Introduction

The generation of mammalian cells commences with a unicellular totipotent zygote, which gives rise to the blastocyst. Within the blastocyst, cells from the inner cell mass (ICM) contribute to the development of the whole organism. Embryonic stem (ES) cells are obtained by in vitro culture from the ICM cells, and are able to generate cells of all the three embryonic germ layers, confirming their pluripotent state (1). In contrast to the cells of the ICM, ES cells have the capacity to proliferate indefinitely in this pluripotent state, and harbour immense therapeutic potential. They could potentially provide an unlimited source of cells to replace those that are lost or impaired as a result of disease. Additionally, pluripotent cells can also be obtained by reprogramming somatic cells through three different techniques: nuclear transfer, cell fusion and direct nuclear reprogramming by enforced expression of transcription factors. From these three techniques, direct nuclear reprogramming, from here on referred to as ‘reprogramming’, has gained major interest since its initial discovery in 2006 by Yamanaka and colleagues (2).

Reprogramming is the culmination of several seminal discoveries made over the last 50 years. The first landmark finding was the pioneering work conducted by John Gurdon and his colleagues, which showed that tadpoles could be generated from unfertilised eggs by nuclear transfer from intestinal cells of adult frogs (3). These studies demonstrated that the genetic information is not lost during development and somatic cells have the capacity to re-enter a pluripotent state. The second milestone was the establishment of the fact that transcription factors play a significant role in cell fate and identity. For instance, expression of MyoD, a transcription factor in myocytes, can convert fibroblasts into myocytes (4). Third, the establishment of ES cell culture conditions and the characterisation of factors that ensure the maintenance of pluripotency. The reprogramming protocol established by Takahashi and Yamanaka revealed that the ectopic expression of the four pluripotent associated transcription factors Oct4, Sox2, Klf4 and c-Myc (OKSM) in mouse fibroblasts was sufficient to trigger their transition into pluripotent cells (2). Importantly, the transition of a differentiated cell into an induced pluripotent stem cell (iPSC) upon ectopic expression of the reprogramming factors is accompanied by changes in the epigenome, which is reset to an ES cell-like state (5).

Epigenetic changes are associated with alterations in gene regulation without altering the DNA sequence, and these changes are related to chromatin modifications. Chromatin is the combination of DNA and proteins found in the nucleus of eukaryotic cells. The nucleosome is the basic structural unit of chromatin, consisting of histone proteins around which the DNA is organised (Fig. 1). Gene activity is ultimately regulated by the state of the chromatin, which is the result of a combination of nucleosome dynamics and positioning, as well as DNA methylation, histone variants and histone modifications. This encompasses the epigenetic regulation of gene transcription. This review will focus on the different epigenetic changes underlying the reprogramming process and the influence of chromatin modifiers involved in this process.

Epigenetic Changes During Reprogramming

DNA Methylation

DNA methylation occurs primarily in the cytosine of the dinucleotide motif CpG (cytosine residues preceding a guanine) by DNA methyltransferases. The CpG dinucleotides are distributed across the genome and are often associated with gene promoter regions. DNA methylation is usually associated with transcriptional inactivation and gene silencing. Indeed, reprogramming of somatic cells into iPSCs is accompanied by significant changes in DNA methylation patterns. Analysis of DNA methylation patterns in the CpG and non-CpG regions of somatic cells and ES cells, revealed that cytosine methylation in the non-CpG region was significantly lower in somatic cells (6), whereas CpG islands that were methylated in somatic cells, were demethylated in iPSCs (7). A detailed analysis of the reprogramming process of fibroblasts to iPSCs revealed that different loci become demethylated and methylated predominantly at the end of reprogramming, when cells commit to pluripotency (8). These genome-wide DNA methylation changes are accompanied by transcriptional upregulation of enzymes associated with DNA demethylation (Apobec2 and Tet1) and DNA methylation (Dnmt3a and Dnmt3L) (8). Interestingly, iPSCs retain an ‘epigenetic memory’, a residual DNA methylation pattern that links iPSCs to their donor cells and can be reset by continued cell culture or using chromatin modulators (9,10).

Histones Variants and Modifications During Reprogramming

Histone Variants

H2AZ, a variant of the nucleosome core histone H2A, is one of the chromatin regulators enriched in the promoter of developmentally prominent genes and has an essential role in ES cell differentiation (11). H2A can be replaced by H2AZ by the ATP-dependent chromatin remodelling complex SWR1 (12). H2AZ and the Polycomb group complex (PcG) occupy the promoters of pluripotency-associated genes, and deletion of H2AZ leads to a rise in the pluripotent state of the cells (11). Histone H3 variants
Epigenetic Changes During Reprogramming

Histone Modifications

Different post-translational modifications of histones have a notable effect on transcription. Three of the most common post-translational histone modifications occurring in chromatin are methylation, acetylation and phosphorylation (14).

Several lysines in the so-called histone tails are the main target for post-translational modification. Methylation of histones occurs by lysine methyltransferases that can modify one single lysine in a single histone which can then lead to activation or repression of transcription depending on the histone and the site. Methylation of histone H3 lysine 4 (H3K4), histone H3 lysine 36 (H3K36) and histone H3 lysine 79 (H3K79) is associated with transcriptional activation; whereas methylation of H3K9, H3K27, H4K20 is related to repression of transcription (15). Demethylation of lysines by histone demethylase LSD1 in H3K4, for example, represses transcription whereas demethylation of H3K9 by demethylase LSD1 and a complex of androgen receptors, activates transcription (16). In ES cells, promoters with CpG rich regions at housekeeping genes are usually enriched in H3K4me3 and are generally active, whereas developmentally associated genes are enriched for both H3K4me3 and H3K27me3, so called bivalent domains, and are usually silent (17). As mentioned above, during reprogramming, the epigenome is reset and becomes similar to that of ES cells (5). The methylation pattern of H3K27 and H3K4 tri-methylation, for example, is similar in ES and iPSCs, whereas fibroblasts have a different histone methylation pattern (Fig. 1) (5). Moreover, investigation of epigenetic changes in intermediate cell populations during reprogramming has shown that changes in the histone methylation pattern of H3K4 and H3K27 follow two waves (8). The genes change their H3K4me3 and H3K27me3 status first in an early stage, followed by a period of minor changes and concludes again with a second wave of changes in histone modification (8). Chromatin changes at enhancers and promoters in the early days of reprogramming revealed that H3K4me2 appears in promoters which are associated with pluripotent genes in the late stages of reprogramming (18).

Histone acetylation is a covalent modification and is primarily associated with transcriptional activation. Three main families of acetyltransferases are involved in this modification: GNAT, MYST, and CBP/P300 (14). Repressive chromatin remodelling factors such as PcG and the nucleosome remodelling and deacetylase complex (NuRD) induce histone deacetylation and trimethylation of H3K27 (H3K27me3) in promoters and lead to suppression of the differentiation pathways in pluripotent cells (19).

Nucleosome Shifting and Chromatin Remodelling

Nucleosome dynamics and positioning are able to modulate DNA accessibility and therefore affect the binding of transcription factors to their target sites (20). Transcription factors can recruit different chromatin regulators to promoter regions to modify and mobilise nucleosomes, in order to increase their access to DNA sequences. It has for instance been shown that the presence of nucleosomes in the Oct4 binding sites in
the pluripotency associated gene *Nanog* prevent Oct4 accessibility to the target in differentiated cells and consequently leads to gene repression even if the DNA is demethylated (21). Recent work has shown that Oct4, Sox2 and Klf4 function as pioneer factors at enhancers of genes that promote reprogramming (22). The ability of these pioneer transcription factors to bind to compacted/closed chromatin enables them to induce a cascade of changes in chromatin, which occurs during the later stages of reprogramming.

Nucleosome positions can be changed by chromatin remodelers and transcription factors (23). The effect of transcription factors is related to their competition with the nucleosome to access their target (23). This competition depends on the affinity of nucleosomes and transcription factors for the DNA sequence and also depends on their concentrations (24). Repositioning of nucleosomes by transcription factors occurs with the assistance of ATP-dependent chromatin remodelling factors. ATP-dependent chromatin remodelling complexes affect the interaction between DNA and histones and sequentially allows accessibility of transcription factors to their targets. The BAF complex, for example, is involved in regulating pluripotency gene expression and facilitates the binding of Oct4 to its target promoters (25). The addition of Brg1 and Baf155, which are components of the BAF complex, to the reprogramming cocktail (OKSM) increases demethylation of the promoters of Oct4, Nanog, and Rex1 which is followed by an increase in reprogramming efficiency (25).

**Targeting the Epigenetic Machinery to Increase Reprogramming Efficiency**

Generation of iPSCs from somatic cells is theoretically simple but reprogramming efficiency is low and varies from 0.01% to 5% depending on the reprogramming system and the cell of origin. However, reprogramming efficiency can be significantly increased by targeting the epigenetic machinery.

DNA methyltransferase inhibitors such as 5-aza-cytidine (AZA) have a positive effect on reprogramming. It has been shown that AZA leads to reactivation of endogenous Oct4 expression in partially reprogrammed cells, and enhances their ability to form ES cell-like colonies (17,26). Histone deacetylase inhibitors such as valproic acid (VPA), trichostatin A (TSA), and suberoylanilide hydroxamic acid (SAHA) increased the reprogramming efficiencies in mouse and human fibroblasts with the transcription factors Oct4, Sox2 and Klf4 (27). Moreover, targeting of the NuRD complex member of the methyl-CpG binding domain family of proteins, Mbd3, by short-hairpin RNA-mediated knockdown has a positive effect on the reprogramming of somatic cells (28). Depletion of Mbd3 increases reversion of epiblast stem cells, primordial germ cells and somatic cells such as fibroblast and haematopoietic and neural cells to a pluripotent state with 100% efficiency (29).

**Conclusion**

iPSCs are pluripotent stem cells that are derived from the ectopic expression of the transcription factors Oct4, Klf4, Sox2 and c-Myc in somatic cells (2). Reprogramming, the process whereby a somatic cell transitions into an iPSC, is the result of the reconfiguration of the somatic cell epigenome to an ES cell-like state (5). Yet there are apparent differences in the genome-wide DNA methylation and histone modification profiles of iPS and ES cells, even amongst different iPSC lines (7,9,10). These few but significant epigenetic differences, that are in part due to an ‘epigenetic memory’, are responsible for the variations in differentiation potential observed between different iPSC lines (9,10). As mentioned previously, cell passaging or chromatin modifying compounds can erase this epigenetic memory (9,10) and/or increase reprogramming efficiency. In the future, it is conceivable that factor-based reprogramming methods could be further optimised by supplementing the reprogramming cocktail with such chromatin modifiers. This could facilitate the derivation of iPSCs with an unbiased differentiation potential. On the other hand, the phenomena of epigenetic memory in donor somatic cells may well lead to the creation of iPSC cells with a desired differentiation bias into a specific lineage or tissue type. Furthermore, if the epigenome reset occurs in two waves, partially reprogrammed intermediate cell populations could be utilised which are biased towards a desired lineage in order to circumvent the full reprogramming and differentiation process. It is evident that the epigenome is at the heart of the reprogramming process, thus, only by further understanding its role in cell fate can the full potential of the iPSC and the regenerative medicine field be unleashed.

**References**