

Chapter 2

Establishment of Tissue-Specific Epigenetic States During Development

Ionel Sandovici

Abstract Complex organisms require tissue-specific transcriptional programs, which are acquired during development through the stepwise activation of transcription factor networks acting in tight coordination with epigenetic mechanisms. Recent progresses in genome-wide mapping of various epigenetic marks across a panel of mammalian cell types and developmental stages, together with a multitude of functional analyses, led to significant advances in our understanding of tissue-specific epigenetic regulation of gene expression. These new developments open at last the opportunity to systematically explore the contribution of epigenetics to human disease.

2.1 Introduction

Complex organisms such as mammals comprise over 200 different cell types, each one expressing specific sets of genes that define their unique functions (Alberts et al. 2002). The tissue-specific transcriptional programs are acquired over the course of development, during which cells transit from a pluripotent state to differentiated cell lineages, in a well-orchestrated spatial and temporal manner. This stepwise process is controlled by the sequential activation of specific transcription factors acting coordinately with epigenetic mechanisms (Reik et al. 2001; Hemberger et al. 2009; Albert and Peters 2009), which particularly target key DNA regulatory sequences such as promoters, enhancers, and insulators (Maston et al. 2006).

I. Sandovici (✉)

Metabolic Research Laboratories, Department of Obstetrics and Gynaecology,
University of Cambridge, Cambridge CB2 0SW, UK

Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK
e-mail: ionel.sandovici@bbsrc.ac.uk

Epigenetics refers to heritable changes in gene expression, caused by mechanisms other than changes in the underlying DNA sequence (Bird 2007). These heritable changes in gene expression are brought about by a complex array of reversible epigenetic marks: DNA modifications (such as 5-methylcytosine (Suzuki and Bird 2008), 5-hydroxymethylcytosine (Tahiliani et al. 2009), 5-formylcytosine, and 5-carboxylcytosine (Ito et al. 2011)), posttranslational modifications of histones (such as methylation, acetylation, phosphorylation, ubiquitination, sumoylation, ADP ribosylation, deimination, proline isomerization) (Bannister and Kouzarides 2011), histone variants (Talbert and Henikoff 2010), and alternative nucleosome positioning (Bai and Morozov 2010). The epigenetic marks are laid on the chromatin by an array of chromatin- and DNA-binding proteins with enzymatic activities (Brenner and Fuks 2007), as well as noncoding RNAs (Grewal 2010), all of which act as epigenetic initiators (epigenators) (Berger et al. 2009).

In the past few years significant advances in our understanding of tissue-specific epigenetic regulation of gene expression have been made possible by loss-of-function studies, as well as genome-wide mapping of different epigenetic marks across a panel of mammalian cell types and developmental stages. A seminal contribution in this direction has been provided by the Encyclopedia of DNA Elements (ENCODE) project. Very recently, the human ENCODE project has achieved the systematic characterization of a large variety of epigenetic marks in 147 different cell types (ENCODE Project Consortium 2012). Moreover, the integration of ENCODE data with other resources such as the genome-wide association studies (GWAS) has started to reveal new insights into human disease (Maurano et al. 2012). In this chapter I summarize the current view on the establishment of tissue-specific epigenetic marks during development, how these epigenetic patterns are correlated with transcription in a cell-specific manner, and how the tissue-specific epigenetic states may be directly linked with human disease.

2.2 Epigenetic Reprogramming During Preimplantation Development

The life of an organism begins at fertilization with the formation of the zygote. Fertilization coincides with a wave of epigenetic reprogramming (Fig. 2.1) that is required for the achievement of developmental totipotency (Reik et al. 2001; Hemberger et al. 2009; Albert and Peters 2009). Interestingly, even before fertilization the oocyte exhibits global hypomethylation, particularly at specific families of long interspersed element 1 (LINE1) and long terminal repeat (LTR) retroelements (Smith et al. 2012). A major initial event in the post-fertilization reprogramming process is the active loss of DNA methylation in the paternal pronucleus (Santos et al. 2002), likely by partial conversion of 5-methylcytosine into 5-hydroxymethylcytosine by TET3 (ten–eleven translocation) protein (Wossidlo et al. 2011; Gu et al. 2011). This is followed by passive DNA

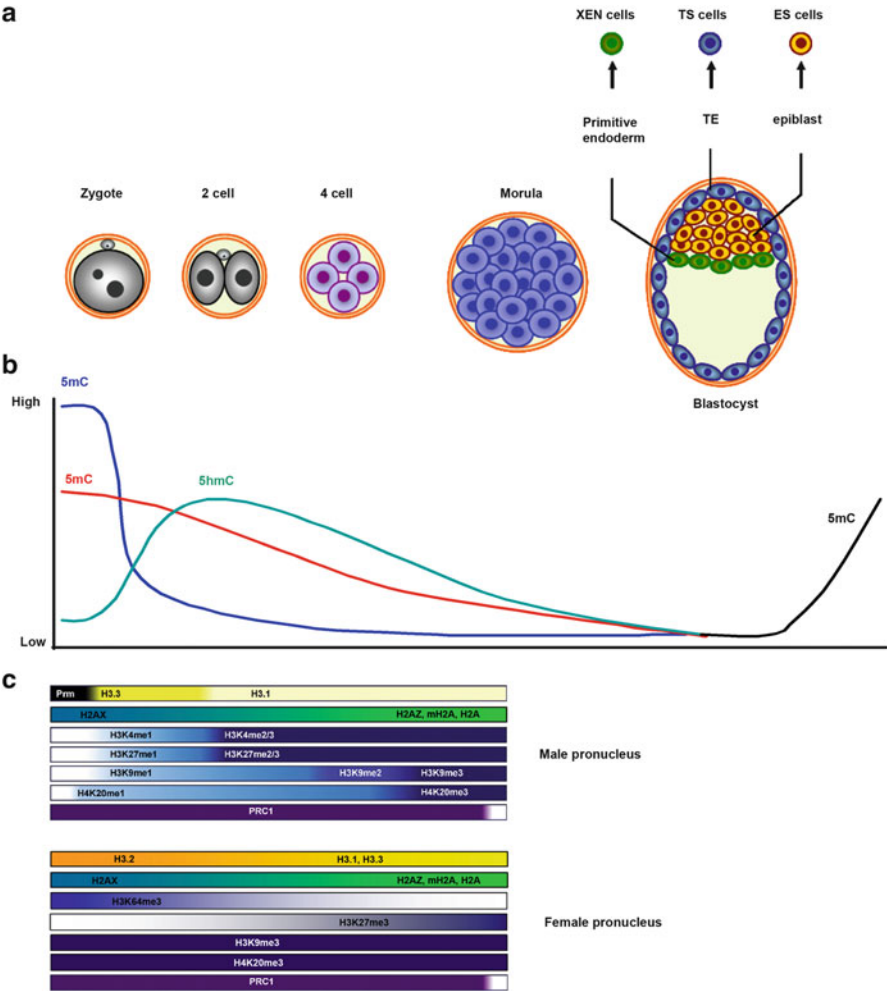


Fig. 2.1 Epigenetic reprogramming in early embryos (see text for details). (a) Diagram of the first events during the preimplantation development. XEN—extraembryonic endoderm stem cells, ES—embryonic stem cells, TE—trophoblast, TS—trophoblast stem cells. (b) Dynamics of DNA modification changes during early development: blue—paternal pronucleus; red—maternal pronucleus. 5mC—5-methylcytosine, 5hmC—5-hydroxymethylcytosine. (c) Histone variants and histone modification dynamics during early development (adapted from Hemberger et al. (2009); Albert and Peters (2009)). Prm—protamines; H3.1, H3.2, H3.3—histone H3 variants; H2AX, H2AZ, mH2A, and H2A—histone H2 variants; K—lysine; me—methylation; PRC1—Polycomb group (PcG) repressive complex 1

demethylation of the maternal pronucleus, facilitated by the exclusion of DNA methyltransferase 1 (DNMT1) from the nucleus (Howell et al. 2001), as well as by TET-mediated hydroxylation (Inoue and Zhang 2011). However, imprinting control regions (ICRs), oocyte, and sperm-contributed differentially methylated

regions (DMRs), as well as several families of repeats, such as class II intracisternal A-particles (IAPs) and L1Md_A elements, retain high levels of DNA methylation throughout the preimplantation development (Lane et al. 2003; Smith et al. 2012). Minimum levels of DNA methylation are reached at the blastocyst stage, followed by postimplantation gain of methylation to typical somatic levels (Smith et al. 2012).

Genome-wide reprogramming of histone modifications also occurs during the preimplantation development (Fig. 2.1). Immediately after fertilization, the paternal pronucleus is stripped of sperm-specific proteins called protamines and repackaged with maternally stored histone variant H3.3 that is usually associated with chromatin regions actively transcribed (Torres-Padilla et al. 2006). Interestingly, deposition of H3.3 into the paternal genome by the histone chaperone regulator A (HIRA) is an important event for the establishment of pericentric heterochromatin, which is required for proper chromosome segregation during the first mitosis (van der Heijden et al. 2005; Santenard et al. 2010). Only a few hours later, during the first DNA replication, the canonical histone H3 variants are incorporated for the first time into the paternal genome (Santenard et al. 2010). Histone H3.3 within the male pronucleus becomes trimethylated at lysine 27 (H3K27me3) and this repressive histone mark, together with H3K9me1 (monomethylation of lysine 9) retained in pericentromeric regions and residual DNA methylation, serves as a substrate for pericentric heterochromatin formation mediated by the Polycomb group (PcG) repressive complex 1 (PRC1) (Puschendorf et al. 2008). In the female pronucleus histone H3.3 transiently disappears and is replaced by histone H3.2 (Akiyama et al. 2011). After the two-cell stage, H3.1 and H3.3 variants re-localize to heterochromatin and euchromatin, respectively (Akiyama et al. 2011). The pericentric heterochromatin in the female pronucleus is marked with the repressive histone marks H3K9me3, H4K20me3, and H3K64me3 and binds HP1 β (heterochromatin 1 beta) protein (Santos et al. 2005; Probst et al. 2007; Daujat et al. 2009). Histone H2A variants are also reprogrammed during the preimplantation development. H2AZ (important for gene silencing), macroH2A (associated with heterochromatic regions and inactive X chromosome in females), and the canonical H2A are not incorporated into chromatin during the early cleavage stages, and are possibly even actively removed after fertilization. In contrast, H2AX (implicated in DNA repair) is particularly enriched in early embryos (Nashun et al. 2010). Together, all the reprogramming events described above are thought to contribute to the efficient acquisition of totipotency during preimplantation development.

The earliest sign of cell differentiation occurs at the blastocyst stage (embryonic day—E3.5 in mouse and embryonic day 5 in human), with the specification of the inner cell mass (ICM) and the trophectoderm (TE). This event coincides with the first wave of de novo DNA methylation. As result, TE is relatively hypomethylated compared with ICM, as revealed by the 5-methylcytosine staining (Santos et al. 2002). Similar to DNA methylation, several histone modifications, including H3K27 methylation, H3K9Ac (histone H3, lysine 9 acetylation), H4 acetylation, and H3K9 methylation, also exhibit asymmetry between ICM and TE, either at

global level or at specific loci (Erhardt et al. 2003; Sarmiento et al. 2004; O'Neill et al. 2006). The ICM is then separated during the late blastocyst stage into epiblast, that will form the future embryo proper, and primitive endoderm that contributes, together with the trophoblast cells derived from TE, to the formation of the extraembryonic tissues (Reik et al. 2001). Some cells derived from the primitive endoderm also contribute to the formation of the definitive embryonic endoderm (Kwon et al. 2008). After the formation of the three lineages (epiblast, primitive endoderm, and trophoblast), the cells undergo successive steps of differentiation to form all cell types of the organism, including placenta.

2.3 Epigenetic Regulation of Pluripotency

The three lineages at the blastocyst stage have been used for derivation of distinct stem cell types that can be maintained in vitro (Fig. 2.1): trophoblast stem (TS) cells from TE (Tanaka et al. 1998), extraembryonic endoderm stem (XEN) cells from the primitive endoderm (Kunath et al. 2005), and embryonic stem (ES) cells from the epiblast (Matsui et al. 1992). Analyses performed on these cell lines enabled the identification of key genetic factors that regulate pluripotency, such as OCT4 (octamer-binding transcription factor 4, also known as POU5F1—POU domain, class 5, transcription factor 1), NANOG (Nanog homeobox), SOX2 (SRY [sex-determining region Y]-box 2), and SALL4 (Sal-like protein 4) for ES cells (Mitsui et al. 2003; Loh et al. 2006; Wu et al. 2006); CDX2 (caudal type homeobox 2), EOMES (eomesodermin), and TEAD4 (TEA domain family member 4) for TS cells (Niwa et al. 2005; Yagi et al. 2007); and GATA4 (GATA-binding protein 4), GATA6 (GATA-binding protein 6), SOX7 (SRY [sex-determining region Y]-box 7), and SOX17 (SRY [sex-determining region Y]-box 17) for XEN (Kunath et al. 2005; Lim et al. 2008). The study of these stem cell lines also revealed an intriguing interplay between pluripotency transcription factors and epigenetic mechanisms. In fact, it is now thought that the dynamic balance between these two regulatory systems may form the basis for the pluripotent state.

2.3.1 DNA Methylation

The promoter regions of many pluripotency genes are unmethylated in pluripotent stem cell lines but methylated in somatic cells (Fouse et al. 2008; Meissner et al. 2008; Farthing et al. 2008; Senner et al. 2012). DNA methylation is thought to be particularly important for the epigenetic regulation of some “gatekeeper” genes that reinforce the commitment of pluripotent stem cells to a certain lineage such as *Elf5*, (E74-like factor 5 [ets domain transcription factor]) which, together with *Cdx2* and *Eomes*, safeguards ES cells from differentiating into trophoblast derivatives (Ng et al. 2008). Moreover, DNA methylation is the only epigenetic

mechanism that represses the activity of some genes implicated in differentiation of ES cells, and lack of DNA methylation in mutant ES cells leads to activation of these genes (Fouse et al. 2008). ES cells also contain substantial levels of non-CpG methylation (Ramsahoye et al. 2000; Lister et al. 2009; Ziller et al. 2011). Increased levels of non-CpG methylation have been found in exons of highly expressed genes, such as *OCT4* (Lister et al. 2009). Despite these characteristics, whether DNA methylation is absolutely necessary to maintain the pluripotency state remains controversial. Indeed, ES cells lacking completely DNA methylation (triple knock-out for *Dnmt1*, *Dnmt3a*, and *Dnmt3b*) can grow robustly and maintain to a large extent their undifferentiated characteristics (Tsumura et al. 2006).

The cytosines in DNA can acquire alternative modifications besides 5-methylcytosine. Tet1 protein, which is highly expressed in ES cells, can further modify 5-methylcytosine into 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxylcytosine (Ito et al. 2011; Wu et al. 2011; He et al. 2011). High levels of 5-hydroxymethylcytosine and 5-formylcytosine in ES cells are associated with actively transcribed genes, as well as with Polycomb-repressed developmental regulators, and were demonstrated to guard against trans-differentiation to extra-embryonic lineages (Ficz et al. 2011; Wu et al. 2011; Booth et al. 2012; Raiber et al. 2012). However, deletion of *Tet1* gene in mice is compatible with embryonic and postnatal development, possibly due to partial compensation by *Tet2*. Accordingly, mutant ES cells display only subtle changes in gene expression and skewed differentiation towards trophectoderm in vitro (Dawlaty et al. 2011).

2.3.2 Histone Modifications

In addition to DNA modifications, histone modifications are also important in controlling gene expression during cell renewal of pluripotent stem cell lines. In agreement with the previously established notion that H3K4me3 is an activating histone modification (Santos-Rosa et al. 2002), peaks of this mark are observed in ES cells in association with promoter regions of key pluripotency genes (Azuara et al. 2006; Barski et al. 2007). However, approximately 2,000 genes that are transcriptionally repressed in ES cells but are required for later differentiation (such as *Sox*, *Hox*, *Fox*, *Pax*, and *Irx* gene families) are concomitantly decorated at their promoter regions with both active H3K4me3 and repressive H3K27me3, pattern dubbed as “bivalent” (Azuara et al. 2006; Bernstein et al. 2006; Pan et al. 2007; Zhao et al. 2007). The bivalent domains are often found at promoters containing CpG islands and many bind OCT4, NANOG, or SOX2 (Bernstein et al. 2006; Mikkelsen et al. 2007). Virtually all bivalent domains bind PcG proteins belonging to the PRC2 complex (embryonic ectoderm development—EED, AE-binding protein 2—AEBP2, SUZ12—suppressor of zeste 12 homolog [*Drosophila*], and the H3K27 methyltransferase EZH2—enhancer of zeste homolog 2 [*Drosophila*]) (Ku et al. 2008). Recently, jumonji, AT-rich interactive domain 2 (JARID2—a member of the Jumonji family of lysine demethylases) and the

Polycomb-like 2 (PCL2) protein were also found to associate with PRC2 in mouse ES cells and are thought to play important roles in pluripotency (Peng et al. 2009; Pasini et al. 2010; Walker et al. 2010). A subset of the bivalent domains also binds PcG proteins of the PRC1 complex (Ku et al. 2008). The PRC1 protein RNF2 (ring finger protein 2, also known as RING1B) is responsible for ubiquitination of histone H2A at lysine 119, which in turn is responsible for RNA polymerase II stalling (a mechanism for transcriptional silencing) at promoters of bivalent genes (Stock et al. 2007). The H3K4me3 mark at the bivalent domains is generated by the H3K4 methyltransferase activity of the MLL/trithorax complex (myeloid/lymphoid or mixed-lineage leukemia [trithorax homolog, *Drosophila*]) (Dou et al. 2006). Similar bivalent chromatin profiles have also been identified in TS cells at promoter regions of silenced, lineage-specific regulatory genes (Santos et al. 2010). However, in XEN cells lineage-specific genes are marked solely by repressive histone modifications, pattern thought to reflect the restricted developmental potential of these cells (Santos et al. 2010).

The repressive histone modifications H3K9me2 and H3K9me3 colonize different regions of the ES cells' genome. H3K9me2 is found in large blocks in the genome (several megabases each), which are highly conserved between mouse and human (Wen et al. 2009). H3K9me3, found in ES cells mostly in partnership with H3K20me3, is required for silencing several classes of endogenous retroviruses (Mikkelsen et al. 2007; Rowe et al. 2010; Matsui et al. 2010; Macfarlan et al. 2011).

In addition to promoters and repetitive DNA, histone modifications are particularly important in regulating the activity of enhancer elements (short regions of DNA often found distant from transcription start sites that bound transcription factors and enhance gene transcription). Based on the patterns of histone modifications, two distinct classes of enhancers can be identified in ES cells. Both classes are characterized by open chromatin, marked by the presence of DNase I hypersensitive sites (DHSs), enrichment in highly mobile nucleosomes containing the specialized histone variants H3.3 and H2A.Z, binding of the histone acetyltransferase P300, and monomethylation of histone H3 at lysine 4 (H3K4me1). Active enhancers, often located in the vicinity of active genes such as the key pluripotency genes, are characterized by acetylation of histone H3 at lysine 27 (H3K27ac). In contrast, the so-called poised enhancers, located near genes involved in controlling early steps of differentiation and marked with bivalent domains at their promoters, are depleted in H3K27ac and instead are enriched in H3K27me3 and H3K9me3 (Creyghton et al. 2010; Rada-Iglesias et al. 2011; Zentner et al. 2011; Buecker and Wysocka 2012).

2.3.3 Chromatin-Modifying Complexes

With the exception of acetylation, most histone modifications do not impose directly changes in the chromatin conformation. Instead, they often bind chromatin-remodeling factors, which utilize the energy released from ATP

hydrolysis to exchange histones and reposition or evict nucleosomes. When compared with differentiated cells, pluripotent stem cells are characterized by a generally open chromatin state (Gaspar-Maia et al. 2011). There are four families of chromatin remodelers (SWI/SNF, CHD/NURD, ISWI, and INO80) and many subunits of these families have been shown to play important roles in pluripotent stem cells (reviewed by Gaspar-Maia et al. 2011). For example, the SWI/SNF family member BGR1 (also known as SMARCA4) opposes PcG proteins by opening the chromatin at LIF/STAT3 (leukemia inhibitory factor/signal transducer and activator of transcription 3) target sites in ES cells. However, BRG1 also facilitates PcG function at classical PcG targets, including all four *Hox* loci, reinforcing their repression in ES cells (Ho et al. 2011). The chromodomain helicase DNA-binding protein 1 (CHD1) member of the CHD family binds globally to active euchromatin and co-localizes with RNA polymerase II (RNAPII) in ES cells and CHD1 depletion by RNA interference leads to accumulation of high levels of heterochromatin (Gaspar-Maia et al. 2009). CHD3 and CHD4 constitute the catalytic subunit of the nucleosome-remodeling (NuRD) complex, which also contains histone deacetylases (HDAC1 and HDAC2) and a methyl-binding protein (MBD3). MBD3 cooperates with BRG1 to maintain the global levels of 5-hydroxymethylcytosine in ES cells (Yildirim et al. 2011). Finally, the TIP60/KAT5–P400 (lysine acetyltransferase 5/E1A-binding protein p400) complex of the INO80 family facilitates transcription by combining nucleosome remodeling with histone acetylase activity. ES cells depleted in different subunits of the TIP60–P400 complex exhibit decreased proliferation rates, reduced pluripotency, and reduced viability (Fazzio et al. 2008). The TIP60–P400 complex also binds H3K4me3 at bivalent domains, an interaction that is facilitated by Nanog (Fazzio et al. 2008).

2.3.4 Mechanisms for Targeting Epigenetic Patterns in Pluripotent Stem Cells

The patterns of chromatin and DNA modifications in pluripotent stem cells cannot be explained only by the genomic distribution of transcription factor-binding sites. Indeed, targeting of PcG proteins at the bivalent promoters in ES cells is only partially explained by the concomitant binding of the core pluripotency transcription factors (Bernstein et al. 2006; Mikkelsen et al. 2007). Therefore, other factors such as the local DNA sequence, noncoding RNAs (ncRNAs), and the higher order chromatin structure may be important in this process.

The direct involvement of DNA sequence was first demonstrated in the case of JARID2, protein that binds directly to DNA and plays a major role in targeting PRC2 complexes to the correct sites (Peng et al. 2009; Pasini et al. 2010). Additionally, a recent study has identified a PcG responsive element in human ES cells, a highly conserved 1.8 kb DNA sequence located between *HOXD11* (homeobox D11) and *HOXD12* (homeobox D11) genes, which is nucleosome depleted and GC-rich and contains YY1 transcription factor-binding sites (Woo et al. 2010).

It has also been demonstrated that short DNA sequences inserted into the mouse ES cells can autonomously induce hypo- or de novo methylation in *cis* (Lienert et al. 2011).

ncRNAs may be another important class of regulators for establishing epigenetic patterns in ES cells. Short ncRNAs (<200 nt) interact with PRC2 and are involved in stabilizing PRC2 association with chromatin, though the importance of direct base pairing at specific sequence motifs is still unknown (Kanhare et al. 2010). Furthermore, over 3,000 large intergenic noncoding RNAs (lincRNAs) have been recently identified in mouse ES cells. At least a third of them are associated with chromatin complexes involved in reading, writing, or erasing histone modifications and are critical for pluripotency maintenance (Guttman et al. 2009, 2011).

Finally, the higher order chromatin structure may also contribute to the appropriate establishment of epigenetic marks in pluripotent stem cells (reviewed by Li et al. 2012). The insulator protein CTCF (CCCTC-binding factor [zinc finger protein]), known to mediate long-range interactions between distant regulatory elements, has been shown to cooperate with Oct4 in organizing the chromatin loops at the *Nanog* locus (Levasseur et al. 2008). A recent genomic analysis in ES cells has revealed that binding sites for cohesin (a key partner for CTCF), mediator, and the cohesin-loading factor NIPBL (nipped-B homolog [*Drosophila*]) overlap at active promoters and enhancers (Kagey et al. 2010). Together, these complexes contribute to chromatin looping between enhancers and promoters in patterns that are specific to ES cells (Kagey et al. 2010).

2.4 Establishment of Epigenetic Patterns During Differentiation

Gene deletion studies in mice or knockdown experiments in ES cells demonstrated for the first time that many epigenetic modifiers play critical roles during differentiation. Accordingly, deletion of *Dnmt1* results in lethality before E10.5 (Li et al. 1992) and disruption of *Dnmt3b* is lethal before E9.5 (Okano et al. 1999). Additionally, ES cells deficient for all three DNA methyltransferases show increased cell death upon differentiation into epiblast lineages, but not during differentiation into extraembryonic lineages and do not contribute to embryonic lineages when injected into blastocysts (Sakaue et al. 2010). Deletion of the histone methyltransferase G9a/EHMT2 (euchromatic histone-lysine N-methyltransferase 2) results in embryonic lethality between E8.5 and E9.5 (Tachibana et al. 2002). Knockdown of various PcG proteins in ES cells affects their ability to differentiate (Azuara et al. 2006; Bernstein et al. 2006; Pasini et al. 2007, 2010). Depletion of the MLL complex component DPY30 (dpy-30 homolog [*C. elegans*]) in ES cells, which decreases H3K4me3 at bivalent domains, results in a significant reduction in the differentiation potential, particularly along the neural lineage (Jiang et al. 2011). Deletions of *Mbd3* or *Hdac1* result in aberrant differentiation of mouse ES cells (Kaji et al. 2006; Dovey et al. 2010).

In addition to these loss-of-function studies, another major approach for improving our understanding of the transitions that occur during differentiation was the use of various “omics” analyses. Most studies that addressed the role of epigenetic modifications during differentiation have compared the genomic distribution of various marks between ES cells and cells differentiated in vitro or with donor-derived somatic cells. These studies have identified several general mechanisms implicated in the establishment of tissue-specific epigenetic patterns during differentiation.

2.4.1 Dynamics of DNA Methylation During Differentiation

Whereas a small number of genes undergo DNA demethylation upon commitment to a cell lineage, many more gain CpG methylation (Fouse et al. 2008; Meissner et al. 2008). Loss of DNA methylation (Fig. 2.2a) is observed especially at lineage-specific gene-regulatory elements. De novo DNA methylation (Fig. 2.2b) is responsible for the active repression of core pluripotency and germline-specific genes, as well as for some lineage choice events. Repression of pluripotency genes is initiated by local binding of the G9a histone methyltransferase, which through its SET domain brings about local methylation of histone H3K9me3 (Feldman et al. 2006; Epsztejn-Litman et al. 2008). Subsequently, H3K9me3 binds HP1/CBX5 (chromobox homolog 5), thus generating a local heterochromatic structure. In parallel, G9a/EHMT2 recruits the DNA methyltransferases Dnmt3a and 3b, which then induce de novo methylation of these genes (Feldman et al. 2006; Epsztejn-Litman et al. 2008). De novo DNA methylation during ES cell differentiation also occurs at CpG island promoters and at sequences outside of promoter regions, many of which act as enhancers (Mohn et al. 2008; Meissner et al. 2008; Stadler et al. 2011). Importantly, a recent study performed on early embryos confirmed that Dnmt3b catalyzes the gain of DNA methylation in E6.5 epiblast cells (Borgel et al. 2010). Similarly with the data obtained in ES cells cultured in vitro, these epigenetic events target promoters of pluripotency and germline-specific genes, as well as genes programmed to be expressed later during development. For the latter category of genes, promoter methylation acquired in epiblast is then erased during terminal cell differentiation (Borgel et al. 2010).

2.4.2 Resolution of Bivalent Domains

In the case of bivalent domains (Fig. 2.2c) it is thought that the concomitant presence of active and repressive modifications in pluripotent stem cells allows rapid resolution of these domains into single H3K27me3 or H3K4me3 marks during differentiation (Mikkelsen et al. 2007). Removal of H3K27me3 is achieved by two H3K27 demethylases: UTX/KDM6A (lysine [K]-specific demethylase 6A) and JMJD3/KDM6B (lysine [K]-specific demethylase 6B) (Agger et al. 2007; Lee et al. 2007; Lan et al. 2007). For example, human ES cells induced to differentiate

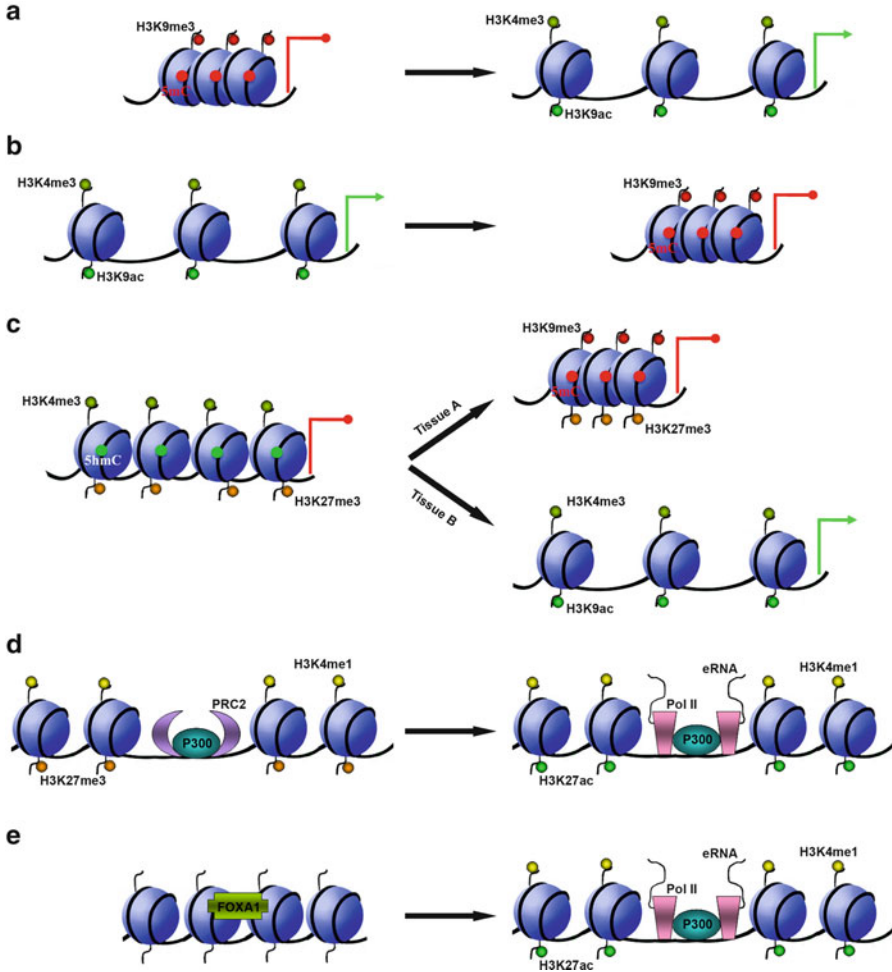


Fig. 2.2 Epigenetic changes during differentiation (see text for details). (a) Loss of DNA methylation at lineage-specific regulatory genes (together with changes in histone marks) leads to gene activation. (b) De novo DNA methylation at promoters of pluripotency genes or germline-specific genes. (c) Epigenetic reprogramming of bivalent promoters allows gene silencing or activation in a tissue-specific manner. (d) Transition between the “poised” and the “active” state at enhancer elements during differentiation. (e) The role of “pioneer” transcription factors (such as FoxA1—forkhead box A1) in generating de novo enhancers during differentiation. K—lysine; me—methylation; ac—acetylation; P300—histone acetyltransferase; PRC2—Polycomb group (PcG) repressive complex 2; Pol II—RNA polymerase II; eRNA—enhancer-associated RNAs

by treatment with retinoic acid recruit KDM6A at the promoters of the anterior genes of *HOXA* and *HOXB* loci. Recruitment of KDM6A to these promoters coincides with disappearance of H3K27me3, decreased occupancy of the PRC2 complex components SUZ12 and EZH2, and gene activation, while knockdown of

KDM6A prevents these events (Agger et al. 2007; Lee et al. 2007). KDM6A has also been demonstrated to activate muscle-specific genes during myogenesis, being targeted to the correct promoters by the transcriptional activator Six4 (SIX homeobox 4) (Seenundun et al. 2010). Interestingly, KDM6A associates with two H3K4 methyltransferases, MLL3/KMT2C (myeloid/lymphoid or mixed-lineage leukemia 3) and MLL4/KMT2D (myeloid/lymphoid or mixed-lineage leukemia 4), suggesting cooperation between H3K4 methylation and H3K27 demethylation (Lee et al. 2007; Issaeva et al. 2007). JMJD3/KDM6B has been demonstrated to resolve the bivalent domain at the *Nes* (nestin) gene promoter and to control the expression of key regulators and markers of neurogenesis during the commitment of ES cells towards the neural lineage (Burgold et al. 2008). Removal of H3K4me3 from the bivalent domains is achieved by the KDM5 demethylases. KDM5A (JARID1A/RBP2) is recruited at the bivalent domains by the PRC2 complex (Pasini et al. 2008). Additionally, KDM5B (JARID1B/PLU1) binds to a substantial fraction of bivalent domains in ES cells and is required for silencing stem cell and germ cell-specific genes during ES cell differentiation into neural progenitor cells (Schmitz et al. 2011).

It is important to stress that bivalent domains play important roles throughout differentiation. Indeed, when ES cells are differentiated into neural cells, the resolution of some bivalent domains is counterbalanced by appearance of new ones at other promoter regions. Moreover, ~41 % of the bivalent domains found in ES cells are preserved after differentiation into terminal pyramidal neurons (Mohn et al. 2008). In hemangioblasts, which are hematopoietic/endothelial precursors, some neuronal genes retain bivalency and require the presence of the PRC1 component RING1B/RNF2 to remain silent (Mazzarella et al. 2011). Adult stem cells, which maintain the natural homeostasis of adult tissues by supplying a continuous pool of differentiated cells in response to external signals, also contain bivalent chromatin domains (Mikkelsen et al. 2007; Cui et al. 2009).

2.4.3 Chromatin Changes at Enhancer Elements

Epigenetic reprogramming at enhancer elements is perhaps one of the most important events in the establishment of tissue-specific gene expression patterns during differentiation. Indeed, during differentiation of human ES cells into a mesendodermal lineage, chromatin modifications at promoters remained largely invariant, with much greater dynamics in chromatin modifications at enhancers, especially for H3K4me1 and H3K27ac (Hawkins et al. 2011). The main event that takes place during differentiation at enhancers is a switch from the poised to the active status (Fig. 2.2d), which coincides with the ability to drive gene expression. Interestingly, recent evidence suggests that the pluripotency factors active in ES cells are not only involved in maintaining the pluripotent state of these cells but also have a direct role during differentiation. Accordingly, SOX2, which binds many poised enhancers in ES cells, is replaced by SOX3 in neural progenitor cells and then by SOX11 in differentiated neurons. Upon binding of activating SOX3 or

SOX11 transcription factors, the poised chromatin state is resolved into an active one (Bergsland et al. 2011). Other poised enhancers are co-occupied by several pluripotency factors in ES cells and this multiple binding is thought to prevent their premature activation. One example is the *EOMES* enhancer, which is bound by OCT4, SOX2, and NANOG in human ES cells (Teo et al. 2011). At the onset of endoderm specification, SOX2 departure and the persistence of NANOG binding lead to activation of this enhancer and increased expression of eomesodermin (Teo et al. 2011).

Not all enhancers that are used during differentiation are established in pluripotent cells as active or poised ones. Many enhancers are generated de novo during differentiation by the intervention of the so-called pioneer factors, which are often lineage-specific transcription factors that have the ability to bind DNA sequences at chromatin-compacted regions (Fig. 2.2e). Pioneer factors such as FOXA (forkhead box A) and GATA (GATA-binding protein) recruit subsequently chromatin remodelers, which establish the characteristic open chromatin structure of active enhancers (Zaret and Carroll 2011).

2.5 Cell Type-Specific Epigenetic Patterns in Differentiated Mammalian Cells: Lessons Learned from the ENCODE Project

A first indication on the complexity of cell type-specific epigenetic patterns came from the analysis of just 1 % of the human genome during the pilot phase of the ENCODE project (ENCODE Project Consortium 2007). This preliminary analysis was then extended to the entire genome during the production stage of the ENCODE project. This stage of the project ended up with the release of a much more extensive set of results (Fig. 2.3) obtained in 147 different cell types (including both immortalized cell lines and primary cell types from a variety of tissues and developmental stages). The most important conclusion of the project was that approximately 80 % of the human genome participates in at least one biochemical function, most of which are related to gene regulation (ENCODE Project Consortium 2012). Some of the most relevant findings of the project are summarized below.

2.5.1 *Transcriptional Landscape*

RNA sequencing was performed in the set of 15 cell types most commonly used across the consortium and showed that, although the 20,687 protein-coding genes cover less than 3 % of the genome, cumulatively, nearly 75 % of the human genome is transcribed (Djebali et al. 2012). Each protein-coding gene associates on

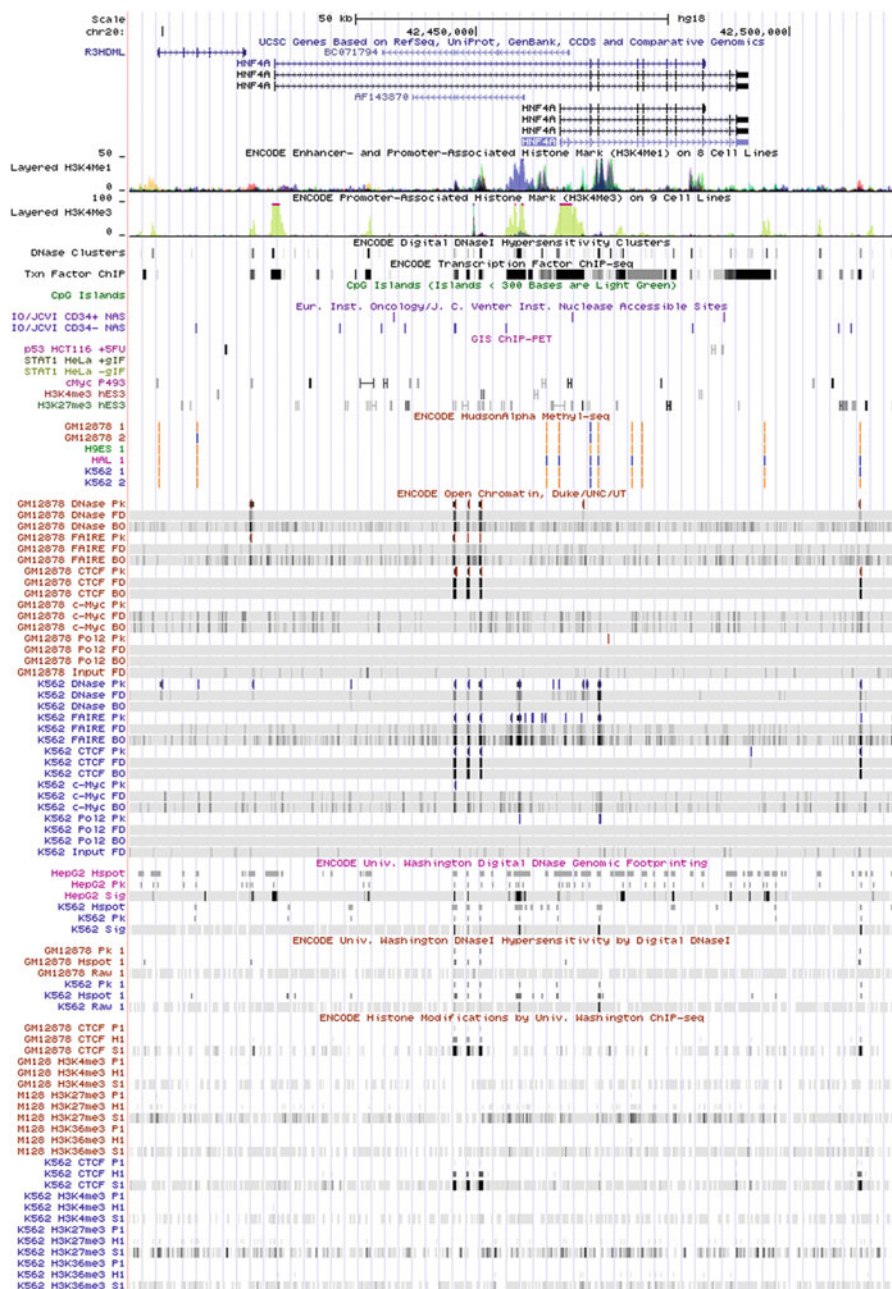


Fig. 2.3 Example of epigenome maps obtained through the ENCODE project. This window is centered around the hepatocyte nuclear factor 4, alpha (*HNF4A*) locus on human chromosome 20, as visualized in the UCSC Genome Browser (<http://genome.ucsc.edu>). The top part of the figure indicates the alternative transcripts identified at this locus. Follow a series of maps indicating the distribution of various histone marks, transcription factor-binding sites, and regions of open chromatin in several cell types, with peaks of enrichments indicated by vertical bars

average 6.3 alternatively transcripts and, although many isoforms are expressed simultaneously in a single cell type, one of these dominates (Djebali et al. 2012). The majority of protein-coding genes (53 %) are constitutively expressed (in all cell lines) and only a small fraction (7 %) are cell line specific (Djebali et al. 2012). ENCODE also identified 9,277 manually curated long noncoding RNA (lncRNA) loci generating ~15,000 transcripts, most of which are associated with chromatin and display more tissue-specific expression patterns than the protein-coding genes (Derrien et al. 2012). Approximately 18 % of the protein-coding and lncRNA genes exhibit allele-specific expression (Djebali et al. 2012). ENCODE also identified 11,216 pseudogenes, of which 876 are transcribed (Pei et al. 2012). Only a small fraction of the transcribed pseudogenes are active in all tissues analyzed, while most are transcribed only in one tissue (Pei et al. 2012). There are also 7,053 annotated small RNAs, which include 1,944 small nuclear (sn)RNAs, 1,521 small nucleolar (sno)RNAs, 1,756 μ (mi)RNAs, and 624 transfer (t)RNAs (Djebali et al. 2012). Other categories of transcripts include unannotated short RNAs such as the promoter-associated short RNAs (PASRs) and the terminus-associated short RNAs (TASRs), transcripts emanating from repeat elements and enhancer-associated RNAs (eRNAs) (Djebali et al. 2012).

2.5.2 DNA Methylation Landscape

DNA methylation was analyzed in 82 cell types (cell lines and primary cells) using the reduced representation bisulfite sequencing (RRBS), which can interrogate 1.2 million CpGs located in intergenic regions, proximal promoters, and gene bodies (8.6 % of non-repetitive genomic CpGs), with a preferential bias towards CpG islands (Meissner et al. 2008). Ninety-six percentage of all analyzed CpGs were found to exhibit differential methylation in at least one cell type, with the highest variability found at gene bodies and intergenic regions, rather than at promoters (ENCODE Project Consortium 2012). In addition, unmethylated intragenic CpG islands were found to associate binding of P300 histone acetyltransferase, a known marker for enhancer activity (Creyghton et al. 2010; ENCODE Project Consortium 2012). Differential DNA methylation associates with tissue-specific binding of CTCF, a ubiquitously expressed regulator of transcription and chromatin structure. Comparison between DNA methylation distribution and CTCF-binding sites for a subset of ~4,000 CTCF peaks indicates that over 40 % of the cell type-specific CTCF binding is associated with local differential DNA methylation (Wang et al. 2012). Additionally, 20 % of the DHSs with cell type-specific accessibility show a significant negative correlation with levels of DNA methylation, while the remaining 80 % of DHSs are constitutively hypomethylated (Thurman et al. 2012; Neph et al. 2012). Moreover, for 70 % of transcription factors, average methylation at cognate binding sites is significantly and negatively correlated with transcript levels of the corresponding transcription factors (Thurman et al. 2012).

2.5.3 *Histone Modification Landscape*

The ENCODE project analyzed systematically 11 histone modifications and 1 histone variant (H2A.Z) in 46 cell types. The main conclusion of this analysis is that histone modification patterns can be reliably used to assign functional attributes to genomic regions (ENCODE Project Consortium 2012; Dong et al. 2012). For example, transcriptionally active GC-rich (and TATA-less) promoters are associated with H2A.Z, H3K9ac, H3K27ac, H3K4me3, and H3K4me2, while repressed promoters are associated with H3K27me3 or H3K9me3. H3K79me2 and H3K36me3 are marks of transcription elongation; however, H3K79me2 occurs preferentially at the 5' end of the gene bodies, while H3K36me3 is enriched at 3' of the first intron (ENCODE Project Consortium 2012). These two last marks can also be used to predict patterns of alternative splicing: H3K36me3 has a positive contribution to exon inclusion, while H3K79me2 has a negative contribution (ENCODE Project Consortium 2012). By overlapping histone patterns with DHS maps 44,853 novel putative promoters were identified, many of which are active in a cell-specific manner, are contained within the gene bodies of previously annotated genes, and show antisense orientation (Thurman et al. 2012). Patterns of histone modifications at enhancer regions are amongst the best associated with cell-specific gene activity. Active enhancers are characterized by the presence of DHSs that bind RNA polymerase II and are enriched in H3K4me1, H3K27ac, H3K9ac, and H3K79me2 and depleted in H3K27me3 (Thurman et al. 2012; Djebali et al. 2012).

2.5.4 *Open Chromatin Landscape*

Regions of open chromatin identified by DNase I hypersensitivity are often found at regulatory DNA regions. Using DNase-seq in 125 cell types ~2.9 million DHSs were identified, most of them being located distal to transcription start sites (TSSs) and highly cell specific. A complementary technique—FAIRE-seq (formaldehyde-assisted isolation of regulatory elements) performed in 25 cell types—also identified ~4.8 million sites depleted in nucleosomes, many of which overlap with DHSs (ENCODE Project Consortium 2012; Thurman et al. 2012). Overlapping DHSs with high-throughput ChIP-seq data for 42 transcription factors in the K562 cell line (immortalized myeloid leukemia cells) showed that over 94 % of the transcription factor-binding sites fall within accessible chromatin. Notable exceptions are transcription factors known to bind to compacted heterochromatin, such as TRIM28 (tripartite motif containing 28), SETDB1 (SET domain, bifurcated 1), and ZNF274 (zinc finger protein 274) (Thurman et al. 2012). Moreover, a correlation between distal DHSs and DHSs located at known promoters across 79 cell types allowed functional connection of ~580,000 distal enhancers with their target promoters. Most promoters are connected with more than one distal DHS and vice versa, indicating a very complex *cis*-regulatory circuit of the human genome. In addition to this synchronized activation between promoters and distal enhancers,

hundreds of enhancers around the genome showed patterns of matched co-activation, suggesting highly choreographed cell type-specific behavior and common functions (Thurman et al. 2012). Finally, micrococcal nuclease (MNase) digestion followed by high-depth sequencing was used to map nucleosome occupancy in two cell types: GM12878 (lymphoblastoid cell line) and K562 (Kundaje et al. 2012). This analysis, combined with 12 histone marks, DNase-seq, and binding sites for 119 DNA-binding proteins, demonstrated that, with the exception of CTCF/cohesion complex, nucleosomes as well as histone marks are deposited asymmetrically around promoters, enhancers, or transcription factor-binding sites (Kundaje et al. 2012).

2.5.5 Long-Range Interaction Landscape

Physical interactions between distant chromosomal regions, which are thought to be important for regulation of gene expression, were assessed using two complementary technologies: 5C (chromosome conformation capture carbon copy) and ChIA-PET (chromatin interaction analysis with paired-end tag sequencing) (ENCODE Project Consortium 2012; Sanyal et al. 2012). The 5C approach was used for an unbiased interrogation of all interactions between TSSs previously identified by the pilot ENCODE project and distal genomic regions. This assay performed in four cell types identified over 1,000 long-range interactions (Sanyal et al. 2012). The most frequent interactions of the assessed TSSs were with enhancers, other promoters, and CTCF-binding sites, and each of these elements was found to be engaged in multiple interactions. The TSS–enhancer and TSS–promoter interactions were often found to be cell type specific, while the interactions of TSS–CTCF were most of the time common to all four cell types (Sanyal et al. 2012). The ChIA-PET approach, which interrogates interactions between chromatin regions that bind RNA polymerase II, has been applied within the ENCODE project for the K562 cell line. This analysis identified over 120,000 promoter-centered interactions, the vast majority of which were intrachromosomal. Similar to the 5C approach, this analysis showed that most promoters are engaged in multiple promoter–enhancer and promoter–promoter interactions (ENCODE Project Consortium 2012). The ChIA-PET has also been used in an independent study performed in five human cell types (including K562) (Li et al. 2012). This study demonstrated widespread promoter–promoter interactions between genes transcribed cooperatively, as well as cell type-specific promoter–enhancer interactions (Li et al. 2012).

2.6 Tissue-Specific Epigenetic States and Human Disease

It is increasingly acknowledged that epigenetic phenomena may be a crucial component in the development of human disease. The importance of epigenetics has been clearly demonstrated in monogenic disorders involving imprinted genes

(such as Beckwith–Wiedemann, Prader–Willi, and Angelman syndromes), in single-gene disorders of the epigenetic machinery (such as Rett, ICF, ATRX, and Rubinstein–Taybi syndromes) and in cancer (reviewed by Feinberg 2007; Portela and Esteller 2010). Since these subjects are being presented in depth elsewhere in the book, in this section I discuss the existing evidence for tissue-specific epigenetic alterations in common diseases and the link between genetic variants, tissue-specific epigenetic patterns, and disease.

2.6.1 Epigenetic Alterations in Common Human Diseases

The tissue specificity of epigenetic patterns makes it less straightforward to extrapolate the epigenetic information obtained in accessible samples such as peripheral white blood or buccal cells to the relevant tissues involved in various human diseases. Additional obstacles in studying epigenetic alterations in the context of human diseases are the observed variation of epigenetic marks between healthy individuals and with advancing age (Sandovici et al. 2003; Bjornsson et al. 2008). Despite these important challenges, in the past decade a number of studies have been able to uncover epigenetic alterations in several major forms of common human diseases in tissues and at loci that are directly involved in the pathogenesis of the studied conditions.

By far, the most studied epigenetic mark until now was DNA methylation, fact explained at least in part by the ease of obtaining DNA samples compared with good-quality chromatin, as well as by the robustness and relatively low cost of new microarray-based technologies. For example, a study performed in monozygous twins (MZ) discordant for type 1 diabetes (T1D) using Illumina Infinium 27 K microarrays to measure DNA methylation in CD14⁺ monocytes identified 58 CpG sites hypermethylated and 78 hypomethylated in the T1D-affected co-twins (Rakyan et al. 2011). Using the same technology to measure DNA methylation in the CD4⁺ T lymphocytes from patients with systemic lupus erythematosus and controls 236 hypomethylated and 105 hypermethylated CpG sites were identified (Jeffries et al. 2011). Another study using Illumina Infinium 27 K microarrays identified 276 CpG loci affiliated to promoters of 254 genes displaying significant differential DNA methylation in islets from type 2 diabetes (T2D) patients compared with controls, 244 of which were hypomethylated. These methylation changes affected many genes implicated in β -cell survival and function, were absent in blood cells from T2D individuals, and could not be induced experimentally in nondiabetic islets exposed to high glucose (Volkmar et al. 2012). As a last example, a study performed in the frontal cortex of patients with schizophrenia or bipolar disorder versus controls using CpG-island microarrays identified DNA methylation differences at dozens of loci, including several involved in glutamatergic and GABAergic neurotransmission, brain development, and other processes functionally linked to these diseases (Mill et al. 2008).

A variety of histone marks associated with transcriptional activation or repression have been studied in several common diseases in a tissue-specific context.

For example, in human prefrontal cortexes from patients with schizophrenia, the decreased *GAD1* (glutamate decarboxylase 1 [brain, 67 kDa]) expression compared to controls was associated with decreased levels of promoter H3K4me3, especially in females (Huang et al. 2007). Altered levels of the repressive histone mark H3K9me2 have been found at many loci implicated in autoimmunity and inflammation in lymphocytes collected from T1D patients versus controls (Miao et al. 2008). Marked differences in H3K9Ac levels were found at the upstream regions of *HLA-DRB1* (major histocompatibility complex, class II, DR beta 1) and *HLA-DQB1* (major histocompatibility complex, class II, DQ beta 1) genes in T1D monocytes relative to controls (Miao et al. 2012). Differential distribution of H3K4me3 and H3K9me3 peaks across the genome has been identified in cardiomyocytes collected from patients with heart failure caused by dilated cardiomyopathy compared to controls and many disease-dependent clusters contained genes implicated in signal transduction pathways for cardiac function (Kaneda et al. 2009).

A number of recent studies have also started to identify important roles for tissue-specific epigenators in the pathogenesis of several common human diseases. One illustrative example was recently published in human pancreatic islets. A comprehensive strand-specific transcriptome analysis identified 1,128 lncRNAs, many of which are cell specific and linked with β -cell differentiation and maturation programs. Using a gene candidate approach, several of these genes were found abnormally expressed in samples collected from T2D patients (Morán et al. 2012).

2.6.2 Genetic Variants, Tissue-Specific Epigenetic Patterns, and Human Disease

The convergence between disease-associated genetic variants emerging from GWAS and epigenetic maps led to the remarkable observation that many single-nucleotide polymorphisms (SNPs) linked with various diseases are located at regulatory DNA sequences. For example, rs7903146, a *TCF7L2* (transcription factor 7-like 2 [T-cell specific, HMG-box]) intronic variant strongly associated with T2D was found to be located in an islet-selective open chromatin region that exhibits enhancer activity. Human islets heterozygous for rs7903146 showed allelic imbalance in the local chromatin organization and altered enhancer activity (Gaulton et al. 2010). Additionally, a systematic analysis of chromatin-state dynamics in several human cell types, which identified cell type-specific enhancers, found that top-scoring disease-associated SNPs are frequently positioned within enhancer regions specifically active in the relevant cell types. Accordingly, SNPs associated with erythrocyte phenotypes are located in enhancers specific to erythroleukemia cells (K562), SNPs associated with systemic lupus erythematosus are located in enhancers specific to lymphoblastoid cells (GM12878), while SNPs

associated with triglyceride and total lipid levels in blood are located in enhancers specific to hepatocellular carcinoma cells (HepG2) (Ernst et al. 2011).

Building on these initial observations, the recent data published by the ENCODE consortium demonstrated unequivocally that over a third of the disease-associated genetic variants emerged from the GWASs performed so far localize within regulatory DNA elements marked by the presence of DHSs (ENCODE Project Consortium 2012; Maurano et al. 2012; Schaub et al. 2012). Beyond this statistically significant concentration of disease-associated genetic variants around regulatory DNA elements, the systematic analysis of a large number of cell and tissue types led to several additional striking observations. First of all, it was observed that genetic variants associated with a certain disease are particularly enriched at DHSs that are active in the cell types implicated in its pathogeny. Examples include the enrichment of genetic variants associated with Crohn's disease at DHSs active in T cells (subtypes T_H17 and T_H1) and the enrichment of SNPs associated with multiple sclerosis at DHSs active in $CD3^+$ T cells from cord blood and $CD19^+/CD20^+$ B cells (Maurano et al. 2012). In many cases common SNPs associated with specific diseases are located at binding sites for key transcription factors and, as a result, their presence induces allelic imbalances of chromatin states (Gaulton et al. 2010; Maurano et al. 2012). Additionally, the disease-associated SNPs can also alter tissue-specific enhancer–promoter interactions (Li et al. 2012; Maurano et al. 2012). More than 80 % of DHSs containing disease-associated SNPs are active in fetal cells and tissues, with the greatest enrichment for SNPs linked with phenotypes for which gestation or early growth have been shown to play major roles (such as cardiovascular disease) and with a relative depletion for SNPs linked with aging-related diseases (Maurano et al. 2012). This finding is in agreement with the recurring theme of chromatin landscape plasticity during early development and the risk for specific common adult diseases (Sandovici et al. 2008, 2011).

2.7 Concluding Remarks and Future Perspectives

The recent epigenomics studies have began to uncover in great details the epigenetic landscape of pluripotent stem cells and the transitions that occur during cell differentiation. Despite the advances made by these studies, our understanding of the functional role of particular epigenetic mechanisms remains relatively poor. This limitation highlights the need for more mechanistic studies. It became apparent that some of the epigenetic features discovered by studying pluripotent stem cells are acquired during their culture in vitro and may not reflect the patterns existing in vivo. For example, ES cells cultured in defined medium with inhibitors of two kinases (MEK [MAP kinase/ERK kinase] and GSK3B [glycogen synthase kinase 3 beta]), a condition known as “2i,” postulated to establish a naive ground state, have reduced prevalence of bivalent domains, despite similar differentiation potential with serum-grown ES cells (Marks et al. 2012). Additionally, genome-wide DNA methylation profiling of a large collection of human pluripotent stem

cell lines has revealed many aberrations of this epigenetic mark, which were specific to the culture conditions used (Nazor et al. 2012). Currently it is still difficult to investigate lineage specification and the associated establishment of global epigenetic patterns for many cell types without expanding them in vitro. However, as the number of cells required for epigenetic analyses continues to decrease, it is likely that these exciting studies will become possible in the near future.

The recent completion of the ENCODE project represents a milestone achievement for our better understanding of the human genome and epigenome. However, the information obtained is still not comprehensive. For example, just 11 of more than 60 known histone modifications were analyzed in only 46 out of 147 cell types included in ENCODE and the real number of histone modifications may be even larger (Tan et al. 2011). Most of the other assays were performed only in small subsets of cell types, suggesting that the data obtained so far may represent only a fraction of the potential functional information encoded in the human genome. An important future goal is therefore to complete these gaps, for example by current complementing international projects such as the NIH Roadmap Epigenomics Mapping Consortium (Bernstein et al. 2010), Alliance for the Human Epigenome and Disease (AHEAD) (Jones et al. 2008), and BLUEPRINT (Adams et al. 2012). Important continuations of the original ENCODE project are also provided by the modENCODE project, set to identify functional elements in selected model organisms (*Drosophila melanogaster* and *Caenorhabditis elegans*) (Celniker et al. 2009) and Mouse ENCODE project (Mouse ENCODE Consortium 2012). A better integration between ENCODE and GWAS data is likely to have a significant impact on our understanding of common human diseases. This will be further enhanced by the recent completion of the 1000 Genomes project (The 1000 Genomes Project Consortium 2012).

Acknowledgements I thank Gabriella Ficiz and Wendy Cooper for critical reading and helpful suggestions.

References

- Adams D, Altucci L, Antonarakis SE, Ballesteros J, Beck S et al (2012) BLUEPRINT to decode the epigenetic signature written in blood. *Nat Biotechnol* 30:224–226
- Agger K, Cloos PA, Christensen J, Pasini D, Rose S et al (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* 449:731–734
- Akiyama T, Suzuki O, Matsuda J, Aoki F (2011) Dynamic replacement of histone H3 variants reprograms epigenetic marks in early mouse embryos. *PLoS Genet* 7:e1002279
- Albert M, Peters AH (2009) Genetic and epigenetic control of early mouse development. *Curr Opin Genet Dev* 19:113–121
- Alberts B, Johnson A, Lewis J, Raff M, Roberts K et al (2002) *Molecular biology of the cell*, 4th edn. Garland Science, New York
- Azuara V, Perry P, Sauer S, Spivakov M, Jørgensen HF et al (2006) Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 8:532–538

- Bai L, Morozov AV (2010) Gene regulation by nucleosome positioning. *Trends Genet* 26:476–483
- Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. *Cell Res* 21:381–395
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE et al (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129:823–837
- Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A (2009) An operational definition of epigenetics. *Genes Dev* 23:781–783
- Bergsland M, Ramsköld D, Zaouter C, Klum S, Sandberg R et al (2011) Sequentially acting Sox transcription factors in neural lineage development. *Genes Dev* 25:2453–2464
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ et al (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125:315–326
- Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A et al (2010) The NIH roadmap epigenomics mapping consortium. *Nat Biotechnol* 28:1045–1048
- Bird A (2007) Perceptions of epigenetics. *Nature* 447:396–398
- Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T et al (2008) Intra-individual change over time in DNA methylation with familial clustering. *JAMA* 299:2877–2883
- Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F et al (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* 336:934–937
- Borgel J, Guibert S, Li Y, Chiba H, Schübeler D et al (2010) Targets and dynamics of promoter DNA methylation during early mouse development. *Nat Genet* 42:1093–1100
- Brenner C, Fuks F (2007) A methylation rendezvous: reader meets writers. *Dev Cell* 12:843–844
- Buecker C, Wysocka J (2012) Enhancers as information integration hubs in development: lessons from genomics. *Trends Genet* 28:276–284
- Burgold T, Spreafico F, De Santa F, Totaro MG, Prosperini E et al (2008) The histone H3 lysine 27-specific demethylase Jmjd3 is required for neural commitment. *PLoS One* 3:e3034
- Celniker SE, Dillon LA, Gerstein MB, Gunsalus KC, Henikoff S et al (2009) Unlocking the secrets of the genome. *Nature* 459:927–930
- Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW et al (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* 107:21931–21936
- Cui K, Zang C, Roh TY, Schones DE, Childs RW et al (2009) Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell Stem Cell* 4:80–93
- Daujat S, Weiss T, Mohn F, Lange UC, Ziegler-Birling C et al (2009) H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. *Nat Struct Mol Biol* 16:777–781
- Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S et al (2011) Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* 9:166–175
- Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S et al (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* 22:1775–1789
- Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T et al (2012) Landscape of transcription in human cells. *Nature* 489:101–108
- Dong X, Greven MC, Kundaje A, Djebali S, Brown JB et al (2012) Modeling gene expression using chromatin features in various cellular contexts. *Genome Biol* 13:R53
- Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW et al (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* 13:713–719
- Dovey OM, Foster CT, Cowley SM (2010) Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation. *Proc Natl Acad Sci U S A* 107:8242–8247
- ENCODE Project Consortium (2007) Identification and analysis of functional elements in 1 % of the human genome by the ENCODE pilot project. *Nature* 447:799–816

- ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74
- Epsztejn-Litman S, Feldman N, Abu-Remaileh M, Shufaro Y, Gerson A et al (2008) De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* 15:1176–1183
- Erhardt S, Su IH, Schneider R, Barton S, Bannister AJ et al (2003) Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. *Development* 130:4235–4248
- Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD et al (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473:43–49
- Farthing CR, Ficiz G, Ng RK, Chan CF, Andrews S et al (2008) Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet* 4:e1000116
- Fazio TG, Huff JT, Panning B (2008) An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* 134:162–174
- Feldman N, Gerson A, Fang J, Li E, Zhang Y et al (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* 8:188–194
- Feinberg AP (2007) Phenotypic plasticity and the epigenetics of human disease. *Nature* 447:433–440
- Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F et al (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* 473:398–402
- Fouse SD, Shen Y, Pellegrini M, Cole S, Meissner A et al (2008) Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* 2:160–169
- Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ et al (2009) Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* 460:863–868
- Gaspar-Maia A, Alajem A, Meshorer E, Ramalho-Santos M (2011) Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol* 12:36–47
- Gaulton KJ, Nammo T, Pasquali L, Simon JM, Giresi PG et al (2010) A map of open chromatin in human pancreatic islets. *Nat Genet* 42:255–259
- Grewal SI (2010) RNAi-dependent formation of heterochromatin and its diverse functions. *Curr Opin Genet Dev* 20:134–141
- Gu TP, Guo F, Yang H, Wu HP, Xu GF et al (2011) The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* 477:606–610
- Guttman M, Amit I, Garber M, French C, Lin MF et al (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458:223–227
- Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK et al (2011) lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 477:295–300
- Hawkins RD, Hon GC, Yang C, Antosiewicz-Bourget JE, Lee LK et al (2011) Dynamic chromatin states in human ES cells reveal potential regulatory sequences and genes involved in pluripotency. *Cell Res* 21:1393–1409
- He YF, Li BZ, Li Z, Liu P, Wang Y et al (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333:1303–1307
- Hemberger M, Dean W, Reik W (2009) Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nat Rev Mol Cell Biol* 10:526–537
- Ho L, Miller EL, Ronan JL, Ho WQ, Jothi R et al (2011) esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. *Nat Cell Biol* 13:903–913
- Howell CY, Bestor TH, Ding F, Latham KE, Mertineit C et al (2001) Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* 104:829–838
- Huang HS, Matevosian A, Whittle C, Kim SY, Schumacher A et al (2007) Prefrontal dysfunction in schizophrenia involves mixed-lineage leukemia 1-regulated histone methylation at GABAergic gene promoters. *J Neurosci* 27:11254–11262

- Issaeva I, Zonis Y, Rozovskaia T, Orlovsky K, Croce CM et al (2007) Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. *Mol Cell Biol* 27:1889–1903
- Inoue A, Zhang Y (2011) Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* 334:194
- Ito S, Shen L, Dai Q, Wu SC, Collins LB et al (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333:1300–1303
- Jeffries MA, Dozmorov M, Tang Y, Merrill JT, Wren JD et al (2011) Genome-wide DNA methylation patterns in CD4⁺ T cells from patients with systemic lupus erythematosus. *Epigenetics* 6:593–601
- Jiang H, Shukla A, Wang X, Chen WY, Bernstein BE et al (2011) Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell* 144:513–525
- Jones PA, Archer TK, Baylin SB, Beck S, Berger S et al (2008) Moving AHEAD with an international human epigenome project. *Nature* 454:711–715
- Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA et al (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467:430–435
- Kaji K, Caballero IM, MacLeod R, Nichols J, Wilson VA et al (2006) The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nat Cell Biol* 8:285–292
- Kaneda R, Takada S, Yamashita Y, Choi YL, Nonaka-Sarukawa M et al (2009) Genome-wide histone methylation profile for heart failure. *Genes Cells* 14:69–77
- Kanhere A, Viiri K, Araújo CC, Rasaiyaah J, Bouwman RD et al (2010) Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Mol Cell* 38:675–688
- Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M et al (2008) Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* 4:e1000242
- Kunath T, Arnaud D, Uy GD, Okamoto I, Chureau C et al (2005) Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development* 132:1649–1661
- Kundaje A, Kyriazopoulou-Panagiotopoulou S, Libbrecht M, Smith CL, Raha D et al (2012) Ubiquitous heterogeneity and asymmetry of the chromatin environment at regulatory elements. *Genome Res* 22:1735–1747
- Kwon GS, Viotti M, Hadjantonakis AK (2008) The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Dev Cell* 15:509–520
- Lan F, Bayliss PE, Rinn JL, Whetstone JR, Wang JK et al (2007) A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* 449:689–694
- Lane N, Dean W, Erhardt S, Hajkova P, Surani A et al (2003) Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* 35:88–93
- Lee MG, Villa R, Trojer P, Norman J, Yan KP et al (2007) Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* 318:447–450
- Levasseur DN, Wang J, Dorschner MO, Stamatoyannopoulos JA, Orkin SH (2008) Oct4 dependence of chromatin structure within the extended Nanog locus in ES cells. *Genes Dev* 22:575–580
- Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69:915–926
- Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M et al (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148:84–98
- Lienert F, Wirbelauer C, Som I, Dean A, Mohn F et al (2011) Identification of genetic elements that autonomously determine DNA methylation states. *Nat Genet* 43:1091–1097
- Lim CY, Tam WL, Zhang J, Ang HS, Jia H et al (2008) Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. *Cell Stem Cell* 3:543–554

- Lister R, Pelizzola M, Downen RH, Hawkins RD, Hon G et al (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462:315–322
- Loh YH, Wu Q, Chew JL, Vega VB, Zhang W et al (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 38:431–440
- Macfarlan TS, Gifford WD, Agarwal S, Driscoll S, Lettieri K et al (2011) Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev* 25:594–607
- Marks H, Kalkan T, Menafra R, Denissov S, Jones K et al (2012) The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149:590–604
- Maston GA, Evans SK, Green MR (2006) Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet* 7:29–59
- Matsui Y, Zsebo K, Hogan BL (1992) Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* 70:841–847
- Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H et al (2010) Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* 464:927–931
- Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E et al (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337:1190–1195
- Mazzarella L, Jørgensen HF, Soza-Ried J, Terry AV, Pearson S et al (2011) Embryonic stem cell-derived hemangioblasts remain epigenetically plastic and require PRC1 to prevent neural gene expression. *Blood* 117:83–87
- Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J et al (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454:766–770
- Miao F, Smith DD, Zhang L, Min A, Feng W et al (2008) Lymphocytes from patients with type 1 diabetes display a distinct profile of chromatin histone H3 lysine 9 dimethylation: an epigenetic study in diabetes. *Diabetes* 57:3189–3198
- Miao F, Chen Z, Zhang L, Liu Z, Wu X et al (2012) Profiles of epigenetic histone post-translational modifications at type 1 diabetes susceptible genes. *J Biol Chem* 287:16335–16345
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448:553–560
- Mill J, Tang T, Kaminsky Z, Khare T, Yazdanpanah S et al (2008) Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am J Hum Genet* 82:696–711
- Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M et al (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* 113:631–642
- Mohn F, Weber M, Rebhan M, Roloff TC, Richter J et al (2008) Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* 30:755–766
- Morán I, Akerman I, van de Bunt M, Xie R, Benazra M et al (2012) Human β cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in Type 2 diabetes. *Cell Metab* 16:435–448
- Mouse ENCODE Consortium (2012) An encyclopedia of mouse DNA elements (Mouse ENCODE). *Genome Biol* 13:418
- Nashun B, Yukawa M, Liu H, Akiyama T, Aoki F (2010) Changes in the nuclear deposition of histone H2A variants during pre-implantation development in mice. *Development* 137:3785–3794
- Nazor KL, Altun G, Lynch C, Tran H, Harness JV et al (2012) Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. *Cell Stem Cell* 10:620–634
- Neph S, Vierstra J, Stergachis AB, Reynolds AP, Haugen E et al (2012) An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 489:83–90
- Niwa H, Toyooka Y, Shimamoto D, Strumpf D, Takahashi K et al (2005) Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* 123:917–929
- Ng RK, Dean W, Dawson C, Lucifero D, Madeja Z et al (2008) Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nat Cell Biol* 10:1280–1290

- Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99:247–257
- O'Neill LP, VerMilyea MD, Turner BM (2006) Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nat Genet* 38:835–841
- Pan G, Tian S, Nie J, Yang C, Ruotti V et al (2007) Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* 1:299–312
- Pasini D, Bracken AP, Hansen JB, Capillo M, Helin K (2007) The Polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol Cell Biol* 27:3769–3779
- Pasini D, Hansen KH, Christensen J, Agger K, Cloos PA et al (2008) Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. *Genes Dev* 22:1345–1355
- Pasini D, Cloos PA, Walfridsson J, Olsson L, Bukowski JP et al (2010) JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature* 464:306–310
- Pei B, Sisu C, Frankish A, Howald C, Habegger L et al (2012) The GENCODE pseudogene resource. *Genome Biol* 13:R51
- Peng JC, Valouev A, Swigut T, Zhang J, Zhao Y et al (2009) Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell* 139:1290–1302
- Portela A, Esteller M (2010) Epigenetic modifications and human disease. *Nat Biotechnol* 28:1057–1068
- Probst AV, Santos F, Reik W, Almouzni G, Dean W (2007) Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* 116:403–415
- Puschendorf M, Terranova R, Boutsma E, Mao X, Isono K et al (2008) PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat Genet* 40:411–420
- Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA et al (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470:279–283
- Raiber EA, Beraldi D, Ficiz G, Burgess H, Branco MR et al (2012) Genome-wide distribution of 5-formylcytosine in ES cells is associated with transcription and depends on thymine DNA glycosylase. *Genome Biol* 13:R69
- Rakyan VK, Beyan H, Down TA, Hawa MI, Maslau S et al (2011) Identification of type 1 diabetes-associated DNA methylation variable positions that precede disease diagnosis. *PLoS Genet* 7:e1002300
- Ramsahoye BH, Biniszkievicz D, Lyko F, Clark V, Bird AP et al (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A* 97:5237–5242
- Reik W, Dean W, Walter J (2001) Epigenetic reprogramming in mammalian development. *Science* 293:1089–1093
- Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S et al (2010) KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* 463:237–240
- Sakaue M, Ohta H, Kumaki Y, Oda M, Sakaide Y et al (2010) DNA methylation is dispensable for the growth and survival of the extraembryonic lineages. *Curr Biol* 20:1452–1457
- Sandovici I, Leppert M, Hawk PR, Suarez A, Linares Y et al (2003) Familial aggregation of abnormal methylation of parental alleles at the IGF2/H19 and IGF2R differentially methylated regions. *Hum Mol Genet* 12:1569–1578
- Sandovici I, Smith NH, Ozanne S, Constância M (2008) The dynamic epigenome: the impact of the environment on epigenetic regulation of gene expression and developmental programming. In: Tost J (ed) *Epigenetics*. Caister Academic, Norfolk, pp 343–369
- Sandovici I, Smith NH, Nitert MD, Ackers-Johnson M, Uribe-Lewis S et al (2011) Maternal diet and aging alter the epigenetic control of a promoter-enhancer interaction at the Hnf4a gene in rat pancreatic islets. *Proc Natl Acad Sci U S A* 108:5449–5454

- Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ et al (2010) Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat Cell Biol* 12:853–862
- Santos F, Hendrich B, Reik W, Dean W (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* 241:172–182
- Santos F, Peters AH, Otte AP, Reik W, Dean W (2005) Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev Biol* 280:225–236
- Santos J, Pereira CF, Di-Gregorio A, Spruce T, Alder O et al (2010) Differences in the epigenetic and reprogramming properties of pluripotent and extra-embryonic stem cells implicate chromatin remodelling as an important early event in the developing mouse embryo. *Epigenetics Chromatin* 3:1
- Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE et al (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* 419:407–411
- Sanyal A, Lajoie BR, Jain G, Dekker J (2012) The long-range interaction landscape of gene promoters. *Nature* 489:109–113
- Sarmiento OF, Digilio LC, Wang Y, Perlin J, Herr JC et al (2004) Dynamic alterations of specific histone modifications during early murine development. *J Cell Sci* 117:4449–4459
- Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M (2012) Linking disease associations with regulatory information in the human genome. *Genome Res* 22:1748–1759
- Schmitz SU, Albert M, Malatesta M, Morey L, Johansen JV et al (2011) Jarid1b targets genes regulating development and is involved in neural differentiation. *EMBO J* 30:4586–4600
- Seenundun S, Rampalli S, Liu QC, Aziz A, Palii C et al (2010) UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. *EMBO J* 29:1401–1411
- Senner CE, Krueger F, Oxley D, Andrews S, Hemberger M (2012) DNA methylation profiles define stem cell identity and reveal a tight embryonic-extraembryonic lineage boundary. *Stem Cells* 30:2732–2745
- Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A et al (2012) A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 484:339–344
- Stadler MB, Murr R, Burger L, Ivanek R, Lienert F et al (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* 480:490–495
- Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M et al (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol* 9:1428–1435
- Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* 9:465–476
- Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T et al (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* 16:1779–1791
- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324:930–935
- Talbert PB, Henikoff S (2010) Histone variants—ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* 11:264–275
- Tan M, Luo H, Lee S, Jin F, Yang JS et al (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146:1016–1028
- Tanaka S, Kunath T, Hadjantonakis AK, Nagy A, Rossant J (1998) Promotion of trophoblast stem cell proliferation by FGF4. *Science* 282:2072–2075
- Teo AK, Arnold SJ, Trotter MW, Brown S, Ang LT et al (2011) Pluripotency factors regulate definitive endoderm specification through *omesodermin*. *Genes Dev* 25:238–250
- The 1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491:56–65
- Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT et al (2012) The accessible chromatin landscape of the human genome. *Nature* 489:75–82

- Torres-Padilla ME, Bannister AJ, Hurd PJ, Kouzarides T, Zernicka-Goetz M (2006) Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int J Dev Biol* 50:455–461
- Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M et al (2006) Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* 11:805–814
- van der Heijden GW, Dieker JW, Derijck AA, Muller S, Berden JH et al (2005) Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech Dev* 122:1008–1022
- Volkmar M, Dedeurwaerder S, Cunha DA, Ndlovu MN, Defrance M et al (2012) DNA methylation profiling identifies epigenetic dysregulation in pancreatic islets from type 2 diabetic patients. *EMBO J* 31:1405–1426
- Walker E, Chang WY, Hunkapiller J, Cagney G, Garcha K et al (2010) Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* 6:153–166
- Wang H, Maurano MT, Qu H, Varley KE, Gertz J et al (2012) Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res* 22:1680–1688
- Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP (2009) Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet* 41:246–250
- Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE (2010) A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* 140:99–110
- Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V et al (2011) 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun* 2:241
- Wu H, D'Alessio AC, Ito S, Xia K, Wang Z et al (2011) Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* 473:389–393
- Wu Q, Chen X, Zhang J, Loh YH, Low TY et al (2006) Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. *J Biol Chem* 281:24090–24094
- Yagi R, Kohn MJ, Karavanova I, Kaneko KJ, Vullhorst D et al (2007) Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. *Development* 134:3827–3836
- Yildirim O, Li R, Hung JH, Chen PB, Dong X et al (2011) Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* 147:1498–1510
- Zaret KS, Carroll JS (2011) Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* 25:2227–2241
- Zentner GE, Tesar PJ, Scacheri PC (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res* 21:1273–1283
- Zhao XD, Han X, Chew JL, Liu J, Chiu KP et al (2007) Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell* 1:286–298
- Ziller MJ, Müller F, Liao J, Zhang Y, Gu H et al (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet* 7:e1002389

Epigenetics and Complex Traits

Naumova, A.K.; Greenwood, C.M. (Eds.)

2013, XXII, 341 p. 43 illus., 33 illus. in color., Hardcover

ISBN: 978-1-4614-8077-8