific programs are activated in a highly orchestrated fashion. To maintain the balance between these states, the genome of ES cells is thought to be for differentiation upon induction (Jorgensen et al., 2006; Pietersen and van Lohuizen, 2008). Thus, ES cells have been widely exploited as a model system to understand how genetic and epigenetic factors can influence pluripotency and the process of differentiation.

The mechanisms by which stem cells are able to execute lineage-specific gene expression programs comprise, in part, a network of transcription factors and chromatin regulators. It is well known that transcription factors play essential roles as drivers of cell fate by controlling gene expression (Niwa, 2007; Silva and Smith, 2008). While chromatin regulation has also emerged as a key mechanism for controlling genome organization and function, deciphering its role in development and cell fate determination has been less obvious and more complex. The basic subunit of chromatin is the nucleosome, which comprises an octamer of the four core histones; two copies each of H2A, H2B, H3 and H4, wrapped by 147 bp of DNA (Luger et al., 1997). The organization of nucleosomes into higher order structures can have both local and global effects on a variety of DNA-mediated processes including gene regulation (Figure 1). For example, specific regions of the genome may be made more or less accessible to DNA binding factors by regulating chromatin states. Chromatin conformation can be altered by different classes of regulators such as histone-modifying enzymes and ATP-dependent remodeling complexes, and by site-specific incorporation of histone variants (Becker and Horz, 2002; Henikoff, 2008; Kusch and Workman, 2007; Margueron et al., 2005). Mutations in many of the genes that code for these regulators can result in embryonic lethality or disease (Bhaumik et al., 2007; Kokavec et al., 2008; Lin and Dent, 2006), underscoring the essential role of chromatin regulation during early mammalian development.

![Figure 1](stembook.org). The basic unit of chromatin organization is the nucleosome, which comprises 147 bp of DNA wrapped around a core of histone proteins. Nucleosomes can be organized into higher order structures and the level of packaging can have profound consequences on all DNA-mediated processes including gene regulation. Euchromatin is associated with an open chromatin conformation and this structure is permissible for transcription whereas heterochromatin is more compact and refractory to factors that need to gain access to the DNA template. Nucleosome positioning and chromatin compaction can be influenced by a wide-range of processes including modification to both histones and DNA.
The idea that epigenetic changes contribute to cell fate transitions during development (Waddington, 1969) is consistent with historic experiments in which the identity of a somatic cell could be erased and reprogrammed by nuclear transplantation into a state resembling the early embryo (Briggs and King, 1957). The recent successes in the reprogramming of somatic cells to an embryonic stem cell fate by ectopic expression of key transcription factors also have now clearly demonstrated that the epigenome of a differentiated cell can be rewired to support embryonic development (Aoi et al., 2008; Lowry et al., 2008; Martin, 1981; Meissner et al., 2007; Nakagawa et al., 2008; Park et al., 2008; Takahashi et al., 2007a; Takahashi et al., 2007b; Takahashi and Yamanaka, 2006; Wernig et al., 2008; Yamanaka, 2008). The underlying mechanisms governing this process, however, remain poorly understood. Technological advances have made it possible to analyze transcriptional and chromatin states across the entire ES cell genome as well as in cells from different tissues and organisms. The consensus from these studies is that each cell type possesses not only a unique expression profile (transcriptome), but also a unique chromatin signature (epigenome) that can be heritably maintained through subsequent cell divisions. Thus, the chromatin landscape of a cell appears to be inextricably linked to its function. As such, a comprehensive analysis of the epigenome in pluripotent and lineage-committed cells may reveal how chromatin influences gene expression programs and ultimately cell fate.

Here, we highlight recent findings that reveal the chromatin signature in ES cells and its distinction from that of lineage-committed cells. We also include studies that have identified the regulatory factors that play critical roles in the establishment and maintenance of these chromatin states. Importantly, such studies have begun to unravel how chromatin regulation contributes to pluripotency and to the maintenance of cell state, while concomitantly priming ES cells for differentiation. We also discuss the implications of this work for understanding the reprogramming process and how insights from these studies may be extrapolated to better understand development and disease.

2. The chromatin landscape of the pluripotent genome

2.1. Histone modification patterns

The post-translational modification of histone proteins has been a topic of intense interest over the last decade as these marks are thought to impart important epigenetic information to allow for interpretation of the genetic code. Histone proteins are subject to a wide variety of modifications including acetylation, methylation, phosphorylation, ubiquitination, and ribosylation (Kouzarides, 2007). These modifications function by altering protein:protein and protein:DNA interactions or by creating binding surfaces for downstream effectors. It has been proposed that unique constellations of histone modifications define chromatin domains such as euchromatin and heterochromatin and may mark functional transcriptional units. For example, transcription initiation is characterized by the presence of H3K4me3 as well as H3K9/K14Ac at the 5' proximal promoter region (Bernstein et al., 2002; Liang et al., 2004; Ng et al., 2003; Pokholok et al., 2005; Santos-Rosa et al., 2002; Schneider et al., 2004; Schubeler et al., 2004); whereas H3K36me3 is enriched throughout the body of transcribed genes (Bannister et al., 2005; Pokholok et al., 2005). Conversely, H3K27me3 and H3K9me3 are thought to be classical markers of more compact, silent chromatin (Cao et al., 2002; Lachner et al., 2001). In many cases, these patterns have been found to be gene and cell type specific and may correlate with the developmental potential of the cell. Thus, genome-wide maps of histone modification patterns along developmental lineages can provide important insight into gene regulation and ultimately cell state.

Interestingly, it appears that ES cells have higher global levels of histone acetylation as compared to differentiated cells, consistent with a more open chromatin configuration and with the observed higher level of transcription, which includes the expression of both protein-coding and non-coding regions of the genome (Efroni et al., 2008). ES cell differentiation is concomitant with a marked decrease in histone acetylation and increase in heterochromatin formation (Meshorer and Misteli, 2006). Thus, ES cell differentiation is accompanied by the transition from a highly dynamic chromatin state to a more stable configuration in lineage-committed derivatives (Meshorer et al., 2006). Consistent with this, a recent study observed an accumulation of large blocks of silent chromatin marked by H3K9 methylation in differentiated cells as compared to ES cells and cancer cells, suggesting that chromatin changes may be important for maintaining the differentiated state (Wen et al., 2009). It is intriguing to speculate that chromatin in ES cells is actively maintained in a promiscuous state that allows for pervasive transcription. It is currently unknown, however, whether this unique chromatin landscape is a cause or consequence of hyper-transcription.
The chromatin signature of pluripotent cells

Figure 2. Bivalent chromatin domains mark the promoters of developmentally important genes in pluripotent ES cells. PRCC2 and TrxG proteins catalyze the tri-methylation of histone H3 on lysine 27 and 4, respectively. PRCC1 is also recruited to many of these genes and can mono-ubiquitynylate histone H2A on lysine 119, a modification that is also thought to be important for gene silencing, possibly through blocking RNAPII elongation. As such, bivalent genes are said to be silent, yet poised for activation. H2AZ is highly enriched in a manner that is remarkably similar to PRCC2 and may also be an important regulatory component at bivalent genes. Upon differentiation, the bivalent histone marks can be resolved to monovalent modifications in which the gene is “ON” or “OFF”. Bivalent domains can also be maintained or newly established in lineage-committed cells.

Nevertheless, the resulting stochastic flux in gene expression or transcriptional noise could allow the cell to respond quickly to a variety of developmental cues and may be important for lineage determination.

2.2. Bivalent chromatin domains

Technological advances have allowed investigators to map chromatin structure throughout the genome by combining chromatin immunoprecipitation with DNA microarray (ChIP-chip) or high throughput sequencing (ChIP-Seq) platforms. These studies have revealed that the distribution of the various histone modifications is not uniform throughout the genome and that these patterns are dynamic across developmental stages and during lineage commitment (Mendenhall and Bernstein, 2008). A remarkable find in both mouse and human is the co-enrichment of functionally opposite chromatin marks at many genes in ES cells (Azuara et al., 2006; Bernstein et al., 2006; Pan et al., 2007; Zhao et al., 2007). The bivalent chromatin structures consist of regions of H3K4me3 overlapping H3K27me3 at the transcription start site (TSS) and are generally associated with genes with low expression levels (Figure 2). These genes comprise a wide range of developmental regulators including transcription factors and signaling proteins (Bernstein et al., 2006). Interestingly, although they are silent, bivalent genes tend to replicate early in S phase, a property normally associated with active loci (Azuara et al., 2006). Further analysis indicated that this might be due to maintenance of histone H3K9 acetylation and H3K4 methylation patterns at these loci, despite the presence of the silent H3K27me3 mark (Azuara et al., 2006; Jorgensen et al., 2006). Based on these observations, it was suggested that the co-localization of opposing chromatin modifications in ES cells keeps development genes in a poised state, priming them for rapid induction upon differentiation.

It appears that the majority of bivalent marks resolve to one or the other upon ES cell differentiation, with the remaining modification corresponding to the expression state of the gene (Bernstein et al., 2006; Mikkelsen et al., 2007; Pan et al., 2007). For example, genes whose expression levels are induced upon ES cell differentiation maintain the H3K4me3 mark; while genes that are silenced upon differentiation harbor the H3K27me3 modification (Figure 2). Further work showed that differentiated cells such as multi-potent neural and hematopoietic stem cells, embryonic fibroblasts, and CD4+ T cells also possess bivalent marks, albeit perhaps fewer than in ES cells (Barski et al., 2007; Cui et al., 2009; Mikkelsen et al., 2007; Pan et al., 2007; Roh et al., 2006), indicating that plasticity may be maintained at some genes in differentiated cells. It also remains possible, however, that the latter observation reflects differing sub-populations in a heterogeneous mix of cells, as the presence of both marks may be spatially coincident but temporally distinct. Nonetheless, the conserved pattern of bivalent modifications at developmental genes in both mouse and human cells suggests an evolutionarily important role for these marks. How the bivalent modification pattern is established and transmitted through cell division during ES cell self-renewal and the cues responsible for rearrangement of this pattern during differentiation remain open questions.
The chromatin signature of pluripotent cells

### Table 1. Representative Polycomb Group and Trithorax Group protein complexes in stem cells

<table>
<thead>
<tr>
<th>Complex</th>
<th>Cellular Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polycomb Group (PcG) proteins</td>
<td>PRC2 complex</td>
<td>Repression of developmental genes in mES cells; together, Suz12, Eed, and Ezh2 catalyze H3K27 trimethylation; essential for early embryogenesis and ES cell differentiation</td>
</tr>
<tr>
<td></td>
<td>PRC1 complex</td>
<td>Catalyzes mono-ubiquitylation of histone H2A implicated in Pol II pausing in mES cells; Ring1b is essential for early embryogenesis and for proper maintenance of ES cells</td>
</tr>
<tr>
<td>Trithorax Group (Trx) proteins</td>
<td>MLL</td>
<td>MLL complex is a H3K4 methyltransferase whose activity counteracts the repressive effect of PcG proteins; essential for cell fate transitions in neuronal (Mll1) and hematopoietic (Mll5) lineages</td>
</tr>
</tbody>
</table>

2.3. Polycomb and Trithorax group proteins

Polycomb Group (PcG) and Trithorax Group (TrxG) proteins are classically defined as regulators of heritable gene expression patterns and cell fate during development (Bantignies and Cavalli, 2006; Pietersen and van Lohuizen, 2008; Ringrose and Paro, 2004; Schuettengruber et al., 2007; Schwartz and Pirrotta, 2004). The proteins that comprise these groups form multi-subunit complexes that maintain the silent or activated state, respectively, through modification of chromatin structure (Table 1). A number of these factors are essential for vertebrate development (Ayton et al., 2001; Faust et al., 1998; Morin-Kensicki et al., 2001; O’Carroll et al., 2001; Pasini et al., 2004; Voncken et al., 2003; Yu et al., 1995), have been implicated in stem cell maintenance (McMahon et al., 2007; Molofsky et al., 2003; Park et al., 2003) as well as in ES cell differentiation (Boyer et al., 2006; Pasini et al., 2007), and have been linked to tumorigenesis (Sparmann and van Lohuizen, 2006). Thus, a more detailed understanding of their roles will be important to elucidate how chromatin regulation contributes to development and disease.

In mammals, studies have identified two major Polycomb repressive complexes (PRC) based on their distinct biochemical activities and genetic interactions. PRC2 catalyzes the tri-methylation of histone H3 lysine 27 (H3K27me3) (Cao et al., 2002; Cao and Zhang, 2004; Czermin et al., 2002; Kirmizis et al., 2004; Kuzmichev et al., 2002). PRC1 possesses E3 ubiquitin ligase activity (de Napoles et al., 2004; O’Carroll et al., 2001; Pasini et al., 2004; Voncken et al., 2003; Yu et al., 1995), have been implicated in stem cell maintenance (McMahon et al., 2007; Molofsky et al., 2003; Park et al., 2003) as well as in ES cell differentiation (Boyer et al., 2006; Pasini et al., 2007), and have been linked to tumorigenesis (Sparmann and van Lohuizen, 2006). Thus, a more detailed understanding of their roles will be important to elucidate how chromatin regulation contributes to development and disease.

In mammals, studies have identified two major Polycomb repressive complexes (PRC) based on their distinct biochemical activities and genetic interactions. PRC2 catalyzes the tri-methylation of histone H3 lysine 27 (H3K27me3) (Cao et al., 2002; Cao and Zhang, 2004; Czermin et al., 2002; Kirmizis et al., 2004; Kuzmichev et al., 2002). PRC1 possesses E3 ubiquitin ligase activity (de Napoles et al., 2004) and may contribute to gene silencing by promoting chromatin condensation and chromatin stabilization, and by counteracting the effects of ATP-dependent remodelers (Francis et al., 2004; Francis et al., 2001). Trithorax group proteins are also found in multi-subunit complexes that play roles in transcription initiation by catalyzing the H3K4me3 modification and through recruitment or stabilization of the transcriptional machinery at target sites (Ringrose and Paro, 2004; Schuettengruber et al., 2007; Schwartz and Pirrotta, 2007). While H3K4me3 can be found at nearly 70% of all gene promoters in ES cells (Guenther et al., 2007), H3K27me3 is restricted to a much smaller subset of genes (~10% of all gene promoters) (Boyer et al., 2006; Lee et al., 2006; Mikkelsen et al., 2007). Interestingly, the vast majority of H3K27me3 in ES cells is found together at promoters with H3K4me3 as discussed above. Thus, the opposing activities of Polycomb and Trithorax result in the formation of bivalent chromatin and transcriptionally competent states.

Genome-wide profiling of several core components of PRC1 and PRC2 in ES cells suggested their potential role in lineage commitment in mammals. Most notably, PcG proteins were enriched at the promoter regions of a large class of genes with important roles in development and whose expression levels are low in ES cells as compared to differentiated cells (Boyer et al., 2006; Lee et al., 2006). While components of both PRC1 and PRC2 occupied discrete intervals at the TSS of target genes, a small fraction of these exhibited extended regions of binding that could span over 100kb such as at the HOX gene clusters (Lee et al., 2006). A similar distribution pattern was also observed for H3K27me3 (Bernstein et al., 2006). Interestingly, the sequences occupied by PcG proteins or H3K27me3 were found to be CpG-rich and to harbor some of the most highly conserved non-coding elements in the mammalian genome.
Consistent with a role in gene repression, the majority of target genes were de-repressed in ES cells that harbored mutations in core components of the PRC2 complex such as Eed and Suz12 (Azuara et al., 2006; Boyer et al., 2006; Chamberlain et al., 2008; Pasini et al., 2004). It is important to note that while loss of PcG proteins leads to gene de-repression, these cells generally maintain ES cell identity. In some cases, however, the ability of these cells to differentiate is compromised suggesting that PcG proteins are necessary for lineage commitment (Pasini et al., 2007). While PRC2 and PRC1 are found together at many promoters in ES cells, this is not an exclusive relationship as a recent study identified two classes of bivalent domains at promoter regions based on co-occupancy of both PRC1 and PRC2 or by PRC2 alone (Ku et al., 2008). It appears that promoters that are targets of both complexes are more efficiently retain Polycomb-mediated chromatin configuration upon differentiation. Collectively, these studies suggest that PcG proteins regulate developmental progression in ES cells by silencing genes whose untimely expression could lead to inappropriate differentiation.

While these studies have provided important insights into the control of developmental gene expression patterns during development, many questions remain. It will be critical to fully decipher how PcG- and TrxG-proteins are recruited to their target sites in mammals and to understand the mechanisms by which their associated modifications are maintained during DNA replication or resolved during lineage specification.

### 2.4. Histone variants

In addition to histone modifications, the replacement of major histones with specific variants has emerged as a key mechanism to modulate nucleosome dynamics and chromatin structure. Histone variants are non-allelic forms of the major histones that can be incorporated into chromatin in a cell cycle independent manner. Incorporation of various types of histone variants can induce localized changes in chromatin structure and form specialized chromosomal domains (Boulard et al., 2007; Henikoff et al., 2004; Jin et al., 2005). These replacement histones play important roles in all eukaryotes by influencing a wide range of DNA-mediated processes such as genome integrity, X-inactivation, and gene regulation (Boulard et al., 2007; Henikoff et al., 2004). Similar to histone modification patterns, the distribution of histone variants is not uniform and sites of enrichment demarcate chromatin into functional domains. H2AZ is a highly conserved variant of H2A that is essential for the development in all metazoans examined to date (Clarkson et al., 1999; Faast et al., 2001; Ridgway et al., 2004; Updike and Mango, 2006; van Dal and Elgin, 1992). This particular variant has been implicated in multiple and sometimes apparently contrasting roles in divergent organisms (Guillemette and Gaudreau, 2006; Zlatanova and Thakar, 2008), suggesting that it may function in a highly context-dependent and cell type specific manner. Thus, it will be important to elucidate the roles of H2AZ in mammalian development and stem cell function.

A recent genome-wide analysis of H2AZ localization by ChIP-chip in ES cells revealed a highly coincident enrichment pattern with Polycomb group proteins (Creyghton et al., 2008). In fact, H2AZ and Suz12, a core component of PRC2, displayed a remarkably similar spatial distribution pattern. This suggested that H2AZ might also be important for the regulation of bivalent genes and developmental gene expression patterns in ES cells. Consistent with this, H2AZ depletion by RNA interference led to de-repression of target genes and loss of PcG proteins from promoter regions. It is not known, however, whether H2AZ incorporation is required for gene repression or for proper gene activation upon induction as has been observed in yeast (Li et al., 2005; Zhang et al., 2005). The co-localization of PcG proteins and H2AZ suggests that H2AZ may be the target of PRC1-catalyzed ubiquitylation and that this might be important for polymerase pausing in ES cells. It has also been proposed that the incorporation of H2AZ favors nucleosome eviction (Mavrich et al., 2008), which would promote histone turnover and chromatin accessibility upon loss of Polycomb-mediated repression. Thus, H2AZ incorporation may be a key mechanism to allow developmental genes to remain silent, yet poised for activation in ES cells.

Surprisingly, H2AZ is not required for maintenance of the ES cell state as its depletion does not alter self-renewal or the expression of pluripotency markers such as Oct4 (Creyghton et al., 2008). Rather, this histone variant appears to be necessary for ES cell differentiation. In contrast to its localization in pluripotent cells, H2AZ appears to be highly enriched at active genes in lineage-committed cells (Barks et al., 2007; Creyghton et al., 2008) suggesting that its re-distribution is important for exit from the pluripotent state and for cell fate determination. The role of this variant in lineage-committed cells and whether it also plays a role at bivalent genes, however, remain open questions. It will also be of interest to investigate the developmental roles of the numerous histone variants. Nonetheless, these results suggest an important functional interaction between H2AZ and PcG proteins that is necessary for the proper execution of developmental programs in ES cells and provide insights into the essential nature of this variant during mammalian development.
3. Histone modifiers and chromatin remodelers shape the stem cell genome

Histone modifications are the result of the activities of an expansive group of histone-modifying enzymes that catalyze the addition and removal of chemical moieties and that often function coordinately to regulate chromatin structure and gene expression patterns. The reversible modification of histones may be one mechanism that could allow for rapid changes in expression states without the need for multiple rounds of DNA replication, enabling the cell to respond appropriately to developmental cues. Loss of many of these regulators results in early embryonic lethality and has been implicated in a wide-range of disease processes suggesting an important role for regulating chromatin states during development (Bhaumik et al., 2007; Ko et al., 2008). Chromatin structure can also be modified by macromolecular machines that utilize the energy from ATP-hydrolysis to alter DNA accessibility by catalyzing the displacement or redistribution of nucleosomes (Henikoff, 2008). In many cases histone modifiers and chromatin remodelers function cooperatively with key transcription factors to establish and maintain proper chromatin and gene expression states. Studies have revealed cell type specific functions and important roles for many of these factors in regulating the balance between the undifferentiated state and lineage commitment as well as in development and disease. While there are a growing number of studies about how chromatin regulators may contribute to our understanding of pluripotency and lineage-commitment in mammals, the following discussion as summarized in Table 2 is restricted to some of the most recent examples.

3.1. Histone lysine methyltransferases and demethylases

Histone lysine methyltransferases typically harbor an evolutionarily conserved SET domain that is responsible for its activity (Dillon et al., 2005; Martin and Zhang, 2005). Polycomb and Trithorax group proteins are two of the most studied complexes that catalyze the methylation of histone proteins. The role of Polycomb and Trithorax group proteins in regulating developmental gene expression patterns in metazoans has been well defined by classical genetics, yet it is only recently that they have been implicated as regulators of gene expression by covalent modification of chromatin structure as discussed in Section 2.3 (Ringrose and Paro, 2004; Schuettengruber et al., 2007; Schwartz and Pirrotta, 2007). It is thought that the concerted action of this class of regulators is important for lineage commitment during ES cell differentiation and for maintenance of adult stem cells.

ES cell differentiation is accompanied by changes in chromatin modification patterns and an increase in heterochromatin formation as characterized in part by enrichment of H3K9 di- and tri-methylation. Thus, the levels and distribution of these modifications would need to be highly regulated in order to maintain ES cells in a state of self-renewal (Figure 3). G9a catalyzes the tri-methylation of histone H3 at lysine 9 in euchromatic regions (Feldman et al., 2006) and is essential for embryonic development (Tachibana et al., 2002). It has recently been shown that G9a facilitates the heterochromatinization of numerous early embryonic genes including Oct4, by methylation of histones and possibly through subsequent recruitment of DNA methyltransferases to protect against reactivation of these genes during differentiation (Epsztejn-Litman et al., 2008). Thus, G9a may be an important gatekeeper that mediates the transition from a pluripotent to a lineage-committed state. Another Histone H3-K9 methyltransferase, ESET/Setdb1, is also essential for early development and ES cell survival and may be required at an earlier stage than G9a (Dodge et al., 2004), yet little else is known about its precise role during mammalian development.

Until recently histone methylation was thought to be a chemically stable modification, but it is now known that a broadening class of jumonji domain-containing histone demethylases can remove this mark (Agger et al., 2008; Swigut and Wysocka, 2007). The identification of demethylases that remove methyl moieties from either modified H3K4 or H3K27 is of particular relevance to understanding the regulation of bivalent domains in ES cells as well as the transition in gene expression states during development. For example, Jmjd3 and Utx are H3K27me2/me3 demethylases (Agger et al., 2007; Hong et al., 2007; Lan et al., 2007; Lee et al., 2007; Xiang et al., 2007) and thus would be predicted to have important roles in counteracting Polycomb-mediated gene silencing and histone H2A ubiquitinylation. Consistent with this, these enzymes are necessary for proper development and differentiation in a variety of systems and are targeted to developmental regulators such as HOX genes during differentiation where they become activated in a spatio-temporal manner (Agger et al., 2007; De Santa et al., 2007; Hong et al., 2007; Lan et al., 2007; Lee et al., 2007; Sen et al., 2008). Moreover, Jmjd3 has been linked to the resolution of bivalent chromatin marks upon activation of a developmental pathway such as neural commitment (Burgold et al., 2008). These findings suggest a mechanism by which silent genes in ES cells can become rapidly activated in response to developmental cues.

RBP2 (Jarid1a) is also a member of the jumonji domain-containing demethylase family that has been implicated in development and in mediating cell fate transitions (Lopez-Bigas et al., 2008; Pasini et al., 2008). Rbp2/Jarid1a
Table 2. Chromatin modifying factors with known roles in stem cells and during mammalian development

<table>
<thead>
<tr>
<th>Class</th>
<th>Representative Examples</th>
<th>Cellular Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histone lysine methylation and demethylation</td>
<td>G9a (H3K9 methyltransferase)</td>
<td>Heterochromatin formation; required for repression of key pluripotency genes during ES cell differentiation; null mice are embryonic lethal</td>
<td>Epsztejn-Litman et al., 2008; Feldman et al., 2006; Tachibana et al., 2002</td>
</tr>
<tr>
<td></td>
<td>Jmjd1a/Jmjd2c (H3K9 demethylases)</td>
<td>Maintenance of pluripotency; loss of function leads to inappropriate differentiation</td>
<td>Loh et al., 2007</td>
</tr>
<tr>
<td></td>
<td>Rbp2 (H3K4me3 demethylase)</td>
<td>Repression of developmental regulators in ES cells; cell fate transitions</td>
<td>Lopez-Bigas et al., 2008; Pasini et al., 2008</td>
</tr>
<tr>
<td></td>
<td>Jmjd3, Utx (H3K27me3 demethylases)</td>
<td>resolution of bivalent domains in neuronal lineage; necessary for cellular differentiation and body axis patterning; cell fate transitions</td>
<td>Agger et al., 2007; Burgold et al., 2008; Hong et al., 2007; Lan et al., 2007; Lee et al., 2007; Sen et al., 2008; Xiang et al., 2007</td>
</tr>
<tr>
<td>Histone acetylation and deacetylation</td>
<td>Tip60/p400 complex</td>
<td>Transcriptional activity; maintenance of ES cell identity</td>
<td>Fazzio et al., 2008</td>
</tr>
<tr>
<td></td>
<td>HDAC1 (histone deacetylase)</td>
<td>Transcriptional repression; developmental roles in early embryogenesis; deficiency leads to defects in organogenesis while gain of function confers a cancerous state</td>
<td>Ma and Schultz, 2008; Weichert et al., 2008</td>
</tr>
<tr>
<td></td>
<td>BAF250A</td>
<td>Stem cell self-renewal and pluripotency; roles in early embryogenesis</td>
<td>Gao et al., 2008; Yan et al., 2008</td>
</tr>
<tr>
<td></td>
<td>BAF250B</td>
<td>Stem cell self-renewal and pluripotency; roles in early embryogenesis</td>
<td>Gao et al., 2008; Yan et al., 2008</td>
</tr>
<tr>
<td></td>
<td>BRG1</td>
<td>Catalytic subunit of mammalian SWI/SNF complex; depletion promotes loss of self-renewal and pluripotency in ES cells and induce differentiation; essential for proper development; implicated in tumorigenesis</td>
<td>Ho et al., 2009a; Ho et al., 2009b; Kidder et al., 2008; Roberts and Orkin, 2004</td>
</tr>
</tbody>
</table>

Demethylates histone H3 on Lysine 4 and is thought to counteract the activity of the JMJD3/UTX-Trithorax system. Consistent with this, Rbp2 is directly recruited by PRC2 to a common set of bivalent promoters in ES cells. Moreover, PRC2 and RBP2 co-enrichment is not observed at these sites upon ES cell differentiation (Pasini et al., 2008). Thus, the coordinate removal of activating modifications and incorporation of repressive ones suggests a mechanism for rapid silencing of target genes during development.

Two other Jumonji-domain containing proteins, Jmjd1a and Jmjd2c, have been identified as histone H3 Lys9 specific demethylases whose activities are important for regulating self-renewal (Loh et al., 2007). Consistent with this, their depletion leads to induction of differentiation programs and loss of pluripotency. At the molecular level, Jmjd2c protects the pluripotency regulator Nanog from H3K9 methylation and heterochromatin formation in ES cells. Moreover, the expression of Jmjd2c appears to be positively regulated by Oct4, an essential pluripotency regulator that becomes inactivated upon differentiation, in part, by G9a-mediated H3K9 methylation. Thus, G9a and Jmjd1a/2c may perform mutually antagonistic functions that are necessary for the proper regulation of self-renewal and differentiation. Notably, the oncogene GASC1/Jmjd2 has also been identified as a histone lysine 9 demethylase similar to Jmjd1a and Jmjd2c and has been suggested to play a role in cancer development (Cloos et al., 2006). Thus, it will be of interest to
The chromatin signature of pluripotent cells

Figure 3. Histone methyltransferases and demethylases function antagonistically to control gene expression patterns during ES cell differentiation. In pluripotent cells, the essential pluripotency gene Oct4 is maintained in an active state by the histone H3 K9 demethylases Jmjd1a and Jmjdc2 whose expression is, in turn, regulated by Oct4 to maintain self-renewal. Upon differentiation, G9a tri-methylates histone H3 lysine 9 in the promoter region of Oct4 which results in recruitment of HP1 and DNA methyltransferases, leading to stable gene repression and exit from pluripotency.

determine whether it also plays a role in development or in ES cell differentiation. Collectively, these studies provide an important functional link between the ES cell transcriptional regulatory circuitry as governed by key transcription factors and chromatin modulation by histone modifiers.

3.2. Histone acetyltransferases (HATs) and deacetylases (HDACs)

Histone acetylation has been implicated in diverse biological processes, including transcriptional activity, cell cycle progression, and DNA damage response (Bartova et al., 2008; Yang and Seto, 2008a). Historically associated with gene activation, histone acetylation is now recognized to have context-dependent effects. For example, the Tip60/p400 complex belongs to the MYST family of histone acetyltransferases, which has been linked to both transcriptional activation and silencing (Pillus, 2008). A recent RNAi screen in ES cells identified several of its subunits as essential for stem cell identity (Fazzio et al., 2008). Genome-wide analysis in ES cells showed that Tip60/p400 is recruited to the promoters of both active and inactive genes, and that the Trithorax-catalyzed H3K4me3 mark is necessary for this localization. The complex appears to be required for repression of genes that are involved in differentiation and development, as depletion of components of the complex caused de-repression of target genes (Fazzio et al., 2008). Notably, the effects of Tip60/p400 depletion are similar to that observed upon loss of the transcription factor Nanog, a master regulator of ES cell function. The nature of the relationship between Tip60/p400 and Nanog, however, remains undefined. Two other notable HATs, p300 and Gcn5, also have essential functions during development (Bu et al., 2007; Phan et al., 2005; Yao et al., 1998). Interestingly, p300 appears to co-occur with the three pluripotency regulators Oct4, Sox2, and Nanog as analyzed by genome-wide locations analysis (Chen et al., 2008). The interaction between p300 with genes that are also targets of the ES cell-specific transcription factor regulatory network suggests that this HAT likely plays an important, albeit unknown, role in stem cell function.

Just as maintenance of histone acetylation patterns is important for ES cell differentiation, the reverse process of histone deacetylation is also necessary for proper development. Removal of the acetyl group from histones is thought to confer a closed chromatin conformation that is not permissive for transcription. In mammals, multiple classes of histone deacetylases (HDACs) exist including the class 1 histone deacetylases (Hdac1). In particular, Hdac1 appears to play a major role in pre-implantation embryogenesis in establishing a repressive chromatin state (Thiagalingam et al., 2003; Yang and Seto, 2008b). Its precise role in ES cells or other pluripotent cell types, however, remains to be defined and it will be of interest to identify Hdac1 specific target genes and to elucidate its effect on chromatin structure during ES cell differentiation.
Methyl binding domain 3 (Mbd3) is a member of a family of methyl-CpG-binding domain proteins that are thought to be downstream effectors of DNA methylation. Mbd3 is a core component of the Nucleosome Remodeling and Deacetylase (NuRD) complex, although Mbd3 itself does not bind the meCpG sites. NuRD harbors both histone deacetylase (Hdac1 and Hdac2) and ATP-dependent chromatin remodeling (Mi-2) activities (Xue et al., 1998), and components of the complex have specific developmental roles rather than broad roles in most cellular processes. Consistent with this, Mbd3 is essential for early embryogenesis (Hendrich et al., 2001; Kaji et al., 2007) and plays a critical role in ES cell differentiation (Kaji et al., 2006; Kaji et al., 2007). Recent studies indicate that Mbd3 may be involved in an epigenetic silencing pathway that is necessary for the proper regulation of gene expression programs in vivo and during ES cell differentiation where it is necessary for lineage commitment (Kaji et al., 2006; Kaji et al., 2007).

An alternate NuRD complex (NODE – Nanog and Oct4 associated deacetylases) that contains Mta1 instead of Mbd3 has been recently isolated based on its interaction with Oct4 and Nanog in ES cells (Liang et al., 2008). While Mbd3- and Mta1-containing complexes exhibit comparable levels of histone deacetylase activity (Liang et al., 2008; Saito and Ishikawa, 2002), Oct4 and Nanog interact preferentially with NODE. It has been shown that Oct4 and Nanog occupy the promoters of developmental genes that are silent in ES cells (Boyer et al., 2005). Consistent with its interaction with the two pluripotency transcription factors, ablation of Mta1 in ES cells led to de-repression of developmental regulators and inappropriate differentiation (Liang et al., 2008). Collectively these studies reveal another layer of regulation necessary for the control of developmental progression and exit from the pluripotent state.

3.3. ATP-dependent chromatin remodeling

Numerous ATP-dependent chromatin-remodeling factors reside in multi-subunit complexes that act on the chromatin template and that have roles in early development and cell fate specification (Ko et al., 2008; Wang et al., 2007). While histone modifications can alter the stability of histone-histone or histone-DNA contacts, chromatin remodelers are capable of mobilizing or displacing nucleosomes, which can lead to profound consequences on gene expression. Consequently, ATP-dependent chromatin remodeling complexes may play an important role in mediating cell fate decisions. On the basis of domain structure, the ATP-dependent remodeling factors can be grouped into four families (SWI/SNF, ISWI, Mi-2/CHD, and INO80), with each family having broad functions in diverse biological processes and cell types (Boyer et al., 2000; Wang et al., 2007). The action of ATP-dependent remodelers has been linked to a variety of DNA-mediated processes including both transcriptional activation and repression and emerging evidence indicates the regulation of nucleosome positioning may have important consequences on transcriptional states (Henikoff, 2008). Interestingly, many components of these complexes have been genetically identified as members of the Trithorax group and mutations are early embryonic lethal, suggesting they function to control developmental gene expression programs. Therefore, ATP-dependent nucleosome remodeling may represent an additional mechanism to regulate developmental programs.

While their biochemical properties have been characterized in vitro, little is known about how ATP-dependent remodelers influence chromatin states in vivo and how their activities may contribute to lineage commitment. For example, it is known that Brg1, the ATPase subunit of the mammalian SWI/SNF-related or BAF (Brg/Brahma-associated factors) complex, is necessary for proper development, but its function during development or during ES cell differentiation has not been defined. Recent studies have revealed that the composition of the BAF complex varies during development (Lessard et al., 2007; Yan et al., 2008) and that an ES cell specific BAF (esBAF) complex is required for pluripotency and self-renewal (Ho et al., 2009a). Furthermore, genome-wide analysis of Brg1 revealed its enrichment at the promoter regions of genes that are also occupied by the three pluripotency regulators Oct4, Sox2, and Nanog, but not with components of the PRC2 complex (Ho et al., 2009b). A second genomic study also showed that Brg1 co-occupied many of the same genes as Oct4, Sox2, and Nanog, but also observed its association with Polycomb group target genes (Kidder et al., 2008). Despite these differences, it is clear that both of the studies connect esBAF to the transcriptional regulatory circuitry in ES cells and possibly to key signaling pathways such as those mediated by Lif and Bmp. Baf250a and Baf250b, which are two related genes that code for subunits of the developmentally regulated BAF complex, also play a role in development and stem cell self-renewal (Gao et al., 2008; Yan et al., 2008). Clearly much work needs to be done to further understand the contribution of these and other remodeling factors to the organization of chromatin states and to elucidate how their collective activities impact development. Nonetheless, these findings are consistent with the role of this class of chromatin remodelers with the maintenance of self-renewal and the ES cell state.
4. Chromatin and reprogramming cell fate

In the past two years, multiple studies have reported the reprogramming of differentiated cells into an ES cell-like state using a combination of only a few transcription factors, including Oct4, Sox2, Klf4, Myc, Nanog, and Lin-28 (Aoi et al., 2008; Hanna et al., 2007; Lowry et al., 2008; Meissner et al., 2007; Nakagawa et al., 2008; Okita et al., 2007; Park et al., 2008; Takahashi et al., 2007a; Takahashi et al., 2007b; Takahashi and Yamanaka, 2006; Wernig et al., 2008). These induced pluripotent stem (iPS) cells exhibit features of ES cells, including a characteristic morphology, the expression of pluripotent markers, self-renewal, and the capacity to give rise to all tissues of the three embryonic germ layers (Jaenisch and Young, 2008; Lewitzky and Yamanaka, 2007; Welstead et al., 2008; Yamanaka, 2008). Aside from the important clinical implications of iPS cells in generating histocompatible tissue for patients, the generation of iPS cells provides a remarkable opportunity to study the interplay between the transcription factor network and chromatin organization in pluripotency and cell fate specification.

Several groups have carried out systematic investigations of the reprogramming process and have characterized the transcriptome and chromatin profiles of iPS cells compared to stable intermediates that exhibit some, but not all of the characteristics of pluripotent cells (Brambrink et al., 2008; Maherali et al., 2007; Mikkelsen et al., 2008; Stadtfeld et al., 2008). For example, it appears that bivalent marks are re-established at the proper gene promoters in iPS cells (Maherali et al., 2007; Mikkelsen et al., 2008). Conversely, while partially reprogrammed cell lines can be stably propagated, these cells are ensnared in an intermediate state due to the incomplete re-activation and repression of pluripotency and lineage-specific programs, respectively. These stable intermediates display differences in their chromatin profiles as compared to both their cell type of origin and that of ES cells. This clearly supports the idea that epigenetic reprogramming is essential for cell fate transitions. Future investigations that clearly define the molecular mechanisms responsible for this process are expected to enhance our ability to direct the differentiation of cells toward specific lineages both in vitro and in vivo.

The recent findings from these studies reaffirm the idea that unique chromatin signatures correlate with cell type and function, and that a given chromatin state is a balance of multiple epigenetic marks and implies that cell fate can be experimentally manipulated through tipping this balance. Indeed, it has recently been demonstrated that reprogramming may be achieved with only Oct4 and Sox2 if fibroblasts are also treated with valproic acid, a potent histone deacetylase inhibitor (Huangfu et al., 2008). Thus, the induction of a hyper-acetylated state (indicative of open chromatin structure) would be rapidly achieved and would likely facilitate transcription factor access and ultimately reprogramming. Also consistent with this notion is the finding that the DNA methyltransferase inhibitor 5-aza-cytidine enhances reprogramming efficiency, presumably by counteracting its gene silencing effects (Mikkelsen et al., 2008). Moreover, it has been reported that the small molecule inhibitor of G9a (BIX-01294), which substantially and reversibly reduces global levels of H3K9me2 (Kubicek et al., 2007), in combination with BayK8644 can substitute for Sox2 and improves reprogramming efficiency (Shi et al., 2008). Despite the successful derivation of iPS cells from both mouse and human cell types, the process is remarkably inefficient, suggesting that the incomplete reprogramming of the somatic genome is a major hurdle for iPS research. While many questions remain, these studies should allow for a better understanding of how dynamic chromatin states contribute to cell identity.

Collectively, the profiling of chromatin states in ES cells, iPS intermediates, and in fully reprogrammed iPS cells has provided insights into how chromatin may influence the impact of the epigenome on cell state.

5. Perspectives and Future Directions

Undifferentiated, pluripotent Pluripotent cells possess the dual capacity for self-renewal and lineage commitment. Recent studies have revealed key differences in chromatin states of pluripotent cells as compared to differentiated cell types (Table 3). It is thought that these fundamental distinctions can provide information about cell type and function. Most genes are associated with one of many patterns of chromatin modifications, but the patterns themselves do not necessarily inform us how that gene is regulated or how its expression state is inherited during differentiation. Most histone modifications are highly dynamic in response to developmental cues and to changes in the cellular environment. One outcome of the post-genomics era has been the capacity for investigators to view the fundamental process of cell fate specification from a global perspective. It is now possible to combine information about chromatin states with global views of regulatory small RNAs, transcription factor networks, signal transduction, and DNA methylation into an integrative picture of the molecular mechanisms that govern cell fate specification. Thus, the nature of chromatin organization and how it mediates cell type and function is but one of many pieces of the puzzle essential for elucidating the fundamental mechanisms that govern developmental processes.
The chromatin signature of pluripotent cells

Table 3. Chromatin states in pluripotent and lineage-committed cells.

<table>
<thead>
<tr>
<th>Pluripotent cells</th>
<th>Lineage-committed cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperdynamic chromatin structure consistent with open chromatin (e.g. hyperacetylation of histone H3)</td>
<td>Increase in heterochromatin formation consistent with less accessible chromatin (e.g. histone deacetylation)</td>
</tr>
<tr>
<td>Transcriptionally permissive</td>
<td>More permanent repression of pluripotency genes and inappropriate developmental programs</td>
</tr>
<tr>
<td>Bivalent chromatin domains (e.g. H3K4me3 and H3K27me3)</td>
<td>Resolution, establishment, or maintenance of bivalent chromatin domains</td>
</tr>
<tr>
<td>Co-enrichment of H2AZ and Polycomb Complexes at bivalent genes</td>
<td>H2AZ is highly enriched at active genes independent of Polycomb Complexes</td>
</tr>
</tbody>
</table>

In the clinical setting, chromatin signature maps may help in understanding the initiation, progression, and diagnosis of diseases with epigenetic etiologies, such as cancer and certain imprinting disorders. In this regard, cancer is particularly relevant as there are important parallels between the biology of stem cells and the emergence of cancer. Combined with existing methodologies such as expression profiling and histology, the chromatin signature of a particular cancer subtype can be better identified and targeted treatments could be rationally designed with this information in hand. In the field of regenerative medicine, chromatin state maps could guide the development of patient-specific stem cells. This would bypass ethical as well as technical issues in the use of stem cells for tissue therapy. Hence, investigation of the relationship between chromatin organization and cell function is relevant for understanding one of the most fundamental problems in biology, that is, how a single cell can specify an entire multi-cellular organism.

6. Acknowledgements

We are grateful to members of the Boyer lab for thoughtful discussions and critical review of the manuscript and especially to Seraphim Thornton for help with the illustrations. We are thankful to Shlomo Meislin and Matthew I. Rasmussen for assistance with Figure 1. We apologize to the authors whose work we were unable discuss due to space limitations. LAB received support from the Pew Scholars Program in the Biomedical Sciences and the Massachusetts Life Sciences Center.

7. References


The chromatin signature of pluripotent cells


The chromatin signature of pluripotent cells


stembook.org
The chromatin signature of pluripotent cells


The chromatin signature of pluripotent cells


The chromatin signature of pluripotent cells


The chromatin signature of pluripotent cells


The chromatin signature of pluripotent cells


