Cell
P R E S S

DNA methylation and cellular reprogramming

Daniel D. De Carvalho[*], Jueng Soo You[*] and Peter A. Jones

Departments of Urology, Biochemistry and Molecular Biology, University of Southern California Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

The recent discovery that a small number of defined factors are sufficient to reprogram somatic cells into pluripotent stem cells has significantly expanded our knowledge of the plasticity of the epigenome. In this review we discuss some aspects of cell fate plasticity and epigenetic alterations, with emphasis on DNA methylation during cellular reprogramming. Recent data suggest that DNA methylation is a major barrier to induced pluripotent stem (iPS) cell reprogramming. The demethylating agent 5-azacytidine can enhance the efficiency of iPS cells generation and the putative DNA demethylase protein activation-induced cytidine deaminase (AID/AICDA) can erase DNA methylation at pluripotency gene promoters, thereby allowing cellular reprogramming. Elucidation of the epigenetic changes taking place during cellular reprogramming will enhance our understanding of stem cell biology and facilitate therapeutic applications.

Introduction
Mammalian development is a continuous process whereby cells gradually become more specialized and restrict their developmental potential [1]. This begins after fertilization, with formation of the only totipotent cell, the zygote. Through a series of divisions, increasingly committed cells are generated, and these give rise to specialized differentiated cells. In mammals, this process is thought to be unidirectional, and the epigenetic machinery is responsible for this developmental loss of plasticity [2]. The epigenetic machinery is composed of molecules that regulate epigenetic modifications, such as DNA methylation and histone modifications, which are involved in heritable gene expression patterns [3].

Although cellular differentiation is usually unidirectional *in vivo*, this process can be reprogrammed and reverted *in vitro*. Several seminal experiments in cell-fusion systems showed that genes silenced during differentiation could be reprogrammed and re-expressed in mammalian cells [4,5]. Over 20 years ago Helen Blau and colleagues generated stable heterokaryons by fusing terminally differentiated human fibroblasts with terminally differentiated mouse muscle cells and detected synthesis of human muscle proteins in the heterokaryons [4]. This showed that human fibroblasts retain some plasticity and, in the presence of the proper factors, differentiated cells from one lineage can express genes from another lineage

(Figure 1a). These experiments suggest that, in mammalian cells, the differentiation process can be reprogrammed. These initial studies did not measure expression of mouse fibroblast genes in the heterokaryons, making it difficult to reach conclusions about the directionality of this process. Subsequent experiments fusing human cancer (HeLa) cells with mouse muscle cells did not exhibit expression of human muscle proteins in heterokaryons [5] (Figure 1a), suggesting that cancer cells have reduced epigenetic plasticity. This reduced plasticity could be explained by acquisition of a more stable repressive mark, and when HeLa cells were treated with the DNA demethylating agent 5-azacytidine (5-Aza-CR) prior to fusion, heterokaryons did express human muscle-specific genes [5] (Figure 1a). These data suggest that DNA methylation (Box 1) is a stable epigenetic mark and a limiting factor for cellular plasticity.

Since these initial experiments, studies have shown that terminally differentiated somatic cells can be reprogrammed using defined factors to generate induced pluripotent stem (iPS) cells [6]. A significant effort is underway to uncover the regulatory mechanisms involved; it is becoming increasingly evident that epigenetic mechanisms play a role. In this review we discuss advances in cellular reprogramming and regulation by the epigenetic machinery.

Transcription factor-mediated reprogramming
The initial mammalian reprogramming experiments raised the question of whether it is possible to reprogram differentiated cells using defined factors, instead of the plethora involved in the cell-fusion approach. Later, transduction of a single transcription factor, *MyoD*, was shown to be sufficient to reprogram fibroblast and adipoblast cell lines into myoblasts [7]. Other groups confirmed that lineage conversion was possible by expressing one or few transcription factors in terminally differentiated cells. *C/EBPα* overexpression allowed lineage conversion from B cells to macrophages [8], and introduction of *Ngn3*, *Pdx1* and *MafA* was sufficient to convert differentiated pancreatic exocrine cells into β-cells [9]. These studies suggest that 'master' genes can drive cell reprogramming. Although lineage conversion was a major step forward, the goal of reprogramming experiments has been to generate pluripotent cells from differentiated cells. This goal was first achieved using nuclear transfer, where the nucleus of a differentiated cell is fused with an enucleated oocyte [10], and demonstrated that, in the presence of myriad factors from the oocyte, nuclei of differentiated cells have sufficient plasticity to generate a viable embryo and all terminally differentiated cell types.

Corresponding author: Jones, P.A. (pjones@med.usc.edu)
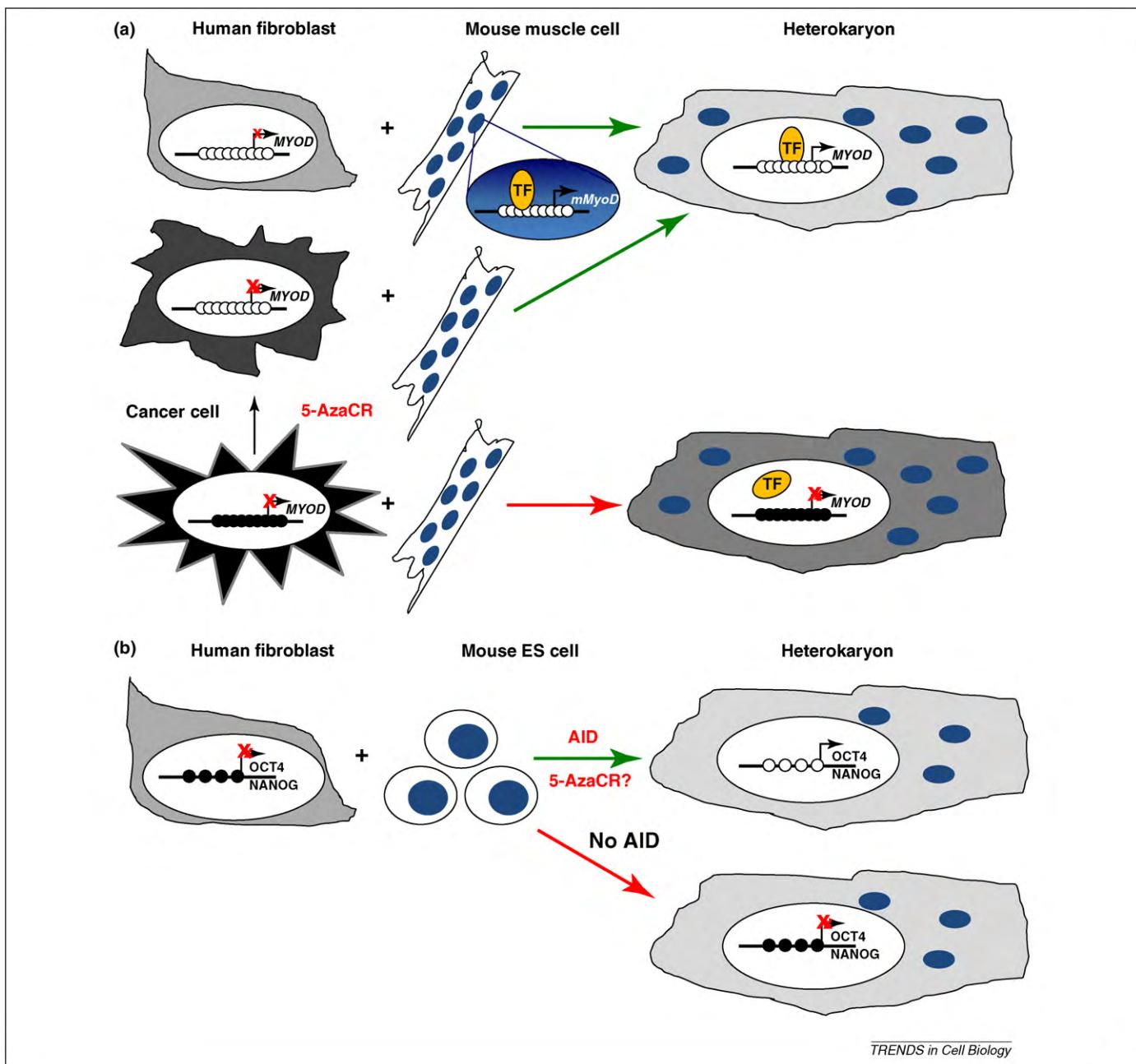[*] These authors contributed equally to this work.

Figure 1. Cell-fusion experiments showing the importance of DNA methylation as an epigenetic barrier to reprogramming. **(a)** Human fibroblasts can be fused with mouse muscle cells to generate heterokaryons; the plethora of muscle transcription factors can induce the expression of human muscle genes that were repressed in fibroblasts. On the other hand, the fusion of HeLa human cancer cells with murine muscle cells cannot induce human muscle genes unless treated with the demethylating drug 5-azacytidine prior to cell fusion [4,5]. TF, transcription factor. **(b)** Human fibroblasts can also be fused with murine ES cells to generate heterokaryons; factors present in the ES cells can reprogram the human *Oct4* and *Nanog* promoters, causing DNA demethylation and gene expression. On the other hand, this reprogramming cannot occur when the mouse putative DNA demethylase, AID, is silenced by RNAi [59].

An exciting breakthrough came with the discovery that four defined transcription factors, *Oct4*, *Sox2*, *Klf4* and *cMyc*, can generate pluripotent stem cells from fibroblasts [6]. These pluripotent cells were called induced pluripotent stem cells (iPSCs). These cells express embryonic stem (ES) cell markers, such as *E-Ras*, *Cripto*, *Dax1*, *Zfp296* and *Fgf4*, exhibit ES cell morphology and growth properties, differentiate into tissues from the three germ layers, and can contribute to embryonic development [6]. The first generation of iPSCs did not generate viable chimeras when injected into blastocysts, and the endogenous *Oct4* and *Nanog* promoters were still methylated [6]. Subsequent

experiments showed that fully reprogrammed iPSCs can be generated by the same four pluripotent factors (*Oct4*, *Sox2*, *Klf4* and *cMyc*) followed by selection for endogenous expression of *Oct4* and *Nanog* [11,12]. Fully reprogrammed iPSCs were unmethylated at their *Oct4* and *Nanog* promoters; these cells generated viable chimeras and could contribute to the germline [11,12].

More recently it has been shown that cell-cycle control checkpoints, mediated by the INK4a-ARF locus and p53, limit reprogramming, and disruption of these pathways enhances iPSC generation [13–17]. More specialized cells are reprogrammed less efficiently than progenitor cells, at

least in the hematopoietic system [18]. Increased iPSC generation efficiency is seen after treating cells with butyrate, vitamin C or exposure to a hypoxic environment [19–21].

Due to their ability to generate virtually any cell type, and potential for use in personalized transplantation therapy, iPSCs hold great promise in regenerative medicine. Proof-of-principle studies established that iPSCs can be used therapeutically. A genetically engineered mouse model of human sickle-cell anemia was successfully treated with hematopoietic progenitor cells derived from autologous iPSCs in which the genetic alteration had been corrected [22]. Similarly, *in vitro* differentiation of iPSCs into dopaminergic neurons followed by transplantation into the brains of adult rats restored dopaminergic function and improved behavioral symptoms in a model of Parkinson's disease [23]. Genetic disorders can complicate clinical use, however, because the genetic alteration must be repaired in the original donor cell before iPS reprogramming. Fibroblasts obtained from patients with Fanconi anemia could not be reprogrammed to generate iPSCs, probably because of genetic instability and predisposition to apoptosis [24]. Once the genetic defect was corrected, by inducing *FANCA* or *FANCD2* expression, fibroblasts from patients with Fanconi anemia generated fully reprogrammed iPSCs [24].

In view of the potential of iPSCs, significant attention has been paid to reprogramming technology (Box 2) and a range of protocols have been employed. Currently, iPSCs can be derived from somatic cells of all three germ layers (including fibroblasts, B lymphocytes, stomach cells, liver cells, intestinal epithelium, neural stem and progenitor cells, and keratinocytes) using combinations of transcription factors (*Oct4*, *Sox2*, *Klf4*, *cMyc*, *Nanog* and *Lin28*) and diverse delivery systems (retroviruses and lentiviruses, adenovirus transduction, plasmid transfection, episomal vectors and protein transduction) [11,12,25–41]. The goal of reprogramming strategies is to generate cells for clinical use. However, a significant impediment is the tumorigenic

potential of iPSCs. Approximately 20% of chimeric mice generated from iPSCs using the original four transcription factors (*Oct4*, *Sox2*, *Klf4* and *cMyc*) develop tumors [12]. Without *cMyc*, iPSCs are obtained and chimeric mice do not develop tumors, but iPS reprogramming efficiency is reduced [42]. Another strategy to minimize the use of tumorigenic transcription factors is to replace them with appropriate small molecules during reprogramming. Treating cells with the histone-deacetylase inhibitor valproic acid compensates for the absence of *cMyc* [43], and treatment with the G9a histone methyltransferase inhibitor BIX-01294 allows reprogramming of somatic cells using only *Oct4* and *Klf4* [40]. In addition, small molecules acting on signaling pathways can enhance reprogramming efficiency. β-Catenin accumulation after treating ES cells with

the glycogen synthase kinase-3 (GSK-3) inhibitor 6-bromoindirubin-3′-oxime enhances the reprogramming of somatic cells after cell fusion [44]. In a small molecule screen, a TGF-β inhibitor, E-616452, replaced *Sox2* in reprogramming and was renamed RepSox, for replacement of *Sox2* [45].

Thus, the development of iPS reprogramming technology represents a major step towards personalized regenerative medicine and improves our knowledge about cellular plasticity and epigenetic changes associated with cell fate.

DNA methylation as an epigenetic barrier to cellular reprogramming

A major challenge of reprogramming is the extremely low efficiency (0.01–0.1%) of iPSC generation [2]. Overcoming the epigenetic repression of pluripotency genes in differentiated cells is difficult, and evidence suggests that the few cells that do overcome these barriers do so due to stochastic events. iPSC reprogramming is a gradual process that takes several weeks, and genetically identical clones selected from somatic cells infected with pluripotency factors display heterogeneity in their ability to generate iPSCs [27]. Treating differentiated cells with drugs to revert repressive epigenetic marks can significantly increase the efficiency of iPS formation [37,40,43].

These data suggest that epigenetic modifications acquired during cell differentiation can lock a cell into a differentiated phenotype. During the artificial process of iPSC reprogramming the robust and constant expression of pluripotency factors could induce initial changes associated with reprogramming, giving rise to partially reprogrammed cells. To generate iPSCs it might then be necessary for some stochastic event to overcome the epigenetic repression of one or more key loci. Treatment of partially reprogrammed cell lines with the DNA methylation inhibitor 5-Aza-CR or reduction of *Dnmt1* expression by siRNA or shRNA induces a rapid and stable transition from partially to fully reprogrammed iPSCs [37].

DNA methylation is an extremely important epigenetic barrier to cellular reprogramming. Demethylation increases iPS reprogramming efficiency and, in hypermethylated cancer cells, allows myogenic reprogramming [5]. This raises the question whether cell types with aberrant DNA methylation at CpG island promoters, such as cancerous [46] or aging cells [47], are more resistant to cellular reprogramming. In mice, nuclei of leukemia, lymphoma, and mammary tumor cells support normal preimplantation development but fail to produce ES cells after nuclear transfer [48]. iPSCs can be generated from human esophageal, stomach, colorectal, liver, pancreatic, and cholangiocellular cancer cells, although it is not clear whether reprogramming was complete [49]. It was recently shown that aberrant DNA methylation of *CDKN2A* induced by immortalization of human embryonic lung fibroblasts followed by extensive cell culture is reverted by iPS reprogramming. Immortalization induces DNA methylation in a set of genes normally methylated in an age-dependent manner. Interestingly, iPS reprogramming induced some DNA demethylation but could not restore methylation levels seen before immortalization [50]. It will be interesting to know whether CpG islands, which are aberrantly methylated in cancer, can be fully demethylated during iPS reprogramming.

The molecular mechanisms of DNA demethylation are largely elusive, although it has been proposed as a passive or active process (Figure 2). Passive DNA demethylation occurs when maintenance methylation is impaired during DNA replication, precluding methylation of newly synthesized DNA [51]. Active DNA demethylation depends on demethylating enzymes and can occur independent of replication [51]. The low efficiency and lengthy time necessary for creating iPSCs are consistent with a passive DNA demethylation model. Transcription factor binding to regulatory elements of pluripotency genes might inhibit binding of DNA methyltransferases and prevent methylation of newly synthesized DNA. Nuclear transfer and cell-fusion experiments, which are not accompanied by cell division, support an active and specifically targeted DNA demethylation model, although the DNA demethylases involved are not known [52].

To date, AID (activation-induced cytidine deaminase, also known as AICDA) is the only enzyme that has been
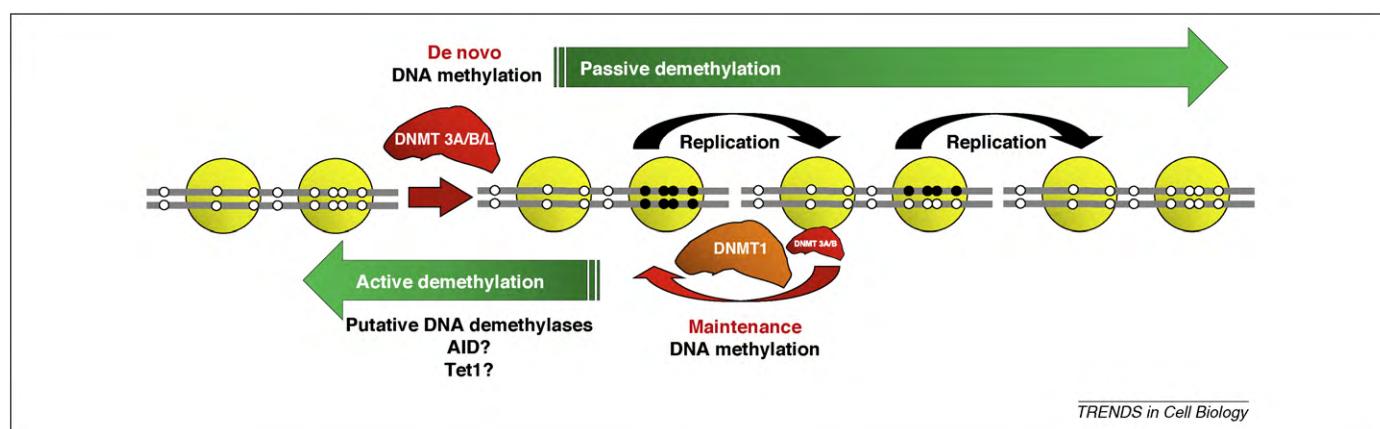
Figure 2. DNA methylation and demethylation. *De novo* DNA methylation patterns are established by the catalytically active methyltransferases DNMT3A and DNMT3B. This process is enhanced in the presence of the catalytically inactive DNMT3L. During replication, the original DNA methylation pattern is maintained largely by the activity of DNMT1, with some participation by DNMT3A and DNMT3B. DNA methylation patterns can be erased by active or passive DNA demethylation. Active demethylation can take place by enzymatic replacement of a methylated cytosine with an unmethylated residue without the requirement of cell division. Several enzymes, such as AID and Tet1, are proposed to play roles as DNA demethylases. On the other hand, passive demethylation can take place during successive replications if the activity of maintenance DNA methylation is abrogated. Black circles represent methylated CpG sites and open circles represent unmethylated CpG sites.

suggested as a DNA demethylase in mammalian reprogramming [53]. This enzyme – a cytidine deaminase – deaminates C to U in the DNA of immunoglobulin genes causing somatic mutations or class-switch recombination depending on the DNA repair pathway [54]. Although AID is expressed mainly in B cells [54], it is also expressed in pluripotent cells including oocytes, embryonic germ cells and ES cells [55]. AID has been implicated in global DNA demethylation in zebrafish after fertilization [56] and in mouse primordial germ cells (PGC) [53]. It is suggested that AID-mediated DNA demethylation occurs due to deamination of methylated cytidine residues in single-stranded DNA, followed by DNA repair. Thus, a methylated cytosine is replaced by an unmethylated one [56]. In the absence of DNA repair, C to T mutation can occur [55]. DNA methylation in PGC from *AID*-null mice is considerably lower than in ES cells, and *AID* knockout mice lack developmental defects, suggesting that other unidentified factors, possibly working through different mechanisms, are also important for demethylation. It was recently shown that XRCC1, a component of the base excision repair (BER) machinery, is associated with DNA demethylation in PGC [57].

AID is implicated in the tumorigenic process, principally in B-cell tumors [58], possibly by inducing chromosomal translocations and mutations in tumor suppressor genes and oncogenes [58]. It was proposed that a high-fidelity repair system in normal B cells might limit the tumorigenic potential of AID and that this system fails during tumorigenesis [58].

Recently, Blau and colleagues used cell fusion to identify new epigenetic regulators [59]. This approach allows rapid, extensive and bidirectional reprogramming. The proportion of each cell type used in the fusion, with the cell in excess being dominant, dictates directionality [60]. The authors fused a higher proportion of mouse ES cells with human fibroblasts and studied reprogramming in the presence of all ES cell factors (Figure 1b). Following cell fusion, rapid increases in human *Oct4* and *Nanog* expression were detected, accompanied by DNA demethylation at their promoters [59]. DNA demethylation and increased *Oct4/Nanog* expression was not observed when *AID* expression was reduced by siRNA [59]. These data highlight the putative DNA demethylating activity of AID and the importance of active DNA demethylation in reprogramming. It would be interesting to know whether treatment of human fibroblasts with demethylating drugs prior to fusion with mouse ES cells allows *Oct4/Nanog* expression even in the presence of *AID* siRNA (Figure 1b). AID appears to induce DNA demethylation at specific target regions rather than globally, although the mechanism regulating this specificity is not clear.

AID is highly expressed in oocytes [55], and DNA demethylation of specific regions of donated nuclei is necessary for functional reprogramming by nuclear transfer [52]. Generating human ES cell lines using nuclear transfer is significantly more efficient than generating iPSCs (4–16% [61] versus 0.01–0.1% [2], respectively). It is tempting to speculate that AID in the oocyte cytoplasm during nuclear transfer demethylates key pluripotency genes in the donated nuclei. If so, *AID* knockdown in oocytes should reduce nuclear transfer efficiency.

DNA methylation is clearly a key and stable mechanism in specifying cell identity. One possibility is that AID could be used to enhance iPS efficiency by removing DNA methylation at pluripotency genes. However, additional experiments are necessary to know whether AID can be used without introducing mutations or translocations [62] and thereby enhancing the tumorigenic potential of iPSCs. DNA demethylation appears necessary in normal development to erase epigenetic memory and allow PGC to acquire fitness for totipotency [63] Even so, iPS reprogramming cannot completely reset genomic methylation. Low passage iPSCs harbor residual DNA methylation patterns from their tissue of origin [64]. This suggests that a better understanding of the epigenetic reprogramming process during normal development will improve artificial reprogramming during iPSC generation.

Another putative DNA demethylase, Tet1, was recently shown to revert DNA methylation through oxidative demethylation [65], the mechanism used by JmjC enzymes to demethylate histone substrates [66,67]. Although all three Tet family members (Tet1, Tet2 and Tet3) convert 5mC to 5hmC, only Tet1 regulates ES-cell self-renewal [68]. Tet1 is necessary to maintain *Nanog* expression in ES cells; *Tet1* knockdown represses *Nanog* expression, and this correlated with increased DNA methylation at its promoter [68]. Conversely, during ES cell differentiation, *Tet1* is repressed and associated with DNA methylation and repression of the *Nanog* promoter [68]. It will be exciting to discover when Tet1 becomes expressed during iPS reprogramming and whether it plays a role in this process. If Tet1 is only expressed when iPSCs are fully reprogrammed it might be involved in iPS self-renewal, whereas expression in partially reprogrammed iPSCs might suggest participation in active demethylation of pluripotency genes.

Epigenome reprogramming
During differentiation, the epigenomic landscape undergoes dramatic alterations that are responsible for maintaining stem cell pluripotency, generating tissue-specific expression profiles and 'locking' differentiated cells into specific cell types (Figure 3) [1]. During iPSC reprogramming this landscape is modified to allow differentiated cells to reacquire 'stemness' potential (Figure 3) [1]. Although iPSCs generate the same range of cell types as ES cells, additional studies were necessary to establish their similarity at the transcriptome and epigenome levels [25,37,69–71]. Fully reprogrammed iPSCs are very similar, although not identical, to ES cells in regard to their gene expression profiles [37,70,71], with high expression of pluripotency and self-renewing genes (*Oct4*, *Nanog*, *Sox2*, *Lin28*, *Zic3*, *Fgf4*, *Tdgf1* and *Rex1*) and low expression of lineage-specifying transcription factors and developmental genes [37]. Even so, iPSCs retain an expression signature distinct from ES cells [70]. iPSCs generated from different tissues using different methods have similar gene expression patterns that distinguish them from ES cells [70]. As iPSCs are passaged their gene expression signatures become closer to that of ES cells [70]. Interestingly, genetically identical mouse ES and iPSCs have indistinguishable gene expression patterns, except for the imprinted *Dlk1–Dio3* gene cluster [71].
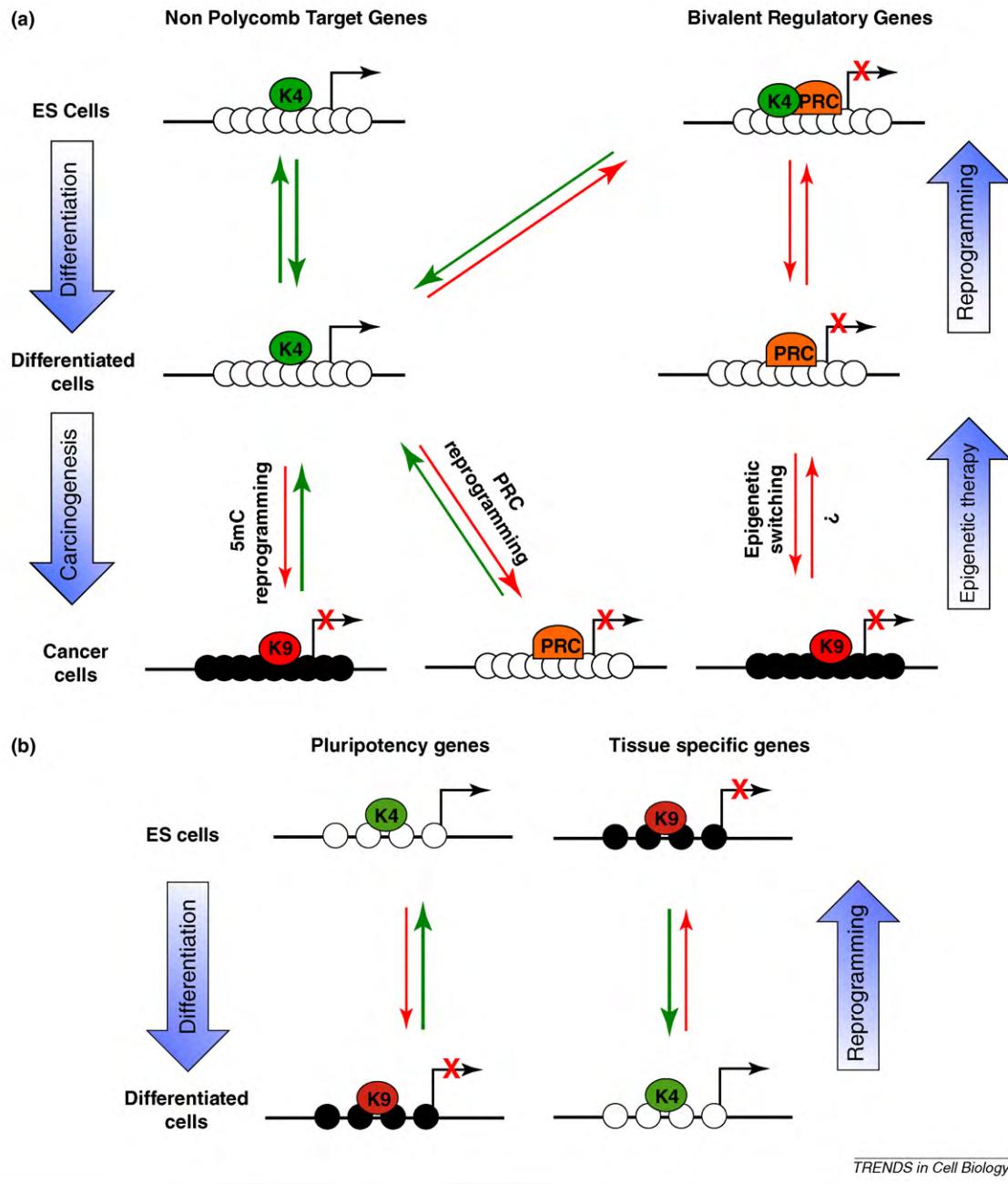
Figure 3. Epigenome landscaping during differentiation and carcinogenesis. **(a)** CpG-rich regions. In embryonic stem cells most CpG-rich regions within promoters are marked by H3K4me3 and can be subdivided into two groups. The first comprises genes that are not Polycomb targets and are usually expressed in ES cells. The second comprises Polycomb target genes. This group adopts a bivalent chromatin structure and is usually repressed in ES cells. During cell differentiation the bivalent chromatin structure becomes monovalent and some genes retain only the H3K4me3 mark (K4), whereas others retain only the H3K27me3 mark (PRC). Interestingly, during reprogramming towards pluripotency, most of these genes reacquire the bivalent structure [72,73]. During carcinogenesis, a group of expressed genes become *de novo* repressed, either by DNA methylation (5mC reprogramming) or by histone methylation (PRC reprogramming). In addition, another group of genes already repressed by Polycomb gains a more stable repression by DNA methylation (epigenetic switching) [84]. The gain of DNA methylation is usually accompanied by H3K9 methylation (K9). Black circles are methylated CpG sites and white circles are unmethylated. **(b)** CpG-poor regions. In ES cells, pluripotency genes located in non-CGI are marked by H3K4me3, the CpG sites are unmethylated, and the genes are expressed. During cell differentiation these genes gain DNA methylation and the repressive H3K9me3 mark, and are not expressed. Tissue-specific genes are silenced by DNA methylation in ES cells, and during differentiation these genes are expressed in some tissues, but remain silenced in others. During reprogramming this group can revert their chromatin structure, and this becomes very similar to that of ES cells [69]. Black circles are methylated CpG sites and white circles are unmethylated.

The similarities and differences between ES and iPS cell gene expression patterns could be explained by their chromatin states [25,69,70]. In ES cells almost all CpG rich promoters are enriched for the activation-associated H3K4 tri-methylation (H3K4me3) mark and a subset of these are associated with the repressive H3K27me3 mark, in a so-called 'bivalent' state [72,73]. During differentiation, most of these 'bivalent' domains lose one of the two marks, becoming 'monovalent' [73] (Figure 3a). After complete iPS reprogramming, approximately 80% of these promoters reacquire the 'bivalent' marks [37] (Figure 3a). Furthermore, 95% of promoters of key developmental transcription factors reacquire the 'bivalent' marks [37]. Another study showed that 957 genes present distinguishable histone

methylation signatures between ES cells and fibroblasts; 94.4% of these genes had identical histone methylation patterns between ES and iPSCs [25].

DNA methylation patterns in fully reprogrammed iPSCs also are similar but not identical to ES cells [69]. Genome-wide DNA methylation analysis was used to identify markers capable of distinguishing between differentiated and pluripotent (ES and iPS) cells [69]. This revealed that the promoters of *Oct4*, *Nanog* and *Dnmt3b* are differentially methylated; they are unmethylated in ES cells, methylated in differentiated cells and partially methylated in iPSCs [69] (Figure 3b). Moreover, global methylation analyses using a methylation-sensitive restriction enzyme assay showed that female iPS and ES cells have similar hypomethylation patterns, and random X chromosome inactivation displays the same dynamics in both cell lines [25].

Another genome-wide analysis found regions of DNA hypo- and hypermethylation in iPSC lines compared to parental fibroblasts [74], suggesting that full reprogramming requires extensive changes in DNA methylation. Partially reprogrammed iPSCs have gene expression and epigenome signatures intermediate between fully reprogrammed iPSCs and differentiated cell lines derived from them [37]. These partially reprogrammed iPSCs can reactivate genes related to stem cell self-renewal and maintenance, but not pluripotency genes. Lineage-specific transcription factors are incompletely repressed and they have incomplete epigenetic remodeling, including persistent DNA hypermethylation on pluripotency promoters [37]. These data indicate that fully reprogrammed iPSCs have a similar, although not identical, chromatin structure to ES cells. More work is necessary to understand the fundamental mechanisms that regulate differentiation, reprogramming and pluripotency.

Concluding remarks

In the time since the first cell-fusion experiments, great improvements have been achieved in our understanding of epigenetic mechanisms underlying cellular plasticity. However, many questions remain. What events are responsible for overcoming DNA methylation of pluripotency genes during reprogramming? Does AID or another putative DNA demethylase play a role? What are the requirements for DNA demethylation during iPSC reprogramming? Although non-CpG island genes such as *Oct4* and *Nanog* can be demethylated, it is unclear what happens at methylated CpG-rich regions. Can CpG islands be demethylated during iPS reprogramming? In cancer cells several CpG islands become *de-novo* methylated (Figure 3a), raising the question of whether cancer cells are more difficult to reprogram.

To date, three approaches have been used to induce reprogramming. Cell fusion can rapidly, easily and efficiently induce reactivation of tissue-specific genes or pluripotency factors by generating non-proliferating mixed-species heterokaryons. This is a powerful system for understanding the molecular mechanisms of nuclear plasticity, but it is not suitable for clinical use. Nuclear transfer, on the other hand, generates ES cells useful in studying early developmental biology and in therapeutic applications, although ethical concerns exist. Finally, iPSC tech-

nology yields pluripotent cells for clinical use without ethical concerns, and can be used to model human diseases and screen potential new treatments [75].

Understanding the molecular mechanisms by which terminally differentiated cells are reprogrammed to generate pluripotent cells will shed light on developmental epigenetic changes, the stability of epigenetic marks, and how they influence the epigenome. Identifying the molecules and events that dynamically change the epigenetic landscape will allow improvement of epigenetic therapies (Box 2).

References

1 Amabile, G. and Meissner, A. (2009) Induced pluripotent stem cells: current progress and potential for regenerative medicine. *Trends Mol. Med.* 15, 59–68

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

3 Editorial (2008) Moving AHEAD with an international human epigenome project. *Nature* 454, 711–715

4 Chiu, C.P. and Blau, H.M. (1984) Reprogramming cell differentiation in the absence of DNA synthesis. *Cell* 37, 879–887

5 Chiu, C.P. and Blau, H.M. (1985) 5-Azacytidine permits gene activation in a previously noninducible cell type. *Cell* 40, 417–424

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

7 Davis, R.L. *et al.* (1987) Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51, 987–1000

8 Xie, H. *et al.* (2004) Stepwise reprogramming of B cells into macrophages. *Cell* 117, 663–676

9 Zhou, Q. *et al.* (2008) *In vivo* reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature* 455, 627–632

10 Wilmut, I. *et al.* (1997) Viable offspring derived from fetal and adult mammalian cells. *Nature* 385, 810–813

11 Wernig, M. *et al.* (2007) *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448, 318–324

12 Okita, K. *et al.* (2007) Generation of germline-competent induced pluripotent stem cells. *Nature* 448, 313–317

13 Hong, H. *et al.* (2009) Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature* 460, 1132–1135

14 Kawamura, T. *et al.* (2009) Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* 460, 1140–1144

15 Utikal, J. *et al.* (2009) Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* 460, 1145–1148

16 Marion, R.M. *et al.* (2009) A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* 460, 1149–1153

17 Li, H. *et al.* (2009) The *Ink4/Arf* locus is a barrier for iPS cell reprogramming. *Nature* 460, 1136–1139

18 Eminli, S. *et al.* (2009) Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nat. Genet.* 41, 968–976

19 Mali, P. *et al.* (2010) Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. *Stem Cells* 28, 713–720

20 Esteban, M.A. *et al.* (2010) Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* 6, 71–79

21 Yoshida, Y. *et al.* (2009) Hypoxia enhances the generation of induced pluripotent stem cells. *Cell Stem Cell* 5, 237–241

22 Hanna, J. *et al.* (2007) Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* 318, 1920–1923

23 Wernig, M. *et al.* (2008) Neurons derived from reprogrammed fibroblasts functionally integrate into the fetal brain and improve symptoms of rats with Parkinson's disease. *Proc. Natl. Acad. Sci. U. S. A.* 105, 5856–5861

24 Raya, A. *et al.* (2009) Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature* 460, 53–59

25 Maherali, N. *et al.* (2007) Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 1, 55–70

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

27 Meissner, A. *et al.* (2007) Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nat. Biotechnol.* 25, 1177–1181

28 Qin, D. *et al.* (2007) Direct generation of ES-like cells from unmodified mouse embryonic fibroblasts by Oct4/Sox2/Myc/Klf4. *Cell Res.* 17, 959–962

29 Stadtfeld, M. *et al.* (2008) Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2, 230–240

30 Carey, B.W. *et al.* (2009) Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc. Natl. Acad. Sci. U. S. A.* 106, 157–162

31 Aoi, T. *et al.* (2008) Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* 321, 699–702

32 Stadtfeld, M. *et al.* (2008) Induced pluripotent stem cells generated without viral integration. *Science* 322, 945–949

33 Kaji, K. *et al.* (2009) Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* 458, 771–775

34 Okita, K. *et al.* (2008) Generation of mouse induced pluripotent stem cells without viral vectors. *Science* 322, 949–953

35 Zhou, H. *et al.* (2009) Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* 4, 381–384

36 Hanna, J. *et al.* (2008) Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* 133, 250–264

37 Mikkelsen, T.S. *et al.* (2008) Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454, 49–55

38 Kim, J.B. *et al.* (2008) Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* 454, 646–650

39 Eminli, S. *et al.* (2008) Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells* 26, 2467–2474

40 Shi, Y. *et al.* (2008) A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* 2, 525–528

41 Kim, J.B. *et al.* (2009) Oct4-induced pluripotency in adult neural stem cells. *Cell* 136, 411–419

42 Nakagawa, M. *et al.* (2008) Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* 26, 101–106

43 Huangfu, D. *et al.* (2008) Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* 26, 795–797

44 Lluis, F. *et al.* (2008) Periodic activation of Wnt/beta-catenin signaling enhances somatic cell reprogramming mediated by cell fusion. *Cell Stem Cell* 3, 493–507

45 Ichida, J.K. *et al.* (2009) A small-molecule inhibitor of tgf-Beta signaling replaces sox2 in reprogramming by inducing nanog. *Cell Stem Cell* 5, 491–503

46 Jones, P.A. and Baylin, S.B. (2007) The epigenomics of cancer. *Cell* 128, 683–692

47 Teschendorff, A.E. *et al.* (2010) Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res.* 20, 440–446

48 Hochedlinger, K. *et al.* (2004) Reprogramming of a melanoma genome by nuclear transplantation. *Genes Dev.* 18, 1875–1885

49 Miyoshi, N. *et al.* (2010) Defined factors induce reprogramming of gastrointestinal cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* 107, 40–45

50 Ron-Bigger, S. *et al.* (2010) Aberrant epigenetic silencing of tumor suppressor genes is reversed by direct reprogramming. *Stem Cells* 28, 1349–1354

51 Zhu, J.K. (2009) Active DNA demethylation mediated by DNA glycosylases. *Annu. Rev. Genet.* 43, 143–166

52 Simonsson, S. and Gurdon, J. (2004) DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. *Nat. Cell Biol.* 6, 984–990

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

54 Delker, R.K. *et al.* (2009) A coming-of-age story: activation-induced cytidine deaminase turns 10. *Nat. Immunol.* 10, 1147–1153

55 Morgan, H.D. *et al.* (2004) Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol. Chem.* 279, 52353–52360

56 Rai, K. *et al.* (2008) DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* 135, 1201–1212

57 Hajkova, P. *et al.* (2010) Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* 329, 78–82

58 Liu, M. and Schatz, D.G. (2009) Balancing AID and DNA repair during somatic hypermutation. *Trends Immunol.* 30, 173–181

59 Bhutani, N. *et al.* (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* 463, 1042–1047

60 Palermo, A. *et al.* (2009) Nuclear reprogramming in heterokaryons is rapid, extensive, and bidirectional. *FASEB J.* 23, 1431–1440

61 Mombaerts, P. (2003) Therapeutic cloning in the mouse. *Proc. Natl. Acad. Sci. U. S. A.* 100 (Suppl 1), 11924–11925

62 Tsai, A.G. *et al.* (2008) Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. *Cell* 135, 1130–1142

63 Hayashi, K. and Surani, M.A. (2009) Resetting the epigenome beyond pluripotency in the germline. *Cell Stem Cell* 4, 493–498

64 Kim, K. *et al.* (2010) Epigenetic memory in induced pluripotent stem cells. *Nature* DOI: 10.1038/nature09342

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

66 Tsukada, Y. *et al.* (2006) Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439, 811–816

67 Klose, R.J. *et al.* (2006) JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* 7, 715–727

68 Ito, S. *et al.* (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466, 1129–1133

69 Deng, J. *et al.* (2009) Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat. Biotechnol.* 27, 353–360

70 Chin, M.H. *et al.* (2009) Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5, 111–123

71 Stadtfeld, M. *et al.* (2010) Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 465, 175–181

72 Bernstein, B.E. *et al.* (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326

73 Mikkelsen, T.S. *et al.* (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553–560

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

75 Yamanaka, S. and Blau, H.M. (2010) Nuclear reprogramming to a pluripotent state by three approaches. *Nature* 465, 704–712

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

77 Epsztejn-Litman, S. *et al.* (2008) *De novo* DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* 15, 1176–1183

78 Feldman, N. *et al.* (2006) G9a-mediated irreversible epigenetic inactivation of *Oct-3/4* during early embryogenesis. *Nat. Cell Biol.* 8, 188–194

79 Jones, P.A. and Liang, G. (2009) Rethinking how DNA methylation patterns are maintained. *Nat. Rev. Genet.* 10, 805–811

80 Shen, L. *et al.* (2007) Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. *PLoS Genet.* 3, 2023–2036

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

82 Laird, P.W. (2010) Principles and challenges of genome-wide DNA methylation analysis. *Nat. Rev. Genet.* 11, 191–203

83 Sun, N. *et al.* (2010) Human iPS cell-based therapy: considerations before clinical applications. *Cell Cycle* 9, 880–885

84 Gal-Yam, E.N. *et al.* (2008) Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. *Proc. Natl. Acad. Sci. U. S. A.* 105, 12979–12984