UC San Diego
UC San Diego Previously Published Works

Title
The epigenome in pluripotency and differentiation

Permalink
https://escholarship.org/uc/item/5pw2h4j0

Journal
Epigenomics, 6(1)

ISSN
1750-1911

Authors
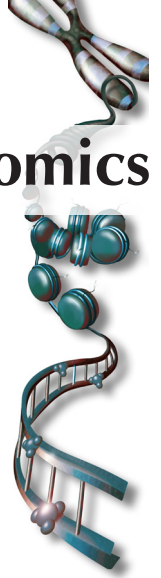Thiagarajan, Rathi D
Morey, Robert
Laurent, Louise C

Publication Date
2014-02-01

DOI
10.2217/epi.13.80

Peer reviewed

The epigenome in pluripotency and differentiation

The ability to culture pluripotent stem cells and direct their differentiation into specific cell types *in vitro* provides a valuable experimental system for modeling pluripotency, development and cellular differentiation. High-throughput profiling of the transcriptomes and epigenomes of pluripotent stem cells and their differentiated derivatives has led to identification of patterns characteristic of each cell type, discovery of new regulatory features in the epigenome and early insights into the complexity of dynamic interactions among regulatory elements. This work has also revealed potential limitations of the use of pluripotent stem cells as *in vitro* models of developmental events, due to epigenetic variability among different pluripotent stem cell lines and epigenetic instability during derivation and culture, particularly at imprinted and X-inactivated loci. This review focuses on the two most well-studied epigenetic mechanisms, DNA methylation and histone modifications, within the context of pluripotency and differentiation.

Rathi D Thiagarajan[1], Robert Morey[1] & Louise C Laurent*[1]
[1]Department of Reproductive Medicine, The University of California, San Diego, La Jolla, CA, USA
*Author for correspondence: llaurent@ucsd.edu

Early mammalian development involves precise orchestration of gene expression in a spatial and temporal manner in order to establish cell lineage fate. Starting from a totipotent state, cells fated to the embryonic lineages pass through a pluripotent state and then branch off into the germline and the three germ cell lineages: the ectoderm, mesoderm and endoderm. Multipotent progenitor cells in these major lineages then differentiate further to produce more than 200 specialized cell types in the fully developed organism.

The differentiation process is accompanied by changes in the transcriptome, and much has been learned about the signals that govern it by performing gene expression analyses of differentiating and differentiated cells. Such experiments have demonstrated the critical role that transcription factors play in the regulation of temporal and spatial gene expression programs by binding to *cis*-regulatory regions in response to environmental cues. There is an increasing appreciation of the role of epi-genetic mechanisms in the regulation of the transcriptome, and recognition that changes in the epigenome during differentiation can point to genomic features that play key roles in the differentiation process [1].

In virtually every cell type and organ system, normal differentiation and development is associated with characteristic changes in epigenetic patterns, which allow for the establishment and stable maintenance of a wide variety of cellular phenotypes without alterations to the genome (recently reviewed in [2]). In terms of disease, epigenetic aberrations have been shown to result in developmental abnormalities, degenerative disease and cancer (also reviewed in [3–6])[7–9]. In the most inclusive definition, epigenetics is the study of mechanisms that change gene activity without altering the DNA sequence, and thereby include not only DNA methylation and histone modifications, but also transcription factors and noncoding RNAs. In fact, there has been rather extensive debate regard-

ing the definition of epigenetics, and in particular how heritable a mark must be to be considered epigenetic. Recent advances in the fields of cell biology and epigenomics have demonstrated that many epigenomic marks are less stable than previously believed, and that certain sequences of epigenetic events that were once thought to be irreversible (e.g., progression from pluripotent or multipotent stem cells to fully differentiated mature cell types) can in fact be reversed using cellular reprogramming techniques [10–12].

Human pluripotent stem cells (PSCs) have tremendous potential for *in vitro* modeling of development and disease due to their self-renewal properties and their ability to differentiate into all cell lineages in the body. Human embryonic stem cells (hESCs) [13], which are derived from the inner cell mass of blastocyst-stage preimplantation embryos, are considered the gold standard PSCs. Two other types of PSCs have been produced: induced pluripotent stem cells (iPSCs), which are generated from somatic cells by overexpression of a small number of reprogramming factors [10,11]; and somatic cell nuclear transfer ESCs (SCNT-ESCs), which are derived from embryos resulting from the introduction of a somatic cell nucleus into an enucleated oocyte [14]. In the blastocyst, the pluripotent state represented by the cells in the inner cell mass is transient. *In vitro* however, the pluripotent state as represented by PSCs can be artificially maintained for an indefinite period of time. This greatly increases the utility of the cells, but also makes them susceptible to genomic and epigenomic aberrations [15–18]. It is therefore important to assess how such aberrations may impact the efficacy of PSCs in accurately modeling development and disease, and serving as sources of material for cell therapy.

The dynamic interplay between the transcriptome and epigenome warrants implementation of genome-wide approaches and high-throughput technologies to identify and characterize mechanisms and molecular factors that regulate genome function during development, lineage-specification and disease pathology. In this article, we describe the current genome-wide profiling platforms and large-scale consortiums that have utilized these technologies to understand the epigenome, and also review the existing literature with an emphasis on publications that have utilized genome-wide approaches to describe: the epigenetic landscape of pluripotency and cellular differentiation, as modeled by undifferentiated and differentiated PSCs; differences among different types of PSCs; differences between *in vitro* maintained ESCs and cells in the developing embryo; and the dynamic changes in the epigenome that occur during reprogramming. We will focus on two of the most well-studied epigenetic mechanisms: DNA methylation and histone modifications, and will only briefly touch on noncoding RNAs (as we have recently written a review on the role of miRNAs in pluripotency [19]).

The epigenome in mammalian development

DNA methylation is an essential mechanism that has been shown to play a key role in both gene regulation in the context of genomic imprinting, X chromosome inactivation and regulation of some autosomal nonimprinted genes and preservation of genomic stability. It is becoming increasingly clear that DNA methylation at different genomic loci is regulated by several different mechanisms and involves the interplay between DNA methyltransferases (DNMTs), DNA dioxygenases, cytidine deaminases and various DNA-binding proteins. These processes are particularly active during early development, and evidence is growing that they are important for the establishment and stabilization of the pluripotent state.

DNMT1 is the primary maintenance DNA methyltransferase, and is responsible for converting the hemi-methylated CpG dinucleotides that are generated during DNA synthesis into fully methylated CpGs. DNMT3A and DNMT3B are *de novo* DNA methyltransferases, which can establish new DNA methylation at completely unmethylated CpG sites, as well as methylate cytosines located at non-CpG sites. Maintenance of DNA methyltranferase activity has been shown to be critical for development. Mice with homozygous mutations in *Dnmt1*, resulting in partial to complete loss of Dnmt1 function, were embryonic lethal by midgestation and were associated with imprinting defects [20]. Mouse *Dnmt3a* knockouts are postnatal lethal and associated with imprinting defects [21], while *Dnmt3b* knockouts are lethal in the late embryonic period and showed defects in methylation of endogenous viral and satellite DNA sequences [21]. *Dnmt3a/b* double knockouts have a more severe phenotype, dying at midgestation, indicating that there is partial redundancy between these two DNA methyltransferases [20,22].

It has been shown that the absence of all three DNA methyltransferases in mouse TKO ESCs results in near-total demethylation of genomic DNA, but does not affect self-renewal, expression of pluripotency markers or global chromatin structure [23]. However, in contrast to wild-type mouse ESCs or mouse ESCs lacking Dnmt1 only, the mouse TKO ESCs do not stably differentiate when induced to form embryoid bodies, and readily revert to a pluripotent phenotype when returned to culture conditions favorable for undifferentiated PSCs [24]. These results indicate that DNA methylation is dispensable for maintenance of the undifferentiated state, but is necessary for lineage commitment.

Differential methylation of CpG sites in CpG islands [25] and those distributed more sparsely across the genome is established during development by global demethylation followed by selective remethylation. Following widespread erasure of DNA methylation marks during early development, *de novo* DNMTs are involved in global remethylation of the genome. During this process, most CpG islands are excluded from DNA methylation [26]. It has been shown that the primary mechanism for protection of CpG islands from methylation is through proximal *cis*-regulatory regions, such as transcription factor binding sites [26,27]. R-loops also protect CpG islands from methylation. R-loops result from the fact that CpG islands are skewed toward having one strand of the double helix rich in guanine, and the complementary strand rich in cytosine (GC skew). During transcription, the nascent G-rich RNA segments bind to the C-rich DNA strand, which forces the G-rich DNA strand to form a loop, thus protecting these sequences from the *de novo* DNA methyltransferases [28].

Early mammalian development, including the period from gametogenesis through the specification of the three major cell lineages, is marked by several major epigenetic changes – imprinting and X chromosome inactivation being the most heavily studied [1]. Studies in this area have been challenging due to the limited numbers of cells in the early embryo and the differences among species, which prevent generalizations from model organisms to the human system [29,30].

Imprinting refers to the inactivation of one of the two alleles at a given genomic locus in a parent-of-origin-specific manner, and is established during gametogenesis. Following fertilization, human zygotes undergo a wave of global DNA demethylation, but DNA methylation status at imprinting control regions (ICRs) remains unchanged, preserving parental allele-specific DNA methylation at these sites and consequently parent-of-origin-specific imprinted gene expression [31]. Imprinted regions are characterized by expression of nearby long noncoding RNAs (lncRNAs) [32], specific histone modifications [33,34], and DNA methylation. Aberrant imprinting in humans has been correlated with several developmental diseases, such as Beckwith-Wiedemann, Prader-Willi and Angelman syndromes, as well as malignancies, such as Wilms' tumor [35–37]. Most imprinted genes are found in clusters, with members of a cluster regulated simultaneously by a common ICR [38]. ICRs were traditionally thought to be established and maintained through DNA methylation, but recent reports suggest that other epigenetic modifiers are involved [39]. These reports include histone modification features that may signal the specificity of imprint acquisition during spermatogenesis and

may be critical for establishing the DNA methylation imprints during oogenesis [33,34].

How imprinted loci retain their parent-of-origin specific gene expression during development has yet to be fully elucidated. In murine models, it appears that the protection and maintenance of imprinted marks during global demethylation events in pre-implantation embryos, and in subsequent differentiation events in development, depends on an epigenetic modifier complex formed by ZFP57 and TRIM28 [40–42]. The zinc-finger protein ZFP57 recognizes and binds to the methylated hexanucleotide motif (TGCCGC) in ICRs, and recruits TRIM28 and the DNMT3s to maintain DNA methylation and heterochromatinization at many imprinted domains across the genome [43]. How accurately the ZFP57/TRIM28/DNMT3 complex prevents demethylation of ICRs during demethylation events, and if other components or complexes are involved in the protection and maintenance of imprinting, has yet to be determined. Mouse models have also shown that the PGC7 complex has a role in protecting the maternal sequences from early development demethylation [44]. Interestingly, PGC7 was also shown to interact with histone modification factor, H3K9me2, highlighting cooperative interplay between the methylome and histone modifications to prevent demethylation [45]. How DNA methylation is initially established at ICRs remains an outstanding question in the field.

Another level of epigenetic control of gene expression is through X chromosome inactivation (XCI). XCI results in the transcriptional silencing of one of the two X chromosomes in female cells, thereby balancing the gene dosage from the X chromosome between female and male cells. The mechanism by which one of the two X chromosomes is inactivated involves extensive changes in histone modifications, DNA methylation changes and several noncoding RNAs [38,46]. The most well characterized of these critical noncoding RNAs is X-inactive specific transcript (XIST), which is expressed from one of the two X chromosomes in female cells. XCI regulation by the XIST transcript is a paradigm for epigenetic silencing mediated by noncoding RNAs. XIST acts in *cis* to coat the X chromosome from which it is expressed, mediating the initiation and maintenance of XCI by recruiting histone-modifying complexes and by inducing nuclear reorganization [47]. In the mouse, XCI is initiated during the pre-implantation period with silencing of the paternal X chromosome. The paternal X chromosome is then reactivated in the ICM at the blastocyst stage resulting in two active X chromosomes. In the trophoectoderm (TE) lineage however, paternally imprinted XCI is maintained. During gastrulation, random XCI in embryonic cells is initiated by XIST. Although it is believed

that XIST works similarly in mouse and human cells, the mechanisms controlling its expression and the developmental stage at which it is expressed is fundamentally different in murine and human development (recently reviewed in [48,49]). Recent evidence points to potential differences between the primate and murine XCI pathway, where the primate TE lineage does not display imprinted paternal XCI [50].

While XCI is a complex process that is still incompletely understood, application of high-throughput sequencing technology is beginning to yield insights and challenge current paradigms. Using allele-specific sequencing of mouse trophoblast stem cells, Calabrese *et al.* discovered DNAse1 hypersensitive sites on the inactive X chromosome, despite the lack of other markers of transcriptional activity [51]. Additionally, the small fraction of genes on the X chromosome (~15%) that escape X inactivation were found to be located both inside and outside of the domain coated by XIST. These results suggest that there are as-yet undiscovered site-specific regulatory elements that act on the inactive X chromosome.

The epigenomic landscape of human PSCs

Epigenomic profiling of human PSCs has been used to address several important questions. First, are human PSCs epigenetically stable? Second, are there differences in the epigenetic stability of human PSCs derived or cultured under different conditions, and do any observed differences correlate with the pluripotency of the cells? Third, what epigenetic changes occur during reprogramming and differentiation of human PSCs (and do they represent the changes that occur during normal cellular differentiation *in vivo*)?

Aberrant DNA methylation at imprinted loci in hESCs and human iPSCs (hiPSCs) has recently been shown to be both widespread and variable [52,53]. Comprehensive DNA methylation profiling of large numbers of PSC lines revealed that hypermethylation of the *PEG3*, *MEG3* and *H19* loci were found to be particularly prevalent and associated with downregulation of gene expression in human PSC lines [52], in contrast to earlier studies using targeted gene expression assays on smaller numbers of cell lines, which reported stable monoallelic expression for these genes in hPSCs [54–56]. It is unclear whether or not *in vitro* culture conditions are to blame for the observed imprinting aberrations, as they were evident even in low passage hESCs [53]. Another study showed that the reprogramming of somatic cells to pluripotency did not affect the stability of the vast majority of differentially methylated regions (DMRs), which are regions that are differentially methylated between different cell types and are often associated with regulatory regions, such as enhanc-

ers [57]. Loss of imprinting in reprogrammed hiPSCs at the *PEG3*, *DIRAS3* and *ZDBF2* loci was also reported, but only the aberrant methylation of *ZDBF2* could be attributed to reprogramming, as the other two genes also exhibited a loss of imprinting in the parental somatic cells [57]. For hPSCs displaying aberrations at imprinted regions, differentiation did not result in correction of the aberrations, or in appearance of new aberrations [52,53].

A majority of human female ESCs show non-random XCI [58]. Female hESCs can be categorized into three different classes based on their XIST expression [59]. Class I hESCs express low levels of XIST, have two active X chromosomes in the pluripotent state and undergo random XCI upon differentiation; this state has been associated with naive pluripotency [52,60–62]. Class II hESCs express XIST at levels similar to those of differentiated cells, and are in the XaXi state (Xa: active; Xi: inactive) with the inactive X being coated by XIST; most reported female hPSCs fall into this class [63–67]. Class III hESCs are frequently in the XaXi state, but have lost XIST expression, so the inactive X chromosome lacks XIST coating. Unlike in Class I cells, XIST expression in Class III cells remains off following differentiation [68]. Currently, a widely accepted model is that newly derived human ESCs start in Class I and contain two active X chromosomes, but due to suboptimal *in vitro* culture conditions, they undergo XCI to become Class II, and some lines later lose XIST expression and fall into Class III [60]. However, experiments in mammalian pre-implantation embryos indicate that the process of XCI in mice and humans is highly divergent, and thus bring this model into question. Unlike in mouse embryos, XIST is expressed in the majority of the cells in the human embryo starting at morula stage, and initially coats both X chromosomes in female cells, even in inner cell mass cells [29].

Several studies detailing the XCI of iPSCs in reprogrammed female human cells have recently been published, with sometimes conflicting results. Some groups have reported that reprogramming does not result in the reactivation of Xi, and that iPSCs receive the Xi from their parent somatic cell [66,67], while others have shown that reprogramming results in the reactivation of Xi [61,69–70]. Furthermore, several recent reports show that hiPSCs are often initially XaXi with XIST expression present, but quickly lose their XIST expression in culture [52,58,66,67,70–72]. When XIST expression is lost, it appears that a subset of Xi-linked genes is activated and this activation is accompanied by the loss of DNA methylation in multiple regions of the X chromosome in a patchy and progressive fashion [52,72]. It has been suggested that certain culture condi-

tions promote the XaXa state. Lengner *et al.* showed that hypoxia could be used to derive hESCs that have not undergone XCI [60]. Tomoda *et al.* reported that reprogramming on LIF-secreting feeder cells often resulted in XaXa hiPSC lines, and that the addition of LIF to the medium for XaXi hiPSCs resulted in reactivation to the XaXa state [70]. Other experiments indicate that HDAC inhibitors [73] or other small molecules [62] may promote conversion of the XaXi state to the XaXa state. Recently, a lncRNA, XACT, that coats the active X chromosome has been identified in humans [74]. It appears as if this lncRNA is not present in mice or in differentiated cell types, and the authors speculate that XACT may be involved in the initiation of XCI in humans.

Although evidence is mounting that reprogramming/derivation methods and culture conditions can influence the XCI status of hPSCs, robust and reproducible methods for control of hPSC XCI have not yet been established. In addition, what the 'normal' state of X chromosome inactivation is in the pluripotent cells in the human embryo still remains to be completely understood. In the meantime, it has been noted that for modeling of diseases caused by X-linked mutations, it is important to monitor the XCI status of the hPSCs over time [52,72].

Soon after iPSC technology was first developed, the question of whether there are systematic differences between iPSCs and hESCs arose. Since the histories of the nuclear genomes of these two types of hPSCs are quite different, it is not unreasonable to think that there might be significant differences in their epigenomes. As will be discussed below, several studies have addressed this issue, but we wish to point out that a common limitation to the existing literature is that sets of well-matched cell lines do not yet exist, for which adequate numbers of hESC and hiPSCs have been genetically matched for potentially confounding variables, such as oxygen tension, media type (including growth factors and small molecule additives), substrate (including feeder cells and type of extracellular matrix), and passage method. The recent success in generation of human SCNT-ESCs has generated further speculation as the history of the nuclear genomes of these cells have similarities with both iPSCs (i.e., being derived from the nucleus of a somatic cell) and ESCs (i.e., being reprogrammed by the cytoplasm of an oocyte, and undergoing preimplantation embryo development from the cleavage to blastocyst stages). In particular, it will be interesting to uncover the magnitude and types of differences between SCNT-ESCs and hESCs, as these may represent the epigenomic changes that take place during gametogenesis and fertilization.

The epigenomics of reprogramming & differentiation

It has been suggested that reprogramming of somatic cells to the pluripotent state occurs in two phases. The first phase involves changes in gene expression and remodeling of histone marks, and the second phase results in the consolidation of pluripotency through changes in DNA methylation and histone modifications [75,76]. Two studies support the notion that remodeling of DNA methylation is the rate-limiting step in the process of complete reprogramming [77,78]. First, the Blau laboratory reported that activation of the *OCT4* and *NANOG* genes occurred much more rapidly during reprogramming accomplished by fusion of somatic cells with PSCs (~1 day) than by the overexpression of standard reprogramming factors (~2 weeks), and that AID was required for this process [77]. Second, 'lifting' of the repression at the *NANOG* promoter by the methyl DNA-binding protein MDB2 was necessary to attain the fully reprogrammed state; paradoxically, MDB2 was shown to be subject to repression by the pluripotency-associated miRNA miR-302 [78]. In terms of histone modifications, a recent genome-wide analysis of the binding sites of pluripotency factors, POU5F1/OCT4, SOX2, KLF4 and c-MYC (OSKM) was conducted during the first 48 h of reprogramming [79]. Genomic regions lacking OSKM binding were identified as loci that were potentially bound by endogenous factors that were impediments to reprogramming. Interestingly, it was found that the repressive H3K9me3 mark was found to be associated with these regions lacking OSKM binding, and was therefore inferred to be an impeding factor [79].

It has been observed that global DNA methylation is higher in hPSCs than in differentiated human cells [9,52,57]. This is true not only at CpG sites, but also for cytosines at non-CpG sites [9,57]. In fact, consistent with an earlier paper reporting non-CpG methylation in mouse ESCs [80] more recent reports have shown that non-CpG methylation was markedly more prevalent in hESCs than somatic cells, accounting for up to 20% of DNA methylation in hESCs [9,57]. It has also been shown that loci that are specifically hypomethylated in certain tissue types are uniformly hypermethylated in hPSCs, and that a subset of these sites become hypomethylated upon *in vitro* differentiation to the relevant lineage [52].

Histone modification patterns can affect the developmental potential of hPSCs. One important histone modification feature among PSCs is bivalent domains. Bivalent domains are marked by poised/active (H3K4me3) and repressive marks (H3K27me3), and are typically associated with transcription factors and genes associated with development and lineage specification [81,82]. This is thought to provide rapid access to lineage speci-

fication factors during differentiation through the loss of H3K27me3. Changes in histone modifications are generally thought to precede DNA methylation during cellular differentiation, as they are more dynamically regulated during differentiation [83]. ESCs have a more open chromatin configuration with abundant active chromatin marks such as H3K4me3 and fewer H3K9 trimethylation regions compared to differentiated cells [84,85]. Other chromatin modifications important for development include histone H3 at lysine 4, 9 and 27 [86].

Chromatin modifications at enhancers are another prevalent feature of differentiation and cell lineage decisions. Beyond the H3K4me3 and H3K27me3 marks, which are predominantly found at promoters, the active enhancer marks, H3K4me1 and H3K27ac, are of increasing interest. Several studies identified the presence of over 7000 enhancers in hESC that can be grouped as either active or poised enhancers and can be distinguished by the H3K27ac mark [87,88]. Recently, it was shown that approximately <1% of enhancers form 50 kb domains, known as 'super-enhancers', which are marked by epigenetic regulatory features such as the Mediator coactivator complex [89], H3K4me1 and H3K27ac. These super-enhancers are frequently associated with transcription factors that have been identified as critical regulators of cell fate [90]. These results highlight an important interaction between transcription factors and global chromatin engagement.

Several studies in a variety of biological systems have suggested that regulation of gene expression by DNA methylation and histone modifications is a complex combinatorial process [91], wherein some histone marks are well-correlated with DNA methylation, while others regulate sets of genes that do not appear to be regulated by DNA methylation [92], and yet others show complex relationships that seem to be context-dependent. One of the most interesting interactions is between DNA methylation and the H3K27me3 mark. As mentioned above, in mouse ESCs, it was shown that repressive H3K27me3 marks were frequently colocalized with activating H3K4me3 marks, resulting in a 'poised' bivalent state [83], which resolves into the fully active or fully repressed state upon differentiation [82,93]. In HeLa cells, the polycomb group protein EZH2, which catalyzes the methylation of H3K27, physically associates with DNA methyltransferases and promotes DNA methylation [94]. Furthermore, studies in mouse ESCs showed that DNA methylation prevents the placement of H3K27me3 marks [95,96], and recently, a study was published that demonstrated a correlation between 5-hydroxymethylcytosine and H3K27me3 in a variety of cell types [97]. Taken together, these studies suggest that there may be a temporal hierarchy, in which the polycomb repressive complex first places the repressive H3K27me3 mark, and then recruits DNA methyltransferases to solidify repression of gene expression. In this repressed state, DNA methylation inhibits the placement of new H3K27me3 marks. However, conversion of 5-methylcytosine to 5-hydroxymethylcytosine appears to again allow the placement of H3K27me3 marks.

Two recent studies have reported on integrative analysis of DNA methylation, histone modification and transcriptome data in hESCs in the undifferentiated state and during early *in vitro* differentiation toward the three germ lineages [98,99]. In the Gifford *et al.* study, it was observed that in hESCs, some CpG-poor regions of the genome switched from a highly methylated state (highly methylated regions [HMRs]) to high enrichment for H3K27me3 when the cells were differentiated to definitive endoderm, but remained HMRs during early differentiation to ectoderm or mesoderm [99]. This suggests that some genomic regions that are activated upon differentiation go from a highly DNA methylated state to a H3K27me3 state, the complement of the findings in mouse ESCs discussed above which showed that regions of the genome that became repressed with differentiation were first marked by H3K27me3, and then switched to a DNA methylation state. The same authors also reported that these regions frequently included binding sites for the endoderm-associated transcription factor FOXA2, suggesting that certain transcription factors may be able to bind to methylated DNA and stimulate demethylation [99].

In contrast to the sequential role for the H3K27me3 marks and DNA methylation in gene repression of the same genomic loci suggested by the earlier studies in mouse ESCs and the Gifford *et al.* study discussed above [99], the Xie *et al.* study found that H3K27me3 and DNA methylation appear to repress different genomic loci [98]. Namely, genes that were differentially expressed during early lineage specification had promoters that were CG-rich and were repressed by H3K27me3, while genes that were differentially expressed at later stages of differentiation had promoters that were CG-poor and were controlled by DNA methylation. These results support the notion that histone chromatin marks are the predominant repressive marks in open chromatin, mediating transcriptional repression in the rapidly changing cellular milieu during early lineage commitment, while DNA methylation mediates a more stable form of repression, and therefore would be better suited for setting long lasting marks necessary for stable maintenance of the cellular phenotypes produced by late stage differentiation. It is important to note that in these studies, the characterization of genomic loci involved in early lineage commitment was performed by studying hESCs undergoing *in vitro* differentiation, while the identification of

loci involved in later differentiation was performed using data from tissue samples. Although *in vitro* differentiation of hESCs appears to recapitulate *in vivo* differentiation in many respects, including sequential changes in cellular morphology and expression of key markers, it should be kept in mind that it has not been rigorously verified that they correspond in all respects.

In addition to interrogating histone modifications, the Gifford *et al.* study also used ChIP-seq to characterize the binding patterns of the pluripotency-associated transcription factors POU5F1, SOX2, and NANOG [99]. The combinatorial binding patterns for these transcription factors were assessed, revealing that sites bound by POU5F1 alone were primarily associated with promoter regions, in contrast to sites bound by all three factors, SOX2 only or NANOG only, which tended to be in intergenic or intragenic regions. Integrating the transcription factor and histone modification data revealed that dual POU5F1/H3K4me1 occupancy in the undifferentiated state was associated with HMRs regions. The study by Gifford *et al.* also discovered an enrichment for predicted binding sites for a number of early lineage-specific transcription factors in regions bound by different combinations of POU5F1, SOX2, NANOG and activating histone marks in hESCs in the undifferentiated and early differentiated states [99]. This finding suggests that pluripotency-associated transcription factors might 'set up' pluripotent cells for differentiation by ensuring that certain genomic regions are permissive for the proper epigenetic modifications involved in differentiation, or even play a direct role in early differentiation of certain lineages.

In another approach, a recent manuscript from the Sander laboratory used integrative analysis of histone modifications (H3K4me3 and H3K27me3) and transcriptome data to compare the later stages of differentiation *in vivo* and *in vitro* [100]. In this study, hESCs were subjected to *in vitro* lineage commitment to the pancreatic endoderm stage, and then either transplanted into mice or maintained *in vitro* for further differentiation. The resulting two populations of cells were analyzed by ChIP-seq and RNA-seq, and compared to human islets. Undifferentiated hPSCs were found to have bivalent marks (H3K4me3 and H3K27me3) for definitive endoderm, pancreatic endoderm and endocrine regulators. The regulators of definitive endoderm were appropriately activated during *in vitro* differentiation at the definitive endoderm stage and repressed at the pancreatic endoderm stage, and these modulations in transcription were mirrored by the appropriate loss of the H3K27me3 mark. Moreover, regulators of pancreatic endoderm were also appropriately activated during *in vitro* differentiation at the pancreatic endoderm stage, with the corresponding loss of H3K27me3. However, the proper derepression of late endocrine markers occurred only in the cells transplanted into mice for *in vivo* maturation; in the cells maintained *in vitro*, late endocrine markers were poorly expressed, and their promoters remained in the bivalent state. From these results, the authors concluded that lack of proper histone remodeling is at least in part responsible for the dysfunction of *in vitro* differentiated pancreatic endocrine cells [100].

Genome-wide profiling methods & epigenome maps

Advances in high-throughput profiling methods, such as microarrays and next-generation sequencing (NGS) greatly advanced our understanding of the epigenome. NGS platforms have provided the capacity to tackle the complexity and multidimensionality of the epigenome. Several methods have been adopted to profile the DNA methylation and histone modification status of pluripotent cells and differentiated cells. We will briefly describe the most commonly used high-throughput methods (reviewed in detail in [101]).

DNA methylation profiling involves bisulfite treatment of DNA, which results in the conversion of unmethylated cytosine residues to uracil residues, while methylated cytosines are protected from bisulfite conversion, and remain as cytosines. Following PCR, the converted cytosines become thymidines. The bisulfite-treated DNA can then be analyzed on microarrays, such as the Infinium HumanMethylation450K BeadChip (Illumina), designed to detect the presence of a cytosine versus thymidine at selected positions in the genome or subjected to NGS to directly read the presence of a cytosine versus thymidine, either in selected regions of the genome (e.g., using reduced representation bisulfite sequencing (RRBS) [102] or across the entire genome [103]. One of the major limitations of the standard microarray or bisulfite sequencing methods is that they do not distinguish between 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) [104,105], the latter being an intermediate during DNA demethylation and shown to mark regulatory regions in fetal brain cells [106] and enriched at enhancers in ESCs [107–109]. However, a new method, called oxidative bisulfite sequencing (oxBS-seq), when compared to standard bisulfite sequencing of the same sample, does allow distinguishing betweenallows one to distinguish between unmethylated, methylated and hydroxymethylated cytosines [110]. Comparisons between genome-wide DNA methylation profiling technologies have been reviewed in detail elsewhere [111,112] and summarized in Figure 1. As an adjunctive approach, RNA-seq or allele-specific qRT-PCR can be used to detect allele-specific expression and thereby infer the DNA methylation status of loci subject to imprinting or XCI [51].

Chromatin immunoprecipitation (ChIP) is the primary method used to detect physical interactions between DNA sequences and proteins. Using antibodies specific for the modified histone or transcription factor of interest, the genomic regions bound by the selected protein are isolated and identified by tiling array (ChIP-chip [82,113]) or NGS analysis. One of the major advantages of ChIP-seq is the ability to profile and detect repetitive sequences. Other derivations of the technology include ChIP-PET, where pair-ended reads are obtained for the ChIP fragments [114]. Integrative, epigenomic profiling approaches such as the combination of ChIP and bisulfite sequencing (ChIP-BS-seq, BisChIP-seq) [96,115] will enable interrogation of the same molecule and yield insights into the crosstalk between DNA methylation and histone modifications.

Large-scale epigenetic project initiatives have capitalized on NGS technologies to create and compile genome-wide, epigenome maps of cell types and tissues important in development and diseases [116]. The Roadmap Epigenomics [117] and the Encyclopedia of DNA Elements (ENCODE) [118] consortiums are both NIH-funded projects. ENCODE represents an international consortium of research groups with the primary aim to discover and identify functional and regulatory elements in the genome. In humans alone, ENCODE hosts over 4000 individual datasets from a wide variety of platforms on 147 cell lines. The goal of the Roadmap Epigenomics Project is to create reference maps of normal and primary cell types and to host these datasets in Human Epigenome Atlas public database. The current release includes 61 complete epigenome datasets as of May 2012. Both consortia host many
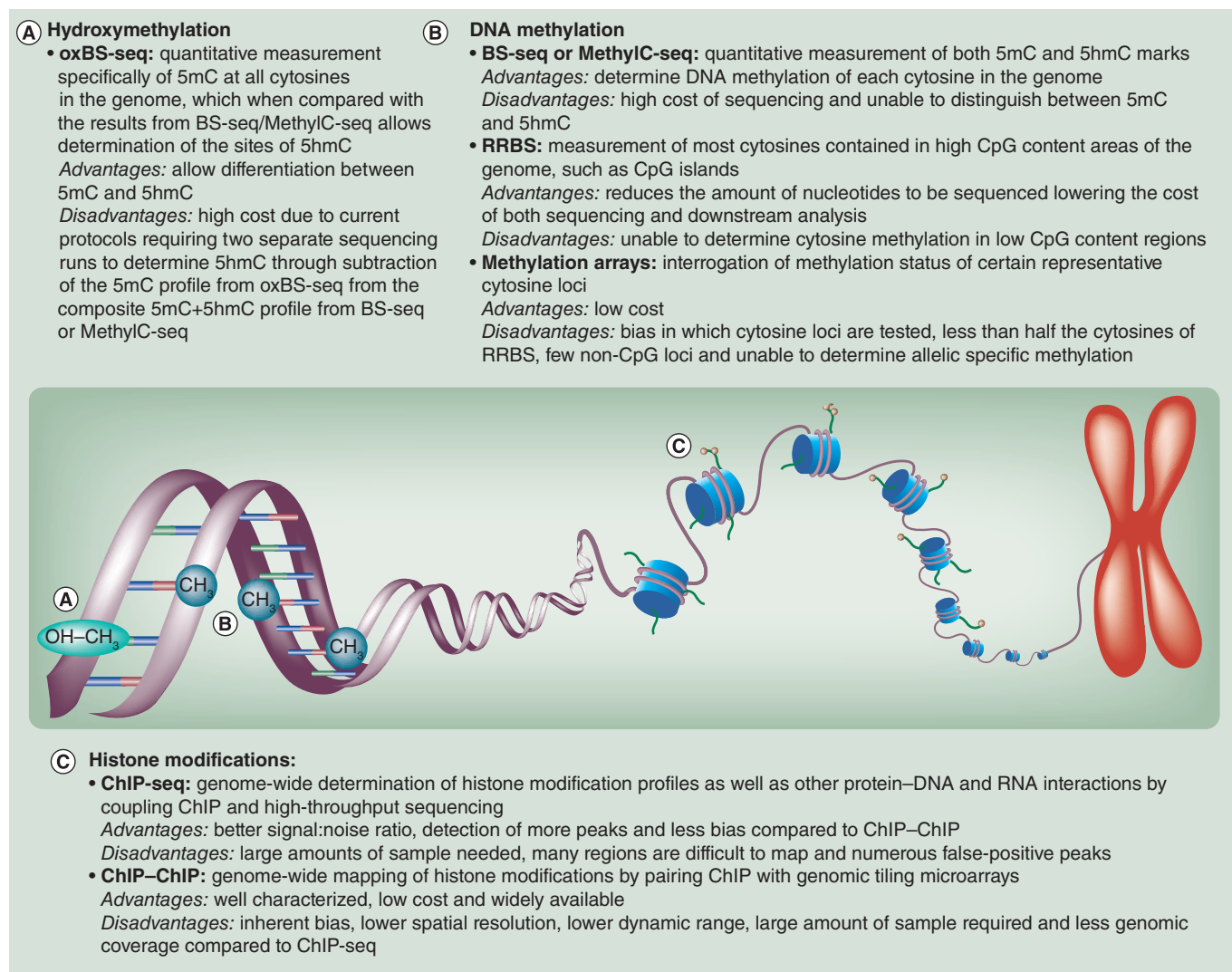
(A) **Hydroxymethylation**
- **oxBS-seq:** quantitative measurement specifically of 5mC at all cytosines in the genome, which when compared with the results from BS-seq/MethylC-seq allows determination of the sites of 5hmC
  Advantages: allow differentiation between 5mC and 5hmC
  Disadvantages: high cost due to current protocols requiring two separate sequencing runs to determine 5hmC through subtraction of the 5mC profile from oxBS-seq from the composite 5mC+5hmC profile from BS-seq or MethylC-seq

(B) **DNA methylation**
- **BS-seq or MethylC-seq:** quantitative measurement of both 5mC and 5hmC marks
  Advantages: determine DNA methylation of each cytosine in the genome
  Disadvantages: high cost of sequencing and unable to distinguish between 5mC and 5hmC
- **RRBS:** measurement of most cytosines contained in high CpG content areas of the genome, such as CpG islands
  Advantanges: reduces the amount of nucleotides to be sequenced lowering the cost of both sequencing and downstream analysis
  Disadvantages: unable to determine cytosine methylation in low CpG content regions
- **Methylation arrays:** interrogation of methylation status of certain representative cytosine loci
  Advantages: low cost
  Disadvantages: bias in which cytosine loci are tested, less than half the cytosines of RRBS, few non-CpG loci and unable to determine allelic specific methylation

(C) **Histone modifications:**
- **ChIP-seq:** genome-wide determination of histone modification profiles as well as other protein–DNA and RNA interactions by coupling ChIP and high-throughput sequencing
  Advantages: better signal:noise ratio, detection of more peaks and less bias compared to ChIP–ChIP
  Disadvantages: large amounts of sample needed, many regions are difficult to map and numerous false-positive peaks
- **ChIP–ChIP:** genome-wide mapping of histone modifications by pairing ChIP with genomic tiling microarrays
  Advantages: well characterized, low cost and widely available
  Disadvantages: inherent bias, lower spatial resolution, lower dynamic range, large amount of sample required and less genomic coverage compared to ChIP-seq

Figure 1. Key genome-wide approaches for profiling the epigenome.
5mC: 5-methylcytosine; 5hmC: 5-hydroxymethylcytosine; ChIP: Chromatin immunoprecipitation; oxBS-seq: Oxidative bisulfite sequencing; RRBS: Reduced representation bisulfite sequencing.

Table 1. Recent epigenomic maps of human stem cells and their derivatives.

Accession ID†	Summary	Platform	Platform details	Organism	Ref.
GSE26386	Identified different chromatin states corresponding to distinct regulatory elements such as repressed and active promoters, enhancers and insulators in diverse cell types including hESCs, cancer cells, fibroblasts, leukemia cells and others	ChIP-Seq, GE	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array, [HT_HG-U133A] Affymetrix HT Human Genome U133A Array, Illumina Genome Analyzer II	Homo sapiens	[7]
SRP000941	Histone modifications mapped in hESCs and fibroblasts showing dramatic redistribution of repressive chromatin marks	ChIP-Seq; mC-Seq (methylC-seq)	Illumina Genome Analyzer, II, IIx, HiSeq 2000, AB SOLiD 4 System	Homo sapiens	[8]
GSE19418	Genome-wide bisulfite sequencing of changes in DNA methylation during differentiation: hESCs, a fibroblastic differentiated derivative of the hESCs and neonatal fibroblasts	BS-Seq	Illumina Genome Analyzer II	Homo sapiens	[9]
GSE36679	Genome-wide transcription profiles during undirected differentiation of severely hypomethylated (Dnmt1-/-) ESCs as well as ESCs completely devoid of DNA methylation (Dnmt1-/-; Dnmt3a-/-; Dnmt3b-/- or TKO) and assay their potential to transit in and out of the ESC state	GE	Affymetrix Mouse Gene 1.0 ST Array	Mus musculus	[24]
GSE35140	CTCF binding sites in mouse brain were compared to ESCs and liver, and analyzed in an allele-specific manner	ChIP-Seq	Illumina Genome Analyzer IIx	Mus musculus	[39]
GSE31183	ZFP57/KAP1 binds all H3K9me3-bearing methylated imprinted control regions in ESCs. The results show a general mechanism for the protection of specific loci against the wave of DNA demethylation during early development	ChIP-Seq; RNA-Seq	Illumina Genome Analyzer II	Mus musculus	[40]
Not available	Maternal Trim28 is essential for embryogenesis to preserve epigenetic marks on a subset of genes	GE	Genomax (Singapore)	Mus musculus	[42]
GSE30654	Specific tissue types were distinguished by unique patterns of DNA hypomethylation, which were recapitulated by DNA demethylation during in vitro directed differentiation, X chromosome inactivation is unstable in pluripotent cells over time in culture, Aberrations in X inactivation and imprinting are maintained during differentiation	Methyl array, GE	Illumina HumanHT-12 V3.0 expression beadchip, Illumina HumanMethylation27 HumanMethylation27 BeadChip (HumanMethylation27_270596_v.1.2), Illumina HumanMethylation450 BeadChip (HumanMethylation450_15017482)	Homo sapiens	[52]

These genome-wide datasets are often integrative and multidimensional, involving profiling of epigenetic features such as histone modifications, DNA methylation and enhancers. Gene expression data complements the epigenetic feature in question and the results of these studies often demonstrate crosstalk between the epigenome and transcriptome.
†Accession IDs listed pertain to Gene Expression Omnibus (GEO) (GSEXXXXX), Short Reads Archive (SRAXXXXX, SRPXXXX), ArrayExpress (E-MTAB-XXXX).
AID: Activation-induced cytidine deaminase; BS-Seq: Bisulfite sequencing; ChIP: Chromatin immunoprecipitation; ESC: Embryonic stem cell; GE: Gene expression array; hESC: Human embryonic stem cell; hiPSC: Human-induced pluripotent stem cell; iPSC: Induced pluripotent stem cell; MeDIP: Methylated DNA immunoprecipitation; nChIP: Native chromatin immunoprecipitation; RRBS: Reduced representation bisulfite sequencing; TKO: Triple knockout; WGBS: Whole-genome bisulfite sequencing.

Table 1. Recent epigenomic maps of human stem cells and their derivatives (cont.).

Accession ID†	Summary	Platform	Platform details	Organism	Ref.
GSM429321-23, GSM432685-92, GSM438361-64, GSE17917, GSE18292 and GSE16256; see [122]	First genome-wide, maps of methylated cytosines in both hESCs and fetal fibroblasts, along with comparative analysis of messenger RNA and small RNA components of the transcriptome, several histone modifications, and sites of DNA–protein interaction for several key regulatory factors	WGBS; RNA-Seq; smRNA-Seq; ChIP-Seq	Illumina Genome Analyzer, II, IIX, HiSeq 2000, AB SOLiD 4 System	Homo sapiens	[57]
Not available	Female (hESCs) exhibit variation (0–100%) for XCI markers, including XIST RNA expression and enrichment of histone H3 lysine 27 trimethylation (H3K27me3) on the inactive X chromosome (Xi). Female hESCs have already completed XCI during the process of derivation and/or propagation	SNP-Array; GE; meDIP-Chip	Not available	Homo sapiens	[58]
GSE20937	Describe the establishment of XaXa hESCs derived under physiological oxygen concentrations	GE	Agilent-014850 Whole Human Genome Microarray 4×44K G4112F	Homo sapiens	[60]
GSE21037	Rett syndrome iPSCs are able to undergo X inactivation and generate functional neurons	GE	[HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array	Homo sapiens	[61]
GSE21222	Convert conventional hESCs into a more immature state that extensively shares defining features with pluripotent mouse ESCs. This was achieved by ectopic induction of Oct4, Klf4, and Klf2 factors combined with LIF and inhibitors of GSK3β and MAPK (ERK1/2) pathway	GE	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Homo sapiens	[62]
GSE22246	Reprogramming of female fibroblasts does not reactivate Xi, however the Xi undergoes chromatin changes. This nonrandom pattern of X chromosome inactivation in female hiPSCs is maintained upon differentiation	GE	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Homo sapiens	[67]
GSE34527	X-inactivation status in female hiPSC lines depends on derivation conditions	GE	Agilent-014850 Whole Human Genome Microarray 4×44K G4112F	Homo sapiens	[70]
GSE36233	X-inactivation markers can be used to separate hiPSC lines into distinct epigenetic classes	GE	Not available	Homo sapiens	[71]
GSE15112	Sodium butyrate supports the extensive self-renewal of mouse and hESCs, while promoting their convergence toward an intermediate stem cell state. In response to butyrate, hESCs regress to an earlier developmental stage characterized by a gene expression profile resembling that of mouse	GE	Agilent-012391 Whole Human Genome Oligo Microarray G4112A; Agilent-014850 Whole Human Genome Microarray 4×44K G4112F; Agilent-014868 Whole Mouse Genome Microarray 4×44K G4122F	Homo sapiens; Mus musculus	[73]

These genome-wide datasets are often integrative and multidimensional, involving profiling of epigenetic features such as histone modifications, DNA methylation and enhancers. Gene expression data complements the epigenetic feature in question and the results of these studies often demonstrate crosstalk between the epigenome and transcriptome.
†Accession IDs listed pertain to Gene Expression Omnibus (GEO) (GSEXXXXX), Short Reads Archive (SRAXXXXX, SRPXXXX), ArrayExpress (E-MTAB-XXXX).
AID: Activation-induced cytidine deaminase; BS-Seq: Bisulfite sequencing; ChIP: Chromatin immunoprecipitation; ESC: Embryonic stem cell; GE: Gene expression array; hESC: Human embryonic stem cell; hiPSC: Human-induced pluripotent stem cell; iPSC: Induced pluripotent stem cell; MeDIP: Methylated DNA immunoprecipitation; nChIP: Native chromatin immunoprecipitation; RRBS: Reduced representation bisulfite sequencing; TKO: Triple knockout; WGBS: Whole-genome bisulfite sequencing.

Table 1. Recent epigenomic maps of human stem cells and their derivatives (cont.).					
Accession ID†	Summary	Platform	Platform details	Organism	Ref.
GSE39757	Discovery of a long noncoding RNA, XACT, which is expressed from and coats the active X chromosome specifically in hPSCs	RNA-Seq	Illumina HiSeq 2000	Homo sapiens	[74]
GSE46700	Compare the ability of cells lacking AID to reprogram from a differentiated murine cell type to an induced pluripotent stem cell. AID removes epigenetic memory to stabilize the pluripotent state	RRBS; RNA-Seq	Illumina HiSeq 2000	Mus musculus	[75]
Not available	Defined intermediate cell populations poised to become iPSCs. Induced pluripotency elicits two transcriptional waves, which are driven by c-Myc/Klf4 (first wave) and Oct4/Sox2/Klf4 (second wave). The establishment of bivalent domains occurs gradually after the first wave, whereas changes in DNA methylation take place after the second wave when cells acquire stable pluripotency	ChIP-Seq, miRNA array	Not available	Mus musculus	[76]
GSE36570	A map of Oct4, Sox2, Klf4, and c-Myc (O, S, K and M) on the human genome during the first 48 h of reprogramming fibroblasts to pluripotency	ChIP-Seq	Illumina Genome Analyzer IIx	Homo sapiens	[79]
GSE12646	H3K4me1, H3K9me1 and H3K27me1 associate with enhancers of differentiation genes prior to their activation and correlate with basal expression during differentiation of multipotent human primary hematopoietic stem cells/progenitor cells into erythrocyte precursors	ChIP-Seq; GE	Illumina Genome Analyzer; [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Homo sapiens	[81]
GSE8024, GSE12241; see [123]	Genome-wide chromatin-state maps of mouse ESCs, neural progenitor cells and embryonic fibroblasts. Define three categories of promoters in ESCs, including bivalent promoters; and demonstrate the use of ChIP for genome-wide annotation active transposable elements, imprinting control regions and allele-specific transcription	ChIP-Seq	Affymetrix Mouse Genome 430 2.0 Array Illumina Genome Analyzer	Mus musculus	[82]
See [124]	Bivalent domains silence developmental genes in ESCs while keeping them poised for activation	ChIP-Chip; GE	Not available	Homo sapiens; Mus musculus; Canis lupus familiaris	[83]
GSE29413	Genes upregulated after deletion of the H3K9 methyltransferase Setdb1 are distinct from those derepressed in mESC deficient in the DNA methyltransferases Dnmt1 and Dnmt3a/b, with the exception of a few germline genes. ETDB1 role in inhibiting aberrant gene expression by suppressing expression of proximal ERVs	RNA-Seq; Nchip-Seq; ChIP-Seq	Illumina Genome Analyzer II	Mus musculus	[92]

These genome-wide datasets are often integrative and multidimensional, involving profiling of epigenetic features such as histone modifications, DNA methylation and enhancers. Gene expression data complements the epigenetic feature in question and the results of these studies often demonstrate crosstalk between the epigenome and transcriptome.
†Accession IDs listed pertain to Gene Expression Omnibus (GEO) (GSEXXXXX), Short Reads Archive (SRAXXXXX, SRPXXXX), ArrayExpress (E-MTAB-XXXX).
AID: Activation-induced cytidine deaminase; BS-Seq: Bisulfite sequencing; ChIP: Chromatin immunoprecipitation; ESC: Embryonic stem cell; GE: Gene expression array; hESC: Human embryonic stem cell; hiPSC: Human-induced pluripotent stem cell; iPSC: Induced pluripotent stem cell; MeDIP: Methylated DNA immunoprecipitation; nChIP: Native chromatin immunoprecipitation; RRBS: Reduced representation bisulfite sequencing; TKO: Triple knockout; WGBS: Whole-genome bisulfite sequencing.

Table 1. Recent epigenomic maps of human stem cells and their derivatives (cont.).

Accession ID†	Summary	Platform	Platform details	Organism	Ref.
GSE8439, GSE8463	Polycomb-associated (H3K27me3) and (H3K4me3) across ESCs. H3K27me3 modifications change during early differentiation, both relieving existing repressive domains and imparting new ones, and that colocalization with H3K4me3 is not restricted to pluripotent cells	ChIP-Chip	UW_Homo_sapiens_condensed_array_36848_v1; NimbleGen microarray	*Homo sapiens*	[93]
Not available	DNAme is acting globally to antagonize the placement of H3K27me3. In the mouse genome, H3K27me3 is aquired when DNAme is diminished	ChIP-Seq; RNA-Seq; MeDIP-methl array	NimbleGen 36720 K CpG Island Plus RefSeq Promoter Array	*Mus musculus*	[95]
GSE28254	H3K27me3 and DNA methylation are mostly compatible except for CpG islands (marks are mutually exclusive). ChIP-BS-seq-based analysis in *Dnmt* TKO ESCs showed total loss of CpG methylation is associated with alteration of H3K27me3 levels throughout the genome	ChIP-BS-Seq	Illumina Genome Analyzer IIx, Illumina HiSeq 2000	*Homo sapiens, Mus musculus*	[96]
SRP000941	Differentiated hESCs into mesendoderm, neural progenitor cells, trophoblast-like cells, and mesenchymal stem cells and systematically characterized DNA methylation, chromatin modifications, and the transcriptome in each lineage	ChIP-Seq; RNA-Seq; WGBS	Illumina Genome Analyzer, II, IIx, HiSeq 2000, AB SOLiD 4 System	*Homo sapiens*	[98]
GSE17312	Transcriptional and epigenetic profiling of populations derived through directed differentiation of hESCs representing each of the three embryonic germ layers	ChIP-Seq; RNA-Seq; WGBS	Affymetrix Human Genome U133 Plus 2.0 Array, Illumina Genome Analyzer II, Illumina Genome Analyzer IIx	*Homo sapiens*	[99]
E-MTAB-1086	Transcriptome and select chromatin modifications at defined stages during pancreatic endocrine differentiation of hESCs	ChIP-Seq; RNA-Seq	Illumina Genome Analyzer IIx	*Homo sapiens*	[100]
SRA000206; see [125]	Genome-wide histone modification maps of 20 histone lysine and arginine methylations as well as histone variant H2A.Z, RNA polymerase II, and the insulator-binding protein CTCF across the human genome	ChIP-Seq	Not available	*Homo sapiens*	[113]

These genome-wide datasets are often integrative and multidimensional, involving profiling of epigenetic features such as histone modifications, DNA methylation and enhancers. Gene expression data complements the epigenetic feature in question and the results of these studies often demonstrate crosstalk between the epigenome and transcriptome.
†Accession IDs listed pertain to Gene Expression Omnibus (GEO) (GSEXXXXX), Short Reads Archive (SRAXXXXXX, SRPXXXXX), ArrayExpress (E-MTAB-XXXX).
AID: Activation-induced cytidine deaminase; BS-Seq: Bisulfite sequencing; ChIP: Chromatin immunoprecipitation; ESC: Embryonic stem cell; GE: Gene expression array; hESC: Human embryonic stem cell; hiPSC: Human-induced pluripotent stem cell; iPSC: Induced pluripotent stem cell; MeDIP: Methylated DNA immunoprecipitation; nChIP: Native chromatin immunoprecipitation; RRBS: Reduced representation bisulfite sequencing; TKO: Triple knockout; WGBS: Whole-genome bisulfite sequencing.

datasets of PSC and differentiated lines. The success of these initiatives have spurred additional projects, such as the latest consortium, BLUEPRINT, an European initiative aimed to generate 100 reference epigenomes from hematopoietic cells and their malignant leukemic counterparts. The epigenome maps from these initiatives and other sources have immensely improved our understanding of the epigenetic regulatory mechanism during pluripotency and its dynamic use during differentiation (Table 1) as described throughout this review. The ENCODE [119] and Roadmap Epigenomics [120] datasets are available online and is also hosted at the UCSC Genome Browser for visualization [121].

Conclusion

The studies discussed here have demonstrated that characterization of early *in vitro* differentiation of hPSCs using genome-wide transcriptome and epigenome datasets can reveal patterns and associations that suggest a complex combinatorial mechanism for epigenomic regulation of early differentiation. It is clear that further work is necessary to verify and determine the functional consequences of the interactions between general epigenomic marks (i.e., DNA methylation and histone modifications), pluripotency-associated transcription factors and lineage-specific transcription factors suggested by these studies. Moreover, as methods for differentiation of hPSCs to more mature cell types are developed and optimized, it will be intriguing to see whether the same patterns of epigenomic regulation will apply, or whether distinct mechanisms will be uncovered.

Future perspective

As noted above, it still remains to be determined which type of hPSC best represents the pluripotent stem cells in the human embryo, and how closely *in vitro* differentiation of hPSCs resembles the *bona fide* cellular differentiation that occurs during development. For some applications, such as high-throughput screening for cell type-specific drug toxicity, modeling of certain diseases, or cell therapy, it may not be critical for hPSC-derived cells to be perfect epigenomic matches of the corresponding tissue-derived cell, as long as the hPSC-derived cells possess specific phenotypic properties (e.g., the ability to secrete insulin in a glucose-dependent manner for cell therapy for Type 1 diabetes mellitus). However, for accurate modeling of development and developmental or degenerative diseases, discrepancies between *in vivo* and *in vitro* differentiation may be problematic. Although technically challenging, it is important that studies delineating the dynamic role of the epigenome in mammalian development continue to be performed.

Executive summary

The epigenome in mammalian development
- Mammalian development is tightly regulated by epigenetic events and can be studied using pluripotent stem cells (PSCs), including human embryonic stem cells and induced PSCs. Although PSCs can be valuable *in vitro* models to study the differentiation process, care must be taken due to epigenetic aberrations, such as instability of X chromosome inactivation and hypo- or hyper-methylation of imprinted loci, which can be induced during derivation and long-term culture.

The epigenomics of reprogramming & differentiation
- Recent studies have shown that regulation of gene expression by DNA methylation and histone modifications is an intricate combinatorial process in which some histone marks are well-correlated (or anti-correlated) with DNA methylation, others regulate sets of genes that do not appear to be regulated by DNA methylation, and yet others show complex relationships with DNA methylation that seem to be context-dependent.

Genome-wide profiling methods & epigenome maps
- Advances in high-throughput profiling methods, such as microarrays and next-generation sequencing (NGS), have greatly advanced our understanding of the epigenome. More recently NGS platforms have provided the capacity to tackle the complexity and multidimensionality of the epigenome. Methods that are becoming increasingly popular for profiling of the DNA methylation and histone modification status of pluripotent cells and differentiated cells include reduced representation and whole-genome bisulfite sequencing and chromatin immunoprecipitation profiling on microarrays and high-throughput sequencing.

References

Papers of special note have been highlighted as:

• of interest

•• of considerable interest

1   Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 16(1), 6–21 (2002).

•   Identified different chromatin states corresponding to distinct regulatory elements such as repressed and active promoters, enhancers and insulators in diverse cell types including human embryonic stem cells (hESCs) and cancer cells.

2   Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* 14(3), 204–220 (2013).

•   Histone modifications mapped in hESC and fibroblasts showing dramatic redistribution of repressive chromatin marks.

3   Portela A, Esteller M. Epigenetic modifications and human disease. *Nat. Biotech.* 28(10), 1057–1068 (2010).

••   First genome-wide maps of methylated cytosines in both hESCs and fetal fibroblasts, along with comparative analysis of messenger RNA and small RNA components of the transcriptome, several histone modifications and sites of DNA–protein interaction for several key regulatory factors.

4   Shen H, Laird PW. Interplay between the cancer genome and epigenome. *Cell* 153(1), 38–55 (2013).

•   Female (hESCs) exhibit variation (0–100%) for XCI markers, including *XIST* RNA expression and enrichment of histone H3 lysine 27 trimethylation (H3K27me3) on the inactive X chromosome (Xi). Female hESCs have already completed XCI during the process of derivation and/or propagation.

5   Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* 3(6), 415–428 (2002).

••   Genome-wide chromatin-state maps of mouse ESCs, neural progenitor cells and embryonic fibroblasts. Defines three broad categories of promoters in ESCs, including bivalent promoters; and demonstrates the use of chromatin immunoprecipitation (ChIP) for genome-wide annotation active transposable elements, imprinting control regions and allele-specific transcription.

6   Bergman Y, Cedar H. DNA methylation dynamics in health and disease. *Nat. Struct. Mol. Biol.* 20(3), 274–281 (2013).

••   Transcriptional and epigenetic profiling of populations derived through directed differentiation of hESCs representing each of the three embryonic germ layers.

7   Ernst J, Kheradpour P, Mikkelsen TS *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473(7345), 43–49 (2011).

•   Specific tissue types were distinguished by unique patterns of DNA hypomethylation, which were recapitulated by DNA demethylation during *in vitro* directed differentiation, X chromosome inactivation is unstable in pluripotent cells over time in culture, Aberrations in X inactivation and imprinting are maintained during differentiation.

8   Hawkins RD, Hon GC, Lee LK *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* 6(5), 479–491 (2010).

9   Laurent L, Wong E, Li G *et al.* Dynamic changes in the human methylome during differentiation. *Genome Res.* 20(3), 320–331 (2010).

10   Takahashi K, Tanabe K, Ohnuki M *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131(5), 861–872 (2007).

11   Yu J, Vodyanik MA, Smuga-Otto K *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318(5858), 1917–1920 (2007).

12   Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126(4), 663–676 (2006).

13   Thomson JA, Itskovitz-Eldor J, Shapiro SS *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* 282(5391), 1145–1147 (1998).

14   Tachibana M, Amato P, Sparman M *et al.* Human embryonic stem cells derived by somatic cell nuclear transfer. *Cell* 153(6), 1228–1238 (2013).

15   Lister R, Pelizzola M, Kida YS *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* 471(7336), 68–73 (2011).

16   Hussein SM, Batada NN, Vuoristo S *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature* 471(7336), 58–62 (2011).

17   Gore A, Li Z, Fung HL *et al.* Somatic coding mutations in human induced pluripotent stem cells. *Nature* 471(7336), 63–67 (2011).

18   Laurent LC, Ulitsky I, Slavin I *et al.* Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell* 8(1), 106–118 (2011).

19   Leonardo TR, Schultheisz HL, Loring JF, Laurent LC. The functions of microRNAs in pluripotency and reprogramming. *Nat. Cell Biol.* 14(11), 1114–1121 (2012).

20   Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69(6), 915–926 (1992).

21   Kaneda M, Okano M, Hata K *et al.* Essential role for *de novo* DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* 429(6994), 900–903 (2004).

22   Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* 99(3), 247–257 (1999).

23   Tsumura A, Hayakawa T, Kumaki Y *et al.* Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* 11(7), 805–814 (2006).

24   Schmidt CS, Bultmann S, Meilinger D *et al.* Global DNA hypomethylation prevents consolidation of differentiation

programs and allows reversion to the embryonic stem cell state. *PLoS ONE* 7(12), e52629 (2012).

25 Bird A, Taggart M, Frommer M, Miller OJ, Macleod D. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* 40(1), 91–99 (1985).

26 Brandeis M, Frank D, Keshet I *et al.* Sp1 elements protect a CpG island from *de novo* methylation. *Nature* 371(6496), 435–438 (1994).

27 Macleod D, Charlton J, Mullins J, Bird AP. Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. *Genes Dev.* 8(19), 2282–2292 (1994).

28 Ginno PA, Lott PL, Christensen HC, Korf I, Chedin F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell* 45(6), 814–825 (2012).

29 Okamoto I, Patrat C, Thepot D *et al.* Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* 472(7343), 370–374 (2011).

30 Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. *Hum. Mol. Genet.* 14(Spec No 1), R47–R58 (2005).

31 Smallwood SA, Kelsey G. *De novo* DNA methylation: a germ cell perspective. *Trends Genet.* 28(1), 33–42 (2012).

32 Sleutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 415(6873), 810–813 (2002).

33 Ciccone DN, Su H, Hevi S *et al.* KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 461(7262), 415–418 (2009).

34 Henckel A, Chebli K, Kota SK, Arnaud P, Feil R. Transcription and histone methylation changes correlate with imprint acquisition in male germ cells. *EMBO J.* 31(3), 606–615 (2012).

35 Soejima H, Higashimoto K. Epigenetic and genetic alterations of the imprinting disorder Beckwith-Wiedemann syndrome and related disorders. *J. Hum. Genet.* 58(7), 402–409 (2013).

36 Cassidy SB, Schwartz S, Miller JL, Driscoll DJ. Prader-Willi syndrome. *Genet. Med.* 14(1), 10–26 (2012).

37 Mabb AM, Judson MC, Zylka MJ, Philpot BD. Angelman syndrome: insights into genomic imprinting and neurodevelopmental phenotypes. *Trends Neurosci.* 34(6), 293–303 (2011).

38 Lee J, Marisas B. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell* 152(6), 1308–1323 (2013).

39 Prickett AR, Barkas N, Mccole RB *et al.* Genomewide and parental allele-specific analysis of CTCF and cohesin DNA binding in mouse brain reveals a tissue-specific binding pattern and an association with imprinted differentially methylated regions. *Genome Res.* doi:10.1101/gr.150136.112 (2013).

40 Quenneville S, Verde G, Corsinotti A *et al.* In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol. Cell* 44(3), 361–372 (2011).

41 Li X, Ito M, Zhou F *et al.* A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev. Cell* 15(4), 547–557 (2008).

42 Messerschmidt DM, De Vries W, Ito M, Solter D, Ferguson-Smith A, Knowles BB. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science* 335(6075), 1499–1502 (2012).

43 Quenneville S, Turelli P, Bojkowska K *et al.* The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development. *Cell Reports* 2(4), 766–773 (2012).

44 Nakamura T, Arai Y, Umehara H *et al.* PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat. Cell Biol.* 9(1), 64–71 (2007).

45 Nakamura T, Liu YJ, Nakashima H *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* 486(7403), 415–419 (2012).

46 Wutz A. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat. Rev. Genet.* 12(8), 542–553 (2011).

47 Augui S, Nora EP, Heard E. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet.* 12(6), 429–442 (2011).

48 Migeon BR, Lee CH, Chowdhury AK, Carpenter H. Species Differences in TSIX/Tsix Reveal the Roles of These Genes in X-Chromosome Inactivation. *Am. J. Hum. Genet.* 71(2), 286–293 (2002).

49 Papp B, Plath K. Epigenetics of reprogramming to induced pluripotency. *Cell* 152(6), 1324–1343 (2013).

50 Tachibana M, Ma H, Sparman ML *et al.* X-chromosome inactivation in monkey embryos and pluripotent stem cells. *Dev. Biol.* 371(2), 146–155 (2012).

51 Calabrese JM, Sun W, Song L *et al.* Site-specific silencing of regulatory elements as a mechanism of X inactivation. *Cell* 151(5), 951–963 (2012).

52 Nazor KL, Altun G, Lynch C *et al.* Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. *Cell Stem Cell* 10(5), 620–634 (2012).

53 Stelzer Y, Ronen D, Bock C, Boyle P, Meissner A, Benvenisty N. Identification of novel imprinted differentially methylated regions by global analysis of human-parthenogenetic-induced pluripotent stem cells. *Stem Cell Reports* 1(1), 79–89 (2013).

54 Rugg-Gunn PJ, Ferguson-Smith AC, Pedersen RA. Status of genomic imprinting in human embryonic stem cells as revealed by a large cohort of independently derived and maintained lines. *Hum. Mol. Genet.* 16(Spec No. 2), R243–R251 (2007).

55 Kim KP, Thurston A, Mummery C *et al.* Gene-specific vulnerability to imprinting variability in human embryonic stem cell lines. *Genome Res.* 17(12), 1731–1742 (2007).

56 Frost J, Monk D, Moschidou D *et al.* The effects of culture on genomic imprinting profiles in human embryonic and fetal mesenchymal stem cells. *Epigenetics* 6(1), 52–62 (2011).

57 Lister R, Pelizzola M, Dowen RH *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462(7271), 315–322 (2009).

58 Shen Y, Matsuno Y, Fouse SD *et al*. X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. *Proc. Natl Acad. Sci. USA* 105(12), 4709–4714 (2008).

59 Silva SS, Rowntree RK, Mekhoubad S, Lee JT. X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proc. Natl Acad. Sci. USA* 105(12), 4820–4825 (2008).

60 Lengner CJ, Gimelbrant AA, Erwin JA *et al*. Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. *Cell* 141(5), 872–883 (2010).

61 Marchetto MCN, Carromeu C, Acab A *et al*. A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* 143(4), 527–539 (2010).

62 Hanna J, Cheng AW, Saha K *et al*. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc. Natl Acad. Sci. USA* 107(20), 9222–9227 (2010).

63 Dvash T, Lavon N, Fan G. Variations of X chromosome inactivation occur in early passages of female human embryonic stem cells. *PLoS ONE* 5(6), e11330 (2010).

64 Hall LL, Byron M, Butler J *et al*. X-inactivation reveals epigenetic anomalies in most hESC but identifies sublines that initiate as expected. *J. Cell. Physiol.* 216(2), 445–452 (2008).

65 Hoffman LM, Hall L, Batten JL *et al*. X-inactivation status varies in human embryonic stem cell lines. *Stem Cells* 23(10), 1468–1478 (2005).

66 Pomp O, Dreesen O, Leong DF *et al*. Unexpected X chromosome skewing during culture and reprogramming of human somatic cells can be alleviated by exogenous telomerase. *Cell Stem Cell* 9(2), 156–165 (2011).

67 Tchieu J, Kuoy E, Chin MH *et al*. Female human iPSCs retain an inactive X chromosome. *Cell Stem Cell* 7(3), 329–342 (2010).

68 Lessing D, Anguera MC, Lee JT. X chromosome inactivation and epigenetic responses to cellular reprogramming. *Ann. Rev. Gen. Hum. Genet.* 14, 85–110 (2013).

69 Kim K-Y, Hysolli E, Park I-H. Neuronal maturation defect in induced pluripotent stem cells from patients with Rett syndrome. *Proc. Natl Acad. Sci. USA* 108(34), 14169–14174 (2011).

70 Tomoda K, Takahashi K, Leung K *et al*. Derivation conditions impact X-inactivation status in female human induced pluripotent stem cells. *Cell Stem Cell* 11(1), 91–99 (2012).

71 Anguera Montserratc, Sadreyev R, Zhang Z *et al*. Molecular signatures of human induced pluripotent stem cells highlight sex differences and cancer genes. *Cell Stem Cell* 11(1), 75–90 (2012).

72 Mekhoubad S, Bock C, Deboer AS, Kiskinis E, Meissner A, Eggan K. Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell* 10(5), 595–609 (2012).

73 Ware CB, Wang L, Mecham BH *et al*. Histone deacetylase inhibition elicits an evolutionarily conserved self-renewal program in embryonic stem cells. *Cell Stem Cell* 4(4), 359–369 (2009).

74 Vallot C, Huret C, Lesecque Y *et al*. XACT, a long noncoding transcript coating the active X chromosome in human pluripotent cells. *Nat. Genet.* 45(3), 239–241 (2013).

75 Kumar R, Dimenna L, Schrode N *et al*. AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature* 500(7460), 89–92 (2013).

76 Polo JM, Anderssen E, Walsh RM *et al*. A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* 151(7), 1617–1632 (2012).

77 Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* 463(7284), 1042–1047 (2010).

78 Lee MR, Prasain N, Chae HD *et al*. Epigenetic regulation of NANOG by miR-302 cluster-MBD2 completes induced pluripotent stem cell reprogramming. *Stem Cells* 31(4), 666–681 (2012).

79 Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* 151(5), 994–1004 (2012).

80 Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl Acad. Sci. USA* 97(10), 5237–5242 (2000).

81 Cui K, Zang C, Roh TY *et al*. Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell Stem Cell* 4(1), 80–93 (2009).

82 Mikkelsen TS, Ku M, Jaffe DB *et al*. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448(7153), 553–560 (2007).

83 Bernstein BE, Mikkelsen TS, Xie X *et al*. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125(2), 315–326 (2006).

84 Kimura H, Tada M, Nakatsuji N, Tada T. Histone code modifications on pluripotential nuclei of reprogrammed somatic cells. *Mol. Cell. Biol.* 24(13), 5710–5720 (2004).

85 Lee JH, Hart SR, Skalnik DG. Histone deacetylase activity is required for embryonic stem cell differentiation. *Genesis* 38(1), 32–38 (2004).

86 Kouzarides T. Chromatin modifications and their function. *Cell* 128(4), 693–705 (2007).

87 Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470(7333), 279–283 (2011).

88 Creyghton MP, Cheng AW, Welstead GG *et al*. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* 107(50), 21931–21936 (2010).

89 Kagey MH, Newman JJ, Bilodeau S *et al*. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467(7314), 430–435 (2010).

90 Whyte WA, Orlando DA, Hnisz D *et al*. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153(2), 307–319 (2013).

91   Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet.* 10(5), 295–304 (2009).

92   Karimi MM, Goyal P, Maksakova IA *et al.* DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* 8(6), 676–687 (2011).

93   Pan G, Tian S, Nie J *et al.* Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* 1(3), 299–312 (2007).

94   Vire E, Brenner C, Deplus R *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439(7078), 871–874 (2006).

95   Hagarman JA, Motley MP, Kristjansdottir K, Soloway PD. Coordinate regulation of DNA methylation and H3K27me3 in mouse embryonic stem cells. *PLoS ONE* 8(1), e53880 (2013).

96   Brinkman AB, Gu H, Bartels SJ *et al.* Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res.* 22(6), 1128–1138 (2012).

97   Haffner MC, Pellakuru LG, Ghosh S *et al.* Tight correlation of 5-hydroxymethylcytosine and polycomb marks in health and disease. *Cell Cycle* 12(12), 1835–1841 (2013).

98   Xie W, Schultz MD, Lister R *et al.* Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* 153(5), 1134–1148 (2013).

99   Gifford CA, Ziller MJ, Gu H *et al.* Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* 153(5), 1149–1163 (2013).

100   Xie R, Everett LJ, Lim HW *et al.* Dynamic chromatin remodeling mediated by polycomb proteins orchestrates pancreatic differentiation of human embryonic stem cells. *Cell Stem Cell* 12(2), 224–237 (2013).

101   Rivera CM, Ren B. Mapping human epigenomes. *Cell* 155(1), 39–55 (2013).

102   Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* 33(18), 5868–5877 (2005).

103   Cokus SJ, Feng S, Zhang X *et al.* Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452(7184), 215–219 (2008).

104   Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 324(5929), 929–930 (2009).

105   Tahiliani M, Koh KP, Shen Y *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324(5929), 930–935 (2009).

106   Lister R, Mukamel EA, Nery JR *et al.* Global epigenomic reconfiguration during mammalian brain development. *Science* 341(6146), 1237905 (2013).

107   Pastor WA, Pape UJ, Huang Y *et al.* Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 473(7347), 394–397 (2011).

108   Szulwach KE, Li X, Li Y *et al.* Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet.* 7(6), e1002154 (2011).

109   Yu M, Hon GC, Szulwach KE *et al.* Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* 149(6), 1368–1380 (2012).

110   Booth MJ, Branco MR, Ficz G *et al.* Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* 336(6083), 934–937 (2012).

111   Bock C, Tomazou EM, Brinkman AB *et al.* Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat. Biotech.* 28(10), 1106–1114 (2010).

112   Harris RA, Wang T, Coarfa C *et al.* Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat. Biotech.* 28(10), 1097–1105 (2010).

113   Barski A, Cuddapah S, Cui K *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* 129(4), 823–837 (2007).

114   Wei CL, Wu Q, Vega VB *et al.* A global map of p53 transcription-factor binding sites in the human genome. *Cell* 124(1), 207–219 (2006).

115   Statham AL, Robinson MD, Song JZ, Coolen MW, Stirzaker C, Clark SJ. Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA. *Genome Res.* 22(6), 1120–1127 (2012).

116   Shakya K, O'connell MJ, Ruskin HJ. The landscape for epigenetic/epigenomic biomedical resources. *Epigenetics* 7(9), 982–986 (2012).

117   Bernstein BE, Stamatoyannopoulos JA, Costello JF *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nat. Biotech.* 28(10), 1045–1048 (2010).

118   Consortium EP. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 306(5696), 636–640 (2004).

119   Encyclopedia of DNA Elements. http://encodeproject.org/ENCODE/

120   Human Epigenome Atlas. www.genboree.org/epigenomeatlas/index.rhtml

121   UCSC Genome Bioinformatics. http://genome.ucsc.edu

122   The Salk Institute for Biological Studies. The Human DNA Methylome. http://neomorph.salk.edu/human_methylome/

123   Broad Institute. ChIP-Seq Data. www.broad.mit.edu/seq_platform/chip/

124   Broad Institute. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. www.broadinstitute.org/cell/hcne_chromatin

125   NIH. NHLBI Division of Intramural Research, Laboratory of Molecular Immunology. High-resolution profiling of histone methylations in the human genome. http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/hgtcell.html