Regenerative Therapy 28 (2025) 246–252

Contents lists available at ScienceDirect

Regenerative Therapy

journal homepage: http://www.elsevier.com/locate/reth

Review

Interplay between pioneer transcription factors and epigenetic modifiers in cell reprogramming

Gerardo Mirizio [a, b, 1], Samuel Sampson [a, b, 1], Makiko Iwafuchi [a, b, *]

[a] Division of Developmental Biology, Center for Stem Cell & Organoid Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, USA
[b] Department of Pediatrics, College of Medicine, University of Cincinnati, OH, 45229, USA

ARTICLE INFO

Article history:
Received 31 October 2024
Received in revised form
5 December 2024
Accepted 20 December 2024

Keywords:
Reprogramming
Differentiation
Pioneer transcription factor
PRC
NuRD
DNA methylation

ABSTRACT

The generation of induced pluripotent stem cells (iPSCs) from differentiated somatic cells by Yamanaka factors, including pioneer transcription factors (TFs), has greatly reshaped our traditional understanding of cell plasticity and demonstrated the remarkable potential of pioneer TFs. In addition to iPSC reprogramming, pioneer TFs are pivotal in direct reprogramming or transdifferentiation where somatic cells are converted into different cell types without passing through a pluripotent state. Pioneer TFs initiate a reprogramming process through chromatin opening, thereby establishing competence for new gene regulatory programs. The action of pioneer TFs is both influenced by and exerts influence on epigenetic regulation. Despite significant advances, many direct reprogramming processes remain inefficient, which limits their reliability for clinical applications. In this review, we discuss the molecular mechanisms underlying pioneer TF-driven reprogramming, with a focus on their interactions with epigenetic modifiers, including Polycomb repressive complexes (PRCs), nucleosome remodeling and deacetylase (NuRD) complexes, and the DNA methylation machinery. A deeper understanding of the dynamic interplay between pioneer TFs and epigenetic modifiers will be essential for advancing reprogramming technologies and unlocking their full clinical potential.

Contents

1. Introduction

The ability to reprogram somatic cells into iPSCs through the forced expression of the Yamanaka factors, OCT4, SOX2, KLF4, and MYC (OSKM), has transformed the fields of cell and developmental biology [1,2]. This breakthrough not only revealed the plasticity of differentiated cells but also demonstrated the remarkable capacity

* Corresponding author. Division of Developmental Biology, Center for Stem Cell & Organoid Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, USA.
E-mail address: makiko.iwafuchi@cchmc.org (M. Iwafuchi).

Peer review under responsibility of the Japanese Society for Regenerative Medicine.

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.reth.2024.12.014

of transcription factors (TFs) to reset cell fate. Notably, reprogramming is not limited to iPSC generation. Direct reprogramming, or transdifferentiation, where one somatic cell type is converted into another without transitioning through a pluripotent intermediate, has emerged as a powerful strategy for generating functional cell types for clinical applications. The core of both iPSC generation and direct reprogramming are pioneer TFs. In the context of iPSC generation, three of the Yamanaka factors, OCT4, SOX2, and KLF4, are classified as pioneer TFs [3,4]. Pioneer TFs are characterized by their unique ability to bind to inaccessible, nucleosome-occupied DNA, thereby facilitating local chromatin opening and recruiting additional transcription factors and co-regulators [5]. This pioneering action enables them to establish competence for new gene regulatory programs, demonstrating their versatility in manipulating cell fate. The functions of pioneer TFs are closely linked to epigenetic regulation, which governs chromatin structure and accessibility. Epigenetic modifiers, such as the Polycomb repressive complexes (PRCs), nucleosome remodeling and deacetylase (NuRD) complexes, and DNA methylation machinery, play critical roles in maintaining or modifying the chromatin landscape during reprogramming. Pioneer TFs collaborate with or counteract these epigenetic modifiers to unlock chromatin and drive new cell fates (Fig. 1).

In this review, we focus on the fundamental roles of pioneer TFs and epigenetic regulation in cellular reprogramming and explore their dynamic interactions. We discuss how they reshape the chromatin landscape during the reprogramming processes, the current limitations of our knowledge, and potential strategies to enhance reprogramming efficiency and fidelity. By gaining a more comprehensive understanding of how pioneer TFs and epigenetic modifiers cooperate or compete, we aim to highlight new avenues for optimizing reprogramming technologies, ultimately advancing their clinical applications.

2. Pioneer transcription factors in cell reprogramming

Pioneer TFs have gained special attention for their ability to overcome chromatin barriers and initiate new gene regulatory programs in cell differentiation, induced pluripotency, and direct reprogramming. These specialized factors represent a distinct subset of TFs that possess the unique ability to access and bind to DNA sequences embedded within closed chromatin [5]. Unlike conventional TFs, which rely on pre-existing chromatin accessibility for binding, pioneer TFs can directly interact with nucleosomal DNA and induce local chromatin opening. This priming function facilitates the recruitment of additional TFs and epigenetic modifiers to initiate new gene regulatory programs.

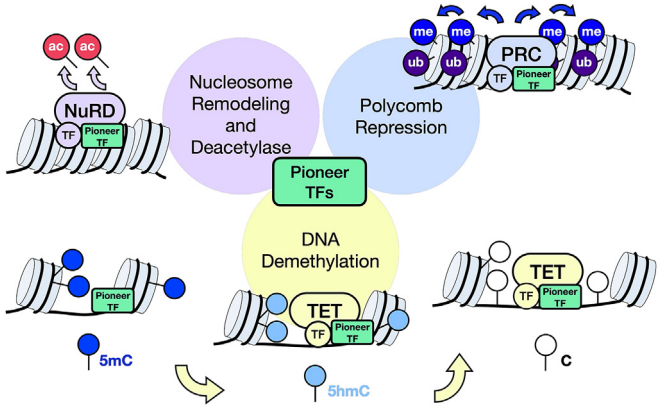

Fig. 1. Interplay between pioneer transcription factors (TFs) and epigenetic modifiers: the Polycomb Repressive Complex (PRC), the Nucleosome Remodeling and Deacetylase (NuRD) complex, and the Ten-Eleven Translocation (TET) methylcytosine dioxygenase. Pioneer TFs can recruit specific epigenetic modifiers through cooperation with distinct partner TFs. For example, the pioneer TF FOXA cooperates with PRDM1 TF to recruit PRC1.

The generation of iPSCs by Takahashi and Yamanaka in 2006 revolutionized the field of developmental biology and regenerative medicine [1]. The reprogramming of somatic cells into iPSCs was achieved by forced expression of four TFs – OCT4, SOX2, KLF4, and MYC (OSKM) – collectively known as the "Yamanaka factors." Notably, three of these factors (OCT4, SOX2, and KLF4) are recognized as pioneer TFs due to their ability to initiate the chromatin changes necessary for reprogramming somatic cells to a pluripotent state [3,4]. These pioneer TFs bind to both pluripotency genes and non-pluripotency-related genes early in the reprogramming process [3,6], possibly reflecting a stochastic nature to the initial stages of iPSC reprogramming [7,8]. At the onset of reprogramming, OSK TFs bind to active somatic gene enhancers and initiate their silencing by redirecting somatic TFs away from these enhancers [6]. Moreover, these pioneer TFs work synergistically to overcome chromatin barriers present in somatic cells and reactivate a pluripotency program, initially through their own actions and later by supporting other pluripotency TFs [3,6].

In direct reprogramming, one somatic cell type is converted directly into another through the forced expression of lineage-specific TFs. In various contexts, combinations of TFs, often involving at least one pioneer TF, drive these transitions [5,9–11]. For instance, the pioneer TF Gata4, together with Mef2c and Tbx5, reprograms fibroblasts into cardiomyocyte-like cells (iCMs) [12]; the pioneer TF Ascl1, in combination with Brn2 (Pou3f2) and Myt1l, reprograms fibroblasts into functional neurons (iNs) [13]; and the pioneer TFs Foxa and Gata4, along with Hnf4α/1α, reprogram fibroblasts into hepatocyte-like cells (iHeps) [14,15]. In these conversions, pioneer TFs act as "first responders" by accessing silent genes within closed chromatin, which are specific to the desired target cell type. For example, during neuronal reprogramming, the pioneer TF Ascl1 binds to silent chromatin and facilitates subsequent binding of Brn2, which is involved in the later stages of cell maturation [16]. Similarly, in hepatic reprogramming, Foxa functions as a pioneer factor by binding closed chromatin and enabling co-factors like Hnf4a to drive hepatocyte-specific gene expression [17]. These hierarchical actions of TFs somewhat recapitulate the normal developmental trajectory, underscoring the importance of understanding normal cell differentiation mechanisms. Interestingly, unlike the initial stochastic period followed by a deterministic phase observed in iPSC reprogramming [7,8], successful direct reprogramming appears to be determined early in the direct conversion process [18–21].

Traditionally, pioneer TFs were thought to control cell fate primarily by promoting epigenetic and gene activation. However, accumulating evidence reveals that pioneer TFs also play a critical role in repressing alternative-lineage programs through cooperation with epigenetic repressors [22]. To enhance the efficiency and fidelity of reprogramming processes, it is essential to fully harness the ability of pioneer TFs to reshape both the active and repressive epigenetic landscapes. We further discuss the interplay between pioneer TFs and epigenetic modifiers in the following sections.

3. Polycomb repression in cell reprogramming

The success of cellular reprogramming depends on a finely orchestrated process of activation and repression of specific transcriptional programs. Initially, this process was believed to rely solely on the combinatorial activity of defined TFs at regulatory elements. However, the low efficiency and reversibility of cell fate

conversions have revealed the presence of an epigenetic barrier to pioneer TF-mediated reprogramming [3,23,24]. Erasing the epigenetic signature of the somatic cell and establishing a new landscape during the reprogramming process is, therefore, a prerequisite for stabilizing a transcriptional network specific to the desired cell type [25]. However, our limited understanding of the interplay between epigenetic modifiers and pioneer TFs during cell reprogramming hinders the development of strategies to improve the efficiency and fidelity of this process.

One of the most evolutionarily conserved epigenetic regulatory systems includes the Polycomb group (PcG) and Trithorax group (TrxG) proteins. In mammals, these proteins form multi-subunit complexes that bind to tissue-specific enhancers and hypomethylated CpG islands at promoters [26]. The two primary PcG complexes, known as Polycomb Repressive Complexes 1 and 2 (PRC1 and PRC2), catalyze the deposition of H2AK119ub1 and H3K27me2/3, respectively, to establish repressive chromatin domains. In contrast, the SWI/SNF and COMPASS family of TrxG complexes counteract PcG-mediated repression by ATP-dependent chromatin remodeling and histone-modifying activities such as H3K4me1/2/3 [27]. Interestingly, in both embryonic and adult stem cells, a large number of genes involved in cell-fate control are simultaneously marked by the repression-associated H3K27me3 and activation-associated H3K4me1/2/3 modifications [28,29]. These “bivalent” regulatory elements resolve into either active or repressive states during differentiation, and their functional relevance is related to the dynamic regulation of tissue-specific gene expression [30]. However, the mechanisms that instruct the dynamic assembly and resolution of bivalent epigenetic states during development have not yet been fully elucidated.

The critical role of PcG complexes in cell reprogramming was first demonstrated through heterokaryon assays, in which differentiated human cells were reprogrammed upon fusion with mouse embryonic stem cells (ESCs). In that study, ESCs deficient in PRC1 or PRC2 catalytic subunits failed to induce the expression of most pluripotency-associated genes following fusion with human lymphocytes [31]. Similarly, shRNA-mediated inhibition of PRC1 (BMI1, RING1) and PRC2 (EZH2, EED, SUZ12) subunits significantly decreased the iPSC reprogramming efficiency, highlighting the importance of PcG complexes in silencing somatic cell gene expression during reprogramming [32]. Furthermore, PRC2 subunits JARID2, MTF2, and esPRC2p48, which are highly expressed in mouse ESC compared to differentiated cells, have been shown to cooperatively silence tissue-specific genes through H3K27 methylation, thereby enhancing OSK-mediated iPSC reprogramming [33]. It has also been shown that H3K27 methylation-mediated silencing appears to be critical specifically for a core subset of target genes during iPSC reprogramming. In this context, the depletion of *Ezh2* at the onset of OSKM-induced reprogramming caused a global reduction of H3K27me3; however, iPSC generation still proceeded [34]. Notably, *Ezh2*-deficient iPSCs retained H3K27me3 marks on specific Polycomb target genes associated with lineage specification, and removing H3K27me3 from these targets significantly impaired iPSC generation [34]. Recent reports have identified the role of additional PcG-associated proteins such as Cbx4/7, Bmi1, Kdm2b, Pcgf6, and RYBP in cell reprogramming [7,23,35–37]. However, the vast functional diversity of PcG subunits and the specificity of reprogramming experiments make it difficult to establish general mechanisms of action. Furthermore, the role of PRCs in direct reprogramming is still largely underexplored, and how PRCs achieve target specificity remains a critical unanswered question.

A novel mechanism of PRC recruitment to cell-type-specific target sites has been recently revealed in the context of human pluripotency and endoderm differentiation, unexpectedly mediated by pioneer TFs. Specifically, the pioneer TFs FOXA and OCT4 cooperate with members of the PRDM family TFs PRDM1 and PRDM14, respectively, to recruit PRC1 to developmentally regulated enhancers [38]. This recruitment leads to the establishment of bivalent enhancers marked by H3K4me1 and H2Aub/H3K27me3, likely through the subsequent recruitment of PRC2. This gain in bivalency prevents the precocious activation of lineage-specific genes in iPSC and restricts alternative-lineage gene expression during endoderm differentiation. Similarly, the PRC variant PRC1.3 and PRDM14 have been shown to jointly repress a set of developmentally regulated genes and promote primed to naïve cell reprogramming [39]. While the mechanisms underlying PRC1.3 recruitment by PRDM14 remain unclear, evidence from the aforementioned study suggests that the pioneer TF OCT4 may play a role in this process [38].

Another major question in the field is how Polycomb complexes are coordinately or hierarchically recruited to promoter-enhancer pairs. Studies using the well-characterized enhancer-promoter pair of the *MYOD1* gene have demonstrated that forced expression of OCT4 in fibroblasts promoted its binding to a nucleosome-depleted region (NDR) at the *MYOD1* enhancer. This binding event converted the *MYOD1* promoter into a bivalent state, resembling its epigenetic state in ESCs [40]. Given that a substantial proportion of Polycomb targets are associated with NDRs at tissue-specific putative enhancers, Polycomb complexes appear to be preferentially recruited to NDRs established by pioneer TF at these permissive enhancers. Altogether, these findings suggest the critical role of pioneer TFs in recruiting PRCs to restrict cell fate.

Despite these strides in the field, critical questions remain unresolved: How do pioneer TF-PRDM-PRC axes contribute to iPSC generation and direct reprogramming? Are there tissue-specific combinations of pioneer and PRDM TFs that are crucial for lineage restriction beyond pluripotent and endoderm cells? How are Polycomb complexes dynamically recruited to enhancer-promoter pairs during reprogramming? Addressing these knowledge gaps will be instrumental in improving the efficiency and fidelity of cellular reprogramming.

4. Nucleosome remodeling and deacetylase complex in cell reprogramming

The Nucleosome Remodeling and Deacetylase (NuRD) complex is a multifunctional epigenetic regulator that plays a critical role in removing permissive epigenetic marks and increasing nucleosome density [41–44]. During ESC differentiation, NuRD, together with LSD1 (KDM1A), decommissions enhancers of pluripotency-related genes, which is essential for fully silencing the pluripotency program and facilitating differentiation [45]. NuRD complex is comprised of two functionally distinct subcomplexes: a histone deacetylase subcomplex containing histone deacetylase 1 or 2 (HDAC1/2), metastasis tumor-associated 1, 2, or 3 (MTA1/2/3), and retinoblastoma-binding protein 4 or 7 (RBBP4/7) proteins, and a nucleosome remodeling subcomplex containing chromatin helicase DNA binding protein 3 or 4 (CHD3/4), GATA zinc finger domain 2A or B (GATAD2A/B), and cyclin-dependent kinase 2 associated protein 1 (CDK2AP1) [42]. The methyl-CpG-binding domain proteins MBD2 and MBD3 physically bridge these subcomplexes to form mutually exclusive MBD2/NuRD and MBD3/NuRD complexes [46]. MBD2 and MBD3 recognize methylated DNA with different affinity, with MBD2 showing high and MBD3 showing low methylation selectivity [47]. Genome-wide binding studies indicate that MBD3 preferentially localizes at unmethylated CpG-rich regions, whereas MBD2 is distributed in a methylation-dependent manner [48]. Intriguingly, *Mbd2* knockout mice were viable and fertile, whereas *Mbd3*-null mice were lethal around midgestation

due to defective epiblast expansion and extraembryonic tissue development starting at E5.5 [49,50]. In mouse ESCs, the loss of *Mbd3* did not compromise cell viability but impaired the silencing of pluripotency genes and developmental lineage commitment [51]. These results suggest that MBD3/NuRD plays a unique, non-redundant role in cell fate control, distinct from MBD2/NuRD.

Mbd3/NuRD has also emerged as a critical regulator in iPSC reprogramming, acting paradoxically as both a barrier and a facilitator of reprogramming depending on context [52–56]. Genetic depletion of *Mbd3* dramatically enhanced the determinism of naïve pluripotency reprogramming in mouse and human somatic cells, achieving nearly 100 % reprogramming efficiency within 7 days [52]. Exogenous expression of *Mbd3*, but not *Mbd2,* profoundly inhibited iPSC generation from *Mbd3*-depleted fibroblasts during early reprogramming (before day 5), while *Mbd3* expression after day 5 had a diminished effect. These findings suggest that Mbd3/NuRD functions as a barrier to iPSC reprogramming in the early stages but has a limited role in maintaining pluripotency once established. Additionally, the complete deletion of *Gatad2a* (NuRD subunit) facilitated deterministic naïve iPSC reprogramming in mouse fibroblasts without compromising somatic cell proliferation [55]. This study suggests that the Gatad2a-Mbd3 axis in the NuRD complex serves as a barrier to the re-establishment of naïve pluripotency by repressing pluripotency circuits during reprogramming. On the other hand, Mbd3/NuRD facilitated iPSC reprogramming from mouse neural stem cells, as complete depletion of *Mbd3* in these cells significantly impaired reprogramming initiation, while *Mbd3* overexpression promoted reprogramming [54]. In a non-Yamanaka iPSC reprogramming method using Sall4, Jdp2, Glis1, and Esrrb TFs, NuRD is essential for efficient iPSC generation from mouse fibroblasts by closing chromatin at somatic genes during the early stages of reprogramming [56]. These context-dependent biological effects likely reflect the intricate, multifaced roles of NuRD in reprogramming.

How do Mbd3/NuRD complexes interact with reprogramming TFs during iPSC generation? NuRD complexes can be recruited by TFs at enhancers, in addition to targeting CpG-rich promoters through MBD [41,43]. OSKM TFs were co-immunoprecipitated with Mbd3 in fibroblasts undergoing iPSC reprogramming, and their induction led to increased recruitment of Mbd3/NuRD complexes to chromatin [52]. These findings suggest that direct OSKM-Mbd3/NuRD interaction restricts OSKM activity and inhibits iPSC reprogramming. In the non-Yamanaka iPSC reprogramming method, Sall4 recruits NuRD to accessible chromatin sites in fibroblasts, leading to the closing of somatic gene loci [56]. These results explain, at least in part, the context-dependent functional variability of NuRD.

Despite these findings, the detailed molecular mechanisms underlying the context-dependent roles of MBD3/NuRD in chromatin and gene regulation remain largely unexplored. Does NuRD function differently depending on co-binding TFs and pre-existing chromatin states? Are both the deacetylase and chromatin remodeling functions of NuRD required for reprogramming? Furthermore, investigating NuRD's role in direct reprogramming could provide valuable insights into its broader role in cell fate determination.

5. DNA methylation machinery in cell reprogramming

Epigenetic barriers, such as DNA methylation and repressive histone modifications, play a critical role in maintaining cellular identity by enforcing the repression of genes specific to alternative-lineage programs [57–59]. These epigenetic barriers are essential for driving cell differentiation with high fidelity and preserving the functional integrity of differentiated cells. However, in the context of cellular reprogramming, they present major obstacles to the generation of new cell fates.

DNA methylation at CpG dinucleotides is a covalent modification of a methyl group to cytosine, and it is generally associated with reduced gene expression [58,60]. This repression occurs through multiple mechanisms, including limiting TF binding to methylated DNA motifs, positioning nucleosomes, and recruiting methyl-binding proteins associated with repressive complexes [58]. DNA-methyltransferases (DNMT1/3a/3b/3l) mediate this modification using the metabolite substrate of S-Adenosyl methionine as a methyl donor [61,62]. DNMT3a/3b are involved in *de novo* methylation, while DNMT1 maintains DNA methylation patterns during cellular replication by recognizing hemimethylated DNA [61,63]. Demethylation can occur passively through DNA replication or actively via enzymatic oxidization of the methyl group [64,65]. The Ten-eleven Translocation (TET1/2/3) methylcytosine dioxygenases catalyze the oxidization of 5-methylcytosine (5 mC) to 5-hydroxymethylcytosine (5hmC) to initiate active demethylation [66]. CpG-rich regions (CpG islands) are typically found at hypomethylated promoters, while the CpG-poor regions tend to be hypermethylated [67,68]. Most CpG-rich promoters remain hypomethylated throughout development. In contrast, distal regulatory regions, such as enhancers, dynamically change in methylation status during development and reprogramming [69–72]. While the loss of either *TET* or *DNMTs* in ESCs minimally impacted pluripotency, it significantly disrupted differentiation, demonstrating the importance of methylation dynamics in lineage commitment [73,74]. For instance, the loss of *DNMT1* or *DNMT3a/b* double knockout in mouse embryos prevented differentiation, leading to embryonic lethality by midgestation [61,63]. *TET1/2/3* triple-knockout ESCs maintained pluripotency but failed to differentiate into embryoid bodies or form teratomas [73,75]. Thus, the precise regulation of DNA methylation and demethylation is essential for directing cell fate and ensuring proper lineage commitment during development.

During iPSC reprogramming, DNA methylation patterns must be restructured to establish a pluripotent state. Improper reprogramming of methylation impairs iPSC generation and limits differentiation potential [23]. The mesenchymal-to-epithelial transition is a critical early step in both iPSC reprogramming and the direct reprogramming of induced hepatocyte-like cells [76,77]. In mouse embryonic fibroblasts (MEFs), *TET1/2/3* knockout caused hypermethylation at genomic loci of microRNAs that suppress TFs promoting epithelial-to-mesenchymal transition, resulting in reprogramming failure [78]. Following OSKM induction, iPSC reprogramming initiates with genome-wide changes in H3K4me2 modifications at enhancers and promoters of pluripotency and developmental genes that are hypomethylated in MEF [79]. In contrast, pluripotency-related enhancers with higher DNA methylation levels in MEF show no increase in H3K4me2 marks. A critical final step in iPSC reprogramming is the demethylation and reactivation of these enhancers [23,80,81]. In this process, TET enzymes are critical for facilitating transcriptional activation of key pluripotency genes [82–85]. iPSCs derived from somatic cells (e.g., fibroblast, neuronal, and blood cells) often retain the DNA methylation profile of their origin, known as “epigenetic memory,” which biases their differentiation toward the original lineage [86]. For instance, iPSCs reprogrammed from hematopoietic cells retain aspects of the original methylation profile, enhancing differentiation into hematopoietic lineages. Thus, successful iPSC reprogramming relies on a comprehensive restructuring of the DNA methylation landscape, and a deeper understanding of this process could enable the complete resetting of cellular memory of the original cell type.

How is DNA demethylation machinery directed to specific loci? Interestingly, certain TFs, such as pioneer TFs, can access and bind

to methylated CpG dinucleotides and initiate DNA demethylation [87–90]. For instance, during pancreatic differentiation, FOXA2 physically interacted with TET1, and loss of TET1 led to a shift from beta-to alpha-cell identity, caused by hypermethylation at the *PAX4* enhancer [91]. Additionally, FOXA1 interacts with DNA repair factors at target sites, which facilitates lineage-specific, active DNA demethylation [92]. Furthermore, FOXA plays a role in passive DNA demethylation. Ectopic expression of FOXA2 in fibroblast induced DNA demethylation at a subset of target sites in a replication-dependent manner [88]. Further research is required to elucidate whether both the active and passive demethylation activities of FOXA are critical for activating its target genes. Understanding the mechanisms by which pioneer TFs interact with DNA demethylation machinery will enable more precise control of DNA methylation landscapes, thereby enhancing the precise resetting of gene regulatory programs.

6. Future of cell reprogramming

The future of cell reprogramming is poised for significant advancement through strategic refinements in the usage of pioneer TFs and epigenetic modifiers. While current reprogramming technologies have already led to groundbreaking achievements in generating pluripotent and somatic cell types, emerging methodologies offer the potential to further enhance efficiency, fidelity, and clinical applicability.

One critical factor impacting reprogramming success is the precise stoichiometric balance of TFs within the reprogramming cocktail. The specific ratios of TFs can significantly improve reprogramming efficiency and fidelity. For instance, a higher level of Oct4 relative to Sox2 and Klf4 increased the number of iPSC colonies [93]. Similarly, in cardiac reprogramming, a higher level of Mef2c relative to Gata4 and Tbx5 dramatically enhanced iCM reprogramming efficiency [94]. Optimizing the balance of TFs also contributes to a more stable or mature cell state. Future reprogramming protocols could benefit from insights into the stoichiometry of TFs and epigenetic modifiers observed during normal development. For instance, absolute quantification of TFs and co-factors in nuclei revealed that, during erythropoiesis, co-repressors such as NuRD were far more abundant at the protein level than co-activators [95]. These insights underscore the potential of dynamic control and quantitative frameworks to advance reprogramming strategies.

Another exciting development in reprogramming involves engineering TFs to maximize functionality. Recently, a chimeric super-SOX, which combines structural elements of Sox2 and Sox17, demonstrated a remarkable ability to induce naïve pluripotency across multiple species [96]. Additionally, a specific point mutation of Sox2 that stabilizes Sox/Oct dimerization markedly enhanced the developmental potential of iPSCs [96]. Furthermore, advances in deep learning models for zinc finger design enable the engineering of TFs that can bind to specific target sequences in the genome, functioning as activators or repressors [97]. These engineered TFs represent a new frontier in reprogramming, with the potential to promote stable and reproducible reprogramming outcomes.

In vivo, direct reprogramming is emerging as an innovative approach with substantial potential for clinical applications [10]. In contrast to traditional reprogramming, *in vivo* reprogramming directly delivers reprogramming TFs into live organs to reprogram specific cells at the site of disease. This approach is particularly promising for regenerative medicine, where damaged or degenerated tissues could be reprogrammed on-site to restore functional cell populations. For instance, in a diabetic mouse model, delivering Ngn3, Pdx1, and Mafa TFs to pancreatic exocrine cells generated insulin-producing beta-cells, which improved glucose tolerance [98]. In myocardial infarction mouse models, iCM TFs reprogrammed scar-forming cardiac fibroblasts into iCMs, leading to scar size reduction and improved heart function after injury [99,100]. In liver fibrosis mouse models, delivery of the iHep TF cocktail reprogrammed hepatic myofibroblasts into hepatocyte-like cells, which mitigated liver fibrosis [101,102]. Direct reprogramming has also shown therapeutic potential for neurodegenerative diseases, where TFs such as Sox2 or NeuroD1 reprogrammed glial cells into functional neurons in mouse models of brain injury, spinal cord injury, and Alzheimer's disease [103–105].

In conclusion, the field of cell reprogramming is advancing toward more precise, efficient, and clinically applicable methods. Enhancing the capability of pioneer TFs in epigenetic and gene regulation could overcome existing limitations and unlock new therapeutic opportunities.

Declaration of competing interest

None.

Acknowledgments

This work was supported by the National Institutes of Health (P30 DK078392 and R01GM143161 to MI) and Cincinnati Children's Research Foundation (Trustee Award and Center for Pediatric Genomics Pilot Awards).

References

[1] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126:663–76.
[2] Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007;131:861–72.
[3] Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell 2012;151:994–1004.
[4] Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. Cell 2015;161:555–68.
[5] Iwafuchi-Doi M, Zaret KS. Pioneer transcription factors in cell reprogramming. Genes Dev 2014;28:2679–92.
[6] Chronis C, Fiziev P, Papp B, Butz S, Bonora G, Sabri S, et al. Cooperative binding of transcription factors orchestrates reprogramming. Cell 2017;168:442–59.
[7] Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, et al. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell 2012;150:1209–22.
[8] Hanna J, Saha K, Pando B, Van Zon J, Lengner CJ, Creyghton MP, et al. Direct cell reprogramming is a stochastic process amenable to acceleration. Nature 2009;462:595–601.
[9] Balsalobre A, Drouin J. Pioneer factors as master regulators of the epigenome and cell fate. Nat Rev Mol Cell Biol 2022;23:449–64.
[10] Wang H, Yang Y, Liu J, Qian L. Direct cell reprogramming: approaches, mechanisms and progress. Nat Rev Mol Cell Biol 2021;22:410–24.
[11] Horisawa K, Suzuki A. Direct cell-fate conversion of somatic cells: toward regenerative medicine and industries. Proc Jpn Acad Ser B Phys Biol Sci 2020;96:131–58.
[12] Ieda M, Fu J-D, Delgado-Olguin P, Vedantham V, Hayashi Y, Bruneau BG, et al. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell 2010;142:375–86.
[13] Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Südhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. Nature 2010;463:1035–41.
[14] Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, et al. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature 2011;475:386–9.
[15] Sekiya S, Suzuki A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. Nature 2011;475:390–3.
[16] Wapinski OL, Vierbuchen T, Qu K, Lee QY, Chanda S, Fuentes DR, et al. Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. Cell 2013;155:621–35.
[17] Horisawa K, Udono M, Ueno K, Ohkawa Y, Nagasaki M, Sekiya S, et al. The dynamics of transcriptional activation by hepatic reprogramming factors. Mol Cell 2020;79:660–76.

[18] Zhou Y, Liu Z, Welch JD, Gao X, Wang L, Garbutt T, et al. Single-cell transcriptomic analyses of cell fate transitions during human cardiac reprogramming. Cell Stem Cell 2019;25:149–64.
[19] Biddy BA, Kong W, Kamimoto K, Guo C, Waye SE, Sun T, et al. Single-cell mapping of lineage and identity in direct reprogramming. Nature 2018;564:219–24.
[20] Kamimoto K, Adil MT, Jindal K, Hoffmann CM, Kong W, Yang X, et al. Gene regulatory network reconfiguration in direct lineage reprogramming. Stem Cell Rep 2023;18:97–112.
[21] Jindal K, Adil MT, Yamaguchi N, Yang X, Wang HC, Kamimoto K, et al. Single-cell lineage capture across genomic modalities with CellTag-multi reveals fate-specific gene regulatory changes. Nat Biotechnol 2024;42:946–59.
[22] Katsuda T, Sussman JH, Zaret KS, Stanger BZ. The yin and yang of pioneer transcription factors: dual roles in repression and activation. Bioessays 2024;46:e2400138.
[23] Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, et al. Dissecting direct reprogramming through integrative genomic analysis. Nature 2008;454:49–55.
[24] Sridharan R, Gonzales-Cope M, Chronis C, Bonora G, McKee R, Huang C, et al. B. a Garcia, K. Plath, Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1γ in reprogramming to pluripotency. Nat Cell Biol 2013;15:872–82.
[25] Nashun B, Hill PW, Hajkova P. Reprogramming of cell fate: epigenetic memory and the erasure of memories past. EMBO J 2015;34:1296–308.
[26] Tanay A, O'Donnell AH, Damelin M, Bestor TH. Hyperconserved CpG domains underlie Polycomb-binding sites. Proc Natl Acad Sci U S A 2007;104:5521–6.
[27] Piunti A, Shilatifard A. Epigenetic balance of gene expression by Polycomb and COMPASS families. Science 2016;352:aad9780.
[28] Azuara V, Perry P, Sauer S, Spivakov M, Jørgensen HF, John RM, et al. Chromatin signatures of pluripotent cell lines. Nat Cell Biol 2006;8:532–8.
[29] Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 2006;125:315–26.
[30] Blanco E, González-Ramírez M, Alcaine-Colet A, Aranda S, Di Croce L. The bivalent genome: characterization, structure, and regulation. Trends Genet 2020;36:118–31.
[31] Pereira CF, Piccolo FM, Tsubouchi T, Sauer S, Ryan NK, Bruno L, et al. ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency. Cell Stem Cell 2010;6:547–56.
[32] Onder TT, Kara N, Cherry A, Sinha AU, Zhu N, Bernt KM, et al. Chromatin-modifying enzymes as modulators of reprogramming. Nature 2012;483:598–602.
[33] Zhang Z, Jones A, Sun CW, Li C, Chang CW, Joo HY, et al. PRC2 complexes with JARID2, MTF2, and esPRC2p48 in ES cells to modulate ES cell pluripotency and somatic cell reprograming. Stem Cell 2011;29:229–40.
[34] Fragola G, Germain PL, Laise P, Cuomo A, Blasimme A, Gross F, et al. Cell reprogramming requires silencing of a core subset of polycomb targets. PLoS Genet 2013;9.
[35] Moon JH, Heo JS, Kim JS, Jun EK, Lee JH, Kim A, et al. Reprogramming fibroblasts into induced pluripotent stem cells with Bmi1. Cell Res 2011;21:1305–15.
[36] Zdzieblo D, Li X, Lin Q, Zenke M, Illich DJ, Becker M, et al. Pcgf6, a polycomb group protein, regulates mesodermal lineage differentiation in murine ESCs and functions in IPS reprogramming. Stem Cell 2014;32:3112–25.
[37] Zhou Z, Yang X, He J, Liu J, Wu F, Yu S, et al. Kdm2b regulates somatic reprogramming through variant PRC1 complex-dependent function. Cell Rep 2017;21:2160–70.
[38] Matsui S, Granitto M, Buckley M, Ludwig K, Koigi S, Shiley J, et al. Pioneer and PRDM transcription factors coordinate bivalent epigenetic states to safeguard cell fate. Mol Cell 2024;84:476–89.
[39] Collier AJ, Bendall A, Fabian C, Malcolm AA, Tilgner K, et al. Genome-wide screening identifies Polycomb repressive complex 1.3 as an essential regulator of human naïve pluripotent cell reprogramming. Sci Adv 2022;8:eabk0013.
[40] Taberlay PC, Kelly TK, Liu CC, You JS, De Carvalho DD, Miranda TB, et al. Polycomb-repressed genes have permissive enhancers that initiate reprogramming. Cell 2011;147:1283–94.
[41] Hu G, Wade PA. NuRD and pluripotency: a complex balancing act. Cell Stem Cell 2012;10:497–503.
[42] Lai AY, Wade PA. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. Nat Rev Cancer 2011;11:588–96.
[43] Leighton G, Williams DC. The methyl-CpG-binding domain 2 and 3 proteins and formation of the nucleosome remodeling and deacetylase complex. J Mol Biol 2020;432:1624–39.
[44] Bornelöv S, Reynolds N, Xenophontos M, Gharbi S, Johnstone E, Floyd R, et al. The nucleosome remodeling and deacetylation complex modulates chromatin structure at sites of active transcription to fine-tune gene expression. Mol Cell 2018;71:56–72.
[45] Whyte WA, Bilodeau S, Orlando DA, Hoke HA, Frampton GM, Foster CT, et al. Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature 2012;482:221–5.
[46] Le Guezennec X, Vermeulen M, Brinkman AB, Hoeijmakers WAM, Cohen A, Lasonder E, et al. MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. Mol Cell Biol 2006;26:843–51.
[47] Hendrich B, Bird A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 1998;18:6538–47.
[48] Baubec T, Ivánek R, Lienert F, Schübeler D. Methylation-dependent and -independent genomic targeting principles of the mbd protein family. Cell 2013;153:480–92.
[49] Kaji K, Nichols J, Hendrich B. Mbd3, a component of the NuRD co-repressor complex, is required for development of pluripotent cells. Development 2007;134:1123–32.
[50] Hendrich B, Guy J, Ramsahoye B, Wilson VA, Bird A. Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev 2001;15:710–23.
[51] Kaji K, Caballero IM, MacLeod R, Nichols J, Wilson VA, Hendrich B. The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. Nat Cell Biol 2006;8:285–92.
[52] Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, et al. Deterministic direct reprogramming of somatic cells to pluripotency. Nature 2013;502:65–70.
[53] Luo M, Ling T, Xie W, Sun H, Zhou Y, Zhu Q, et al. NuRD blocks reprogramming of mouse somatic cells into Pluripotent stem cells. Stem Cell 2013;31:1278–86.
[54] Dos Santos RL, Tosti L, Radzisheuskaya A, Caballero IM, Kaji K, Hendrich B, et al. MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. Cell Stem Cell 2014;15:102–10.
[55] Mor N, Rais Y, Sheban D, Peles S, Aguilera-Castrejon A, Zviran A, et al. Neutralizing gatad2a-chd4-mbd3/NuRD complex facilitates deterministic induction of naive pluripotency. Cell Stem Cell 2018;23:412–25.
[56] Wang B, Li C, Ming J, Wu L, Fang S, Huang Y, et al. The NuRD complex cooperates with SALL4 to orchestrate reprogramming. Nat Commun 2023;14:2846.
[57] Barrero MJ, Boué S, Izpisúa Belmonte JC. Epigenetic mechanisms that regulate cell identity. Cell Stem Cell 2010;7:565–70.
[58] Suelves M, Carrió E, Núñez-Álvarez Y, Peinado MA. DNA methylation dynamics in cellular commitment and differentiation. Brief Funct Genomics 2016;15:443–53.
[59] Nicetto D, Donahue G, Jain T, Peng T, Sidoli S, Sheng L, et al. H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. Science 2019;363:294–7.
[60] Kim M, Costello J. DNA methylation: an epigenetic mark of cellular memory. Exp Mol Med 2017;49:e322.
[61] Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999;99:247–57.
[62] Hata K, Okano M, Lei H, Li E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development 2002;129:1983–93.
[63] Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 1992;69:915–26.
[64] Yang J, Bashkenova N, Zang R, Huang X, Wang J. The roles of TET family proteins in development and stem cells. Development 2020;147.
[65] Rasmussen KD, Helin K. Role of TET enzymes in DNA methylation, development, and cancer. Genes Dev 2016;30:733–50.
[66] He Y-F, Li B-Z, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 2011;333:1303–7.
[67] Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nat Rev Genet 2013;14:204–20.
[68] Deaton A, Bird A. CpG islands and the regulation of transcription. Genes Dev 2011;25:1010–22.
[69] Meissner A. Epigenetic modifications in pluripotent and differentiated cells. Nat Biotechnol 2010;28:1079–88.
[70] Rönnerblad M, Andersson R, Olofsson T, Douagi I, Karimi M, Lehmann S, et al. FANTOM consortium, Analysis of the DNA methylome and transcriptome in granulopoiesis reveals timed changes and dynamic enhancer methylation. Blood 2014;123:e79–89.
[71] Cui X-L, Nie J, Ku J, Dougherty U, West-Szymanski DC, Collin F, et al. A human tissue map of 5-hydroxymethylcytosines exhibits tissue specificity through gene and enhancer modulation. Nat Commun 2020;11:6161.
[72] Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. Role of Tet proteins in enhancer activity and telomere elongation. Genes Dev 2014;28:2103–19.
[73] Dawlaty MM, Breiling A, Le T, Barrasa MI, Raddatz G, Gao Q, et al. Loss of tet enzymes compromises proper differentiation of embryonic stem cells. Dev Cell 2014;29:102–11.
[74] Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J, et al. Severe global DNA hypomethylation blocks Differentiation and Induces Histone Hyperacetylation in embryonic stem cells. Mol Cell Biol 2004;24:8862–71.
[75] Verma N, Pan H, Doré LC, Shukla A, V Li Q, Pelham-Webb B, et al. TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. Nat Genet 2018;50:83–95.
[76] Li R, Liang J, Ni S, Zhou T, Qing X, Li H, et al. A mesenchymal-to-Epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell 2010;7:51–63.
[77] Lim KT, Lee SC, Gao Y, Kim K-P, Song G, An SY, et al. Small molecules facilitate single factor-mediated hepatic reprogramming. Cell Rep 2016:814–29.
[78] Hu X, Zhang L, Mao SQ, Li Z, Chen J, Zhang RR, et al. Tet and TDG mediate DNA demethylation essential for mesenchymal-to- epithelial transition in somatic cell reprogramming. Cell Stem Cell 2014;14:512–22.

[79] Koche RP, Smith ZD, Adli M, Gu H, Ku M, Gnirke A, et al. Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell 2011;8:96–105.
[80] Jaenisch R, Young R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 2008;132:567–82.
[81] Buckberry S, Liu X, Poppe D, Tan JP, Sun G, Chen J, et al. Transient naive reprogramming corrects hiPS cells functionally and epigenetically. Nature 2023;620:863–72.
[82] Doege CA, Inoue K, Yamashita T, Rhee DB, Travis S, Fujita R, et al. Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. Nature 2012;488:652–5.
[83] Costa Y, Ding J, Theunissen TW, Faiola F, Hore TA, Shliaha PV, et al. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. Nature 2013;495:370–4.
[84] Gao Y, Chen J, Li K, Wu T, Huang B, Liu W, et al. Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. Cell Stem Cell 2013;12:453–69.
[85] Sardina JL, Collombet S, Tian TV, Gómez A, Di Stefano B, Berenguer C, et al. Transcription factors drive tet2-mediated enhancer demethylation to reprogram cell fate. Cell Stem Cell 2018;23:727–41.
[86] Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, et al. Epigenetic memory in induced pluripotent stem cells. Nature 2010;467:285–90.
[87] Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell 2010;143:470–84.
[88] Donaghey J, Thakurela S, Charlton J, Chen JS, Smith ZD, Gu H, et al. Genetic determinants and epigenetic effects of pioneer-factor occupancy. Nat Genet 2018;50:250–8.
[89] Zhu H, Wang G, Qian J. Transcription factors as readers and effectors of DNA methylation. Nat Rev Genet 2016;17:551–65.
[90] Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol 2019;20:590–607.
[91] Li J, Wu X, Ke J, Lee M, Lan Q, Li J, et al. TET1 dioxygenase is required for FOXA2-associated chromatin remodeling in pancreatic beta-cell differentiation. Nat Commun 2022;13:3907.
[92] Zhang Y, Zhang D, Li Q, Liang J, Sun L, Yi X, et al. Nucleation of DNA repair factors by FOXA1 links DNA demethylation to transcriptional pioneering. Nat Genet 2016;48:1003–13.
[93] Tiemann U, Sgodda M, Warlich E, Ballmaier M, Schöler HR, Schambach A, et al. Optimal reprogramming factor stoichiometry increases colony numbers and affects molecular characteristics of murine induced pluripotent stem cells. Cytometry 2011;79 A:426–35.
[94] Wang L, Liu Z, Yin C, Asfour H, Chen O, Li Y, et al. Stoichiometry of Gata4, Mef2c, and Tbx5 influences the efficiency and quality of induced cardiac myocyte reprogramming. Circ Res 2015;116:237–44.
[95] Gillespie MA, Palii CG, Sanchez-Taltavull D, Shannon P, Longabaugh WJR, Downes DJ, et al. Absolute quantification of transcription factors reveals principles of gene regulation in erythropoiesis. Mol Cell 2020;78:960–74.
[96] MacCarthy CM, Wu G, Malik V, Menuchin-Lasowski Y, Velychko T, Keshet G, et al. Highly cooperative chimeric super-SOX induces naive pluripotency across species. Cell Stem Cell 2024;31:127–47.
[97] Ichikawa DM, Abdin O, Alerasool N, Kogenaru M, Mueller AL, Wen H, et al. A universal deep-learning model for zinc finger design enables transcription factor reprogramming. Nat Biotechnol 2023;41:1117–29.
[98] Zhou Q, Brown J, Kanarek A, Rajagopal J. D. a Melton, in vivo reprogramming of adult pancreatic exocrine cells to beta-cells. Nature 2008;455:627–32.
[99] Qian L, Huang Y, Spencer CI, Foley A, Vedantham V, Liu L, et al. In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. Nature 2012;485:593–8.
[100] Song K, Nam Y-J, Luo X, Qi X, Tan W, Huang GN, et al. Heart repair by reprogramming non-myocytes with cardiac transcription factors. Nature 2012;485:599–604.
[101] Song G, Pacher M, Balakrishnan A, Yuan Q, Tsay H-C, Yang D, et al. Direct reprogramming of hepatic myofibroblasts into hepatocytes in vivo attenuates liver fibrosis. Cell Stem Cell 2016;18:797–808.
[102] Rezvani M, Español-Suñer R, Malato Y, Dumont L, Grimm AA, Kienle E, et al. In Vivo hepatic reprogramming of myofibroblasts with AAV vectors as a therapeutic strategy for liver fibrosis. Cell Stem Cell 2016;18:809–16.
[103] Heinrich C, Bergami M, Gascón S, Lepier A, Viganò F, Dimou L, et al. Sox2-mediated conversion of NG2 glia into induced neurons in the injured adult cerebral cortex. Stem Cell Rep 2014;3:1000–14.
[104] Su Z, Niu W, Liu M-L, Zou Y, Zhang C-L. In vivo conversion of astrocytes to neurons in the injured adult spinal cord. Nat Commun 2014;5:3338.
[105] Guo Z, Zhang L, Wu Z, Chen Y, Wang F, Chen G. In vivo direct reprogramming of reactive glial cells into functional neurons after brain injury and in an Alzheimer's disease model. Cell Stem Cell 2014;14:188–202.