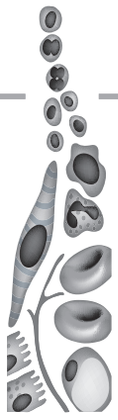

DNA methylation in cell differentiation and reprogramming: an emerging systematic view

Embryonic stem cells have the unique ability to indefinitely self-renew and differentiate into any cell type found in the adult body. Differentiated cells can, in turn, be reprogrammed to embryonic stem-like induced pluripotent stem cells, providing exciting opportunities for achieving patient-specific stem cell therapy while circumventing immunological obstacles and ethical controversies. Since both differentiation and reprogramming are governed by major changes in the epigenome, current directions in the field aim to uncover the epigenetic signals that give pluripotent cells their unique properties. DNA methylation is one of the major epigenetic factors that regulates gene expression in mammals and is essential for establishing cellular identity. Recent analyses of pluripotent and somatic cell methylomes have provided important insights into the extensive role of DNA methylation during cell-fate commitment and reprogramming. In this article, the recent progress of differentiation and reprogramming research illuminated by high-throughput studies is discussed in the context of DNA methylation.

KEYWORDS: 5-aza-cytidine ■ differentiation ■ DNA methylation ■ embryonic stem cell ■ ESC ■ induced pluripotent stem cells ■ iPS ■ pluripotency ■ reprogramming

Kevin Huang[1] & Guoping Fan[†1]

[1]Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095-7088, USA
[†]Author for correspondence:
Tel.: +1 310 267 0439
Fax: +1 310 794 5446
gfan@mednet.ucla.edu

Epigenetic modifications stably influence gene expression without changing the underlying DNA sequence [1]. Epigenetic mechanisms include: DNA methylation; histone tail modifications; chromatin remodeling; and noncoding RNA expression. Of the known epigenetic processes, DNA methylation has been extensively studied across phyla. In mammals, DNA methylation involves the covalent attachment of a methyl group to the 5′ carbon position of cytosine, conferring an additional capacity for signaling and regulatory function [2]. DNA methylation plays critical roles in many biological processes, including gene silencing, genomic imprinting and X chromosome inactivation. In addition, DNA methylation contributes to genomic stability by suppressing transposable and repetitive elements [3].

A family of DNA methyltransferases, including Dnmt1, -3a and -3b, catalyzes DNA methylation and plays key roles in establishing and maintaining methylation patterns across the genome. Dnmt3a and -3b are considered to be primarily associated with *de novo* methylation and establishing methylation patterns [4]. Dnmt1 is suggested to maintain methylation patterns by binding hemimethylated DNA during replication and copying the established methylation patterns onto newly synthesized daughter strands [5–7]. Although these three DNA methyltransferases appear to be functionally distinct, loss-of-function studies suggest that these three enzymes have some overlapping functions [8–11]. In addition, mutant mice lacking any of these enzymes exhibit global hypomethylation and die at mid-gestation or immediately after birth, indicating an essential role for methylation during development [4,5].

Cell differentiation is a process characterized by the progressive loss of developmental potential and gain in functional specialization. During this process, changes in gene expression are accompanied or caused by epigenetic programming. DNA methylation plays an important role in epigenetic programming by silencing developmental genes and activating tissue-specific genes, thus establishing a cellular memory that defines both cell lineage and cell type. By contrast, epigenetic programming can be reversed in a process called reprogramming [12]. Epigenetic marks are reset during this process, usually resulting in the gain of developmental potential. Recently, several groups have demonstrated that fully differentiated somatic cells can be reprogrammed into induced pluripotent stem (iPS) cells by forced expression of key transcription factors [13–15]. These iPS cells appear to closely resemble embryonic stem cells (ESCs) in both the transcriptome and epigenome; however, the exact epigenetic features that define and facilitate pluripotency remain elusive.

Recent advances in high-throughput technology have provided powerful tools to analyze global epigenetic features of pluripotent cells and

future medicine part of fsg

differentiated tissues in both human and mouse paradigms. In this article, the recent progress of stem cell differentiation and somatic cell reprogramming in the context of epigenetic dynamics will be discussed, focusing on the recent insights into how DNA methylation underlies differentiation and reprogramming events between the pluripotent and differentiated state.

The landscape of genome-wide DNA methylation

CpG islands

CpG islands are short regions (~1–2 kb) of DNA within the genome that are high in CG dinucleotide density, generally found in the 5′ regulatory regions of genes, and often unmethylated to enable constitutive expression [16]. A CpG island was first defined as a 200 bp moving window with CG content greater than 50% and a observed/expected CpG dinucleotide frequency of greater than 0.6 [17]. Later classifications defined a CpG island as a 500 bp moving window with greater than 55% CG content and a greater than 0.65 observed/expected CpG frequency [18]. This latter definition is more closely associated with CpG islands within promoters of genes and excludes most Alu repeat regions. These observations demonstrate a clear asymmetrical distribution of CG dinucleotides across the genome, suggesting DNA methylation is targeted for specific regions throughout the genome.

Methylation status at gene promoters

DNA methylation patterns at gene promoters reveal a correlation between promoter methylation status, gene activity and promoter CpG density. Using genome-wide approaches, several groups present consistent observations that low CpG content promoters (LCPs) are often methylated, whereas high CpG content promoters (HCPs) are frequently unmethylated; intermediate CpG content promoters (ICPs) can be either unmethylated or methylated [19–21]. Using RNA Pol-II occupancy as a metric for gene activity, LCPs were determined to be transcriptionally active regardless of methylation status, whereas HCPs appear to be generally inactive when methylated. ICPs appear to have an inverse correlation with methylation level and promoter activity [22]. Consistently, targeted bisulfite sequencing and microarray-based analysis of gene expression shows that regardless of expression, LCPs tend to be highly methylated and HCPs tend to have low methylation. On the other hand, ICP methylation is inversely correlated with gene expression [23]. Gene ontology analysis reveals that LCPs are associated with tissue-specific genes, and HCPs are associated with both ubiquitous housekeeping genes and tightly regulated developmental genes [21,24]. Although these studies have identified a correlation between methylation level, gene expression and promoter CpG density, other studies have suggested that the correlation between methylation level and gene activity is stronger with histone methylation patterns [19,21,25]. Overall, these results suggests promoter CpG density has some capacity to control the level of DNA methylation and modulate gene expression.

Methylation status of gene bodies

Apart from gene promoters, DNA methylation is also found in gene bodies, the transcribed region of genes. Gene body methylation has been identified in plants and invertebrates, and is often associated with active genes [26]. Likewise in mammals, gene body methylation has also been linked with transcriptional activity, particularly on the active X chromosome [27]. Targeted bisulfite sequencing and methylation-sensitive cut counting profiling consistently demonstrate a strong correlation between gene body methylation level and gene expression; weakly expressed genes show considerably less gene body methylation, while highly expressed genes show threefold more gene body methylation [23,28]. Furthermore, genome-wide bisulfite sequencing coupled with RNA-Seq found that the gene bodies enriched in non-CpG methylation positively correlates with gene expression in ESCs [29].

Selective sites are unmethylated

Although DNA methylation is observed in many different genomic contexts, some regions are preferentially undermethylated. Whole genome bisulfite sequencing reveals a dramatic depletion of methylated cytosines near the transcription start site (TSS) in comparison to adjacent promoters and gene bodies [29,30]. Consistently, methylated DNA immunoprecipitation on genome tiling array chip (MeDIP-chip) experiments show that unmethylated CpG islands are typically positioned within a few hundred base pairs from a known TSS, whereas methylated CpG islands are distributed over more distal positions. Bioinformatic analyses of these undermethylated regions identified several conserved *cis*-elements that may actively

protect these islands from *de novo* methylation, including Sp1 and Stat1 [31]. Indeed, Sp1 along with FoxD3 and E-box proteins have been reported for establishing unmethylated marks in select tissue-specific enhancers in ESCs [32]. Interestingly, chromatin immunoprecipitation on histone-3-lysine-4-methyl shows enrichment in unmethylated islands (92%), suggesting CpG undermethylation may be directed by histone cues [31]. These data support previous findings that Dnmt3l (an enzymatically inactive member of the DNA methyltransferase family and regulator of Dnmt3a and -3b activity) interaction with histone H3 is strongly inhibited by mono-, di-, or tri-methylation at H3K4, thereby abrogating recruitment of *de novo* methyltransferase to histones. This indicates that *de novo* methylation is inhibited as a consequence of histone modifications [33]. Furthermore, ChIP-Seq experiments on DNA-binding proteins reveal a marked decrease in methylation density at the site of interaction [29]. Intriguingly, enrichment of enhancer-binding proteins are inversely correlated to the methylation status of tissue-specific enhancer regions, suggesting DNA methylation can regulate tissue-specific genes by sterically hindering protein–DNA interactions. Peculiarly, whereas DNA methylation in promoters and enhancers appears to repress transcription, cytosine methylation on the template strand of gene bodies does not seem to disrupt RNA polymerase activity and is frequently associated with increased transcriptional activity [29]. This puzzling dichotomous role of DNA methylation remains to be fully understood.

DNA methylation in the early embryo

Over the course of mammalian development, the genome undergoes many major DNA methylation remodeling events. Immediately after fertilization and prior to the first cell division, the paternal genome undergoes a massive wave of demethylation. In fact, genome-wide paternal demethylation is one of the strongest supporting evidence for active demethylation in mammals [34–37]. However, how this wave of zygotic demethylation occurs is largely unknown. A recent study demonstrates an important role for Elp3, an elongator complex factor, in mediating active paternal demethylation [38]. The maternal genome also undergoes genome-wide depletion of methylation. However, loss of methylation occurs passively over the next several cell divisions, partially due to Dnmt1 exclusion in the cytoplasm by an unknown mechanism [39,40]. By the eight cell stage of development, both parental and maternal genomes have comparably low levels of DNA methylation [40,41]. DNA methylation patterns are then progressively re-established, marking gradual commitment towards lineage-specific differentiation [42].

DNA methylation in early cell-fate decisions

The earliest cell-fate decision is established by DNA methylation during the blastocyst stage, when the trophectoderm (TE) delineates from the inner cell mass (ICM). The TE, along with the primitive endoderm, becomes fixed in its developmental potential toward placental tissue, whereas the ICM is committed to form the embryo proper [42]. A 5-methylcytosine staining reveals that the TE has drastically reduced global methylation levels compared with the ICM [37]. Intriguingly, loss of methylation in ESCs confers an ability to take on trophoblast lineage and form TE in chimeras, suggesting a loss of lineage restriction within the ICM [43]. This lineage restriction appears to be controlled by the transcription factor Elf5, which is hypomethylated and highly expressed in the trophoblast lineage, but hypermethylated and silenced in the embryonic lineage. Loss of Elf5 results in defective extra-embryonic ectoderm development and failure to derive self-renewing trophoblast stem cells *in vitro*, suggesting Elf5 contributes to the trophoblast lineage identity [44]. Interestingly, Elf5 methylation timing corresponds to increased Dnmt3a expression during the critical time period when *de novo* methylation establishes methylation patterns for downstream development [45]. Together, these results indicate that DNA methylation plays an essential role during the first few cell divisions and early cell-fate commitment by temporally regulating key genes.

DNA methylation in ESCs

Embryonic stem cells are a special population of pluripotent cells derived from the ICM of a blastocyst during mammalian development [46]. ESCs retain the ability to indefinitely self-renew and differentiate into all cell types found in the adult body. For these reasons, human ESCs have received considerable attention for their therapeutic potential in regenerative medicine. Emerging data from genome-wide studies have identified critical features that address both ESC maintenance and differentiation. The advent of microarray and high-throughput sequencing

technology has greatly assisted the scrutiny of global methylation patterns in both ESCs and somatic cells, allowing the identification of broad and novel patterns of DNA methylation.

DNA methylation is essential for pluripotency but not self-renewal

Embryonic stem cells deficient in Dnmt1 and/or Dnmt3a/3b maintain the ability to self-renew and survive in extended culture; however, these ESCs show a loss of pluripotency and severe impairment of differentiation potential [47–49]. A similar phenomenon is observed in ESCs lacking CpG-binding proteins [50]. These results suggest that DNA methylation establishes the epigenetic environment, which facilitates the transcriptional network necessary for cell differentiation, but not self-renewal.

Embryonic stem cells may maintain pluripotency by protecting pluripotent genes from being silenced via *de novo* DNA methylation. For example, the promoter of Oct4, a critical transcription factor involved in pluripotency, is hypomethylated in ESCs but becomes hypermethylated in somatic cells. In addition, genome-wide CpG island analysis in human ESCs reveals various other pluripotency genes that are unmethylated in ESCs but become methylated in somatic cells [31,51]. The search for *cis*-elements involved in protecting *de novo* promoter methylation identified six significant motifs across all unmethylated CpG islands; two of these motifs are known consensus binding sequences for transcription factors Sp1 and Stat1 [31]. Sp1 elements have previously been reported to protect CpG islands from *de novo* methylation [52]. In particular, the Sp1 site in the Oct4 promoter has been shown to direct local demethylation in ESCs [53]. Thus, ESCs can maintain pluripotency by protecting genes involved in stemness from undergoing *de novo* methylation silencing. The regulatory elements involved in protecting these promoters have yet to be fully elucidated and experimentally confirmed. Overall, loss of DNA methylation and hypomethylation of pluripotency genes provide supporting evidence for the critical role of DNA methylation in maintaining pluripotency.

CpG methylation contributes to the silencing of differentiation genes in ESCs

In addition to the requirement for pluripotent gene promoters to be hypomethylated to enable expression, hypermethylation of differentiation genes is also required for suppressing the expression of tissue-specific genes. Genomic analysis provides supporting evidence for the CpG methylation of gene promoters to selectively silence differentiation genes. Gene expression profiles in wild-type and Dnmt1/3a/3b triple knockout (TKO) mouse ESCs identified 337 upregulated and 113 downregulated genes in TKO cells compared with wild-type cells [21]. Gene ontology of the upregulated genes in TKO cells shows over-representation of tissue-specific genes, in particular genes associated with the germ lineage. In addition, many of the upregulated genes (>threefold more than expected) mapped to the X-chromosome. These results demonstrate that DNA methylation is involved in repressing differentiation genes in ESCs, and global DNA demethylation is mostly linked with the upregulation of tissue-specific genes. Comparison between ESC and fibroblast promoter methylation reveals very few hypermethylated genes that were associated with gene activation [51]. Together, CpG methylation within the promoters contributes to gene silencing, whereas hypomethylation enables constitutive expression.

ESCs are enriched in non-CpG methylation

Non-CpG methylation has been extensively studied in plants and invertebrates, but has rarely been studied in mammalian systems. Until recently, DNA methylation was believed to occur predominantly at CpG dinucleotides and non-CpG methylation has been mostly overlooked. An early study examining non-CpG methylation reports that non-CpG methylation is prevalent only in mouse ESCs, but not somatic cells, estimating 15–20% of methylated cytosines to be at non-CpG sites, with CpA as the major non-CpG motif [54]. Several studies have implicated Dnmt3a in catalyzing non-CpG methylation [54–58]. In addition, a recent study demonstrated that Dnmt3b can catalyze non-CpG methylation in human skeletal muscle [59]. In mouse ESCs, knockout of both Dnmt3a and -3b, but not Dnmt1a, show a drastic loss of *de novo* non-CpG methylation on Maloney murine leukemia virus proviral DNA [56]. Although both Dnmt3a and -3b appear to contribute to non-CpG methylation in mouse ESCs, it is still unclear which enzyme is the major contributing one and whether these enzymes have an overlapping function.

A recent major study using Methyl-Seq technology reports significant non-CpG methylation in human ESCs, estimating nearly 25%

of total cytosine methylation to be non-CpG sites, with CHG and CHH as the major motifs (where H = A, C, or T) [29]. By contrast, fibroblast cells show only 0.02% non-CpG methylation. Interestingly, reprogrammed fibroblast cells regain non-CpG methylation patterns, but at abnormal levels, indicating non-CpG methylation may be a general feature exclusive to pluripotent cells. The majority of non-CpG methylation in ESCs have enriched regions within the gene body and show positive correlation with gene activity. Genomic regions enriched in non-CpG methylation are associated with genes involved in RNA processing, RNA splicing and RNA metabolic processes. Interestingly, enrichment of non-CpG methylation in gene bodies correlates with significantly more intronic RNA. Furthermore, non-CpG methylation appears to be biased for the antisense strand of gene bodies, which serves as the template strand for RNA polymerization [29]. A more recent methylome study confirms earlier findings [30]. In particular, the authors note that the major non-CpG motif is CpA, with no significant nucleotide bias on the +2 position of the methylated cytosine [30]. Interestingly, some non-CpG marks in ESCs are conserved in fibroblasts, suggesting non-CpG methylation patterns are nonrandom [30]. Overall, these results indicate that non-CpG methylation may be a general feature pervasive in pluripotent cells occuring in a nonstochastic process. However, the functional role of non-CpG methylation is still unclear.

DNA methylation during differentiation

Global DNA methylation levels remain largely unchanged between ESCs and somatic cells, suggesting differentiation is dictated by the redistribution of methylation patterns across the genome [19,20,60]. DNA methylation is important for temporal control of differentiation and loss of methylation usually results in premature differentiation [61–65]. In addition, DNA methylation appears to control only a subset of critical tissue-specific genes that define cell lineage and cell type (FIGURE 1A). Recent studies using high-throughput technologies provide high-resolution details on global DNA methylation changes during differentiation.

DNA methylation in multipotent stem cells

Multipotent stem cells provide a unique intermediate between pluripotent ESCs and unipotent differentiated cells. Mesenchymal stem cells (MSCs) are a population of adult stem cells derived from the bone marrow and restricted to forming preferentially mesodermal cell types [66]. Adipose tissue-derived stem cells (ADSCs) are isolated from liposuction material, which provide an abundant source of MSCs [67,68]. A few studies show that select adipogenic and nonadipogenic promoters in MSCs, ADSCs and hESCs are hypomethylated and hypermethylated, respectively, suggesting DNA methylation controls adipogenic differentiation by activating adipogenic-related genes and silencing nonspecific lineage genes [69,70]. A study using high-throughput analysis compared human embryonic carcinoma stem cells (ECCs) with multipotent MSCs and ADSCs. ECCs closely resemble ESCs and have been proposed to be suitable surrogates for ESC research [71]. Another study reports that 83 genes (~4% of the total analyzed gene promoters) are differentially methylated between ECCs and all multipotent adult stem cells, 82 of which were hypomethylated in the adult stem cells and hypermethylated in the ECCs [72]. These hypomethylated genes were highly expressed in the adult stem cells and corresponded to genes involved in differentiation. Together, these results indicate that cell differentiation involves select demethylation to enable the expression of key differentiation genes.

Hematopoietic stem cells (HSCs) are a special population of multipotent stem cells that are derived from the bone marrow and give rise to all blood types. DNA methylation plays important roles in hematopoietic self-renewal and differentiation. Hematopoietic lineage-directing genes have hypomethylated promoters and are actively transcribed, indicating that activation of specific lineage-commitment genes is coupled with selective DNA demethylation [73,74]. Interestingly, HSCs deficient in both Dnmt3a and -3b show a loss of proliferative ability but retain differentiation potential, suggesting *de novo* methylation is important for self-renewal in HSCs [75]. Recent studies of HSCs deficient in Dnmt1 also show a loss of self-renewal ability; however, developmental potential was also affected [61,62]. In one study, lymphopoiesis was severely compromised, but not myelopoiesis or erythropoiesis [61]. Expression analysis has demonstrated the downregulation of genes involved in stemness and lymphopoiesis, and premature upregulation of myeloid and erythroid regulators, indicating the methylation environment was permissive to myelo-erythroid, but not lymphoid, differentiation.

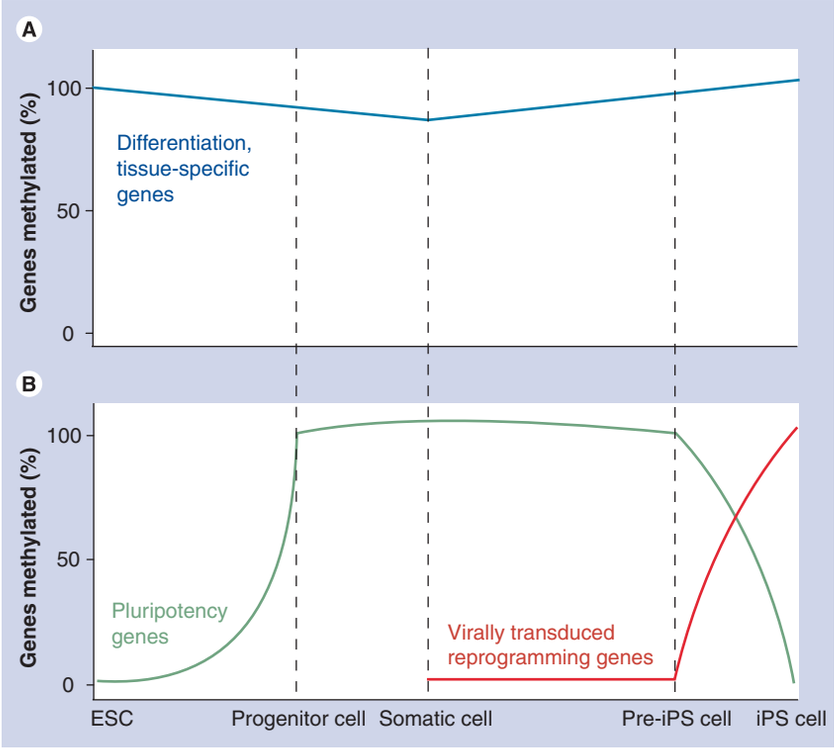

Figure 1. Dynamic regulation of DNA methylation during *in vitro* differentiation of embryonic stem cells or derivations of induced pluripotent stem cells. (A) A subset of lineage-specific genes undergoes dynamic methylation. **(B)** Pluripotency genes undergo dramatic changes in DNA methylation during differentiation and reprogramming.
ESC: Embryonic stem cell; iPS: Induced pluripotent stem.

Surprisingly, a separate expression analysis of Dnmt1-deficient HSCs reveals a upregulation of self-renewal genes, suggesting a novel regulatory mechanism may be involved in controlling proliferation in HSCs [62]. In addition, the study revealed that a loss of Dnmt1 enhances myeloid differentiation, which is associated with precocious upregulation of genes found in further differentiated states. Consistent with the role of DNA methylation during differentiation, Dnmt1 maintains stemness in myeloid progenitors by silencing downstream differentiation genes. Collectively, these results demonstrate that proper DNA methylation patterning is critical for both self-renewal and lineage commitment in HSCs. However, a critical missing piece of information is the global methylation map to examine the relationship between methylation patterns and genes involved in self-renewal and differentiation.

Multipotent neural progenitor cells (NPCs) can give rise to neurons, astrocytes and oligodendrocytes [76]. Although methylation levels between ES-differentiated NPCs and ESCs are highly similar, there are clear differences. Global promoter analysis identified 343 hypermethylated and 22 hypomethylated promoters in NPCs compared with ESCs [20]. Close inspection of the hypomethylated promoters reveal genes activated during brain-specific differentiation, consistent with selective demethylation of tissue-specific genes. Comparisons between ESCs, NPCs and terminally differentiated neurons demonstrate that the majority of *de novo* methylated genes are already present in NPCs, suggesting that the bulk of DNA methylation changes during differentiation is associated with a loss of pluripotency and a commitment to a multipotent state, rather than terminal differentiation [20]. Reduced representation bisulfite sequencing reveal similar observations; approximately 8% of unmethylated CpG sites in ESCs become methylated in NPCs, whereas only approximately 2% of methylated CpGs in ESCs become umethylated in NPCs [19].

Although the general trend appears to be a gain of DNA methylation during differentiation, a study has found abnormal hypermethylation of CpG islands in ES-derived NPCs compared with primary NPCs [77]. This indicates that *in vitro* differentiation involves artificial CpG island hypermethylation. Interestingly, when NPCs further differentiate into astroglial lineage, selective promoter demethylation occurs in glial marker genes, including GFAP and S100β [63,64]. Loss of methylation using 5-aza-cytidine (AZA) also triggers premature glial differentiation [63]. Consistently, Dnmt1-deficient NPCs precociously differentiate into astroglial cells, which have been linked to increased JAK-STAT signaling and demethylation of the STAT1 and GFAP promoters [64]. Thus, NPC differentiation involves the controlled demethylation of select genes.

More recently, a study in epidermal progenitors further supports the role of DNA methylation in maintaining self-renewal and regulating proper differentiation [65]. Using a xenograft model of control and DNMT1-knockdown human keratinocytes, the authors demonstrate how the loss of Dnmt1 leads to cell-intrinsic failure of tissue self-renewal and premature differentiation. Global gene expression profiles of Dnmt1-deficient cells reveals upregulation of cell-cycle inhibitors and differentiation genes. MeDIP-chip experiments reveal an overabundance of methylated differentiation promoters, indicating that methylation controls the expression of differentiation genes. Moreover, this study demonstrates growth arrest and DNA damage-inducible proteins, Gadd45a/b, can promote differentiation by actively demethylating

promoters of differentiation genes [65]. These results show that DNA methylation regulates differentiation by restricting expression and selectively demethylating nonlineage and lineage-specific genes, respectively. Overall, DNA methylation is dynamically regulated upon cell differentiation and facilitates proper temporal expression of differentiation genes.

DNA methylation in germ cells & tissue-specific cells

During the course of germ cell development, the germ line undergoes massive epigenetic remodeling, including genome-wide erasure of DNA methylation marks. This process has been proposed to be necessary to confer primordial germ cells (PGCs) with the ability to eventually generate a totipotent state when the gametes come together [78]. Although PGC reprogramming also involves extensive chromatin modifications, DNA demethylation appears to precede chromatin remodeling [79]. Consistent with conventional knowledge, genome-wide bisulfite sequencing in mouse sperm, ESCs and PGCs reveal a dramatic loss of global methylation in PGCs, whereas sperm and ESCs share similar global methylation levels [80]. The same study also examined the role of AID, a deaminase implicated in demethylation, during PGC reprogramming. Bisulfite sequencing on AID-deficient PGCs reveals a marked increase in global methylation levels in all genomic elements, suggesting that AID plays a significant role in PGC reprogramming [80]. Embryonic germ (EG) cells, derived from PGCs, are a special type of unipotent cell and share many characteristics with pluripotent stem cells [81]. A study comparing ESCs, EG cells and sperm cells reveals remarkably similar methylation patterns [51]. As expected, there is a strong correlation between ESCs and EG cells, with both sharing pluripotent features. Surprisingly, sperm cells also share a high correlation with ESCs and EG cells. Upon closer scrutiny, many of the similarities are found within the promoters; most promoters hypermethylated in ESCs are also hypermethylated in sperm cells, but not fibroblast cells. Overall, this suggests promoter methylation patterns in sperm closely resemble pluripotent ESCs and EG cells [51]. Recent genome-wide studies on sperm DNA reveal a large number of new sperm-specific methylation patterns [31]. Methylation status prediction of these spermatogenesis-related islands overlap with 96% of regions methylated in ESCs, further supporting sperm cell resemblance to ESCs. Interestingly, over 1500 islands are specifically undermethylated in sperm, but not other somatic tissues, suggesting tissue-specific gene expression via selective demethylation [31].

Tissue-specific differentially methylated regions (TDMRs) have been proposed as distinguishing profiles between cell types. Previous studies have found considerable association between tissue-specific promoter methylation and gene expression [82–84]. A recent report provides methylation profiles of DNA (mPod) for human TDMRs in 16 different human tissues [85]. This study approximates that 18% of the genomic regions analyzed were classified as TDMRs, suggesting tissue-specific DNA methylation in promoter CpG islands is relatively common. Interestingly, a motif search on all promoter TDMRs identified Sp1 and Klf4 as putative transcription factors involved in modulating methylation levels during differentiation. Gene ontology analysis on these TDMRs reveals enrichment for tissue-specific function [85]. Additional studies comparing different tissue samples report that a large number of CpG islands are fully methylated in all cell types but unmethylated in a single tissue [31]. These results echo the theme of *de novo* methylation and selective demethylation of promoter-TDMRs in a tissue-specific manner.

Although promoter DNA methylation has been a primary focus on identifying TDMRs, increasing evidence suggests that many TDMRs are found away from the proximal promoter, including exons, introns and intergenic regions. A study examining nonpromoter CpG islands predicted only half to function similar to canonical promoter CpG islands [85]. Virtual restriction landmark genomic scanning and Sequenom MassARRAY also identified that the majority of TDMRs were located in nonpromoter intragenic regions [86]. CpG island array studies also report that TDMRs are located disproportionately distant from the TSS [84]. Comprehensive high-throughput array-based relative methylation (CHARM) analysis between iPS cells and fibroblasts identified 4401 differentially methylated regions (DMRs), 70% of which were associated with CpG island shores and more than half of the DMRs were located 2kb outside of the TSS, indicating that the majority of DMRs are located in distal regions [87]. Interestingly, gene ontology analysis of hypomethylated DMRs shows significant overlap with known binding sites for Oct4, Nanog and Sox2, but not hypermethylated DMRs, suggesting specific promoter demethylation is tightly linked

to pluripotency factor-binding accessibility [87].

Whole genome bisulfite sequencing comparisons between H1 ESCs and IMR90 fibroblast cells identified 491 DMRs, defined by hypermethylation in fibroblasts compared with ESCs [29]. DMRs hypomethylated in ESCs are associated with factors involved in ESC function. In addition, a large proportion of the fibroblasts' genome has lower levels of CpG methylation than ESCs. These large contiguous regions (mean length = 153 kb) map to partially methylated domains, which comprise 38.4% of the autosome. Lower levels of methylation of partially methylated domains in fibroblasts correspond to downregulated genes and overlap with repressive histone marks [29]. Overall, DNA methylation patterns are drastically different between pluripotent cells and differentiated cells. These results demonstrate that DNA methylation helps define cellular identity by globally altering the gene expression of lineage-specific differentiation genes.

DNA methylation in reprogrammed cells

Recent advances in reprogramming provide new tools for evaluating the epigenetic changes during dedifferentiation. Although reprogramming entails massive epigenetic changes, much of our understanding of the epigenetic changes underlying the reprogramming process is limited. Our understanding of gradual epigenetic changes during reprogramming is restricted to the beginning and end states. Although partially reprogrammed cells provide an opportunity to evaluate epigenetic changes at an intermediate state, the stability and definition of these cells remain to be fully elucidated [12]. During the course of reprogramming, *de novo* methylation and selective promoter demethylation is required for silencing transduced retroviral promoters and activating endogenous pluripotent genes (**Figure 1B**) [88].

Selective demethylation

It is still unclear whether demethylation during reprogramming occurs passively or actively. During passive demethylation, reprogramming factors may bind directly to promoter or enhancer regions, to prevent Dnmt1 from binding hemimethylated DNA during replication. Alternatively, Dnmt1 may be directly inhibited or excluded from the nucleus during replication. Since the full course of reprogramming normally takes weeks, this suggests that numerous cell divisions are required for reprogramming, providing evidence to support passive demethylation. Additional evidence comes from increased reprogramming efficiency when Dnmt1 is inhibited using AZA [89,90]. In these experiments, partially reprogrammed cells were treated briefly (48 h) with AZA, which resulted in an over 30-fold increase in reprogramming efficiency [89]. siRNA knockdown of *Dnmt1* also improved reprogramming efficiency.

DNA demethylation may also occur through active demethylation, where enzymes directly remove the methyl-group from methylated cytosines independent of DNA replication [91]. In mammals, there are several putative enzymes that appear to have demethylase activity, although their activity remains controversial [92–95]. Both Dnmt3a and -3b have been proposed to carry deamination activity and are thought to be candidate players in active demethylation [96]. A study demonstrated the requirement for AID-dependent demethylation in reprogramming heterokaryons [97]. AID-dependent demethylation involves deamination of 5-methylcytosine to thymidine by AID, followed by G/T mismatch repair by DNA glycosylases [94]. siRNA knockdown of AID greatly reduced active demethylation at the Oct4 and Nanog promoter, concomitant with reduced Oct4 and Nanog expression. Ectopic expression of human AID in AID-knockdown cells completely rescued demethylation at the Nanog promoter and partially rescued demethylation at the Oct4 promoter, indicating the specificity of AID to Oct4 and Nanog promoters. Interestingly, AID can directly bind to the methylated promoters of Oct4 and Nanog in fibroblast cells but not the unmethylated promoters in reprogrammed cells, suggesting additional factors may be required for AID activity during reprogramming [97]. These results demonstrate that AID is required for actively demethylating promoters of pluripotency genes during reprogramming. Altogether, it is clear that selective demethylation is required for reprogramming.

Do iPS cells fully recapitulate embryonic cell state?

One major question in the reprogramming field is whether fully reprogrammed somatic cells truly mimic ESCs. A recent study comparing human ESCs and iPS cells generated from different laboratories report a consistent genetic signature that distinguishes ESCs and iPS cells, arguing that reprogramming follows a nonstochastic process but does not fully recapitulate the ES state [98]. In the same report, late-passage iPS cells appear to better resemble ESCs than early iPS cells, indicating extended culture confers iPS

cells with more ES-like features. Peculiarly, ESCs that are held in prolonged culture accumulate abnormal methylation patterns [77,99–102]. One study comparing early- and late-passage ESCs found significantly higher methylation levels in the tumor suppressor gene, *Rassf1*, indicating extended culture may induce genetic changes [101]. Another study reported that methylation changes are dissimilar between different human ESC lines and that the degree of change was proportional to the number of passages, suggesting that the methylation changes are acquired progressively and stochastically [99]. It seems puzzling that extended culture of iPS cells more resembles ESCs, whereas prolonged culture of ESCs causes a drift away from the original ES state. Intriguingly, extended culture of human iPS cells has been reported to result in chromosomal abnormalities [103]. However, whether the process of reprogramming has inherent drift and whether the extent of drift is comparable to ESC maintenance is unclear.

Another paradoxical observation comes from the use of AZA both in differentiation and reprogramming. As mentioned earlier, AZA greatly improves the reprogramming efficiency to a pluripotent state; however, AZA treatment of ESCs promotes differentiation to cardiomyocytes [104]. AZA is a general inhibitor of Dnmts and should have global effects consistent in both systems; however, AZA seems to have a bidirectional effect. This may suggest that global demethylation alone is not sufficient in determining the decision towards differentiation or reprogramming, but may involve the current state of the chromatin. Indeed, other chemical drugs that directly affect the chromatin such as valproic acid (VPA), a histone deacetylase inhibitor, contribute greater reprogramming efficiency than AZA [105–108]. Furthermore, previous studies report that the addition of AZA or VPA to bone marrow stromal cells results in the transdifferentiation to neuronal stem cells, adding more complexity to the effects of AZA and VPA [109,110].

Although ESCs and iPS cells have very similar DNA methylation patterns, comparisons between the two methylomes reveal a subset of DMRs that distinguishes ESCs from iPS cells [87]. One study reports that 71 regions were differentially methylated between ESCs and iPS cells, 51 of which were hypermethylated in iPS cells. Gene ontology analysis determined that these hypermethylated genes in iPS cells correspond to genes involved in the developmental processes. When comparing ESCs and iPS cells to fibroblasts, both ESCs and iPS cells change in the same direction; however, iPS cells tend to have much more exaggerated alterations. Intriguingly, iPS cells had DMRs that were distinct from both the ESC and parental fibroblast cell, indicating that iPS cells exist in some unique epigenetic state. Overall, these results demonstrate that although small in number, there are clear methylation differences between ESCs and iPS cells. Whether these methylation differences are significant and biologically meaningful remain unclear.

Future perspective

We are only beginning to get a glimpse of the complexity of DNA methylation in gene regulation. The emerging picture from genome-wide studies reveals widespread epigenetic changes that support the hypothesis that differentiation involves the progressive reduction of developmental potential determined by increasing restrictive epigenetic marks. DNA methylation is considerably prevalent throughout the genome and its effect on gene expression varies depending on the genomic context. How the widespread presence of DNA methylation affects pluripotency, differentiation and reprogramming has yet to be investigated in detail. Although not discussed in detail in this article, differentiation and reprogramming depend on a myriad of other epigenetic marks that act in concert with DNA methylation. In particular, various histone modifications have been strongly correlated with the occurrence of DNA methylation [21,111,112]. Interestingly, different epigenetic mechanisms can have opposing effects on the same locus; these 'bivalent domains' add to the complexity of epigenetic control on gene expression. Bivalent domains have been proposed to poise developmental genes for rapid changes in gene expression during differentiation [19,87,113]. Furthermore, recent studies have uncovered novel facets of DNA methylation in mammals, including methylation at non-CpG contexts and hydroxylation of methylated cytosines [114–116], both of which appear to be prevalent to ESCs. How these new players are involved in regulating pluripotency is a burgeoning field ripe for comprehensive exploration.

The rapid pace of next-generation sequencing will undoubtedly greatly assist the effort to extract the methylome changes during differentiation and reprogramming. As whole genome sequencing becomes more routine, studies following the methylomic changes during differentiation and reprogramming will divulge the relationship between the two processes. Much is still unclear about the process of reprogramming;

whether reprogramming is differentiation in a reverse order or a novel dedifferentiation path is largely unknown. By thoroughly examining the methylomes of intermediate stages between ES, somatic and iPS cells, we will learn how to manipulate the methylation environment to improve efficiencies in both differentiation and reprogramming. One major concern for the biomedical application of stem-cell therapeutics is the acquired methylation aberrations of pluripotent stem cells in extended culture or upon differentiation [101–103]. Genome-wide studies of prolonged cultured cells will allow proper evaluation of whether these methylation changes significantly alter the utility of stem cells in therapy. In addition, studies at single base-pair resolution will shed light on whether gene regulation by DNA methylation is determined by the methylation of specific cytosines or the methylation status of an entire region. Moreover, genome-wide bisulfite sequencing may provide detailed insights on the targets of methyltransferases and demethylases. Collectively, advances in sequencing technology will become invaluable tools for addressing major questions in DNA methylation and stem cell biology.

In summary, DNA methylation plays extensive roles in stem cell maintenance, differentiation and reprogramming. Although much progress has been achieved in the last year, high-resolution genomic analysis is still in its early stages. The combination of unbiased high-throughput technology, bioinformatics and biology will resolve many questions associated with DNA methylation during development and reprogramming. The comprehensive and detailed panorama of DNA methylation dynamics will greatly contribute to understanding the mechanisms of differentiation and reprogramming, and expedite the progress of stem cell-based therapeutics in regenerative medicine.

Financial & competing interests disclosure

The authors are supported by NIH and NINDS grants RO1 051411. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Landscape of genome-wide DNA methylation

- DNA methylation is prevalent throughout the genome, found in promoters, gene bodies, and intergenic and repetitive regions.
- Although DNA methylation is targeted to specific regions within the genome, no evidence supports the existence of *cis*-acting elements directing DNA methylation.
- CpG methylation in gene promoters is frequently associated with gene silencing.
- Gene body methylation in mammals shares functional similarities with plants and correlates with transcriptional activity.

DNA methylation in early embryo & embryonic stem cells

- Early embryogenesis is characterized by a wave of demethylation of both paternal and maternal strands, followed by a wave of *de novo* methylation.
- DNA methylation silences key factors during early-stage lineage commitment, such as *Elf5*.
- Loss of DNA methylation impairs embryonic stem cell (ESC) developmental potential but not self-renewal.
- Non-CpG methylation is prevalent in ESCs but not somatic cells; the function of non-CpG methylation is still unknown.

DNA methylation in differentiated cells

- Differentiation involves the progressive gain of DNA methylation that marks the loss of developmental potential and the commitment to a more differentiated state.
- DNA methylation defines cellular identity by silencing nonrelated lineage-specific genes and selective demethylation in tissue-specific genes is associated with gene activation and lineage differentiation.
- Tissue-specific differential methylated regions are often found away from transcriptional start sites and at shores of CpG islands.

DNA methylation in reprogrammed cells

- Methylation patterns in reprogrammed cells are highly similar to ESCs but show clear differences.
- Reprogramming requires global DNA methylation changes, including the selective demethylation of pluripotency genes.
- Recent evidence suggests that *de novo* methylation contributes to silence retrovirally transduced reprogramming factors and differentiation genes.

Bibliography

Papers of special note have been highlighted as:
- ▪ of interest
- ▪▪ of considerable interest

1 Wu C, Morris JR: Genes, genetics, and epigenetics: a correspondence. *Science* 293(5532), 1103–1105 (2001).

2 Reik W: Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447(7143), 425–432 (2007).

3 Miranda TB, Jones PA: DNA methylation: the nuts and bolts of repression. *J. Cell. Physiol.* 213(2), 384–390 (2007).

4 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).

▪ **First *in vivo* loss-of-function study of *de novo* methyltransferases, establishing the critical role of methylation for development.**

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

▪▪ **First *in vivo* loss-of-function study of major methyltransferase, Dnmt1, establishing the critical role of methylation for development.**

6 Leonhardt H, Page AW, Weier HU, Bestor TH: A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71(5), 865–873 (1992).

7 Liu Y, Oakeley EJ, Sun L, Jost JP: Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. *Nucleic Acids Res.* 26(4), 1038–1045 (1998).

8 Kim GD, Ni J, Kelesoglu N, Roberts RJ, Pradhan S: Co-operation and communication between the human maintenance and *de novo* DNA (cytosine-5) methyltransferases. *EMBO J.* 21(15), 4183–4195 (2002).

9 Chen T, Ueda Y, Dodge JE, Wang Z, Li E: Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* 23(16), 5594–5605 (2003).

▪ **Demonstrates partial redundant roles of DNA methylation between the methyltransferases in embryonic stem cells (ESCs).**

10 Hattori N, Abe T, Suzuki M *et al.*: Preference of DNA methyltransferases for CpG islands in mouse embryonic stem cells. *Genome Res.* 14(9), 1733–1740 (2004).

11 Feng J, Zhou Y, Campbell SL *et al.*: Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat. Neurosci.* 13(4), 423–430 (2010).

▪ **Demonstrates overlapping function of DNA methyltransferases in postmitotic neurons.**

12 Hochedlinger K, Plath K: Epigenetic reprogramming and induced pluripotency. *Development* 136(4), 509–523 (2009).

13 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).

14 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).

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

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

17 Gardiner-Garden M, Frommer M: Cpg islands in vertebrate genomes. *J. Mol. Biol.* 196(2), 261–282 (1987).

18 Takai D, Jones PA: Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl Acad. Sci. USA* 99(6), 3740–3745 (2002).

19 Meissner A, Mikkelsen TS, Gu H *et al.*: Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454(7205), 766–770 (2008).

▪ **Genomic map of both DNA methylation and histone markers.**

20 Mohn F, Weber M, Rebhan M *et al.*: Lineage-specific polycomb targets and *de novo* DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* 30(6), 755–766 (2008).

21 Fouse SD, Shen Y, Pellegrini M *et al.*: Promoter Cpg methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* 2(2), 160–169 (2008).

▪ **Genomic map of both DNA methylation and histone markers.**

22 Weber M, Hellmann I, Stadler MB *et al.*: Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* 39(4), 457–466 (2007).

23 Ball MP, Li JB, Gao Y *et al.*: Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat. Biotechnol.* 27(4), 361–368 (2009).

24 Saxonov S, Berg P, Brutlag DL: A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl Acad. Sci. USA* 103(5), 1412–1417 (2006).

25 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).

26 Suzuki MM, Bird A: DNA methylation landscapes: Provocative insights from epigenomics. *Nat. Rev.* 9(6), 465–476 (2008).

27 Hellman A, Chess A: Gene body-specific methylation on the active X chromosome. *Science* 315(5815), 1141–1143 (2007).

28 Deng J, Shoemaker R, Xie B *et al.*: Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat. Biotechnol.* 27(4), 353–360 (2009).

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

▪▪ **Detailed methylation map by deep genome-wide bisulfite sequencing in ESCs and fibroblast cells with cross reference to various histone ChIP-seq experiments.**

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

31 Straussman R, Nejman D, Roberts D *et al.*: Developmental programming of CpG island methylation profiles in the human genome. *Nat. Struct. Mol. Biol.* 16(5), 564–571 (2009).

32 Xu J, Watts JA, Pope SD *et al.*: Transcriptional competence and the active marking of tissue-specific enhancers by defined transcription factors in embryonic and induced pluripotent stem cells. *Genes Dev.* 23(24), 2824–2838 (2009).

33 Ooi SK, Qiu C, Bernstein E *et al.*: Dnmt3l connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* 448(7154), 714–717 (2007).

34 Oswald J, Engemann S, Lane N *et al.*: Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* 10(8), 475–478 (2000).

35 Mayer W, Niveleau A, Walter J, Fundele R, Haaf T: Demethylation of the zygotic paternal genome. *Nature* 403(6769), 501–502 (2000).

36 Rougier N, Bourc'his D, Gomes DM *et al.*: Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* 12(14), 2108–2113 (1998).

37 Santos F, Hendrich B, Reik W, Dean W: Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* 241(1), 172–182 (2002).

38 Okada Y, Yamagata K, Hong K, Wakayama T, Zhang Y: A role for the elongator complex in zygotic paternal genome demethylation. *Nature* 463(7280), 554–558 (2010).

▪ **First implication of transcription elongator components in active paternal demethylation.**

39 Carlson LL, Page AW, Bestor TH: Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev.* 6(12B), 2536–2541 (1992).

40 Cardoso MC, Leonhardt H: DNA methyltransferase is actively retained in the cytoplasm during early development. *J. Cell Biol.* 147(1), 25–32 (1999).

41 Mayer W, Fundele R, Haaf T: Spatial separation of parental genomes during mouse interspecific (*Mus musculus* x *M. spretus*) spermiogenesis. *Chromosome Res.* 8(6), 555–558 (2000).

42 Kwon GS, Viotti M, Hadjantonakis AK: The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Dev. Cell* 15(4), 509–520 (2008).

43 Ng RK, Dean W, Dawson C *et al.*: Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nat. Cell. Biol.* 10(11), 1280–1290 (2008).

▪▪ **Demonstrates that DNA methylation is critical for earliest cell-fate commitment.**

44 Donnison M, Beaton A, Davey HW, Broadhurst R, L'huillier P, Pfeffer PL: Loss of the extraembryonic ectoderm in Elf5 mutants leads to defects in embryonic patterning. *Development* 132(10), 2299–2308 (2005).

45 Hirasawa R, Sasaki H: Dynamic transition of Dnmt3b expression in mouse pre- and early post-implantation embryos. *Gene Expr. Patterns* 9(1), 27–30 (2009).

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

47 Tucker KL, Beard C, Dausmann J *et al.*: Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. *Genes Dev.* 10(8), 1008–1020 (1996).

48 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).

49 Jackson M, Krassowska A, Gilbert N *et al.*: Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol. Cell. Biol.* 24(20), 8862–8871 (2004).

50 Carlone DL, Lee JH, Young SR *et al.*: Reduced genomic cytosine methylation and defective cellular differentiation in embryonic stem cells lacking CpG binding protein. *Mol. Cell. Biol.* 25(12), 4881–4891 (2005).

51 Farthing CR, Ficz G, Ng RK *et al.*: Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet.* 4(6), e1000116 (2008).

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

53 Gidekel S, Bergman Y: A unique developmental pattern of Oct-3/4 DNA methylation is controlled by a *cis*-demodification element. *J. Biol. Chem.* 277(37), 34521–34530 (2002).

54 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).

▪ **Early study identifying enrichment of non-CpG methylation in mouse ESCs.**

55 Gowher H, Jeltsch A: Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *J. Mol. Biol.* 309(5), 1201–1208 (2001).

56 Dodge JE, Ramsahoye BH, Wo ZG, Okano M, Li E: *De novo* methylation of MMLV provirus in embryonic stem cells: CpG versus non-CpG methylation. *Gene* 289(1–2), 41–48 (2002).

57 White GP, Watt PM, Holt BJ, Holt PG: Differential patterns of methylation of the IFN-γ promoter at CpG and non-CpG sites underlie differences in IFN-γ gene expression between human neonatal and adult CD45ro$^-$ T cells. *J. Immunol.* 168(6), 2820–2827 (2002).

58 Suetake I, Miyazaki J, Murakami C, Takeshima H, Tajima S: Distinct enzymatic properties of recombinant mouse DNA methyltransferases Dnmt3a and Dnmt3b. *J. Biochem.* 133(6), 737–744 (2003).

59 Barres R, Osler ME, Yan J *et al.*: Non-CpG methylation of the pgc-1α promoter through Dnmt3b controls mitochondrial density. *Cell Metab.* 10(3), 189–198 (2009).

60 Biniszkiewicz D, Gribnau J, Ramsahoye B *et al.*: Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Mol. Cell. Biol.* 22(7), 2124–2135 (2002).

61 Broske AM, Vockentanz L, Kharazi S *et al.*: DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat. Genet.* 41(11), 1207–1215 (2009).

62 Trowbridge JJ, Snow JW, Kim J, Orkin SH: DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell Stem Cell* 5(4), 442–449 (2009).

63 Takizawa T, Nakashima K, Namihira M *et al.*: DNA methylation is a critical cell-intrinsic determinant of astrocyte differentiation in the fetal brain. *Dev. Cell* 1(6), 749–758 (2001).

64 Fan G, Martinowich K, Chin MH *et al.*: DNA methylation controls the timing of astrogliogenesis through regulation of JAK-STAT signaling. *Development* 132(15), 3345–3356 (2005).

65 Sen GL, Reuter JA, Webster DE, Zhu L, Khavari PA: Dnmt1 maintains progenitor function in self-renewing somatic tissue. *Nature* 463(7280), 563–567 (2010).

66 Pittenger MF, Mackay AM, Beck SC *et al.*: Multilineage potential of adult human mesenchymal stem cells. *Science* 284(5411), 143–147 (1999).

67 Zuk PA, Zhu M, Mizuno H *et al.*: Multilineage cells from human adipose tissue: implications for cell-based therapies. *Tissue Eng.* 7(2), 211–228 (2001).

68 Boquest AC, Shahdadfar A, Fronsdal K *et al.*: Isolation and transcription profiling of purified uncultured human stromal stem cells: alteration of gene expression after *in vitro* cell culture. *Mol. Biol. Cell* 16(3), 1131–1141 (2005).

69 Noer A, Sorensen AL, Boquest AC, Collas P: Stable CpG hypomethylation of adipogenic promoters in freshly isolated, cultured, and differentiated mesenchymal stem cells from adipose tissue. *Mol. Biol. Cell* 17(8), 3543–3556 (2006).

70 Sorensen AL, Timoskainen S, West FD *et al.*: Lineage-specific promoter DNA methylation patterns segregate adult progenitor cell types. *Stem Cells Dev.* (2009) (Epub ahead of print).

71 Josephson R, Ording CJ, Liu Y *et al.*: Qualification of embryonal carcinoma 2102ep as a reference for human embryonic stem cell research. *Stem Cells* 25(2), 437–446 (2007).

72 Aranda P, Agirre X, Ballestar E *et al.*: Epigenetic signatures associated with different levels of differentiation potential in human stem cells. *PloS ONE* 4(11), e7809 (2009).

73 Attema JL, Papathanasiou P, Forsberg EC, Xu J, Smale ST, Weissman IL: Epigenetic characterization of hematopoietic stem cell differentiation using minichip and bisulfite sequencing analysis. *Proc. Natl Acad. Sci. USA* 104(30), 12371–12376 (2007).

74 Xu J, Pope SD, Jazirehi AR *et al.*: Pioneer factor interactions and unmethylated CpG dinucleotides mark silent tissue-specific enhancers in embryonic stem cells. *Proc. Natl Acad. Sci. USA* 104(30), 12377–12382 (2007).

75 Tadokoro Y, Ema H, Okano M, Li E, Nakauchi H: *De novo* DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. *J. Exp. Med.* 204(4), 715–722 (2007).

76 Temple S: Stem cell plasticity – building the brain of our dreams. *Nat. Rev. Neurosci.* 2(7), 513–520 (2001).

77 Shen Y, Chow J, Wang Z, Fan G: Abnormal CpG island methylation occurs during *in vitro* differentiation of human embryonic stem cells. *Human Mol. Genet.* 15(17), 2623–2635 (2006).

78 Hemberger M, Dean W, Reik W: Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nat. Rev.* 10(8), 526–537 (2009).

79 Hajkova P, Ancelin K, Waldmann T *et al.*: Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* 452(7189), 877–881 (2008).

80 Popp C, Dean W, Feng S *et al.*: Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 463(7284), 1101–1105 (2010).

81 Surani MA, Durcova-Hills G, Hajkova P, Hayashi K, Tee WW: Germ line, stem cells, and epigenetic reprogramming. *Cold Spring Harb. Symp. Quant. Biol.* 73, 9–15 (2008).

82 Eckhardt F, Lewin J, Cortese R *et al.*: DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* 38(12), 1378–1385 (2006).

83 Schilling E, Rehli M: Global, comparative analysis of tissue-specific promoter CpG methylation. *Genomics* 90(3), 314–323 (2007).

84 Illingworth R, Kerr A, Desousa D *et al.*: A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* 6(1), e22 (2008).

85 Rakyan VK, Down TA, Thorne NP *et al.*: An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res.* 18(9), 1518–1529 (2008).

86 Song F, Mahmood S, Ghosh S *et al.*: Tissue specific differentially methylated regions (TDMR): changes in DNA methylation during development. *Genomics* 93(2), 130–139 (2009).

87 Doi A, Park IH, Wen B *et al.*: Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* 41(12), 1350–1353 (2009).

▪ **Compares methylation signatures distinguishing ESCs from induced pluripotent stem cells.**

88 Maherali N, Sridharan R, Xie W *et al.*: Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 1(1), 55–70 (2007).

89 Mikkelsen TS, Hanna J, Zhang X *et al.*: Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454(7200), 49–55 (2008).

▪ **First study demonstrating 5-aza-cytidine treatment improves reprogramming efficiencies.**

90 Ichida JK, Blanchard J, Lam K *et al.*: A small-molecule inhibitor of TGF-β signaling replaces sox2 in reprogramming by inducing nanog. *Cell Stem Cell* 5(5), 491–503 (2009).

91 Zhu JK: Active DNA demethylation mediated by DNA glycosylases. *Annu. Rev. Genet.* 43, 143–166 (2009).

92 Ma DK, Jang MH, Guo JU *et al.*: Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* 323(5917), 1074–1077 (2009).

93 Barreto G, Schafer A, Marhold J *et al.*: Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445(7128), 671–675 (2007).

94 Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK: Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol. Chem.* 279(50), 52353–52360 (2004).

95 Jin SG, Guo C, Pfeifer GP: GADD45A does not promote DNA demethylation. *PLoS Genet.* 4(3), e1000013 (2008).

96 Metivier R, Gallais R, Tiffoche C *et al.*: Cyclical DNA methylation of a transcriptionally active promoter. *Nature* 452(7183), 45–50 (2008).

97 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 (2009).

▪ **First study demonstrating contribution of active demethylation in reprogramming.**

98 Chin MH, Mason MJ, Xie W *et al.*: Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5(1), 111–123 (2009).

99 Bibikova M, Chudin E, Wu B *et al.*: Human embryonic stem cells have a unique epigenetic signature. *Genome Res.* 16(9), 1075–1083 (2006).

100 Allegrucci C, Wu YZ, Thurston A *et al.*: Restriction landmark genome scanning identifies culture-induced DNA methylation instability in the human embryonic stem cell epigenome. *Human Mol. Genet.* 16(10), 1253–1268 (2007).

101 Maitra A, Arking DE, Shivapurkar N *et al.*: Genomic alterations in cultured human embryonic stem cells. *Nat. Genet.* 37(10), 1099–1103 (2005).

102 Rugg-Gunn PJ, Ferguson-Smith AC, Pedersen RA: Epigenetic status of human embryonic stem cells. *Nat. Genet.* 37(6), 585–587 (2005).

103 Aasen T, Raya A, Barrero MJ *et al.*: Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat. Biotechnol.* 26(11), 1276–1284 (2008).

104 Choi SC, Yoon J, Shim WJ, Ro YM, Lim DS: 5-azacytidine induces cardiac differentiation of p19 embryonic stem cells. *Exp. Mol. Med.* 36(6), 515–523 (2004).

105 Huangfu D, Osafune K, Maehr R *et al.*: Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat. Biotechnol.* 26(11), 1269–1275 (2008).

106 Huangfu D, Maehr R, Guo W *et al.*: Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* 26(7), 795–797 (2008).

107 Shi Y, Desponts C, Do JT, Hahm HS, Scholer HR, Ding S: Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 3(5), 568–574 (2008).

108 Shi Y, Do JT, Desponts C, Hahm HS, Scholer HR, Ding S: A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* 2(6), 525–528 (2008).

109 Kohyama J, Abe H, Shimazaki T *et al.*: Brain from bone: efficient "meta-differentiation" of marrow stroma-derived mature osteoblasts to neurons with noggin or a demethylating agent. *Differentiation* 68(4–5), 235–244 (2001).

110 Woodbury D, Reynolds K, Black IB: Adult bone marrow stromal stem cells express germline, ectodermal, endodermal, and mesodermal genes prior to neurogenesis. *J. Neurosci. Res.* 69(6), 908–917 (2002).

111 Bibikova M, Laurent LC, Ren B, Loring JF, Fan JB: Unraveling epigenetic regulation in embryonic stem cells. *Cell Stem Cell* 2(2), 123–134 (2008).

112 Ramirez-Carrozzi VR, Braas D, Bhatt DM *et al.*: A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* 138(1), 114–128 (2009).

113 Pietersen AM, Van Lohuizen M: Stem cell regulation by polycomb repressors: postponing commitment. *Curr. Opin. Cell Biol.* 20(2), 201–207 (2008).

114 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).

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

116 Loenarz C, Schofield CJ: Expanding chemical biology of 2-oxoglutarate oxygenases. *Nature Chem. Biol.* 4(3), 152–156 (2008).